



The value of personal information in vertically differentiated markets with privacy concerns

Yongrui Duan¹ · Yuncong Xie¹ · Yihong Hu¹

Accepted: 9 September 2020 / Published online: 6 October 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Information technology and data science have enabled firms to practice price discrimination on an unprecedented scale, arousing privacy concerns among their customers. When consumers know a firm is practicing price discrimination, they may take costly measures to conceal their identities so as to avoid being targeted. Governments, in turn, may require firms to disclose their price discrimination practices in order to protect consumers' interests. In this paper, we consider a pricing game in which two competitive, vertically differentiated firms may implement price discrimination using information purchased from a third-party data supplier. We determine (1) the firms' optimal pricing strategies when consumers can (or cannot) safeguard their personal information by paying a "privacy cost"; (2) the data supplier's optimal sales strategy and the value of the data; and (3) the effects of the cost of consumer privacy and of the disclosure of price discrimination practices on firms and consumers. We find that for the data supplier, the optimal sales strategy is always to sell exclusively to one firm, regardless of whether consumers are aware that the firm practices "personalized pricing". The question of which firm the data broker should sell to depends on what we term the "quality-adjusted cost"—the ratio between the additional cost of the high-quality product and the magnitude of the quality difference. If this ratio is smaller than $1/2$, the data broker will sell to the high-quality firm; if greater, to the low-quality firm. Second, by comparing two scenarios involving the disclosure or non-disclosure of price discrimination, we find, somewhat counter-intuitively, that mandatory transparency increases industry profits and decreases consumer surplus when only the high-quality firm has access to consumer data. When only the low-quality firm has such access, transparency lowers industry profits once the quality-adjusted cost exceeds a certain threshold. When the quality-adjusted cost is in the intermediate range, mandatory transparency decreases social welfare. This means that the disclosure of price discrimination practices may have unfavorable consequences from a social planning standpoint. Thus, the new insights our findings offer into competitive personalized pricing in vertically differentiated markets will be useful not only to managers in the industry but also to regulators.

Keywords Price discrimination · Privacy · Vertically differentiated market · Transparency policy

✉ Yihong Hu
yhhu@tongji.edu.cn

¹ School of Economics and Management, Tongji University, Shanghai 200092, China

1 Introduction

The explosion of information technology and data science has enabled firms to collect data about consumers in novel ways. Various data-gathering tools, such as cookies, web beacons, and loyalty cards, make it much easier for retailers or data collectors to track consumers on the Internet and implement targeted pricing and marketing. Amazon, the giant retail platform, uses cookies to track consumers' purchasing history and applies algorithms so as to recommend products that may interest individual consumers. Orbitz Worldwide, a travel agency, used consumers' login information to practice price discrimination, recommending more expensive hotels to Mac users than to PC users. In the online market, some retailers even use consumers' personal data to price-discriminate (Jing 2017; Esteves et al. 2017; Choe et al. 2018; Li et al. 2020a). As these examples show, information technology has greatly enhanced the profiling and targeting ability of enterprises.

The wide use of personal data has spawned a data brokering industry. Some firms are either limited by their ability to collect and analyze data or do not find it cost-efficient to collect information on their own. Instead, these firms may purchase information from third-party data brokers (often known as data suppliers, DS), such as Oracle, Experian and Teradata. The customers of a DS may be competitors. For example, in the online market for athletic shoes, many brands (Nike, Adidas, PUMA, etc.) compete for limited market share. These firms can all benefit from a deeper understanding of consumer preferences. The DS can choose to sell consumer information exclusively to a single firm or to offer it to competing firms, but there is a tradeoff. The DS can charge a higher price for information that it provides on an exclusive basis to one firm, but it may achieve higher overall profits by selling the data at a lower price to several competing firms. From the standpoint of the purchasers, on the other hand, exclusive access to consumer data is conducive to higher profits, while sharing it with competitors tends to intensify competition and lower profits.

At the same time, the widespread use of personal information arouses serious concerns among consumers about their privacy. After their data is collected, consumers may face aggressively targeted advertisements and personalized pricing. Consumers resent this because they may be charged a higher price than their anonymous peer buyers (Choe et al. 2018). Therefore, consumers may take costly measures to protect their privacy. They can remove cookies from their browser or use anonymous payment systems to avoid being tracked (Acquisti and Varian 2005). Some third-party platforms provide consumers with privacy protection services for consumers willing to pay a fee. For instance, Reputation.com charges individuals \$9.95 per month to delete personal data from data markets (Valletti and Wu 2019).

Price discrimination based on consumers' personal data has drawn the attention of legislative authorities (Goldfarb et al. 2011). Consumer advocates call for regulations that require companies to clearly inform consumers of their personalized pricing policies. With this information, consumers could decide whether to protect their information and whether to buy from these companies. To this end, a bill called E-STOP, the Ensuring Shoppers Transparency in Online Pricing Act, was proposed in the U.S. Congress in 2012. It would have required online merchants to disclose whether they collect consumer information and practice price discrimination. Policymakers also have concerns about the lack of transparency of companies that buy and sell consumer data largely without consumer awareness. The lack of transparency led to the adoption of the Fair Credit Reporting Act ("FCRA"), a statute the Commission of United States has enforced since its enactment in 1970. In Europe, Directive 2002/58/EC of the European Parliament and of the Council, also known as ePrivacy Directive (ePD), is an EU directive on data protection and privacy in the digital age. It presents a continuation

of earlier efforts, most directly the Data Protection Directive. The effect of transparency requirements on firms, consumers and social welfare in vertically differentiated markets has thus far attracted little research attention, although there is considerable interest in the matter.

This study explores the effects of consumers' privacy concerns and of mandated transparency in vertically differentiated markets. We consider two competitive firms, such as PUMA and Adidas, that separately market products which are innately differentiated in quality. Each firm decides whether to adopt price discrimination and whether to buy consumer information from a third-party data supplier. The data supplier decides which firm(s) it will sell to. We ask the following questions: (a) How does a firm with a high- or low-quality product price its goods when it has purchased consumer information? (b) What is the equilibrium of the two firms' strategies in the competitive context? (c) What is the optimal information-selling strategy for the data supplier? (d) How does consumers' privacy protection behavior affect the data supplier's strategy? (e) How does transparency in personalized pricing affect firms, consumers and social welfare?

In order to answer these questions, we consider a model in which two asymmetric firms whose products differ in quality compete in markets through price discrimination. The firms have to decide whether or not to buy consumer information from a monopoly data supplier. Consumers can take costly actions to conceal and protect their personal information. We focus primarily on the strategic decisions of the data supplier and the firms.

The main contributions of this paper are as follows.

First, we show that the data supplier's optimal sales strategy is to sell exclusively to one of the firms, regardless of whether the price discrimination policy is transparent and whether consumers choose to conceal their identities from information collectors. This conclusion is consistent with findings in the literature on symmetric competing firms.

Second, we highlight the role of the quality-adjusted cost—that is, the ratio between the cost difference and the quality difference—in determining the equilibrium outcome. In particular, when the firms do not disclose price discrimination and the quality-adjusted cost difference is smaller than $1/2$, the optimal strategy for the data supplier is to sell information exclusively to the high-quality firm. When the quality-adjusted cost ratio is larger than $1/2$, the data supplier sells information exclusively to the low-quality firm.

Third, we show that the privacy cost—the cost consumers must pay if they wish to hide their identities—has a strong impact on the profits of the firms, on consumer surplus, and on the value of consumer information. When the information is available exclusively to one of the firms and the privacy cost is not too high, the profits of the firm that has the information increase with the privacy cost. On the other hand, the consumer surplus decreases as the privacy cost grows. It is intuitive that the information price charged by the DS increases with the privacy cost. As this cost rises, fewer consumers are willing to pay for privacy and firms can achieve higher profits, so consumer data increases in value.

Fourth, we examine the effect of requiring transparency in personalized pricing. Counter-intuitively, we find that transparency is always detrimental to consumer surplus, while in most cases it is beneficial to the firms. When only the high-quality firm has consumer information, transparency in personalized pricing improves the total profit of the two firms at the expense of consumer surplus. When the quality-adjusted cost is in the middle range, transparency always diminishes social welfare. Therefore, from the standpoint of a social planner, requiring transparency in personalized pricing is not beneficial to consumers and should be regarded with circumspection.

The remainder of this paper is organized as follows. Section 2 summarizes relevant studies in the literature. Section 3 presents the model settings. Section 4 examines the case in which consumers are not informed about personalized pricing discrimination and do not pay for

privacy. Section 5 considers the opposite case, in which consumers have an endogenous privacy choice and firms are required to disclose price discrimination. Section 6 compares the two cases (with and without transparency) and Sect. 7 presents concluding remarks. All proofs are provided in the “Appendix”.

2 Literature review

This paper relates to two broad streams in the literature.

The first is the literature on behavior-based price discrimination. In a seminal work, Villas-Boas (1999) uses a Hotelling model to evaluate the effects of firms’ ability to identify repeat customers on competition. Firms are able to recognize returning customers’ preferences and attract some of the competitor’s repeat customers by changing prices for overlapping generations of consumers. Villas-Boas finds that firms always set lower prices for new customers because their repeat customers have already shown a preference for their products. Research on behavior-based pricing has been extended to various settings. Villas-Boas (2004) finds it may be disadvantageous for a monopoly to implement behavior-based pricing (BBP). Pazgal and Soberman (2008) study the firms’ endogenous choices on whether to engage in BBP. They find that the equilibrium profit when both firms engage in BBP is always lower than when neither firm uses it. Esteves (2010) studies the effects of myopic consumers on the firms’ BBP decisions. Shin and Sudhir (2010) investigate a firm’s use of a behavior-based price discrimination strategy in relation to two important customer features: the heterogeneity of consumer value and consumer preference change. A comprehensive survey of the literature on BBP is provided in Fudenberg and Villas-Boas (2006) and Esteves (2009).

Recent BBP research has explored more complex applications. Li and Jain (2016) study the effects of consumers’ fairness concerns on BBP. They show that firms can obtain higher total discounted profits from engaging in BBP than from forgoing consumer recognition. Colombo (2016) investigates behavior-based price discrimination when firms do not have complete purchase information about consumers. Esteves (2014) studies behavior-based price discrimination (BBPD) when firms implement a retention strategy to discourage customers from switching their allegiance. Esteves and Reggiani (2014) analyze the effect of demand elasticity on profit, consumer surplus and social welfare when firms implement behavior-based price discrimination. Rhee and Thomadsen (2017) consider BBP in a vertically differentiated setting. They find that both high- and low-quality firms may offer discounts to repeat customers. Esteves and Cerqueira (2017) study behavior-based advertising in a horizontally differentiated market and show how consumer awareness affects industry profits and consumer welfare. De Nijs (2017) uses a two-period model to study the effects of rival firms’ information-sharing behavior on their behavior-based pricing strategies. Choe, King and Matsushima (2018) consider a two-period model of dynamic completion between two firms. They examine the effects of personalized pricing on the firms’ profits and prices in two periods. Studying BBP in a channel setting, Li (2018) investigates how the adoption of BBP affects the profits of channel members and social welfare. Amaldoss and He (2019) investigate the practice of BBP in a horizontally differentiated market where consumers have diverse tastes and limited consideration sets. They find that consumer valuation will affect the difference between the prices that old and new customers are charged under BBP. In contrast to these studies of BBP, we do not assume that there are two periods and that firms acquire consumer preference information from the first period. Instead, we study personalized pricing based on

consumer preferences in a single period when firms can purchase the necessary preference information or analytics from a data supplier.

