



# The Role of Consumer Privacy Concerns in Shaping Platform Strategy for Online Markets

Jin-Kun Yang<sup>1</sup> · Yong-Rui Duan<sup>1</sup> · Jia-Zhen Huo<sup>1</sup>

Received: 13 January 2021 / Revised: 15 November 2021 / Accepted: 15 February 2022 /  
Published online: 27 April 2022

© Operations Research Society of China, Periodicals Agency of Shanghai University, Science Press, and Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

We investigate the effects of consumer privacy concerns on the pricing and personal data collection strategy of an online platform. The online platform derives revenues from disclosing consumer information to firms. Firms compete for the information in order to enable them to price discriminate and thus derive revenues from consumer purchases. A novel aspect of our research is that we allow the online platform to sell only a subset of consumer data. We develop analytical models where consumers can/cannot protect their privacy. Our analysis yields three main conclusions. First, in the monopoly case, we find that when the consumer privacy disclosure aversion cost is relatively low, it is optimal for the platform to sell all consumer information to the firm. Second, in the duopoly case, we illustrate that when the consumer privacy disclosure aversion cost is relatively low, the platform will sell all consumer information to only one firm; when the cost is moderate, the platform will choose to sell the information of only some consumers and to only one firm; when the cost is relatively high, the platform will select only some of the consumers and sell their information to both firms. Third, it will be better for the platform to provide the information protection service for free when the privacy cost is low.

**Keywords** Price discrimination · User privacy · Consumer information · Two-sided market · Privacy disclosure aversion cost

---

This work was supported by the National Natural Science Foundation of China (Nos. 71771179, 72171176 and 72021002).

---

✉ Yong-Rui Duan  
yrduan@tongji.edu.cn

Jin-Kun Yang  
yangjinkun12@126.com

Jia-Zhen Huo  
huojiazhen@163.com

<sup>1</sup> School Economics and Management, Tongji University, Shanghai 200092, China

**Mathematics Subject Classification** 91B24 · 91B42 · 90B50 · 91A80

## 1 Introduction

In recent years, with the radical advancement in information technology, the cost of collecting, storing and analyzing massive amounts of data has been greatly reduced. In the modern information economy, online platforms routinely track the daily activities of consumers and build consumer profiles. Individuals' traits and attributes, such as a person's age, address, gender, preferences, and reservation prices are increasingly regarded as business assets that can be used to target services or offers, provide relevant advertising, or be traded with other parties. Large amounts of user data have been collected by numerous Internet firms for sale [1] and consumer information has become an important element of firm strategy [2]. Online platforms can gain revenue from selling consumer data to firms, which also enables firms to set tailored prices to different consumers.

Main examples of this trend can be found among e-commerce platforms. In the context of e-commerce, platforms usually hold a large amount of consumer data, while the data available to firms on the platform is very limited. Taking taobao.com as an example, taobao.com can obtain consumers' personal information, browsing history, transaction data, geographic location and other data information, but sellers on taobao.com can only obtain limited data such as order information during the trading process. In view of this, some e-commerce platforms disclose some consumer information to sellers on the platform in order to help them understanding consumers' preferences and catering to clientele, so as to achieve competitive advantage in the market. For example, taobao.com has launched a data analysis tool that sellers on taobao.com can pay to use the tool to improve their marketing ability. What's more, Amazon, the largest online e-commerce company in America, exploits consumer information such as browsing history and consumption records to profile users and derives revenue by disclosing that information to third-party merchants [2].

The acquisition and use of consumer information empower firms to price discriminate in their products [3–6]. In recent years, some online agencies have been accused of using big data and other techniques to actualize price discrimination. As early as 2000, an Amazon user complained that after erasing the “cookies” from the browser, the price of previously viewed DVD products had dropped from \$26.24 to \$22.74 on Amazon.com.<sup>1</sup> Coincidentally, the Wall Street Times reported that Expedia's Orbitz, an online booking website, had been playing this trick since 2012: booking a room on a Mac is more expensive than on a PC.<sup>2</sup> More recently, a consumer complained that when he bought a ticket on Ctrip, after exiting and returning to the website, the price of the ticket increased from 17,548 to 18,987 yuan while the equivalent ticket only cost 16,890 yuan on the HNA official website.<sup>3</sup> Uber eat, Tmall supermarket and Lyft

<sup>1</sup> <https://baijiahao.baidu.com/s?id=1593715378653221392&wfr=spider&for=pc>.

<sup>2</sup> <https://baijiahao.baidu.com/s?id=1647006045547598929&wfr=spider&for=pc>.

<sup>3</sup> [https://baike.baidu.com/reference/3148245/3942GuBE7BL1p5sK-xvKGRRgY0HRT1SytOZLY\\_W3-9swjXQmLWUPZUoTfLQp0S2IQ4Ph4Vr4V5Z5jzOL1V50ozFmbQ17bLmHbfBYe2yD](https://baike.baidu.com/reference/3148245/3942GuBE7BL1p5sK-xvKGRRgY0HRT1SytOZLY_W3-9swjXQmLWUPZUoTfLQp0S2IQ4Ph4Vr4V5Z5jzOL1V50ozFmbQ17bLmHbfBYe2yD).

have also been impeached for price discrimination.<sup>4</sup> As shown in these examples, big data technology has greatly enhanced the profiling and targeting capability for online platforms.

In March 2019, Beijing Consumers Association issued their survey results about using big data to punish loyalty. The results showed that 56.92% of respondents believed that they had experienced price discrimination, while others reported no such experiences.<sup>5</sup> In their practical experiment, among the 57 groups of simulated consuming experience samples conducted on 14 APPs or websites, prices of the new and old accounts in 35 groups of samples are identical. Only a few samples, such as Qunar.com and Fliggy, are suspected of price discrimination among different consumers.<sup>6</sup> As these results show, price discrimination on the Internet may exist, but it is not obvious. It seems that the platform did not place all consumers on the price discrimination list.

Price discrimination is one of the means by which operators obtain excess profits. The reason why platforms do not actualize price discrimination to all consumers is that tracking and collecting information about consumers may raise privacy concerns [7, 8], this, in turn, may make purchasing the product less appealing. On April 10, 2018, iiMedia Research, the global leading new economic data mining and analysis organization, released “2018 China Big Data ‘killing’ Internet users’ attitude behavior report”. According to the report, 77.8% of respondents claimed that the enforcement of price discrimination by service applications is unacceptable.<sup>7</sup> In addition to price variations, consumers may be concerned about the associated loss of privacy. Consumers may feel unease when they discover that their personal information such as their gender, age, address, consumption records and browsing history may be recorded. What’s more, it is difficult for consumers to confirm how the platform uses the information. Consumers may face the risk of price discrimination, phone harassment, unsolicited mail or identity theft [9, 10]. Empirical studies on privacy concerns also provide ample evidence that privacy concerns may lead consumers to avoid making purchases [11]. For example, according to the “2019 Cyber Safety Insights Report Global Results,” more than one-third of the respondents indicated that they would not buy intelligent household products due to privacy concerns.<sup>8</sup> Therefore, as consumers become more concerned about their privacy, platforms must weigh potential losses when using consumer information to make profits.

In order to alleviate privacy concerns among consumers, some platforms have emphasized higher protections for user privacy by providing various services. For example, Internet Explorer 10 and 360 Security Browser have introduced the function of DNT (Do Not Track), which can effectively prevent cookie tracking and cross-site tracking behavior of some websites.<sup>9</sup> The search engine DuckDuckGo commits that

<sup>4</sup> <https://new.qq.com/omn/20190315/20190315A0NO8J.html?pc>.

<sup>5</sup> <http://www.fjii.com/cj/ssjj/2019/0328/210307.shtml>.

<sup>6</sup> <http://www.techweb.com.cn/internet/2019-03-28/2729632.shtml>.

<sup>7</sup> <https://www.iimedia.cn/c400/61120.html>.

<sup>8</sup> <https://www.nortonlifelock.com/tw/zh/newsroom/press-kits/2019-norton-lifelock-cyber-safety-insights-report/>.

<sup>9</sup> [http://china.cnr.cn/ygxw/201303/t20130324\\_512216806.shtml](http://china.cnr.cn/ygxw/201303/t20130324_512216806.shtml).

it will never collect or share any individual information and search history [12]. In contrast, other platforms adopt different strategies. For instance, after the 10th anniversary edition of Chrome was released, many users complained that cookies recording browsing history cannot be deleted.<sup>10</sup> These examples highlight the fact that different platforms protect consumer privacy in different levels. Also, consumers' willingness to pay to protect their data differs widely in different contexts and scenarios. Tsai et al. [13] found that most respondents preferred to pay a premium for goods to enjoy more protective privacy policies. Preibusch et al. [14] found that the vast majority of participants chose to buy a DVD from a cheaper but more privacy-invasive merchant, than from a costlier (1 euro more) but less invasive merchant. Tsai et al. [13] suggested that businesses may be able to leverage privacy protection as a selling point. Therefore, whether to provide privacy protection services to consumers and how to charge service fees are issues that any e-commerce platform needs to consider.