Another stream of literature is related to the implications of privacy. Privacy is an important topic in economics for a long time. Posoner (1981) believes that hiding consumer identities reduces economic efficiency. Acquisti, Taylor and Wagman (2016) provide a comprehensive literature review on the economic value of privacy. Given the rise of big data and new marketing technology practice, privacy has greatly influenced pricing, product design, advertising, and government regulation. Our work is particularly related to research into the relationship between consumer privacy and pricing. Taylor (2004) studies how the ability of firms to collect and sell consumer information, as well as the right of consumers to anonymity, impact market competition. Acquisti and Varian (2005) examine the circumstances under which it is profitable for firms to implement BBP when consumers can take measures to protect their privacy. They find that it is feasible but never optimal for firms to distinguish between high- and low-value consumers through price discrimination. Hann et al. (2008) study the effects of consumers' information-concealing behavior on market outcomes. They discover that concealment by high-benefit consumers leads to a reduction in the seller's market share, while concealment by low-benefit consumers may increase a seller's market share. Conitzer, Taylor and Wagman (2012) consider a model with a monopolist firm and heterogeneous customers who can conceal their identity for free or by paying a cost. They show that an increase in the cost of concealment may benefit consumers but does not always do so. When the privacy cost is high enough, the effect may be reversed. Tucker (2014) empirically demonstrates that consumers' perception of control over their privacy increases their acceptance of personalized advertising. Casadesus-Masanell and Hervas-Drane (2015) study a duopoly model in which consumers can endogenously choose how much information is disclosed to firms. Firms benefit from this arrangement in two ways: through consumer buying behavior and by selling the information at a profit. Shy and Stenbacka (2016) examine how different degrees of privacy protection affect industry outcomes when there are switching costs. They find that weak privacy protection is more beneficial for firms than strong protection or no protection.

Our work is also related to firms' consumer profiling and information decisions concerning asymmetric information selling, acquisition (Li et al. 2020b; Li et al. 2020c) and utilization. Koh et al. (2017) examine the effect of a voluntary profiling policy on the seller's profits, on consumer surplus and on social welfare. Consumers are heterogeneous and their sensitivity to privacy varies. The authors show that neither consumer surplus nor social welfare is necessarily larger when consumers choose voluntary profiling instead of no profiling. Choi, Jeon and Kim (2019) consider a model of privacy with information externalities. In this model, data collection requires consumers' consent. They find that, even in the market equilibrium, the collection of personal information exceeds the social optimum. Valletti and Wu (2019) consider the effects of price discrimination with consumer data profiling. Firms can invest in increasing the precision of consumer profiling, and consumers can protect their privacy by paying a cost. They find that socially optimal privacy policies exist when data protection is very easy or very costly. Li et al. (2020d) examine the effect of the transparency of firms' BBP practices on firms and consumers. They find that BBP transparency increases a monopolist's profit but decreases consumer surplus and social welfare. Finally, the study most closely related to our own is by Montes et al. (2019), who study the effects of price discrimination through a duopoly Hotelling model in which consumers can make endogenous choices on privacy protection. Two symmetric firms can purchase consumer information from a data supplier and base their pricing on it. Our study differs from the above literature in that we extend the model to a vertically differentiated market and introduce the quality-adjust cost to explain the final equilibrium results. In reality, it is not easy to find two identical

competitors, and there are usually significant quality differences among companies in the market. Therefore, competition between two quality-differentiated firms is more common in practice, and needs further research. Our paper also considers the policy transparency issue and therefore provides meaningful suggestions for the government.

3 Model settings

Our model includes three classes of agents: consumers, two competing firms selling goods of different quality, and a data supplier (DS) that can collect information about consumers' preferences. We begin by describing the two competing firms. They market substitutable products differentiated by quality. H denotes the high-quality firm and L the low-quality firm. We use q_H, q_L to denote the product quality of the two firms and define the quality difference $\Delta, \Delta = q_H - q_L > 0$. We assume that each firm has different marginal costs, with $c_L < c_H$. Without loss of generality, we assume that H's marginal cost is equal to c and L's is normalized to zero. Thus, c represents the cost difference between the products of the two firms. To make sure the products are profitable, we assume that $0 \leq c \leq \Delta$. The upper bound is the cost level at which the additional utility that quality-sensitive consumers obtain by buying high-quality product is equal to the additional cost. Without loss of generality, we denote $\mu = \frac{c}{\Delta}, \mu \in [0, 1]$, which can be interpreted as the quality-adjusted cost.

The market contains a variety of consumers, each of whom purchases one unit of product. Consumer i 's utility function from buying one unit of product j is represented as $U(\theta_i) = V + \theta_i q_j - p_j$, where $j \in \{H, L\}$. p_j is the price that consumer i pays for product j . θ_i represents consumer i 's quality preference, and without loss of generality we assume that $\theta \sim U[0, 1]$. V is a constant utility that every consumer receives from purchasing the product. Following Rhee and Thomadsen (2017), we assume that V is sufficiently large that each consumer will choose one of the products.

Suppose there are two market segments, the "old market" and the "new market". The difference between the two markets depends on whether consumer preference information is available. In the "new market", firms only know the aggregate distribution of the market and have no way to obtain information about individual consumers. Therefore, firms can only set a basic, undifferentiated price for all consumers. In the "old market", a firm can acquire preference information about consumers and offer tailored prices based on each consumer's taste for quality θ . We can describe the "new market" as anonymous and the "old market" as personalized. The distinction is easy to understand in the context of online retailing. Consumers in the "new" segment are recent entrants to the market, so firms have little or no information about their preferences and behavior. On the other hand, consumers in the "old" segment have a record of activity on the Internet: searching, buying products, writing reviews, and so on. Such data enables retailers to analyze the individual preferences of each "old" consumer. Our model assumes that both the new and the old markets are uniformly distributed over the unit interval.

The widescale use of consumer information may arouse privacy concerns among consumers. In order to protect their personal information, consumers in the old market may choose to conceal their identities (thereby preventing firms from implementing personalized pricing) by paying a privacy cost c_0 . The privacy cost is related to the degree of hardship that consumers face in concealing their identities—for example, the need to delete tracking cookies that various websites place on the consumer's computer. Some people go so far as to create new accounts to avoid being identified as "old" customers. The more informative

the consumer's personal data, the higher the privacy cost. Some consumers are so concerned about privacy that they are even willing to pay a monetary cost to protect their data. For example, Reputation.com charges individuals \$9.95 per month to remove personal data from online markets. GDPR (General Data Protection Regulation) enforces service providers to erase any personal data if a consumer request. For instance, Acxiom, allow consumers to opt out of the use of their personal information in the database. Consumers may need to submit request form and communicate with many data brokers to remove their personal information. These efforts can also be regarded as privacy protection cost.

When the privacy cost is too large, no consumer will pay it. In order to consider more interesting cases, we set $0 < c_0 \leq \frac{\Delta}{2}$, a parameter that is easily derived in the following discussion. When the consumer i pays the privacy cost c_0 and purchases product j , his utility is $V + \theta_i q_j - p_j - c_0$. Hereafter, c_0 is a constant cost independent of the type of consumer.

The third class of agent is the monopoly data supplier, who collects consumer preference information and sells it on to other firms. Without loss of generality, we assume that the firm can perfectly ascertain a consumer's personal tastes θ with the help of the data.

The sequence of the game is as follows. First, the data supplier posts the data price K . Second, both firm s decide whether or not to buy the data. In the third stage, consumers in the "old market" decide whether or not to pay the privacy cost c_0 . Next, firms make their price decisions to compete for customers. They first simultaneously determine the basic prices p_H, p_L , then offer the tailored prices. Finally, each consumer makes a purchase decision.

4 Equilibrium results without transparency of price discrimination

First we discuss the benchmark case in which firms do not implement a policy of transparency and consumers are unaware of the price discrimination strategy of the firms, meaning that consumers don't pay for privacy. This case is equivalent to the one in which the privacy cost c_0 is extremely large. We use the subgame Nash equilibrium as the solution concept.

4.1 The subgames

Before deriving the equilibrium of the firms' information strategy, we first deal with the subgames under different data strategies. In stage 2, both firms make their information decisions. The strategy set is $S_j = \{B(\text{buy}), N(\text{not buy})\}$, $j \in \{H, L\}$. Thus, there are four combinations of strategies, i.e., (N, N), (B, N), (N, B) and (B, B).

4.1.1 Neither firm has information: (N, N)

We first study the case in which neither firm has consumer preference information. In this case, the firms only need to set their basic prices, which are the same for both market segments (i.e., old and new customers).

Now we consider consumers' purchase decisions. Buying from H leads to the utility level

$$V + \theta q_H - p_H, \quad (1)$$

whereas buying from L leads to

$$V + \theta q_L - p_L. \quad (2)$$

Here, p_H and p_L are the prices that the two firms charge for their products. In both the anonymous market and the personalized market, there are two consumer indifference points θ_N and θ_O . The market shares for H are $[0, \theta_N]$ and $[0, \theta_O]$ while the market shares for L are $[\theta_N, 1]$ and $[\theta_O, 1]$. Then we have

$$\theta_N(p_H, p_L) = \theta_O(p_H, p_L) = \frac{p_H - p_L}{\Delta}. \tag{3}$$

Next we consider the firms' pricing decisions. Firm H's profit is

$$\pi_H = \int_{\theta_N}^1 (p_H - c)d\theta + \int_{\theta_O}^1 (p_H - c)d\theta \tag{4}$$

and firm L's profit is

$$\pi_L = \int_0^{\theta_N} p_L d\theta + \int_0^{\theta_O} p_L d\theta \tag{5}$$

We can easily prove that both π_H and π_L are concave functions. By the first order conditions, we derive the optimal prices $p_H = \frac{2(c+\Delta)}{3}$ and $p_L = \frac{c+\Delta}{3}$. The two firms' equilibrium profits are $\pi_H = \frac{2(c-2\Delta)^2}{9\Delta}$ and $\pi_L = \frac{2(c+\Delta)^2}{9\Delta}$, respectively. The total consumer surplus is $CS = 2V_0 + \frac{c^2}{9(q_H - q_L)} + \frac{-2q_H + 11q_L - 10c}{9}$.