Our aim in this paper is to study how consumer privacy cost should affect the platform's information disclosure strategy. In the case of duopoly competition, how will the disclosure of consumer information affect market competition? There are several questions we intend to deal with: (1) How should the platform formulate their information disclosure strategies? (2) How should a firm set their prices with or without consumer information? (3) Should an e-commerce platform offer privacy protection services to consumers?

## 1.1 Literature Review

Our paper relates to three broad streams in the literature. First, our paper belongs to the literature on behavior-based price discrimination. Several decades ago, Narasimhan [15] argued that coupons can serve as a price discrimination device. Subsequently, the research on behavior-based price discrimination has been extended to various situations (see, for example, [16–19]). However, these papers only cover the monopoly case and do not discuss the different ways of obtaining consumer information. Suppose firms want to profile their consumers, the degree of precision in capturing consumer's information will be an important factor. As a complement to their dataset, some firms invest resources in acquiring data from other sources to improve precision. For instance, supermarkets (such as Tesco) not only gather data directly from their consumers, but also often supplement consumer information by spending money on third-party complementary data [12]. Our paper differs from the literature on the way firms access information. In our research, firms pay for the platform in order to get consumer information and firms may obtain the information through auctions.

Some papers study price discrimination strategies in oligopoly markets (see, for example, [20–22]). These papers focus on two-stage dynamic pricing strategies in which the company can identify new and old consumers and distinguish price changes in the second period by observing consumers' purchase choices in the first period. Most of these papers consider the case of a one-sided market, whereas online platforms always have more access to consumer information than online retailers, which grants them more chances to take advantage of the information. Our paper differs from the

<sup>10</sup> <https://www.huxiu.com/article/264468.html>.

present literature in that we do not take the period in which firms have no way to get any information into account. This assumption is based on the case that as for firms, even if consumers make repeat purchases, they are still unable to infer much about the consumers, as a consequence of which they need to buy additional information from a data seller [23]. We consider the case of two-sided market where the platform can obtain consumer information and firms can buy the information directly. For new consumers, firms will set a uniform price to them, which will have no influence on price discrimination, so we ignore these consumers in this paper.

The second stream of literature that we relate to concerns consumer privacy. The issue of consumer privacy has attracted widespread attention in recent years. For example, Conitzer et al. [24] studied a monopolist's pricing problem where consumers can maintain their anonymity and avoid being identified as past customers. Taylor and Wagman [25] showed that the winners, losers and social welfare from potential privacy regulation largely depend on the specific economic setting. They suggest that it may be more appropriate to design a nuanced approach about consumer privacy, which is tailored to specific markets regulations. Casadesus-Masanell and Andres [2] considered a duopoly setting and analyze the effect of privacy considerations for competition in the marketplace. In their paper, platforms derive revenues from consumer purchases and disclose consumer information in a secondary market. Consumers decide how much personal information to provide. As with our setting, online platforms can make profits from disclosing consumer data to firms. Our paper, however, provides a different angle to their work; instead of considering how much of the information to disclose, i.e., the depth of disclosure, we focus on the situation where the platform strategically decides the number of consumer information 'units' to sell, i.e., the width of disclosure. The papers referred to developing the notion of consumer privacy and personalized pricing but do not explicitly consider the impact of information leakage on consumer utility. Tsai et al. [13] designed an empirical study and showed that consumers tended to purchase from online retailers who can better protect their privacy. Goldfarb and Tucker [26] demonstrated that the revealing of consumer information will alter consumers' purchasing decisions. Also, consumers are increasingly unwilling to disclose their information online [27]. The work most closely related to ours is Montes et al. [23], which analyzed the impact of consumer privacy and price discrimination on product pricing, company profits and consumer surplus on two-sided market. As a complement to their research, we assume that if a consumer expects that his personal information has been disclosed, his utility will be reduced. This assumption is related to findings that consumers do incur substantial monetary costs and disutility from violations of their privacy [28]. For example, in 2012, the overall societal cost of spam in the USA is as high as \$20 billion [29]. Moreover, in Montes et al. [23], the privacy cost is an exogenous variable of the platform. By contrast, in our setting, the platform can set a price for a privacy protection service, by which we mean that the platform is able to control the cost of maintaining anonymity. For instance, an online platform can determine how easy it is for consumers to delete cookies planted on their computers [12]. Some online platforms (such as maichawang.com and fengnianxia.com) allow

consumers to make purchases without registration or a login, while some (such as Amazon.com and Tmall.com) do not.<sup>11</sup>

Our model is also related to the literature on two-sided markets. The pioneering works of Caillaud and Jullien [30], Rochet and Tirole [31] and Parker and Van Alstyne [32] provided analytical frameworks for pricing models in a two-sided market. Armstrong [33] used a “competitive bottlenecks” model to analyze the competition in two-sided markets, which makes a great contribution to previous research. Subsequent studies began to focus on the role of non-price variables in shaping the competitive strategies of platforms. For example, Hagiu and Spulber [34] Anderson et al. [35] and Brahem and Jebbi [36] examined the impact of non-price variables such as investment in first-party content, investment in platform performance and privacy issues on platform strategy. In recently years, some studies have linked the pricing issues of two-sided markets to consumer privacy concerns. Kox et al. [8] pointed out that targeting advertisements can increase the possibility of purchase, however, better targeting may raise consumer concerns about the associated loss of privacy. They find that targeting increases competition among platforms and reduces their profit. Dimakopoulos and Sudaric [37] studied the efficiency of data provision while taking users privacy costs into consideration. They demonstrate that data provision will be efficient if platforms engage in two-sided pricing. Gal-Or et al. [38] considered the impact of user privacy on competition between online advertising platforms. They show that offering consumers control over personal information can reduce targeting differentiation between platforms and reduce advertising costs. The platform in our model can be interpreted as an intermediate, where firms show and sell their products on one side and consumers make purchases on the other side. In contrast to the above studies in which firms or advertisers can gain consumer information freely, in our model, firms pay for the platform in order to obtain consumer data.

The rest of the paper is organized as follows. Section 2 explains the model. In Sect. 3, we characterize the equilibrium outcomes under monopoly and duopoly cases. In Sect. 4, we endogenize the privacy of consumers. Section 5 offers conclusions and puts forward some future research directions.

## 2 The Model

There are two competing retail firms who launch products on a platform with zero production cost. They provide products to consumers and compete on price. The platform raises revenue by selling consumer information to firms.

We consider a unit mass of consumers uniformly distributed along the Hotelling line of horizontal consumer preference, parametrized by location  $x \in [0, 1]$ . The platform randomly chooses the proportion  $\alpha \in [0, 1]$  of the consumers then gathers and sells their personal information. So,  $\alpha$  can be interpreted as the possibility that information has been disclosed to at least one firm for a consumer. If the platform sells that information to firm  $i$  ( $i = A, B$ ), then the firm can offer tailored prices to these consumers depending on  $x$ . If a consumer expects that their personal information

<sup>11</sup> <http://violet.blog.techweb.com.cn/archives/146.html>.

has been sold to a firm, then their utility will be reduced by  $k$  [8]. We designate  $k$  as the privacy disclosure aversion cost of each consumer. Assume  $k \geq 0$  without loss of generality. The variable  $k$  indicates the privacy disclosure aversion cost from two aspects. For one thing, consumers may feel unease when they discover that their personal information is sold to online retailers; for another, if the firm knows the location of a consumer exactly, it can set a discriminative price to them. This price is usually high, which will also cause aversion for the consumer. For simplicity, we call  $k$  the privacy cost in the following sections. If firm  $i$  is located at point  $d \in [0, 1]$ , then the expected utility of a consumer is

$$u = v - t|x - d| - p_i - \bar{\alpha}k, \quad (1)$$

where  $v$  is the product value and  $p_i$  designates the price of the product charged by firm  $i$ .  $t \geq 0$  is the transportation cost. The consumer expects that the platform has disclosed his personal information with probability  $\bar{\alpha}$ , so that the expected privacy disclosure aversion cost of the consumer is  $\bar{\alpha}k$ . The focus of our model is how much information the platform should sell to the firm, that is the proportion  $\alpha$ . We focus on rational expectations equilibrium where the consumers' expectation of the platform's proportion choice corresponds to the actual equilibrium proportion choice of the platform, i.e.,  $\bar{\alpha} = \alpha$  in equilibrium. We assume throughout that  $v \geq 2t + k$  to ensure that all the consumers will buy the product even if there is only one monopoly firm in the market.

According to the fraction of consumer information the platform reveals, the market can be divided into two parts. For the consumers whose information has been revealed to firm  $i$ , the firm will offer tailored price to them based on each of their location  $x$ , so we refer to this market as the personalized market and we call the consumers in this market 'personalized consumers'. For the other consumers that firm  $i$  does not recognize, the firm will offer them a uniform price. So, we refer to this market as the anonymous market and we refer to consumers in this market as anonymous consumers.

### 3 Main Analysis

In this section, we first analyze the benchmark case where there is a monopolist seller in the market, followed by the case in which two firms compete.

The model's timing proceeds as follows:

Stage 1 The platform declares the proportion  $\alpha$  of the consumers whose information has been gathered and the price  $I$  for the information.

Stage 2 Firm  $i$  decides whether to purchase the information.

Stage 3 Firm  $i$  determines its uniform price  $p_i$ .

Stage 4 Firm  $i$  offers tailored prices  $p_i(x)$ .

Stage 5 Consumers buy and consume.

### 3.1 Monopoly

Assume the monopoly firm is located at  $d = 0$ . First, we discuss the condition that the monopoly firm did not buy information in Stage 2. When the monopolist has no information about consumers' types, it sets uniform price  $p_0$ . Thus, the consumers who makes a purchase in Stage 5 has utility  $u = v - tx - p_0$ . To obtain the optimal price of the firm, we use  $p_0^* = v - t$  and the optimal profits are  $\pi_0^* = v - t$ .

Now assume the monopoly firm bought the additional consumer information in Stage 2. For the personalized market, the market size is  $\alpha$  and these consumers are scattered randomly in  $[0, 1]$  because they are selected randomly by the platform. The firm offers tailored prices  $p_2(x)$  to each  $x$ , which will make consumers' gross utility equal to 0. According to first-degree price discrimination, the monopoly firm can capture all the surplus. Therefore,  $p_2(x) = v - tx - \alpha k$ . For the anonymous market, the market size is  $1 - \alpha$ . In this market, when the firm has no information about consumers, it sets uniform price  $p_1$ . The monopoly firm's market share comprises of all the consumers located at  $x \in [0, x_0]$  such that  $v - tx_0 - p_1 - \alpha k = 0$ . The consumer located in  $x_0$  is indifferent between buying the product and taking the outside option, where  $x_0 = \min\{\frac{v-p_1-\alpha k}{t}, 1\}$ . Thus, the profits of the monopolist can be expressed as

$$\pi = \alpha \int_0^1 p_2(x)dx + (1 - \alpha)x_0 p_1. \tag{2}$$

To maximize the profit, we obtain  $p_1^* = v - t - \alpha k$  and  $x_0 = 1$  given the assumption  $v \geq 2t + k$ . The price  $p_1^*$  is the equilibrium basic price and is offered only to anonymous consumers. Consumers whose types are known by the firm are offered the tailored price  $p_2(x)$  and left with no surplus. The optimal profits of the monopolist are  $\pi^* = v - t + (\frac{1}{2} - k)\alpha$ . The result implies that the market is fully covered.

Then we discuss the pricing strategy of the platform. Note that without information on consumers, the monopolist's profit is  $\pi_0^*$ . As a result, when  $\pi^* > \pi_0^*$ , the maximum price  $I$  that the firm is willing to pay for the information is

$$I = \pi^* - \pi_0^* = \left(\frac{t}{2} - k\right)\alpha. \tag{3}$$

Otherwise, the monopoly firm will not buy the information. The platform's profits  $\pi_{PF}$  are made from selling the information, thus

$$\pi_{PF} = \begin{cases} I, & \text{if } \pi^* > \pi_0^*, \\ 0, & \text{if } \pi^* \leq \pi_0^*. \end{cases} \tag{4}$$

We designate a parameter  $m = k/t$  to reflect the relative sizes of both consumer privacy costs and transportation costs.  $m < 1$  means that consumer privacy cost is lower than the transportation cost, and vice versa.

**Proposition 1** *For the monopoly case,*



- (i) if  $m \leq \frac{1}{2}$ , then  $\alpha^* = 1$  and the platform will sell all consumers' information to the monopoly firm with price  $I = t/2 - k$ . The platform's profits are  $\pi_{PF}^* = t/2 - k$ . The firm's profits are  $\pi^* = v - t$ . Consumer surplus is equal to zero.
- (ii) if  $m > \frac{1}{2}$ , then the platform will have no incentive to collect and sell any consumers' information. The platform's profits are  $\pi_{PF}^* = 0$ . The firm's profits are  $\pi^* = v - t$ . Consumer surplus is given by  $CS = t/2$ .

If consumer privacy cost  $k$  is relatively low, the firm can make more profit if it has information about consumers. Thus, the platform has the incentive to sell all consumers' information in order to earn profits. However, if consumer privacy cost is relatively high, then tailored prices must be low enough to ensure consumers will buy the product. So, the monopolist prefers to set a uniform price to all consumers in order to eliminate the negative effects of privacy invasion, which means that the information has no value for the firm.

### 3.2 Competition

We now study the case where there are two competing firms, A and B, selling homogeneous products. Without loss of generality, we assume that firm A is located at  $d = 0$  and firm B is located at  $d = 1$ .

As a benchmark, we consider the case in which neither firm has information and consumers do not have privacy cost. In this Hotelling model, it's logical to derive that the optimal price of each firm is equal to the transportation cost  $t$ . The indifferent consumer is located at  $x = 1/2$  while the optimal profits for each firm are  $\pi_A = \pi_B = t$ . Consumer surplus  $CS = v - 5/4t$ . Next, we consider the cases in which only one firm has the information and both firms have the information.

#### 3.2.1 Only One Firm has Information

First, we focus on the case where only one firm can access the information. We assume it is firm A without loss of generality. We assume  $k \leq 3t$  here in order to guarantee that there exists a personalized market. After buying the information, firms A and B will set their uniform price  $p_A$  and  $p_B$  for the  $1 - \alpha$  anonymous market and then firm A will set tailored prices  $p_A(x)$  for the  $\alpha$  personalized market based on consumer information.

Consumers will compare the expected utility of making purchases from the two firms to decide whether to buy and, if so, where to buy. The expected utility derived by anonymous consumers and personalized consumers from firm A are

$$v - tx - p_A - \alpha k \text{ and } v - tx - p_A(x) - \alpha k, \tag{5}$$

whereas the expected utility of buying from firm B is

$$v - t(1 - x) - p_B. \tag{6}$$

In the anonymous market, the indifferent consumer who is strictly indifferent between buying from firm A and firm B is located at

$$x_1(p_A, p_B) = \frac{1}{2} + \frac{p_B - p_A - \alpha k}{2t}. \tag{7}$$

In the personalized market, for purpose of gaining the market share, firm A will set a price  $p_A(x)$  that leaves consumers indifferent between making purchase from firm A and firm B. In consequence,

$$p_A(x) = p_B + (1 - 2x)t - \alpha k. \tag{8}$$

From  $p_A(x) = 0$ , we obtain that the last consumer buying from A is located at

$$x_2(p_B) = \frac{p_B - \alpha k + t}{2t}. \tag{9}$$

In Stage 3, firm A and firm B choose uniform prices  $p_A$  and  $p_B$  to maximize their profits according to

$$\pi_A = (1 - \alpha) \int_0^{x_1} p_A dx + \alpha \int_0^{\min\{x_2, 1\}} p_A(x) dx, \tag{10}$$

$$\pi_B = (1 - \alpha) \int_{x_1}^1 p_B dx + \alpha \int_{\min\{x_2, 1\}}^1 p_B dx. \tag{11}$$

The first term on the right side of the two equations represents the revenue gain from the anonymous market, and the second term represents the revenue of the personalized market. The following lemma shows that, in the anonymous market, firm B has the higher market share, which enables it to choose a higher uniform price and thus earn greater profits than firm A.

**Lemma 1** *Assume that only firm A buys the information. Then the equilibrium uniform prices on the anonymous market are  $p_A^* = \frac{3t-3k\alpha}{3+\alpha}$  and  $p_B^* = \frac{3t-t\alpha+2k\alpha^2}{3+\alpha}$ , while the tailored price on the personalized market is  $p_A^*(x) = \frac{3t-t\alpha+2k\alpha^2}{3+\alpha} + (1 - 2x)t - \alpha k$ . Profits are  $\pi_A^* = \frac{18t^2(1+\alpha)+k^2\alpha^3(3+\alpha^2)+6kt\alpha(-3-2\alpha+\alpha^2)}{4t(3+\alpha)^2}$  and  $\pi_B^* = \frac{(t(-3+\alpha)-2k\alpha^2)^2}{2t(3+\alpha)^2}$  and consumer surplus is given by  $CS = \frac{-k^2\alpha^3(12+\alpha(3+\alpha))+t^2(-45+\alpha(-21+2\alpha))+2t(2v(3+\alpha)^2-3k\alpha(3+\alpha(3+2\alpha)))}{4t(3+\alpha)^2}$ .*

We now analyze how much information the platform (PF) should gather under different circumstances to create an optimal information disclosure strategy. Suppose that the platform can post a price  $I$  for consumers information and sell it in Stage 2. Let  $I_j$  ( $j = \{A, AB\}$ ) designate the price of information when the platform sells to only one firm ( $j = A$ ) or to both firms ( $j = AB$ ).