4.1.2 Both firms have information: (B, B)

When both firms choose to buy consumer preference information, firm H and firm L will compete for every customer in the personalized market. We use $p_H(\theta)$ and $p_L(\theta)$ to denote the tailored prices set by the two firms. Following Choudhary (2005) and Tayler (2014), the tailored prices in the old market are given by

$$\begin{aligned} p_L(\theta) &= c - \theta\Delta, \\ p_H(\theta) &= \theta\Delta. \end{aligned}$$

In the anonymous market, the indifference point between the two firms is given by

$$\theta_N(p_H, p_L) = \frac{p_H - p_L}{\Delta}.$$

Similarly, in the personalized market,

$$\theta_O(p_H(\theta), p_L(\theta)) = \frac{p_H(\theta) - p_L(\theta)}{\Delta} = \frac{c}{\Delta}. \tag{6}$$

The profits of the two firms are given by

$$\pi_H = \int_{\theta_N}^1 (p_H - c)d\theta + \int_{\mu}^1 (p_H(\theta) - c)d\theta. \tag{7}$$

$$\pi_L = \int_0^{\theta_N} p_L d\theta + \int_0^{\mu} p_L(\theta) d\theta. \tag{8}$$

We can easily prove that both π_H and π_L are concave functions.

Proposition 1 Assume that both firms purchase the information and that consumers do not pay for privacy. Then the basic prices are $p_H = \frac{2(c+\Delta)}{3}$ and $p_L = \frac{c+\Delta}{3}$, while the tailored prices are $p_L(\theta) = c - \theta\Delta$ and $p_H(\theta) = \theta\Delta$. The profits are $\pi_H = \frac{11c^2-26c\Delta+17\Delta^2}{18\Delta}$ and $\pi_L = \frac{11c^2+4c\Delta+2\Delta^2}{18\Delta}$. And the total consumer surplus is $CS = 2V_0 - \frac{4c^2}{9(q_H-q_L)} + \frac{-q_H+10q_L+5c}{9}$.

Proposition 1 gives the equilibrium pricing strategy when both firms choose to buy information. We find that the possession of consumer information does not influence the firms’ basic price. The only change in profits occurs in the personalized market. A comparison with the no-information case yields the following results.

Corollary 1 There exist two thresholds $\underline{\mu}$ and $\bar{\mu}$.

- (1) When $\mu \in \left[0, \underline{\mu}\right)$, $\pi_H^{BB} > \pi_H^{NN}$ and $\pi_L^{BB} < \pi_L^{NN}$;
- (2) When $\mu \in \left[\underline{\mu}, \bar{\mu}\right)$, $\pi_H^{BB} \leq \pi_H^{NN}$ and $\pi_L^{BB} \leq \pi_L^{NN}$;
- (3) When $\mu \in \left[\bar{\mu}, 1\right]$, $\pi_H^{BB} \leq \pi_H^{NN}$ and $\pi_L^{BB} > \pi_L^{NN}$.

Corollary 1 suggests that the profits are not always lower when both firms have bought consumer information than when they have no information. A traditional perspective would suggest that the possession of information by both sides engenders fierce competition and therefore reduce profits. As Corollary 1 shows, when μ is in the middle range, profits in the BB case are indeed less than in the no-information case. However, when the quality-adjusted cost μ is low enough, firm H’s profits rise when both firms possess consumer information. Conversely, when μ is high enough, firm L’s profits rise. This is explained by the fact that when μ is low, firm L’s price adjustment range is narrow, so firm H can occupy a greater share of the personalized market; but when μ is large enough, firm L gains a competitive advantage and reaps higher profits in the personalized market. Increased profits compensate for the losses caused by the fierce competition between the two firms.

4.1.3 Only firm H has information: (B, N)

Here we focus on the case in which only firm H chooses to purchase the consumer preference information. In this case, firm H sets a basic price in the anonymous market and a tailored price in the personalized market, while firm L offers the same basic price in all market segments.

The utility function of purchasing from H or L is the same as in the no-information case. In the anonymous market, the indifference point between the two firms is given by

$$\theta_N(p_H, p_L) = \frac{p_H - p_L}{\Delta}.$$

In the personalized market, firm H offers a tailored price $p_H(\theta)$ which leaves consumers indifferent as to which firm’s product they should purchase. Firm H’s tailored price is given by

$$p_H(\theta) = p_L + \theta\Delta. \tag{9}$$

We denote the last consumer purchasing from firm H as θ_1 , which is reached when $p_H(\theta) = c$. Then we have $\theta_1(p_L) = \frac{c-p_L}{\Delta}$.

Next we consider firms’ profit maximization decisions. Firm H’s profits are given by

$$\pi_H = \int_{\theta_N}^1 (p_H - c)d\theta + \int_{\max\{\theta_1, 0\}}^1 (p_H(\theta) - c)d\theta \tag{10}$$

while firm L’s profits are

$$\pi_L = \int_0^{\theta_N} p_L d\theta + \int_0^{\max\{\theta_1, 0\}} p_L d\theta. \tag{11}$$

There are two possible cases. In the first case, Firm L sets its basic price p_L low enough to satisfy $\theta_1(p_L) \geq 0$. In this way, firm L maintains its market share in the old market. When $\theta_1(p_L) < 0$, firm L chooses to abandon the personalized market. This means that firm L can decide whether or not to keep a share of the personalized market. It can set a lower price to stay in this market or set a higher price to focus on the anonymous market.

Proposition 2 *Assume that only firm H buys the information and that consumers do not pay for privacy. There exists a threshold $\mu_1 < \frac{1}{2}$ for μ .*

(1) *When $\mu \in [0, \mu_1)$, the equilibrium basic prices are*

$$p_H = \frac{2(c + \Delta)}{3} \text{ and } p_L = \frac{c + \Delta}{3}.$$

The profits of the two firms are given respectively by

$$\pi_H = \frac{2c^2 - 20c\Delta + 23\Delta^2}{18\Delta} \text{ and } \pi_L = \frac{(c + \Delta)^2}{9\Delta}.$$

Total consumer surplus is

$$CS = 2V_0 + \frac{c^2}{18(q_H - q_L)} + \frac{-8q_H + 26q_L - 16c}{18}.$$

(2) *When $\mu \in [\mu_1, 1]$, the equilibrium basic prices are*

$$p_H = \frac{5c + 4\Delta}{7} \text{ and } p_L = \frac{3c + \Delta}{7},$$

The two firms’ profits are given respectively by

$$\pi_H = \frac{12(c - 2\Delta)^2}{49\Delta} \text{ and } \pi_L = \frac{2(3c + \Delta)^2}{49\Delta},$$

Total consumer surplus is

$$CS = 2V_0 + \frac{4c^2}{49(q_H - q_L)} + \frac{-6q_H + 55q_L - 50c}{49}.$$

Proposition 2 gives the equilibrium pricing strategy when only firm H chooses to buy consumer information. We find that there are two equilibrium outcomes, depending on the value of the quality-adjusted cost μ . When μ is not too large, the equilibrium strategy is reached when the firm L does not operate in the personalized market. The equilibrium prices exactly equal those in the no-information case. When μ is high enough, firm L maintains a share in the personalized market in equilibrium. Firm L sets a lower basic price, while firm H

charges more than in the no-information case. The respective changes in profits are intuitive. Firm H’s equilibrium profit is greater than in the NN case, while firm L’s profit decreases for all possible values of μ . The possession of consumers’ personal data endows firm H with the ability to extract more consumer surplus and to gain a competitive advantage over firm L. As a result, firm L loses much of its share in personalized market.

4.1.4 Only firm L has information: (N, B)

Now we consider the case in which only firm L chooses to purchase the consumer preference information. In this case, firm L sets a basic price in the anonymous market and a tailored price in the personalized market, while firm H offers the same basic price in all market segments.

The utility function of purchasing from H or L is the same as in the no-information case. And the consumer’s indifference point between H and L in the anonymous market is still given by

$$\theta_N(p_H, p_L) = \frac{p_H - p_L}{\Delta}.$$

In the personalized market, firm L offers a tailored price $p_L(\theta)$ that just undercuts firm H. Then, firm H’s tailored price is given by

$$p_L(\theta) = p_H - \theta \Delta q. \tag{12}$$

Similarly, we denote L’s last consumer as θ_2 , which can be derived through $p_L(\theta) = 0$. Then we have $\theta_2(p_H) = \frac{p_H}{\Delta}$.

Next, we consider the firms’ basic price decisions. Firm H’s profits are

$$\pi_H = \int_{\theta_N}^1 (p_H - c) d\theta + \int_{\min\{\theta_2, 1\}}^1 (p_H - c) d\theta \tag{13}$$

while firm L’s profits are

$$\pi_L = \int_0^{\theta_N} p_L d\theta + \int_0^{\min\{\theta_2, 1\}} p_L(\theta) d\theta. \tag{14}$$

Here too there are two possible cases. When firm H’s price is high enough that $\theta_2 > 1$, H may choose to focus on the anonymous market and abandon the personalized market. When p_H is not too large, such that $\theta_2 \leq 1$, firm H can still occupy some personalized market share. Obviously, firm H’s different choices have different market outcomes. The firms maximize their profits as expressed in Eqs. (13) and (14), and we have the next proposition.

Proposition 3 *Assume that only firm L buys the information and that consumers do not pay for privacy. There exists a threshold $\mu_2 > \frac{1}{2}$ for μ .*

(1) *When $\mu \in [0, \mu_2)$, the equilibrium basic prices are*

$$p_H = \frac{4(c + \Delta)}{7} \text{ and } p_L = \frac{2(c + \Delta)}{7},$$

The two firms’ profits are given respectively by

$$\pi_H = \frac{2(3c - 4\Delta)^2}{49\Delta} \text{ and } \pi_L = \frac{12(c + \Delta)^2}{49\Delta}.$$

The total consumer surplus is

$$CS = 2V_0 + \frac{2c^2}{49(q_H - q_L)} + \frac{-5q_H + 54q_L - 52c}{49}.$$

(2) When $\mu \in [\mu_2, 1]$, the equilibrium basic prices are

$$p_H = \frac{2(c + \Delta)}{3} \text{ and } p_L = \frac{c + \Delta}{3},$$

The two firms' profits are given respectively by

$$\pi_H = \frac{(c - 2\Delta)^2}{9\Delta} \text{ and } \pi_L = \frac{4c^2 + 20c\Delta + 7\Delta^2}{18\Delta},$$

Total consumer surplus is

$$CS = 2V_0 + \frac{c^2}{18(q_H - q_L)} + \frac{-5q_H + 23q_L - 22c}{18}.$$

Proposition 3 gives the equilibrium pricing strategy when only firm L chooses to buy consumer information. There still exist two equilibrium outcomes, depending on the value of the quality-adjusted cost μ . When μ is large enough, firm H finds it advantageous to give up its share in the personalized market and set a correspondingly high basic price to guarantee its profits. In this case, the equilibrium prices are exactly the same as in the no-information case. When μ is not too high, firm H finds it more profitable to remain in the personalized market. It is correspondingly required to reduce its profit margin in exchange for a share in the "old market". Thus, firm H sets a lower price than in the first case. Furthermore, comparing the firms' profits in the NN and NB cases, we find that the equilibrium profit for firm H is lower and for firm L is higher than in the no-information case. The use of consumer data enables firm L to extract more consumer surplus and increase its profits. Correspondingly, firm H loses part of its market share in the personalized market under the pressure of firm L's competition.

4.2 Data supplier's sales strategy

Now we can study the data supplier's information sales strategy. The data supplier (DS) can use various information technologies to collect old consumers' personal information and sell it to retailers who can take full advantage of it. The DS may post a price K to maximize its profits. As a monopoly agent here, we suppose that this DS has the exclusive power to bargain for the entire transaction surplus.

We define the strategy set $M = \{H, L, HL\}$ and the corresponding price $K_m, m \in M$. This means that the DS chooses one of three sales strategies: to sell exclusively to firm H, exclusively to firm L or to both firms.

When the DS sells information exclusively to firm H or firm L, the price is given by the difference between the firm's profit when it has exclusive possession of consumer information and its profit when its competitor has this advantage. Hence, when the DS sells exclusively to firm H, the data price is given by

$$K_H = \pi_H^{BN} - \pi_H^{NB},$$

and when it sells exclusively to firm L, the data price is given by

$$K_L = \pi_L^{NB} - \pi_L^{BN}.$$

If the DS chooses to sell information to both firms, the price is given by the sum of the two firms’ profit differentials when they both acquire the information and each firms’ profit when its competitor acquires the information exclusively. Thus, the data price is

$$K_{HL} = (\pi_H^{BB} - \pi_H^{NB}) + (\pi_L^{BB} - \pi_L^{BN}).$$

By comparing these prices, we derive the DS’s optimal information selling strategy. The optimal selling price is given by

$$\pi_{DS} = \max\{K_H, K_L, K_{HL}\}.$$

Proposition 4 *The optimal information-selling strategies for DS are as follows:*

- (1) *When $\mu \leq \frac{1}{2}$, the optimal strategy is to sell exclusively to firm H. The information prices in this case are*

$$K_H = \begin{cases} -\frac{226c^2+116c\Delta-551\Delta^2}{882\Delta}, & 0 \leq \mu < \mu_1 \\ \frac{16\Delta^2-6c^2}{49\Delta} & \mu_1 \leq \mu \leq \frac{1}{2} \end{cases},$$

- (2) *When $\mu > \frac{1}{2}$, the optimal strategy is to sell exclusively to firm L. The information prices in this case are*

$$K_L = \begin{cases} -\frac{6c^2+12c\Delta+10\Delta^2}{49\Delta}, & \frac{1}{2} < \mu < \mu_2 \\ -\frac{128c^2+764c\Delta+307\Delta^2}{882\Delta}, & \mu_2 \leq \mu \leq 1 \end{cases}.$$

Proposition 4 gives the optimal sales strategy for the DS. It is always optimal to sell exclusively to a single firm. When μ is relatively small, it is advantageous for the DS to sell exclusively to firm H. When μ is high enough, the DS should choose to offer information services only to firm L. This can be explained by the fact that the provision of information to both firms would intensify competition in the personalized market. To compete for customers, each firm would set a correspondingly low price, minimizing the surplus that the DS could potentially extract. To maximize the value of the information it offers for sale, the DS must identify the firm that can extract more profit from consumers. For this purpose, it might conduct an auction in which firms bid for an exclusive supply of consumer information.

5 Equilibrium results with transparency of price discrimination

Now we consider the case in which consumers are aware of the firms’ personalized pricing behavior because the firms apply a policy of transparency and inform consumers of their price discrimination. This may give consumers an incentive to pay a monetary cost to conceal their identity. With the information offered by data suppliers, firms can acquire information about consumer preferences and implement personalized pricing in the old market. However, consumers who know about this information strategy can decide whether or not to frustrate it by paying to hide their identities. To understand this case, we analyze the Subgame Nash Equilibrium in the following sections.

5.1 Subgames

The strategy sets for the two firms are the same as Sect. 4. In game stage 2, both firms decide whether or not buy consumer data from the DS. However, in stage 3, consumers can decide whether or not to make efforts to avoid firm’s price discrimination in the personalized market. So one of the four subgames, the case in which neither firm has information, yields the same results as in Sect. 4.1.1. Here we ignore the NN case here and focus on the other three cases.

5.1.1 Both firms have information: (B, B)

In this section, we consider a scenario in which both firms choose to buy consumer preference information and compete in the two market segments. We find that there is no equilibrium when consumers choose to pay a concealing cost $c_0 > 0$. In other words, no one would pay the concealing cost knowing that both firms purchase information. We first denote that $\mu_0 = \frac{c_0}{\Delta}$.

Lemma 1 *When both firms have bought the information and set a tailored price for customers, no one will choose to pay for privacy.*

Note that, in the personalized market, the competition between two vertically differentiated firms intensifies when both firms possess the consumer data because the data enables them to compete for every single customer. As a result, the tailored price set for old consumers is relatively low. It follows that consumers seeking to maximize their utility have no incentive to conceal their identity but in fact have an incentive to reveal it in order to benefit from the low price. Lemma 1 shows that the competitive advantages of the two companies offset each other when both firms possess the valuable consumer data. Proposition 1 still holds in this scenario, as no consumers pay for privacy.

5.1.2 Only firm H has information: (B, N)

Here, as in Sect. 4.1.3, we consider the case in which the DS sells information exclusively to firm H. Some repeat consumers may want to pay a privacy cost in order to avoid being tracked and charged a personalized price. Since only firm H has consumer preference information, if consumers choose to buy firm H’s product, they need to decide whether to accept a tailored price or to pay the basic price and an additional privacy cost. If consumers choose to buy firm L’s product, they simply pay the basic price.

Following Sect. 4.1.3, in the anonymous market, the consumer indifference point between the two firms is given by

$$\theta_N(p_H, p_L) = \frac{p_H - p_L}{\Delta}.$$

In the personalized market, firm H’s tailored price is $p_H(\theta) = p_L + \theta\Delta$. We denote the last consumer purchasing from H as θ_1 , which is reached when $p_H(\theta) = c$. Then we obtain

$$\theta_1(p_L) = \frac{c - p_L}{\Delta}. \tag{15}$$

However, the utility of a consumer who receives the tailored price is $V + \theta q_H - p_H(\theta)$ and the utility for one who chooses to conceal is $V + \theta q_H + p_H - c_0$. The consumer indifference point between concealing and revealing is given by

$$V + \theta q_H + p_H - c_0 = V + \theta q_H - p_H(\theta),$$

which can be derived as

$$\theta_H^c(p_H, p_L) = \frac{p_H - p_L + c_0}{\Delta}. \tag{16}$$

Consumers who have $\theta > \theta_c^H$ choose to pay for privacy. When $\theta \leq \theta_c^H$, consumers may choose to reveal their identity.

Next, we consider the firms’ profit maximization decisions. Firm H’s profits are given by

$$\pi_H = \int_{\theta_N}^1 (p_H - c) d\theta + \int_{\max\{\theta_1, 0\}}^{\min\{\theta_c^H, 1\}} [p_H(\theta) - c] d\theta + \int_{\min\{\theta_c^H, 1\}}^1 (p_H - c) d\theta \tag{17}$$

and firm L’s profits are

$$\pi_L = \int_0^{\theta_N} p_L d\theta + \int_0^{\max\{\theta_1, 0\}} p_L d\theta. \tag{18}$$

Both firms maximize the profits expressed in Eqs. (17) and (18), and we obtain the equilibrium results in the following proposition.

Proposition 5 *Assume that only firm H buys the information and that consumers can pay for privacy. The equilibriums are as follows:*

- (1) *When $0 \leq \mu < \frac{2}{3}$ and $0 < \mu_0 \leq \frac{1}{2} - \frac{1}{4}\mu$, the equilibrium basic prices are*

$$p_H = \frac{c + 2\Delta}{2} \text{ and } p_L = \frac{c + 2\Delta}{4}.$$

The firms’ profits are

$$\pi_H = \frac{3c^2 - 36c\Delta + 44\Delta^2 + 16c_0^2}{32\Delta} \text{ and } \pi_L = \frac{(c + 2\Delta)^2}{16\Delta}.$$

Total consumer surplus is

$$CS = \frac{c^2 + 8c_0^2 + 4c_0(c - 2q_H + 2q_L)}{16(q_H - q_L)} + \frac{1}{4}(-3c - 3q_H + 7q_L) + 2V_0.$$

- (2) *When $\frac{2}{3} \leq \mu \leq 1$ and $0 < \mu_0 \leq \frac{2}{5} - \frac{1}{5}\mu$, the equilibrium basic prices are*

$$p_H = \frac{3c + 4\Delta}{5} \text{ and } p_L = \frac{2c + \Delta}{5}.$$

The firms’ profits are

$$\pi_H = \frac{12(c - 2\Delta)^2 + 25c_0^2}{50\Delta} \text{ and } \pi_L = \frac{2(2c + \Delta)^2}{25\Delta}.$$

Total consumer surplus is

$$CS = \frac{2c^2 + 25c_0^2 + 10c_0(c - 2q_H + 2q_L)}{50(q_H - q_L)} - \frac{1}{25}(24c + 6q_H - 31q_L) + 2V_0.$$

If μ_0 is larger than the upper bound, the equilibrium is as shown in Proposition 2.