If the platform sells to only one firm, and that firm can earn higher profits than the other firm which has not got the information (i.e.,  $\pi_A^* > \pi_B^*$  in lemma 1), then the

two firms will compete for the information in order to make more profits. Such an allocation can be implemented via an auction [39]. In this case, the maximum price  $I_A$  that the platform can set is the maximum price any bidder is willing to pay, that is, the profit gap between the winner and loser in the auction. Thus,

$$I_A = \pi_A^* - \pi_B^* = \frac{\alpha(-2t^2(-15 + \alpha) + k^2\alpha^2(3 - 8\alpha + \alpha^2) + 2kt(-9 - 18\alpha + 7\alpha^2))}{4t(3 + \alpha)^2}. \tag{12}$$

In contrast, if owning information cannot help the firm to gain more profits (i.e.,  $\pi_A^* \leq \pi_B^*$  in lemma 1), then neither of the two firms will buy the information, which means that the information has no value to the firms. Thus, the platform’s profits are:

$$\pi_{PF}^A = \begin{cases} I_A, & \text{if } \pi_A^* > \pi_B^*, \\ 0, & \text{if } \pi_A^* \leq \pi_B^*. \end{cases} \tag{13}$$

The following proposition states the optimal information disclosure strategies for the platform under different circumstances.

**Proposition 2** *Assume the platform sells information to only one firm, then*

- (i) *if  $0 \leq m \leq \frac{\sqrt{97}-7}{8}$ , then the platform will collect and sell all consumers’ information;*
- (ii) *if  $\frac{\sqrt{97}-7}{8} < m < \frac{5}{3}$ , then there exists a unique  $\alpha^A \in (0, 1)$  such that the platform will choose consumers with amount  $\alpha^* = \alpha^A$  to collect and sell their information.*
- (iii) *if  $\frac{5}{3} \leq m \leq 3$ , then the platform will not collect any consumer information.*

When the platform sells information to only one firm, the amount of consumers the platform chooses depends on both the privacy cost and the transportation cost. When consumer privacy cost is much lower than the transportation cost, it is optimal for the platform to gather all consumers’ information. If we assume that transportation costs are fixed, Proposition 2 indicates that when consumers are not too concerned about information leaks, firms are able to extract increased profits by setting higher tailored prices.

**Corollary 1** *Assume that the platform sells information to only one firm. Then the optimal amount of information the platform discloses to the firm and the price of information are decreasing in  $k$  and increasing in  $t$ .*

Corollary 1 indicates that the value of consumer information decreases when privacy cost increases. That is because when consumers care about their privacy, the firm owning the consumer information must set low prices to them to ensure that they will make purchase, and therefore, the firm will be less competitive. When consumer privacy cost is high enough, it will also be unprofitable for the firm to buy information, which means that consumer information has no value at all. With the growth of consumer privacy cost, information becomes less valuable, and thus the platform will choose reducing proportions of consumers whose information is gathered and sold until the information becomes worthless.

### 3.2.2 Both Firms have Information

We assume that the platform randomly chose the proportion  $\alpha$  of consumers and sell their information in the first stage. These consumers constitute the personalized market while the others constitute the anonymous market. As for the anonymous market, the optimal uniform prices each firm makes are the same as with the benchmark, that is  $p_A = p_B = t$ . As for the personalized market, firms can make decisions according to consumer information and their competitor’s decisions. Take firm A for example, it must make sure that consumers prefer its tailored price to the offer of firm B. Thus, firm A will set its tailored price satisfying  $v - tx - p_A(x) - \alpha k \geq v - t(1 - x) - p_B(x) - \alpha k$ , which means that consumer’s expected utility from buying product from firm A must be no less than that of firm B. When  $p_B(x) = 0$ , we have  $p_A(x) \leq t(1 - 2x)$ . In consequence, the tailored price firm A chooses is given by  $p_A(x) = \max\{t(1 - 2x), 0\}$ . Firm B will make decisions in the same way due to the symmetry of the market.

**Lemma 2** *When both firm A and firm B buy the information, then the uniform prices are  $p_A = p_B = t$ , while the tailored prices are  $p_A(x) = \max\{t(1 - 2x), 0\}$  and  $p_B(x) = \max\{t(2x - 1), 0\}$ . Profits are  $\pi_A = \pi_B = \frac{t}{2} - \frac{t}{4}\alpha$  and consumer surplus is given by  $CS = v - k\alpha + \frac{1}{4}t(-5 + 2\alpha)$ .*

Because firm A and firm B are symmetrical, and both have the information, the two firms will set prices to ensure that it’s the same for any consumer to buy a product from firm A or firm B. On account of the fierce competition, the profits they earn are lower than in the no-information case.

If the platform sells information to both firms, and firms’ profits under this condition are higher than the case when its rival has the information but it does not (i.e.,  $\pi_B$  in lemma 2 is higher than  $\pi_B^*$  in lemma 1), then both two firms will buy information from the platform. In this case, the maximum price  $I_{AB}$  that the platform can set is

$$I_{AB} = \pi_B - \pi_B^* = \frac{t}{2} - \frac{t}{4}\alpha - \frac{(t(-3 + \alpha) - 2k\alpha^2)^2}{2t(3 + \alpha)^2}. \tag{14}$$

Otherwise, neither firm will buy the information. Thus, the platform will choose a strategy to maximize its profits, which is given by

$$\pi_{PF}^{AB} = \begin{cases} 2I_{AB}, & \text{if } \pi_B > \pi_B^*, \\ 0, & \text{if } \pi_B \leq \pi_B^*. \end{cases} \tag{15}$$

The next proposition formalizes this section’s main result.

**Proposition 3** *Assume that the platform sells information to both firms. Then there exists a unique  $\alpha^{AB} \in [0, 1]$  such that the platform will choose consumers with amount  $\alpha^* = \alpha^{AB}$  from whom to collect and sell information.*

**Corollary 2** *Assume that the platform sells information to both firms. Then the optimal amount of information the platform discloses to the firms and the price of information are decreasing in  $k$  and increasing in  $t$ .*

The incentive of proposition 3 is like that above. The difference between proposition 2 and 3 is that when the platform sells information to both firms, it can always make positive margins, because both firms want to gain competitive advantage from buying information.

### 3.2.3 Platform Decision

We now analyze how consumer privacy cost should affect the platform's information disclosure strategy. The platform chooses an information disclosure strategy to maximize its profits, that is

$$\pi_{PF} = \max\{\pi_{PF}^A, \pi_{PF}^{AB}\}. \quad (16)$$

We assume that when  $\pi_{PF}^A = \pi_{PF}^{AB}$ , the platform will sell information to only one firm. The following proposition shows the optimal information disclosure strategy for the platform under different situation.

**Proposition 4** *There exists a threshold  $m^0 \in (\frac{\sqrt{97}-7}{8}, \frac{5}{3})$  such that*

- (i) *if  $0 \leq m \leq m^0$ , then the platform will sell information to only one firm. In this case,*
  - (a) *if  $0 \leq m \leq \frac{\sqrt{97}-7}{8}$ , then the platform will choose to collect and sell all consumer information;*
  - (b) *if  $\frac{\sqrt{97}-7}{8} < m \leq m^0$ , then the platform will choose consumers with amount  $\alpha^* = \alpha^A$  to collect and sell their information.*
- (ii) *if  $m^0 < m \leq 3$ , then the platform will choose consumers with amount  $\alpha^* = \alpha^{AB}$  to collect and sell their information to both firms.*

Proposition 4 states that the platform's optimal strategy depends on the variation of  $m$ . When consumer privacy cost is relatively low, it is optimal for the platform to grant only exclusive rights for one firm to get the information. The reason is that when consumers do not care about information disclosure, the tailored price will be higher, and the personalized pricing market will expand. Therefore, the paying firm can extract more profit from the personalized market if its competitor can only have access to the anonymous market. However, if consumers care about their personal privacy, the firm with information cannot gain a competitive edge in the personalized market. That is to say, the value of information will decrease. In the meantime, when both firms have information, the profits they earn are independent of  $m$ , making it more profitable for the platform to sell information to both firms.

This conclusion is similar to that in Montes et al. [23]. They pointed out that it is optimal for the owner of information to grant only exclusive rights over the full database in the duopoly case where the privacy cost is not considered. In this paper, we take consumer privacy cost into consideration and allow the online platform to sell only part of consumer data. We find that consumer privacy cost has an impact on the platform's optimal information disclosure strategy. Proposition 4 also establishes that

when considering consumer privacy cost, it's not always optimal for the platform to disclose all consumers' information. When consumer privacy cost is relatively high, differentiated pricing is harmful to firms, so the platform may choose to reduce the possibility of information disclosure in order to encourage consumer consumption.

**Corollary 3** *The platform's profits are decreasing in  $k$  and increasing in  $t$ .*

It is clear that if consumer privacy cost is relatively high, consumers will be less willing to make purchases. Thus, firms may be well advised to lower their prices to stimulate consumer spending. When the firms cannot make sufficient profit from buying information, their willingness to buy decreases, and the information is no longer as valuable as before, which leads to a reduction in platform profits.

## 4 Extension

In the base model, we assume that consumers have no way to conceal their identities. In this section, we extend the model by considering the option that the platform can provide an information protection service to consumers, that is, consumers can pay an information protection fee  $c$  to avoid information leaks. This fee can also be interpreted as the difficulty of maintaining anonymity. For example, consumers can erase browsing histories, create new accounts using temporary e-mail addresses or spread purchases among numerous unrelated vendors to circumvent tracking. Some evidence has shown that some privacy-wary consumers are willing to pay for privacy. For example, Savage et al. [40] found that consumers may be willing to make a one-time payment of \$2.28 to conceal their browser history; \$4.05 to conceal their contacts list; \$1.19 to conceal their location and \$1.75 to conceal their phone's identification number. What's more, individuals can pay \$9.95 per month to Reputation.com to remove personal data from online data markets [12].

In this case, the platform is able to earn profits by selling consumer information to firms and charging information protection fees from consumers. We want to discuss how this policy will influence platform strategy and consumer purchase behavior. To simplify calculations, we assume that the platform will sell all consumers' information. The model's timing proceeds as follows:

- Stage 1 The platform declares the information protection fee  $c$ .
- Stage 2 Consumers decide whether to pay the fee  $c$ .
- Stage 3 The platform declares the price  $I$  for consumer information.
- Stage 4 Firm  $i$  decides whether to purchase the information.
- Stage 5 Firm  $i$  determines its uniform price  $p_i$ .
- Stage 6 Firm  $i$  offers tailored prices  $p_i(x)$ .
- Stage 7 Consumers buy and consume.

### 4.1 Monopoly

We still assume that the monopolist is located at  $d = 0$ . As in the case with no protection policy, consumers will derive zero utility. Thus, if the utility of making a

purchase can be higher than zero after paying  $c$ , the consumer will pay the fee. In the game, consumers should make their decisions before observing the real price  $p$ , so they expect that the uniform price of the product is  $p^a$ . A consumer located at  $x$  will pay the information protection fee when  $v - t\theta - p^a - c \geq 0$ , that is,  $x < x_3 = \frac{v - p^a - c}{t}$ . We assume that consumers have rational expectations about prices, that is,  $p^a = p$ .

Assume that the monopoly firm buys the information in Stage 4. The firm's profits are

$$\pi(p) = \int_0^{\max\{0, x_3\}} p dx + \int_{\max\{0, x_3\}}^1 p(x) dx, \tag{17}$$

where the tailored price  $p(x) = v - tx - k$ . The first term on the right-hand side represents the profits from selling to consumers who pay the information protection fee, i.e., the anonymous market. The second term represents the profits from the other consumers being charged a tailored price, i.e., the personalized market. The purpose of the firm is to maximize its profits  $\pi(p)$  for  $p > 0$  subject to  $0 \leq x_3 \leq 1$ .

The platform can make profits from providing information protection services to consumers and selling information to the firm. In Stage 1, the platform chooses the fee  $c$  to maximize profit  $\pi_{PF}$  according to

$$\pi_{PF} = cx_3 + I. \tag{18}$$

The maximum price of information that the platform can charge can be calculated similarly to Sect. 3.1. The following proposition shows the platform's optimal strategy in different circumstances.

**Proposition 5** *For the monopoly case,*

- (i) *if  $0 \leq m \leq 1$ , then the platform will provide information protection services to consumers for free. The consumers located at  $[0, m]$  will protect their information. The price of information is  $I = \frac{(k-t)^2}{2t}$ . The platform's profits are  $\pi_{PF}^* = \frac{(k-t)^2}{2t}$ .*
- (ii) *if  $m > 1$ , then all consumers will pay the information protection fee with  $c^* = k - t$ . The platform will provide information to the monopolist for free. The platform's profits are  $\pi_{PF}^* = k - t$ .*

We found that when the consumer privacy cost is lower than the transportation cost, consumers can protect their information for free. However, not all consumers will choose this service. The reason is that if a consumer is far from the monopolist while the firm can access that consumer's information, the monopolist will provide them with a lower price than the uniform price. Thus, consumers who are located far from the monopolist have no incentive to protect their personal information. Platform profits are all gain from selling information to the firm. To the contrary, when consumer privacy cost is relatively high, the platform can set a positive information protection fee to consumers. In order to make sure that all consumers will pay for the service, the platform will state that it provides information to the monopolist for free. Thus

a privacy protection service will make the information useless for the firms, and the platform generates all its profit by charging privacy protection fees from consumers.

### 4.2 Competition

Now we consider the duopoly case where consumers can pay for privacy. In this section, we will analyze consumers' choices, firms' pricing strategies and the platform's decision. We still assume that that firm A is located at  $x = 0$  and firm B is located at  $x = 1$ .

#### 4.2.1 Firms Decisions

First, consider the case where neither of the two firms buy the information. Observing that no firms will be aware of their locations, consumers will not feel a need to pay the fee  $c$ , and there is no privacy cost in such a situation. Thus, the equilibrium prices that the two firms set, and their profits are the same as in the case where the platform does not provide privacy protection services. Then, consider the case where both firm A and firm B buy the information at Stage 4. According to the analyzation in Sect. 3.2.1, fierce competition results in tailored prices so low that it makes no economic sense for any consumer to pay the fee. Thus, the equilibrium prices the two firms set, and their profits are the same as the results in Sect. 3.2.1.

Now, consider the case where only one firm can get the information—we still assume that it is firm A and firm A buys the information. For firm B, knowing from Sect. 3.2.1, consumers located at  $[x_2(p_B), 1]$  will buy the product from firm B, where  $x_2(p_B) = \frac{p_B - k + t}{2t}$ . Therefore, firm B's profits are  $\pi_B = (1 - x_2(p_B))p_B$ . To maximize its profits, firm B will set price  $p_B^* = \frac{k+t}{2}$  and thus  $x_2 = \frac{3}{4}(1 - \frac{k}{t})$ . If  $k > t$ , firm A will be squeezed out of the market and it will not buy the information. Thus, we only consider the case of  $k \leq t$ . In this case, the tailored price is  $p_A(x)^* = (\frac{3}{2} - 2x)t - k/2$ .

For the consumer who buys a product from firm A, he or she will pay the fee to protect their privacy if  $v - tx - p_A^a - c \geq v - tx - p_A^a(x) - k$ , where  $p_A^a$  and  $p_A^a(x)$  are the uniform price and tailored prices consumers expect that firm A will set. Thus, consumers located at  $x \leq x_3 = \frac{-c+t-p_A^a+p_B^a}{2t}$  will pay the fee. We still assume that consumers can form rational decisions about prices, that is  $p_A^a = p_A$  and  $p_A^a(x) = p_A(x)$ . If  $x_3 > x_2$ , then buying information will make no sense. Therefore, firm A will set the uniform price satisfying  $0 \leq x_3 \leq x_2 \leq 1$ . In this case, consumers located between  $[0, x_3]$  will pay the privacy protection fee and therefore can not be recognized by firm A. Consumers located between  $(x_3, x_2]$  will be charged a tailored price by firm A. Consumers between  $(x_2, 1]$  will make purchase from firm B. Hence, firm A solves

$$\max_{p_A} \pi_A = \int_0^{x_3} p_A dx + \int_{x_3}^{x_2} p_A(x) dx \quad \text{s.t. } 0 \leq x_3 \leq x_2 \leq 1. \tag{19}$$

To solve the problem, we obtain that when  $0 \leq c \leq k$ , the optimal uniform price of firm A is  $p_A^* = \frac{1}{2}(-k + 3t)$ ; when  $c > k$ , the optimal uniform price is



$p_A^* = \frac{1}{2}(-2c + k + 3t)$ . Assume that only firm A buys the information, only the optimal uniform price of firm A depends on the information protection fee.

### 4.2.2 Platform Decision

The platform can gain profit from selling consumers' information to firms and providing privacy protection service to consumers.

First consider the case where the platform sells the information to both firms. Denote firm B's profits under the condition where only firm A has information and both firms have information by  $\pi_B^{A*}$  and  $\pi_B^{AB*}$ . The price of the information can be calculated as in Sect. 3.2.1 When  $\pi_B^{AB*} > \pi_B^{A*}$ , the maximum price of information is  $I_{AB} = \pi_B^{AB*} - \pi_B^{A*} = \frac{t}{4} - \frac{(k-t)^2}{8t}$ . The platform's profits are  $\pi_{PF}^{AB} = \begin{cases} 2I_{AB}, & \text{if } \pi_B^{AB*} > \pi_B^{A*}, \\ 0, & \text{if } \pi_B^{AB*} \leq \pi_B^{A*}. \end{cases}$

Then consider the case where the platform sells the information to only one firm. Denote firm A's profits under the condition that only firm A has information by  $\pi_A^{A*}$ . When  $\pi_A^{A*} > \pi_B^{A*}$ , the maximum price of the information is  $I_A = \pi_A^{A*} - \pi_B^{A*}$ . The platform's profits are  $\pi_{PF}^A = \begin{cases} cx_3 + I_A, & \text{if } \pi_A^{A*} > \pi_B^{A*}, \\ 0, & \text{if } \pi_A^{A*} \leq \pi_B^{A*}. \end{cases}$

The platform compares what it will obtain from selling all information to both firms with what it can obtain from selling to one firm. Proposition 6 shows the optimal strategy of the platform in different situations.