Proposition 5 shows the result of the tradeoff by firms when only firm H has consumers’ personal information and consumers can pay for privacy. Figure 1 shows the regions in which

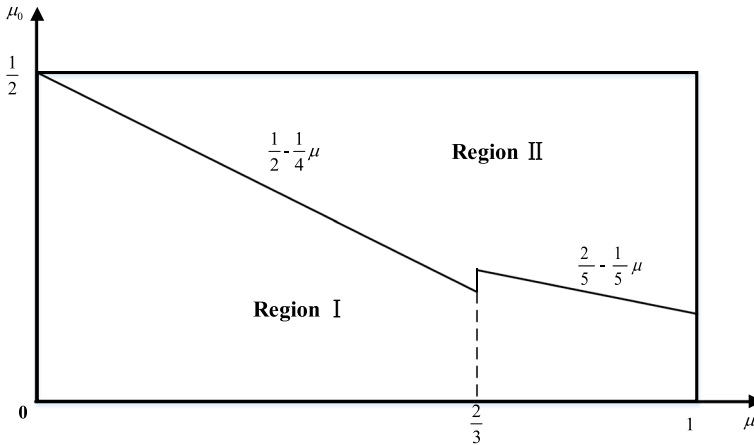


Fig. 1 The regions with and without transparency in the BN case

consumers will or will not pay for privacy in the BN case. In region I, consumers may choose to pay for privacy, while in region II, no one will pay for it. There are two equilibriums in the BN case. When μ is not too large, it is better for firm L to abandon the old market and post a relatively high basic price. Firm H then holds the entire personalized market. When μ is large enough—for example, when $\mu > \frac{2}{3}$ —in equilibrium firm L still retains a share of the old market. As for the impact of consumer privacy, there are upper bounds to the equilibriums. When c_0 exceeds a certain range, the equilibrium is the same as in Proposition 2 because no consumers will pay the privacy cost.

5.1.3 Only firm L has information: (N, B)

Next we consider the case in which the DS sells information exclusively to firm L. Following Sect. 4.1.4, we use the same denotation. In the anonymous market, the consumer indifference point is given by

$$\theta_N(p_H, p_L) = \frac{p_H - p_L}{\Delta}.$$

And in the personalized market, firm L’s tailored price is $p_L(\theta) = p_H - \theta \Delta q$. The last consumer choosing to buy product from firm L is given by the equation $p_L(\theta) = 0$, which is such that

$$\theta_2(p_H) = \frac{p_H}{\Delta}.$$

However, consumers have to decide whether to pay for privacy. The indifference point between concealing and revealing is given by

$$V + \theta q_L - p_L - c_0 = V + \theta q_L - p_L(\theta),$$

which can be derived as

$$\theta_L^c(p_H, p_L) = \frac{p_H - p_L - c_0}{\Delta}. \tag{19}$$

Next, we consider firms’ profit maximization decisions. The profits of the two companies are

$$\pi_H = \int_{\theta_N}^1 (p_H - c)d\theta + \int_{\min\{\theta_2, 1\}}^1 (p_H - c)d\theta. \tag{20}$$

$$\pi_L = \int_0^{\theta_N} p_L d\theta + \int_{\max\{\theta_L^c, 0\}}^{\min\{\theta_2, 1\}} p_L(\theta)d\theta + \int_0^{\max\{\theta_L^c, 0\}} p_L d\theta. \tag{21}$$

Then both firms maximize the profits expressed in Eqs. (20) and (21), and we obtain the equilibrium results in the following proposition.

Proposition 6 *Assume that only firm L buys the information and consumers can pay for privacy. The equilibriums are as follows:*

(1) *When $0 \leq \mu \leq \frac{1}{3}$ and $0 < \mu_0 \leq \frac{1}{5} + \frac{1}{5}\mu$, the equilibrium basic prices are*

$$p_H = \frac{3(c + \Delta)}{5} \text{ and } p_L = \frac{2(c + \Delta)}{5}.$$

The two firms’ profits are

$$\pi_H = \frac{2(2c - 3\Delta)^2}{25\Delta} \text{ and } \pi_L = \frac{12(c + \Delta)^2 + 25c_0^2}{50\Delta}.$$

Total consumer surplus is

$$CS = \frac{2c^2 + 25c_0^2 - 10c_0(c + q_H - q_L)}{50(q_H - q_L)} - \frac{1}{25}(28c + 4q_H - 29q_L) + 2V_0.$$

(2) *When $\frac{1}{3} < \mu \leq 1$ and $0 < \mu_0 \leq \frac{1}{4} + \frac{1}{4}\mu$, the equilibrium basic prices are*

$$p_H = \frac{3(c + \Delta)}{4} \text{ and } p_L = \frac{c + \Delta}{2}.$$

The two firms’ profits are

$$\pi_H = \frac{(c - 3\Delta)^2}{16\Delta} \text{ and } \pi_L = \frac{3c^2 + 30c\Delta + 11\Delta^2 + 16c_0^2}{32\Delta},$$

Total consumer surplus is

$$CS = \frac{c^2 + 8c_0^2 - 4c_0(c + q_H - q_L)}{16(q_H - q_L)} - \frac{1}{16}(22c + 7q_H - 23q_L) + 2V_0.$$

If μ_0 is larger than the upper bound, then the equilibrium is as shown in Proposition 3.

Proposition 6 shows the result of the tradeoff by firms when only firm H has consumers’ personal information and consumers can pay for privacy. Figure 2 shows the regions in which consumers will or will not pay for privacy in the NB case. In region I, consumers may choose to pay for privacy while in region II, no one will pay for it. As in Proposition 5, there are two equilibriums. When μ is not large—for example, $\mu \leq \frac{1}{3}$ —firm H still competes for market share in the personalized market. As μ becomes larger, it is more beneficial for firm H to focus on the anonymous market in equilibrium. Consumer privacy costs also play an important role in these equilibriums, which have upper bounds. When c_0 exceeds the upper bound, results are as described in Proposition 3.

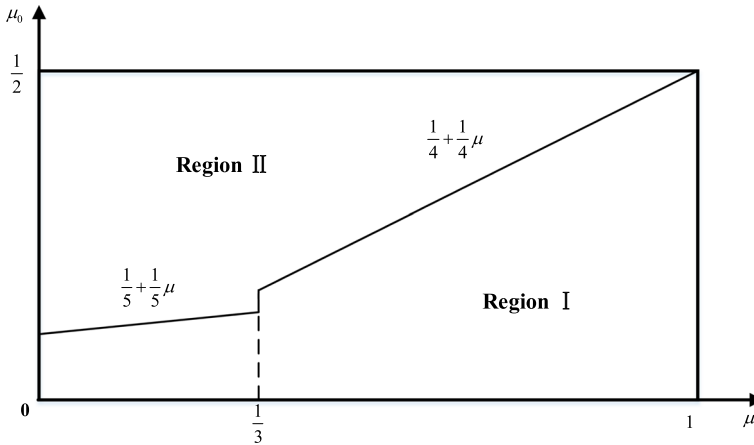


Fig. 2 The regions with and without transparency in the NB case

Our analysis of the cases in which the DS implements an exclusive sales strategy demonstrates that in these cases, the consumer privacy cost influences the profits of the firm that purchases the consumer data, while it makes no difference to the profits of the other firm. This can be explained by the fact that when a firm has exclusive access to consumer data, it sets a tailored price based on its competitor’s basic price and on the preferences of individual consumers. Thus, each consumer in the old market will be offered a personalized price at which he or she is indifferent as to whether to purchase H’s product or L’s product. No matter which product such consumers choose, they will obtain the same utility. On the other hand, a consumer who chooses to pay for privacy and buys from a certain firm should obtain a greater utility than if they bought from the firm’s competitor. Hence the magnitude of the privacy cost affects only the firm that has exclusive access to consumer information.

5.2 Data supplier’s sales and pricing strategy

Now we examine the data supplier’s sales strategy when consumers can pay for privacy. As in Sect. 4.2, the DS has three different strategies to choose from, so the strategy set $M = \{H, L, HL\}$. The DS will compare the profits it can obtain using one of the three and decide which is optimal. Summarizing the discussion in Sect. 5.1, we divide the pricing regions into four, as shown in Fig. 3.

In region I, the privacy cost c_0 is low enough that both the BN and the NB cases are -optimal when consumers can pay for privacy. In region II, the BN case is still optimal, while the NB case becomes equivalent to the case in which no consumers pay for privacy. In region III, the NB case is optimal when consumers may pay for privacy while the BN case becomes equivalent to that in which no consumers pay for privacy. In region IV, the privacy cost is large enough that no consumers would pay for privacy in any case. We can derive different information prices for the different regions. The pricing decisions are given by the next proposition.

Proposition 7 *The optimal sales strategies for the DS are as follows.*

- (1) *When $0 \leq \mu \leq \frac{1}{2}$, the optimal strategy is to sell consumer data exclusively to firm H. The optimal prices are the following:*

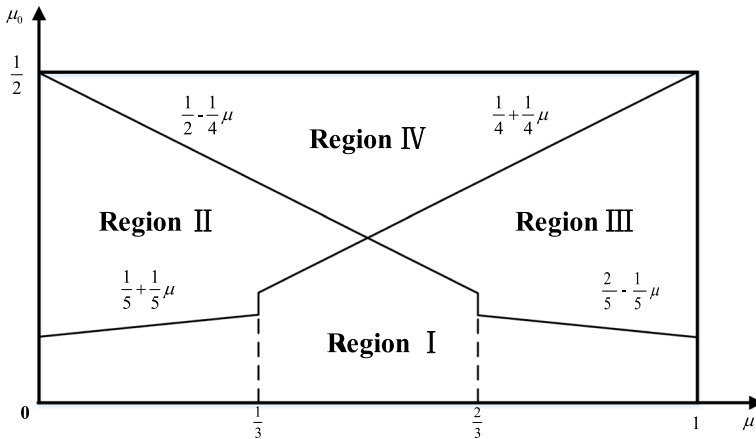


Fig. 3 Pricing Regions

(i) when $0 \leq \mu < \frac{1}{3}$,

$$K = K_H = \begin{cases} \frac{-181c^2 - 132c\Delta + 524\Delta^2 + 400c_0^2}{800\Delta}, & 0 < \mu_0 < \frac{1}{5} + \frac{1}{5}\mu \\ \frac{-429c^2 - 228c\Delta + 1132\Delta^2 + 784c_0^2}{1568\Delta}, & \frac{1}{5} + \frac{1}{5}\mu \leq \mu_0 \leq \frac{1}{2} - \frac{1}{4}\mu \end{cases}$$

(ii) when $\frac{1}{3} \leq \mu \leq \frac{1}{2}$,

$$K = K_H = \begin{cases} \frac{c^2 - 24c\Delta + 26\Delta^2 + 16c_0^2}{32\Delta}, & 0 < \mu_0 < \frac{1}{4} + \frac{1}{4}\mu \\ \frac{-429c^2 - 228c\Delta + 1132\Delta^2 + 784c_0^2}{1568\Delta}, & \frac{1}{4} + \frac{1}{4}\mu \leq \mu_0 \leq \frac{1}{2} - \frac{1}{4}\mu \end{cases}$$

(2) When $\frac{1}{2} < \mu \leq 1$, the optimal strategy is to sell exclusively to firm L. The information prices are

(i) when $\frac{1}{2} < \mu < \frac{2}{3}$,

$$K = K_L = \begin{cases} \frac{71c^2 + 92c\Delta - 4\Delta^2 + 200c_0^2}{400\Delta}, & 0 < \mu_0 < \frac{1}{2} - \frac{1}{4}\mu \\ \frac{-429c^2 + 1086c\Delta + 475\Delta^2 + 784c_0^2}{1568\Delta}, & \frac{1}{2} - \frac{1}{4}\mu \leq \mu_0 \leq \frac{1}{4} + \frac{1}{4}\mu \end{cases}$$

(ii) when $\frac{2}{3} \leq \mu \leq 1$,

$$K = K_L = \begin{cases} \frac{-181c^2 + 494c\Delta + 211\Delta^2 + 400c_0^2}{800\Delta}, & 0 < \mu_0 < \frac{2}{5} - \frac{1}{5}\mu \\ \frac{-429c^2 + 1086c\Delta + 475\Delta^2 + 784c_0^2}{1568\Delta}, & \frac{2}{5} - \frac{1}{5}\mu \leq \mu_0 \leq \frac{1}{4} + \frac{1}{4}\mu \end{cases}$$

When μ_0 exceeds the upper bound, the result is shown in Proposition 3.