**Proposition 6** *For the case with information protection service,*

- (i) *if  $0 \leq m \leq 1$ , then the platform will sell the information to only one firm with  $I = \frac{3k^2 - 10kt + 7t^2}{16t}$  and it will provide an information protection service to consumers for free. The profits of the platform are  $\pi_{PF}^{A*} = \frac{3k^2 - 10kt + 7t^2}{16t}$ .*
- (ii) *if  $m > 1$ , then the platform cannot make profits from disclosing information or providing an information protection service.*

Proposition 6 establishes that when the privacy cost is relatively low, it is optimal for the platform to grant exclusive rights for the information. The reason is that when both firms have information, the competition is too intense to earn profits for the firms, which results in the low value of information. Although consumers can protect their privacy for free, some consumers will still not choose the service. This is because for the consumers who are located far from firm A, the tailored price is lower than the uniform price.

Now, we will analyze the effectiveness of providing such a service. When the platform sells information to only one firm and consumers cannot pay for privacy, the platform can earn profits  $\pi_{PF}^{A*} = \frac{-k^2 - 10kt + 7t^2}{16t}$  according to Sect. 3.2.1 When providing the service, the platform's profits are  $\pi_{PF}^{A*} = \frac{3k^2 - 10kt + 7t^2}{16t}$ . It is clear that when providing information protection services to consumers, the platform can earn more profits even through the service fee is zero. The reason is that regardless of whether the platform provides a privacy protection service, the profits of firm B and the tailored prices of firm A are the same. However, when the platform provides the service, the uniform price firm A set can be higher than the tailored prices in the no

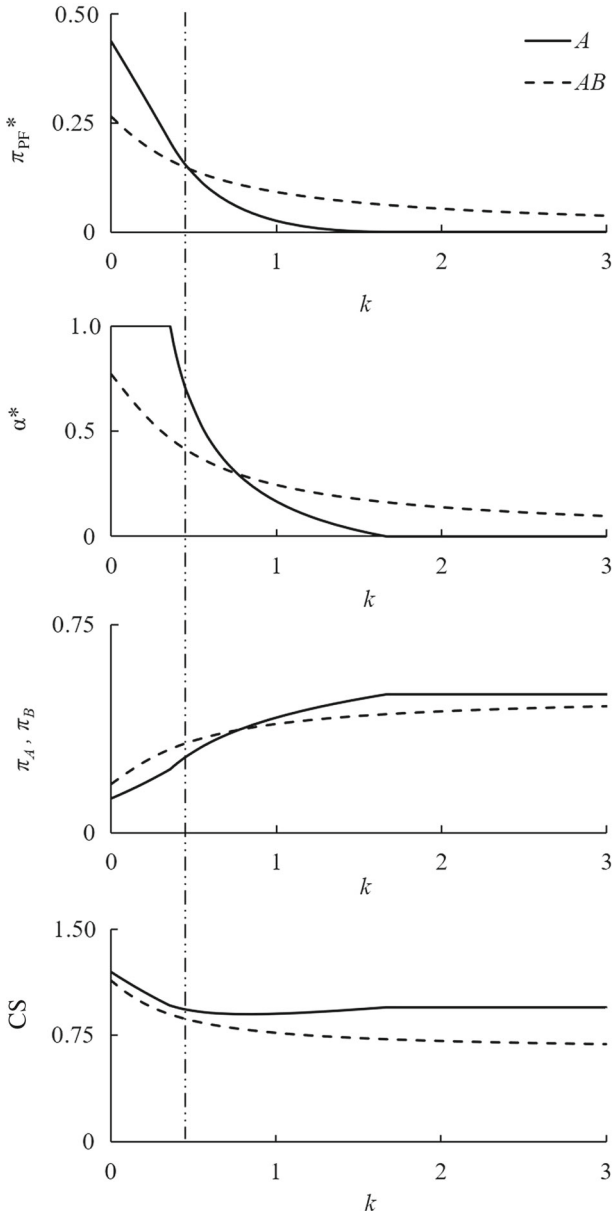
privacy case for the consumers located at  $[0, x_3]$ , because consumers do not have any privacy costs. This result is like the experimental research that showed for products costing about \$15, most participants are willing to pay a roughly \$0.5 premium to purchase goods from merchants with more protective privacy policies [13]. Benefitting from higher selling prices, firm A can make more profits from selling these products at a uniform price, which gives an opportunity to the platform to raise the price of information.

## 5 Main Highlights and Conclusion

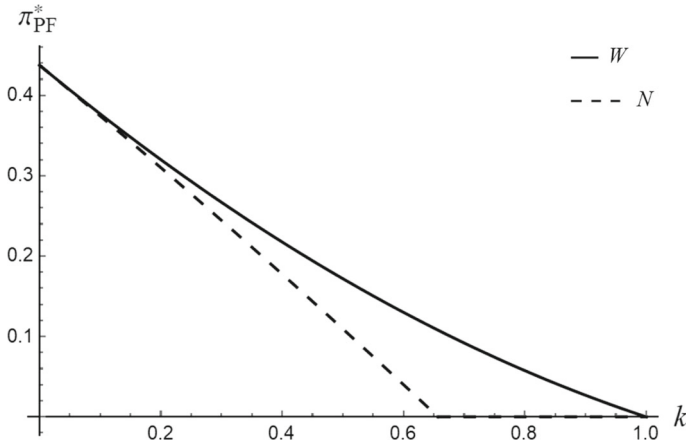
The growing prevalence of big data has caused extensive concern for society. Decision-making in business, economics and other fields increasingly rely on data and analysis. In the field of e-commerce, big data technology has been widely used. With the development of big data and cloud computing technology, platforms can use consumer personal information left on the Internet, such as consumption records and travel traces to build detailed consumer profiles. While the arrival of the Information Age makes it possible for consumers to benefit from targeted product recommendations [41], it also brings new problems. The use of data may expose consumers to risks such as information leakage and price discrimination, which may cause aversion costs. In this paper, we attempt to analyze the effect of this cost on platform and firm strategies. The novelty of our approach is in allowing the online platform to sell only a subset of their consumer data. Our model can be used to answer questions related to the value of consumer information for the platform.

When taking consumer privacy concerns into consideration, we distinguish between the case of monopoly and the case of competition. In the monopoly case, whether the platform sells all their consumer information depends on the relative size of the privacy and transportation costs. In the duopoly case, when the privacy cost is lower than the transportation cost, it is optimal for the platform to sell consumer information to only one firm (denoted by  $A$ ); otherwise it is optimal to sell consumer information to both firms (denoted by  $AB$ ). We found that the platform's profits decrease with the privacy cost. Considering the optimal strategy of the platform, the profits of both firms are increasing in consumer privacy cost, whereas consumer surplus is decreasing in consumer privacy cost. Figures 1 illustrates these results. The solid lines plot the case in which the platform sells information to only one firm and the dashed lines plot the case in which the platform sells information to both firms.

We extend the model by considering that the platform has the option to provide an information protection service to consumers. We find that in the monopoly case, when consumer privacy cost is relatively low, the platform will provide the service for free, otherwise it will set a positive fee. In the duopoly case, when the privacy cost is relatively low, the platform will provide the service for free while it can still earn more profits than the case of no such services. Figure 2 illustrates these results. The solid lines plot the case in which the platform provides an information protection service (denoted by  $W$ ) and the dashed lines plot the case in which consumers cannot protect their privacy (denoted by  $N$ ).



**Fig. 1** Platform profits, the optimal information disclosure level, firms' profits and consumer surplus for  $t = 1$  and  $v = 2.2$  in the duopoly case



**Fig. 2** Platform profits for  $t = 1$  and  $\alpha = 1$  in the duopoly case

This research provides some managerial implications to e-commerce platforms. First, although the platform makes profit by disclosing consumer information, it's not always optimal for the platform to disclose all consumers' information. In particular, when consumer privacy cost is small, price discrimination can bring excess profits to firms, thus the platform tends to disclose more consumer information. When consumer privacy cost becomes larger, price discrimination has less advantage to firms and the amount of consumer information disclosed by the platform is correspondingly reduced. Second, the degree of consumer privacy concerns has impacts on the platform's information disclosure and pricing strategies. At the same time, the price of information decreases as the cost of consumer privacy increases. Therefore, when devising information disclosure and pricing strategies, the platform needs to take consumer privacy concerns into account. Third, consumers' aversion to personal information leakage will have an adverse impact on platform profits and consumer surplus. Therefore, platforms can introduce some information protection services to reduce consumer privacy concerns, such as providing better personalized services, accurate personalized recommendations, and privacy protection tools.

Consumer privacy in online platforms is a developing subject and there could still be many changes as web services continue to evolve. Our analysis may be limited by the assumption that the privacy cost is homogeneous across consumers. If it is heterogeneous, we may draw some different results. We assume that the leakage of information can only cause negative responses for consumers, however, in practical application, consumer data is always being used for personalized recommendation services, which can enhance consumer satisfaction.

There are several directions to extend our work. Some examples for future research include issues related to potentially positive roles for the use of consumer information, such as providing customized services and more complex pricing strategies. Platforms and firms may face a trade-off between the privacy cost and potential benefits, and the trade-off between convenience and privacy is likely to influence consumer choice.

We hope that our research might be instructive, enlightening new paths for future analytical studies.

## Appendix

### Proof of Proposition 2

$$\frac{\partial I_A}{\partial \alpha} = \frac{6t^2(15 - 7\alpha) + k^2\alpha^2(27 - 93\alpha - \alpha^2 + 3\alpha^3) + 2kt(-27 - 99\alpha + 63\alpha^2 + 7\alpha^3)}{4t(3 + \alpha)^3}$$

$\frac{\partial^2 I_A}{\partial \alpha^2} = \frac{6t^2(-33+7\alpha)+72kt(-3+8\alpha)+k^2\alpha(81-432\alpha-6\alpha^2+22\alpha^3+3\alpha^4)}{2t(3+\alpha)^4} \leq 0$  as  $0 \leq \alpha \leq 1$  and  $0 \leq m \leq 3$ . Thus,  $I_A$  is a concave function.

Designate the numerator of  $\frac{\partial I_A}{\partial \alpha}$  as  $(\alpha) = 6t^2(15 - 7\alpha) + k^2\alpha^2(27 - 93\alpha - \alpha^2 + 3\alpha^3) + 2kt(-27 - 99\alpha + 63\alpha^2 + 7\alpha^3)$ .

$\frac{\partial M(\alpha)}{\partial \alpha} = -42t^2 + k^2\alpha^2(-93 - 2\alpha + 9\alpha^2) + 2kt(-99 + 126\alpha + 21\alpha^2) + 2k^2\alpha(27 - 93\alpha - \alpha^2 + 3\alpha^3) < 0$  when  $0 \leq \alpha \leq 1$ . So  $M$  is decreasing in  $\alpha$ .

When  $\alpha = 0$ ,  $M(0) = 18t(5t - 3k)$ ; When  $\alpha = 1$ ,  $M(1) = 16(3t^2 - 7kt - 4k^2)$ . If  $0 \leq m\alpha \leq \frac{\sqrt{97}-7}{8}$ , then  $M(1) \geq 0$  and  $\frac{\partial I_0}{\partial \alpha} \geq 0$ , so  $\alpha^* = 1$ . If  $\frac{\sqrt{97}-7}{8} < m < \frac{5}{3}$ , then  $M(0) > 0$  and  $M(1) < 0$ , which means that there exists a unique  $\alpha^A \in (0, 1)$  to maximize  $I_0$ . If  $\frac{5}{3} \leq m \leq 3$ , then  $M(0) \leq 0$  and  $\frac{\partial I_0}{\partial \alpha} \leq 0$ , so  $\alpha^* = 0$ .

When  $\alpha = 1$  or  $\alpha = \alpha^A$ , it's easy to prove that  $I_A > 0$ , thus the platform will earn positive profit from selling information. When  $\alpha_A = 0$ ,  $I_A \leq 0$ , thus the platform will not collect and sell consumer information.

### Proof of Proposition 3

$$\frac{\partial I_{AB}}{\partial \alpha} = \frac{-16k^2\alpha^3(6 + \alpha) + 8kt\alpha(-18 + \alpha(9 + \alpha)) - t^2(-45 + \alpha(51 + \alpha(9 + \alpha)))}{4t(3 + \alpha)^3}$$

$\frac{\partial^2 I_{AB}}{\partial \alpha^2} = \frac{-4(-3t^2(-6+\alpha)-9kt(-3+5\alpha)+k^2\alpha^2(54+12\alpha+\alpha^2))}{t(3+\alpha)^4} \leq 0$  as  $0 \leq \alpha \leq 1$  and  $0 \leq m \leq 3$ . Thus,  $I_{AB}$  is a concave function.

Designate the numerator of  $\frac{\partial I_{AB}}{\partial \alpha}$  as  $N(\alpha) = -16k^2\alpha^3(6 + \alpha) + 8kt\alpha(-18 + \alpha(9 + \alpha)) - t^2(-45 + \alpha(51 + \alpha(9 + \alpha)))$ . It's easy to prove that  $\frac{\partial N(\alpha)}{\partial \alpha} \leq 0$  when  $0 \leq \alpha \leq 1$ , thus  $N(\alpha)$  is decreasing in  $\alpha$ .

When  $\alpha = 0$ ,  $N(0) = 45t^2 \geq 0$ ; When  $\alpha = 1$ ,  $N(1) = 16(t^2 + 4kt + 7k^2) \leq 0$ . So, there exists a unique  $\alpha^{AB} \in (0, 1)$  to maximize  $I_{AB}$ .

When  $\alpha = 0$ ,  $I_{AB} = 0$ . So, when  $\alpha^* = \alpha^{AB}$ ,  $I_{AB}^* \geq 0$ .

**Proof of Proposition 4 Condition 1** When  $m \geq 5/3$ , according to proposition 2 and proposition 3, the platform will choose to sell information to both firms.

**Condition 2** When  $0 \leq m \leq \frac{\sqrt{97}-7}{8}$ , if the platform sells information to only one firm, according to proposition 2, the platform will set  $\alpha^* = 1$  and  $\pi_{PF}^{A*} = I_A^* = \frac{-k^2-10kt+7t^2}{16t}$ . If the platform sells information to both firms, according to proposition 3,

the platform will set  $\alpha^* = \alpha^{AB}$  and  $\pi_{PF}^{AB*} = 2I_{AB}^* = t - \frac{t}{2}\alpha^{AB} - \frac{(t(-3+\alpha^{AB})-2k\alpha^{AB^2})^2}{t(3+\alpha^{AB})^2}$ .

We have  $\pi_{PF}^{A*} > \pi_{PF}^{AB*}$  as  $\alpha_{AB} \in [0, 1]$ . Thus, the platform will sell information to only one firm.

We have  $\frac{\partial I_A^*}{\partial k} \leq 0, \frac{\partial I_{AB}^*}{\partial k} \leq 0, \frac{\partial I_A^*}{\partial t} \geq 0, \frac{\partial I_{AB}^*}{\partial t} \geq 0$ , thus there exists a unique  $m^0 \in (\frac{\sqrt{97}-7}{8}, \frac{5}{3})$  to make sure that if  $m \in [0, m^0)$ , then  $\pi_{PF}^{A*} > \pi_{PF}^{AB*}$ ; if  $m = m^0$ , then  $\pi_{PF}^{A*} = \pi_{PF}^{AB*}$ ; if  $m \in (m^0, 3]$ , then  $\pi_{PF}^{A*} < \pi_{PF}^{AB*}$ .

**Proof of Proposition 5 Step 1** The monopolist solves

$$\max_{p \geq 0} \pi(p) = \int_0^{x_3} p dx + \int_{x_3}^1 (v - tx - k) dx \text{ s.t. } 0 \leq x_3 \leq 1.$$

The optimal solution is  $p^* = v - k$  and  $x_3^* = \frac{-c+k}{t}$  regardless of the constraint. Now, take the constraint into consideration.

**Condition 1** When  $c \leq k - t$ , we have  $x_3^* \geq 1$ . So all the consumers will pay for privacy. Thus  $p^* = v - c - t$  and  $\pi^* = v - c - t$ .

**Condition 2** When  $k - t < c \leq k$ , we have  $0 \leq x_3^* < 1$ . So consumers located at  $[0, x_3^*]$  will pay for privacy. Thus  $p^* = v - k, x_3^* = \frac{-c+k}{t}$  and  $\pi^* = \frac{(c-k)^2-t^2}{2t} + v - k$ .

**Condition 3** When  $c > k$ , we have  $x_3^* < 0$ . So nobody will pay for privacy. Thus  $p^* = v - c$  and  $\pi^* = \frac{-t}{2} + v - k$ .

**Step 2** The platform solves  $\max_{c \geq 0} \pi_{PF} = cx_3 + I$ .

**Condition 1** When  $c \leq k - t$ , if  $k \leq t$ , there is no solution within the constraints; if  $k > t$ , then  $c^* = k - t, I = 0$  and  $\pi_{PF}^* = k - t$ .

**Condition 2** When  $k - t < c \leq k$ , if  $k \leq t$ , then  $c^* = 0, I = \frac{(k-t)^2}{2t}$  and  $\pi_{PF}^* = \frac{(k-t)^2}{2t}$ ; if  $k > t$ , then  $c^* = k - t, I = 0$  and  $\pi_{PF}^* = k - t$ .

**Condition 3** When  $c > k$ , if  $k \leq t/2$ , then  $I = \frac{t}{2} - k$  and  $\pi_{PF}^* = \frac{t}{2} - k$ ; if  $k > t/2$ , then  $I = 0$  and  $\pi_{PF}^* = 0$ .

We always have  $\frac{(k-t)^2}{2t} \geq \frac{t}{2} - k$ .