Proposition 6 shows that when consumers can pay for privacy, the DS’s optimal strategy is still to sell exclusively to one firm, whether firm H or firm L. When the quality-adjusted cost is not high—i.e., when $\mu \leq \frac{1}{2}$ —it is always better for the DS to offer consumer data exclusively to firm H. When μ exceeds $\frac{1}{2}$, selling to firm L is always preferable. Remarkably, the DS’s optimal selling price is always positively related to the consumer privacy cost c_0 . Intuitively, when the ability to conceal their identity is available to consumers, only those who have a greater preference for quality will choose concealment. As the cost c_0 becomes

larger, fewer consumers are willing to pay for privacy. Thus, the firm with access to consumer data can extract more consumer surplus, and the price of the data offered by the DS will be correspondingly higher. When c_0 exceeds the upper bounds of the constraints, the results will be the same as in Proposition 3.

5.3 Sensitivity analysis and managerial discussion

In this section, we investigate the effects of the privacy cost c_0 on data pricing decisions in equilibrium, as well as on the firms' profits and on consumer surplus.

5.3.1 Firms' Profit

First, we examine the effects of c_0 on the firms' equilibrium profits for different cases in which consumers can pay for privacy. In the second stage of the game, two vertically differentiated firms decide whether to purchase information from the DS. We extend the investigation into four different cases: NN, BN, NB and BB. When no firm has consumer information (the NN case), it is obvious that profits are not related to c_0 . As for the BB case, we have proved in Sect. 5.5.1 that no one will pay for privacy when both firms have information. Therefore, in this section, we focus on the profit implications of the BN and NB cases, in which a single firm has purchased exclusive access to consumer data.

Corollary 2 *Assuming that only firm H/L has the data and that consumers can pay to protect their privacy, the profit of firm H/L increases with c_0 .*

As this corollary indicates, if only one firm has purchased consumer preference information and consumers can pay for privacy, the firm's profits always increase with the privacy cost, no matter whether the products of the firm are high-quality or low-quality. This is attributable to the fact that when access to consumer information is exclusive—for example, when only firm H has such data—repeat customers who choose to buy from firm H have to decide whether to pay the privacy cost. Only those who can increase their utility by paying this cost will do so. As the privacy cost grows larger, fewer consumers will pay it. The firm with access to data—and only that firm—will be able to extract the largest possible surplus from each consumer based on its knowledge of their behavior. The firm's profits will increase correspondingly. To illustrate our results in Corollary 2, we assume that $c = 0.4$, $\Delta = 1$, and plot Figs. 4 and 5 to show the effects of the privacy cost on the profits of the two firms.

The implications drawn from Figs. 4 and 5 are consistent with the results in Corollary 2. When consumers can pay for privacy and the cost of doing so is not too large, the firm with access to consumer information will achieve higher profits as the privacy cost increases, while its competitor's profits are always linear. Moreover, when the privacy cost c_0 is too large, the result is similar to the BN case in Sect. 4.1.3: no consumers choose to pay for privacy.

5.3.2 Consumer surplus

Next, we investigate how the privacy cost c_0 influences the consumer surplus in four equilibrium cases.

Corollary 3 *If only firm H/L has consumer data and consumers have the option of paying a privacy cost, the consumer surplus always declines as c_0 increases.*

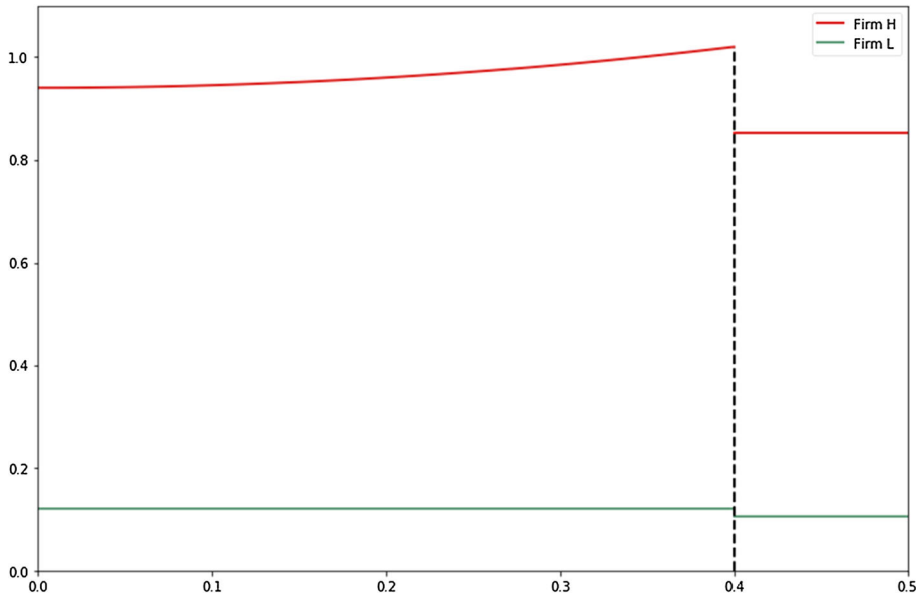


Fig. 4 The effects of c_0 on firms' profits in BN case

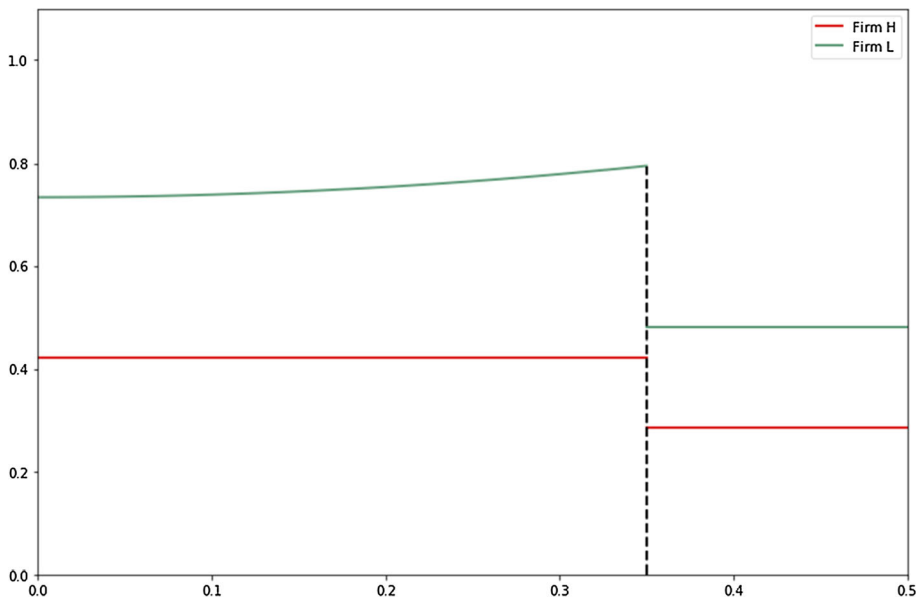


Fig. 5 The effects of c_0 on firms' profits in NB case

A higher privacy cost means that it is harder for consumers to protect their personal information and conceal their identities. The firm with exclusive data access sets a personalized price for each customer that exactly matches the customer's preference. In this way, the firm is able to extract more consumer surplus than it would if it set a uniform basic price. As

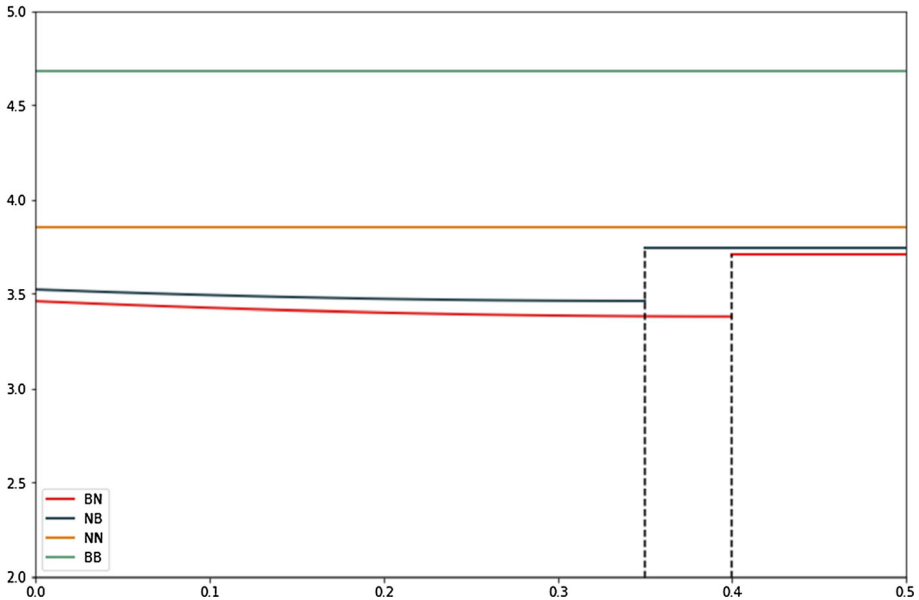


Fig. 6 The impact of c_0 on Consumer Surplus

the privacy cost increases, fewer consumers find it profitable to conceal their identities. The firm with exclusive data access will then be able to set more personalized prices and extract more consumer surplus so as to increase its profits. Thus, the consumer surplus decreases. To illustrate the analytic result in Corollary 3, we perform a numerical analysis with the settings $c = 0.4$, $\Delta = 1$, $q_H = 1.5$, $q_L = 0.5$ and $V = 2$. The results, which are consistent with Corollary 3, are presented in Fig. 6.

Figure 6 illustrates the impact of the privacy cost c_0 on consumer surplus in four different cases. In the NN and BB cases, the consumer surplus is linear because no consumers pay the privacy cost. In the cases where one firm has exclusive data access, BN and NB, we can readily determine that the consumer surplus decreases as c_0 rises, so long as c_0 does not exceed a certain threshold. Beyond this threshold, the result is the same as for the case in which no consumers pay for privacy: consumer surplus is linear.

5.3.3 Social welfare

We next investigate the effects of the privacy cost c_0 on a utilitarian social welfare function. We weight the firms' profits and the consumer surplus equally to evaluate the total social welfare. From the previous discussion, we know that the profits and the consumer surplus are unaffected by the privacy cost in the NN and BB cases because in these cases no one is willing to pay for privacy. Thus, the social welfare (SW) here is also linear with respect to c_0 . As in the BN and NB cases, when the privacy cost is not too large, the profits of the firm with exclusive access to consumer data increase with c_0 , while the consumer surplus (CS) declines as c_0 grows. So the SW here will vary in a different way with respect to c_0 .

Figure 7 gives us a general understanding of the effect of the privacy cost on social welfare. We plot a graph using the parameter values in 5.3.2 and perform a numerical analysis.

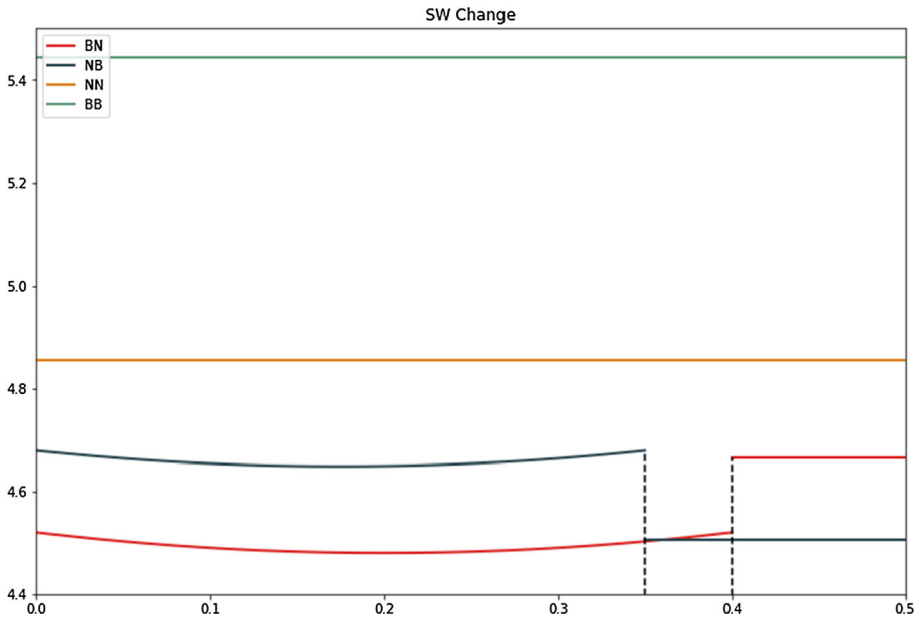


Fig. 7 The impact of c_0 on Social Welfare

For the cases in which only one firm has consumer data, the SW is a fixed value if the privacy cost c_0 exceeds a certain threshold, as no consumers will pay to conceal their identity. However, when the privacy cost is not too large, the welfare is U-shaped with respect to c_0 . There are significant managerial implications here. From the perspective of a policymaker seeking to promote total social welfare, it is always preferable to make privacy protection very cheap or very costly. The reason is not difficult to understand. When the privacy cost is low enough, consumers can protect themselves easily and retain a higher consumer surplus, which raises social welfare by compensating for the firm’s reduced profit. When the privacy cost is high enough, consumers will seldom choose to conceal their identities, and firms will achieve higher profits, which also raises total social welfare.

5.3.4 Information prices

In Sect. 5.2, we determined the information prices with the privacy cost constraint in different regions. The data supplier may choose to sell information to either the high-quality or the low-quality firm. In this section, we examine how the privacy cost influences the information prices offered by the DS.

Corollary 4 *When consumers can pay a cost to protect their privacy, the information price offered by the DS increases with c_0 .*

This result is easily understood. From the discussion in Sect. 5.3.1, we know that the profits of the firm that has exclusive access to consumer data increase with the privacy cost. Fewer consumers will choose to pay for privacy as the privacy cost increases, which means that the firm with exclusive data access will acquire higher profits by charging tailored prices. Similarly, the difference in this firm’s profits between the BN and NB cases also increases,

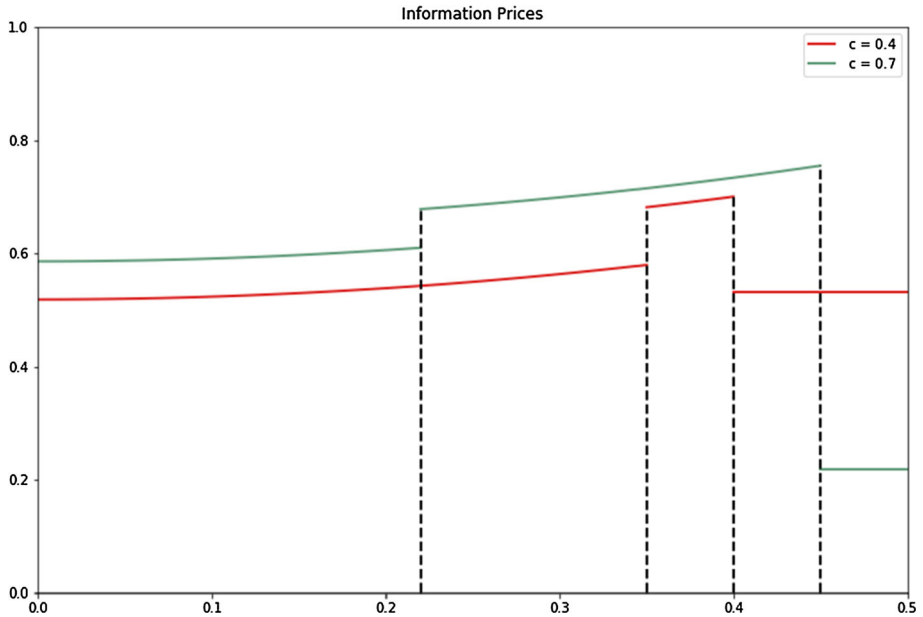


Fig. 8 The impact of c_0 on information prices

which means that the DS can set a higher information price. To illustrate our analytic result in Corollary 4, we plot Fig. 8, showing the effect of the privacy cost on information prices. We use the parameter values in 5.3.2 and perform the numerical analysis for two different cases: $c = 0.4$ and $c = 0.7$.

Figure 4 gives the information price change for the two cases. Following Proposition 7, when $c = 0.4$, $\mu = 0.4$ and the DS sells information exclusively to firm H; when $c = 0.7$, $\mu = 0.7$ and the DS sells the information exclusively to firm L. As shown in Fig. 4, when c_0 is not too large, the information prices always rise as the privacy cost c_0 increases. When c_0 is large enough, Proposition 4 applies and the prices are linear.

6 Effects of transparency in price discrimination

In this section, we compare two cases, with and without mandatory transparency of personalized pricing, to explore how consumers’ awareness of firms’ price discrimination behavior affects industry profits, consumer surplus and social welfare. Following Sects. 4 and 5, we focus on the cases in which c_0 is not very large, because two cases will give the same result when c_0 exceeds the upper bound of the privacy cost. Since transparency makes no difference to the NN and BB cases, we focus solely on the cases in which one firm has exclusive access to information.

6.1 Only firm H has information

First we study the case in which only firm H has access to information.

Proposition 8 *The effects of the transparency of personalized pricing on the BN case are as follows:*

- (1) *Transparency improves industry profits;*
- (2) *Transparency decreases the consumer surplus;*
- (3) *When $0 < \mu \leq \frac{12}{13}$, the transparency of personalized pricing decreases the social welfare.*

Proposition 9 shows that, in the BN case, the transparency of personalized pricing improves the total profits of this industry because it alleviates price competition among companies. Requiring firms to disclose their personalized pricing behavior means that the firm which purchases consumer data cannot fully realize the commercial value of the information by implementing personalized pricing, since consumers will act to protect their privacy. In this way, the industry benefits from data transparency and achieves a better overall outcome. As for consumers, the transparency of firms' pricing behavior results, counterintuitively, in a decreased consumer surplus. On the one hand, it gives some consumers an incentive to pay to protect their privacy. On the other hand, the weakened market competition makes the firms more likely to set higher prices for their products. Similarly, although requiring the disclosure of personalized pricing benefits the overall outcome for the industry, it tends to reduce social welfare, especially when μ is not very large. Although requiring transparency for personalized pricing behavior meets the desire of some consumers to protect their privacy, it also eases competition among companies, so consumers may pay a higher price.

6.2 Only firm L has information

Next we consider the case in which only firm L has access to consumer data.

Proposition 9 *Effects of the transparency of personalized pricing on the NB case:*

- (1) *When $0 \leq \mu < \mu_2$, transparency of personalized pricing improves industry profits; when $\mu_2 \leq \mu \leq 1$, transparency of personalized pricing decreases industry profits;*
- (2) *Transparency of personalized pricing decreases consumer surplus;*
- (3) *When $\frac{1}{13} \leq \mu \leq 1$, transparency of personalized pricing decreases the social welfare.*

These results are similar to those obtained in the NB case. However, unlike Propositions 8, 9 indicates that the transparency of personalized pricing does not always improve total industry profits. There exists a threshold μ_2 . When $\mu < \mu_2$, transparency increases industry profits, while total profits are lower when $\mu \geq \mu_2$. This can readily be explained. From the discussion in Sect. 4.1.4, we know that when $\mu < \mu_2$, firm H keeps competing for “old” consumers while focusing on the new market. Requiring transparency enlarges the anonymous market while diminishing the personalized market. Both firm H and firm L will set correspondingly higher basic prices. Thus, firm L, which has purchased consumer information, will record lower profits, while firm H can realize relatively large profits. The increase in the profits of firm H is greater than the loss of profits experienced by firm L, so overall industry profits will rise. However, when $\mu \geq \mu_2$, firm H competing with firm L in the personalized market. At the same time, the transparency requirement causes firm L to lose a large portion of the personalized market. Although both firms set higher basic prices, the decline in firm L's profits exceeds the growth of firm H's. Meanwhile, the relaxation of competition in the industry makes the two companies more likely to set higher prices so as to achieve higher profits. As a result, consumers face both privacy costs and higher product

prices. Interestingly, we find that unless μ is very small, transparency in personalized pricing also reduces social welfare. In sum, when firms are required to disclose their personalized pricing practices, industry achieves higher profits, while consumers face losses.