**Proof of Proposition 6 Case 1** The platform sells the information to both firms. When  $0 \leq k/t \leq -1 + \sqrt{2}, \pi_{PF}^{AB} = \frac{t}{2} - \frac{(t+k)^2}{4t} \geq 0$ ; when  $k/t > -1 + \sqrt{2}, \pi_{PF}^{AB} = 0$ .

**Case 2** The platform sells the information to only one firm.

**Condition 1** When  $0 \leq c \leq k, \pi_A^{A*} = \frac{4c^2-8ck+5k^2-6kt+9t^2}{16t}$  and  $\pi_B^{A*} = \frac{(k+t)^2}{8t}$ . Assume that  $\pi_A^{A*} > \pi_B^{A*}$ , thus the optimal solution is  $c^* = 0$  and  $\pi_{PF1}^{A*} = \frac{3k^2-10kt+7t^2}{16t}$ . Taking the size of  $k$  and  $t$  into consideration. When  $0 \leq k/t \leq 1, \pi_{PF1}^{A*} \geq 0$ .

**Condition 2** When  $c > k, \pi_A^{A*} = \frac{(k-3t)^2}{16t}$  and  $\pi_B^{A*} = \frac{(k+t)^2}{8t}$ . Assume that  $\pi_A^{A*} > \pi_B^{A*}$ , thus the optimal solution is  $c^* = 0$  and  $\pi_{PF2}^{A*} = \frac{-k^2-10kt+7t^2}{16t}$ . Taking the size of  $k$  and  $t$  into consideration. When  $0 \leq k/t \leq -5 + 4\sqrt{2}, \pi_{PF2}^{A*} \geq 0$ .

It is obvious that  $\pi_{PF1}^{A*} \geq \pi_{PF2}^{A*}$ , thus the platform will provide information protection service to consumers located at  $[0, \frac{k}{2t}]$  with  $c^* = 0$ .

Now, the platform will compare the profits in the two cases. When  $0 \leq k/t \leq -1 + \sqrt{2}$ , we have  $\pi_{PF1}^{A*} > \pi_{PF}^{AB}$ ; when  $-1 + \sqrt{2} < k/t \leq 1$ , we still have  $\pi_{PF1}^{A*} > \pi_{PF}^{AB}$ .

## References

- [1] Acquisti, A., Taylor, C., Wagman, L.: The economics of privacy. *J. Econ. Literat.* **52**(2), 442–492 (2016)
- [2] Casadesus-Masanell, R., Hervás-Drane, A.: Competing with privacy. *Manage. Sci.* **61**(1), 229–246 (2015)
- [3] Jing, B.: Behavior-based pricing, production efficiency, and quality differentiation. *Manage. Sci.* **63**(7), 2365–3237 (2017)
- [4] Esteves, R.B., Cerqueira, S.: Behavior-based pricing under imperfectly informed consumers. *Inf. Econ. Policy* **40**, 60–70 (2017)
- [5] Choe, C., King, S., Matsushima, N.: Pricing with cookies: Behavior-based price discrimination and spatial competition. *Manage. Sci.* **64**(12), 5669–5687 (2018)
- [6] Li, G., Lim, M.K., Wang, Z.: Stakeholders, green manufacturing, and practice performance: empirical evidence from Chinese fashion businesses. *Ann. Oper. Res.* **290**, 961–982 (2020)
- [7] Goldfarb, A., Tucker, C.: Shifts in privacy concerns. *Am. Econ. Rev.* **102**(3), 349–353 (2012)
- [8] Kox, H., Straathof, B., Zwart, G.: Targeted advertising, platform competition, and privacy. *J. Econ. Manage. Strat.* **26**(3), 557–570 (2017)
- [9] Hann, H., Hui, K.L., Lee, S.Y.T., Png, I.P.L.: Consumer privacy and marketing avoidance: a static model. *Manage. Sci.* **54**(6), 1094–1103 (2008)
- [10] Harrell, E., Langton, L.: *Victims of Identity Theft*, 2012. US Department of Justice Bureau of Justice Statistics, Washington DC (2014)
- [11] Xu, H., Teo, H.H., Tan, B.C., Agarwal, R.: Effects of individual self-protection, industry self-regulation, and government regulation on privacy concerns: a study of location-based services. *Inf. Syst. Res.* **23**(4), 1342–1363 (2012)
- [12] Valletti, T., Wu, J.: Consumer profiling with data requirements: structure and policy implications. *Prod. Oper. Manag.* **29**(2), 309–329 (2019)
- [13] Tsai, J.Y., Egelman, S., Cranor, L., Acquisti, A.: The effect of online privacy information on purchasing behavior: an experimental study. *Infor. Sys. Res.* **22**(2), 254–268 (2011)
- [14] Preibusch, S., Kubler, D., Beresford, A.: Price versus privacy: an experiment into the competitive advantage of collecting less personal information. *Electron. Commer. Res.* **13**(4), 423–455 (2013)
- [15] Narasimhan, C.: A price discrimination theory of coupons. *Mark. Sci.* **3**(2), 128–147 (1984)
- [16] Anderson, E.T., Dana, J.D.: When is price discrimination profitable? *Manage. Sci.* **55**(6), 980–989 (2009)
- [17] Liu, Q.H., Shuai, J.: Multi-dimensional price discrimination. *Int. J. Ind. Organ.* **31**(5), 417–428 (2013)
- [18] Bergemann, D., Brooks, B., Morris, S.: The limits of price discrimination. *Am. Econ. Rev.* **105**(3), 921–957 (2015)
- [19] Herbon, A.: Potential additional profits of selling a perishable product due to implementing price discrimination versus implementation costs. *Int. Trans. Operat. Res.* **26**(4), 1402–1421 (2019)
- [20] Chen, Y.: Dynamic price discrimination with asymmetric firms. *J. Ind. Econom.* **56**(4), 729–751 (2008)
- [21] Carroni, E.: Competitive customer poaching with asymmetric firms. *Int. J. Ind. Organ.* **48**(6), 173–206 (2016)
- [22] Carroni, E.: Behaviour-based price discrimination with cross-group externalities. *J. Econ.* **125**(2), 137–157 (2018)
- [23] Montes, R., Sand-Zantman, W., Valletti, T.: The value of personal information in online markets with endogenous privacy. *Manage. Sci.* **65**(3), 1342–1362 (2019)
- [24] Conitzer, V., Taylor, C.R., Wagman, L.: Hide and seek: Costly consumer privacy in a market with repeat purchases. *Mark. Sci.* **31**(2), 277–292 (2012)
- [25] Taylor, C., Wagman, L.: Consumer privacy in oligopolistic markets: Winners, losers, and welfare. *Int. J. Ind. Organ.* **34**, 80–84 (2014)
- [26] Goldfarb, A., Tucker, C.: Online display advertising: targeting and obtrusiveness. *Mark. Sci.* **30**(3), 389–404 (2011)
- [27] Tucker, C.E.: The economics of advertising and privacy. *Int. J. Ind. Organ.* **30**(3), 326–329 (2012)
- [28] Stone, B.: Ads Posted on Facebook Strike Some as Off-Key. *New York Times*. March 4, B1 (2010)
- [29] Rao, J.M., Reiley, D.H.: The economics of spam. *J. Econ. Perspect.* **26**(3), 87–110 (2012)
- [30] Caillaud, B., Jullien, B.: Chicken and egg: Competition among intermediation service providers. *Rand J. Econ.* **34**(2), 309–328 (2003)

- [31] Rochet, J.C., Tirole, J.: Platform competition in two-sided markets. *Eur Econ Associat.* **1**(4), 990–1029 (2003)
- [32] Parker, G.G., Van Alstyne, M.W.: Two-sided network effects: A theory of information product design. *Manage. Sci.* **51**(10), 1494–1504 (2005)
- [33] Armstrong, M.: Competition in two-sided markets. *J. Econ.* **37**(3), 668–691 (2006)
- [34] Hagiu, A., Spulber, D.: First-party content and coordination in two-sided markets. *Manage. Sci.* **59**(4), 933–949 (2013)
- [35] Anderson, E.G., Parker, G.G., Tan, B.: Platform performance investment in the presence of network externalities. *Inf. Syst. Res.* **25**(1), 152–172 (2014)
- [36] Brahem, A., Jebisi, K.: Access pricing to an operating system platform when piracy matters. *J. Syst. Sci. Syst. Eng.* **29**(2), 184–202 (2020)
- [37] Dimakopoulos, P.D., Sudaric, S.: Privacy and platform competition. *Int. J. Ind. Organ.* **61**, 686–713 (2018)
- [38] Gal-Or, E., Gal-Or, R., Penmetsa, N.: The role of user privacy concerns in shaping competition among platforms. *Inf. Syst. Res.* **29**(3), 698–722 (2018)
- [39] Jehiel, P., Moldovanu, B.: Auctions with downstream interaction among buyers. *Rand J Econ.* **31**(4), 768–791 (2000)
- [40] Savage, S., Donald M.W.: The Value of online privacy. (2013). <https://doi.org/10.2139/ssrn.2341311>
- [41] Anand, B.N., Shachar, R.: Targeted advertising as a signal. *Quant. Mark. Econ.* **7**(3), 237–266 (2009)