From the above discussion, it is clear that, in both the BN and NB cases, a requirement for transparency in personalized pricing affects industry profits, consumer surplus and social welfare. Transparency always increases industry profits when firm H has consumer information. However, in the NB case (firm L has consumer information), there exists a threshold μ_2 . When μ is large enough, transparency decreases industry profits. Furthermore, transparency regulation diminishes consumer surplus in both the BN and NB cases. And when μ is not too large or small—the most common situation in real life—a transparency requirement always leads to lower social welfare. Therefore, from the standpoint of a social planner, mandating transparency in personalized pricing may not be efficient and should be considered circumspectly.

7 Conclusions

The digital revolution has created excellent opportunities for firms to gain a profound understanding of consumer preferences. Technology enables them to collect consumer information and practice price discrimination (or “personalized pricing”) on individual consumers. But the excessive use of information sparks privacy concerns among consumers. Consumers may choose effective ways to protect their personal information, but this comes at a cost. Shoppers in online markets face a trade-off between privacy and the benefits of competitive markets with consumers’ preference information. In this paper, we attempt to understand this trade-off by examining the competition between two vertically differentiated firms and the value for them of having access to consumer information. We also compare the market equilibriums in the cases where both firms have access to consumer data or one firm has exclusive access in order to determine the data supplier’s optimal sales strategy. Finally, we investigate the effect of requiring firms to disclose their personalized pricing on firms’ profits, on consumer surplus and on social welfare. Our results provide new insights into competitive price discrimination between asymmetric retailers and may assist social planners in formulating policies regarding price discrimination, transparency and the protection of consumer privacy.

Three findings are particularly interesting. The first of these concerns the data supplier’s sales strategy. Our analysis shows that, no matter whether consumers can pay for privacy or not, it is always optimal for the DS to sell its consumer information exclusively to one firm. The reason is that when both firms have access to consumer data, they both practice price discrimination, which intensifies competition between them and drives their prices down to each firm’s marginal cost. When the firms’ profits decline, the price of information decreases accordingly, to the detriment of the DS. Of equal importance to data suppliers is our finding that the magnitude of the quality-adjusted cost is the key factor in determining which firm the DS should sell information to. In particular, when the firms do not disclose price discrimination and the quality-adjusted cost difference is smaller than $1/2$, the optimal strategy for the data supplier is to sell information exclusively to the high-quality firm. When the quality-adjusted cost ratio is larger than $1/2$, the data supplier sells information exclusively to the low-quality firm.

Our second salient finding concerns the effect of privacy costs. We find that, when consumers pay a privacy cost which is not too large, the profits of the firms that have exclusive access to consumer information increase with the privacy cost. At the same time, the con-

sumer surplus decreases as the privacy cost grows. This is because the firm with exclusive access will make full use of the data on consumer preferences to practice price discrimination and extract more consumer surplus. As the privacy cost grows, fewer consumers will pay for privacy and the firm with data access can achieve a higher profit in a larger personalized market segment. Correspondingly, as more consumers do not protect their privacy and firms are able to extract more consumer surplus and the total consumer surplus ultimately declines. In addition, the information price set by the DS increases with the privacy cost. As fewer consumers are willing a high privacy cost, more and better consumer data is collected by the DS for the profitable use of the firm that buys it. Therefore the data grows in value and commands a high price.

Finally, we answer the question of how a requirement for the disclosure of firms' personalized pricing behavior affects the total industry profit, consumer surplus and social welfare. By comparing the cases in which consumers are or are not aware of the companies' information collection and price discrimination, we find that transparency improves the total profit of the firms at the cost of consumer surplus when only the high-quality product firm H has access to consumer data. This is because personalized pricing actually intensifies price competition between companies, while requiring transparency alleviates it. As a result, firms that do not purchase consumer information do not need to distort prices downward to maintain their market share. When the low-quality firm L has exclusive data access, transparency does not always increase the total industry profit. There exists a threshold for the quality-adjusted cost. When the cost is large enough, total industry profits will be lower. The reason is that the low-quality firm loses a greater share of the personalized market with the implementation of transparency. We also find that a transparency requirement reduces social welfare when the gap between two vertically differentiated firms is not too large or too small. In other words, our results suggest that mandatory transparency in personalized pricing, a measure intended to protect consumer privacy and enhance social welfare, may not be efficient in most common situations.

This paper has several limitations, and future research may extend it in several ways. First, we assume that firms can implement price discrimination perfectly with the help of third-party information, but in fact, limited information cannot give a comprehensive understanding. Therefore it is worthwhile to study the impact of the accuracy of consumer data on firms' price discrimination decisions. Second, we only consider the case in which firms acquire information from third party data suppliers. Further research could examine the case in which when firms can collect information for themselves and how the operating costs involved may affect the decisions of the firms.

Acknowledgements The authors sincerely thank the editor-in-chief and anonymous reviewers for the constructive comments, which have improved the paper significantly. This research is supported by the National Natural Science Foundation of China (Grant Nos. 71771179 and 71532015).

Appendix

Proof of Proposition 1 When both firms choose to buy information, the tailored price is similar to that found in previous studies (Choudary 2005; Tayler 2014): $p_L(\theta) = c - \theta\Delta$ and $p_H(\theta) = \theta\Delta$. The equilibrium price is given by the following two functions:

$$\begin{aligned} & \operatorname{argmax}_{p_H} \left[[(1 - \theta_N)(p_H - c) + \int_{\mu}^1 (p_H(\theta) - c) d\theta] \right]. \\ & \operatorname{argmax}_{p_L} \left[\theta_N p_L + \int_0^{\mu} p_L(\theta) d\theta \right]. \end{aligned}$$

It is easily proved that both π_H and π_L are concave functions. By the first-order derivatives, we have $\frac{\partial \pi_H}{\partial p_H} = \frac{c - 2p_H + p_L + c}{\Delta} = 0$, $\frac{\partial \pi_L}{\partial p_L} = \frac{p_H - 2p_L}{\Delta} = 0$. Solving the two equations, we $p_H = \frac{2(c + \Delta)}{3}$ and $p_L = \frac{c + \Delta}{3}$. The firms' equilibrium profits are $\pi_H = \frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta}$ and $\pi_L = \frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta}$, respectively. Finally, the total consumer surplus is given by

$$\begin{aligned} CS &= \int_0^{\theta_N} (V + \theta q_L - p_L) d\theta + \int_{\theta_N}^1 (V + \theta q_H - p_H) d\theta \\ &+ \int_0^{\mu} (V + \theta q_L - p_L(\theta)) d\theta + \int_{\mu}^1 (V + \theta q_H - p_H(\theta)) d\theta \\ &= 2V - \frac{4c^2}{9(q_H - q_L)} + \frac{-q_H + 10q_L - 5c}{9}. \end{aligned}$$

Proof of Corollary 1 Compare the profits between the NN case and the BB case in equilibrium. For firm H, we have

$$\begin{aligned} \pi_H^{BB} - \pi_H^{NN} &= \frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta} - \frac{2(c - 2\Delta q)^2}{9\Delta} \\ &= \frac{7c^2 - 10c\Delta + \Delta^2}{18\Delta} \\ &= \frac{\Delta}{18} (7\mu^2 - 10\mu + 1). \end{aligned}$$

We solve the inequality $7\mu^2 - 10\mu + 1 \geq 0$ which is due to $\mu \in [0, 1]$. There exists a threshold $\underline{\mu} = \frac{1}{7}(5 - 3\sqrt{2}) \approx 0.11$. When $\mu < \underline{\mu}$, $\pi_H^{BB} - \pi_H^{NN} > 0$ and when $\mu \geq \underline{\mu}$, $\pi_H^{BB} - \pi_H^{NN} \leq 0$.

For firm L,

$$\begin{aligned} \pi_L^{BB} - \pi_L^{NN} &= \frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta} - \frac{2(c + \Delta)^2}{9\Delta} \\ &= \frac{7c^2 - 4c\Delta - 2\Delta^2}{18\Delta} \\ &= \frac{\Delta}{18} (7\mu^2 - 4\mu - 2). \end{aligned}$$

We also solve the inequality $7\mu^2 - 4\mu - 2 \geq 0$. There exists a threshold $\bar{\mu} = \frac{1}{7}(2 + 3\sqrt{2}) \approx 0.89$. When $\mu \geq \bar{\mu}$, $\pi_L^{BB} \geq \pi_L^{NN}$ and when $\mu < \bar{\mu}$, $\pi_L^{BB} < \pi_L^{NN}$. We now obtain the announced result.

Proof of Proposition 2 When only firm H has information, firm H has the exclusive power to price-discriminate and set a tailored price $p_H(\theta) = p_L + \theta\Delta$. Correspondingly, firm L has to decide whether to set a new price so as to retain its market share.

Case 1. $\theta_1(p_L) \geq 0$. Firm L sets a relatively low price in order to defend its market share. To maximize the two firms' profits, we have the following functions:

$$\begin{aligned} & \operatorname{argmax}_{p_H} [(1 - \theta_N)(p_H - c) + \int_{\theta_1}^1 (p_H(\theta) - c)d\theta]. \\ & \operatorname{argmax}_{p_L} [\theta_N p_L + \int_0^{\theta_1} p_L d\theta]. \end{aligned}$$

We can easily prove that both π_H and π_L are concave functions. By the first-order derivatives, we have $\frac{\partial \pi_H}{\partial p_H} = \frac{c - 2p_H + p_L + c}{\Delta} = 0$, $\frac{\partial \pi_L}{\partial p_L} = \frac{p_H - 4p_L + c}{\Delta} = 0$. Solving the two equations simultaneously, we obtain $p_H = \frac{5c + 4\Delta}{7}$ and $p_L = \frac{3c + \Delta}{7}$. To ensure the necessary condition $\theta_1(p_L) = \frac{c - p_L}{\Delta} \geq 0$, we have $c - \frac{3c + \Delta}{7} \geq 0$ and $\mu = \frac{c}{\Delta} \geq \frac{1}{4}$. The firms' profits are, respectively, $\pi_H = \frac{12(c - 2\Delta)^2}{49\Delta}$ and $\pi_L = \frac{2(3c + \Delta)^2}{49\Delta}$. And the total consumer surplus is given by

$$\begin{aligned} CS &= \int_0^{\theta_N} (V + \theta q_L - p_L) d\theta + \int_{\theta_N}^1 (V + \theta q_H - p_H) d\theta \\ &+ \int_0^{\theta_1} (V + \theta q_L - p_L) d\theta + \int_{\theta_1}^1 (V + \theta q_H - p_H(\theta)) d\theta \\ &= 2V + \frac{4c^2}{49(q_H - q_L)} + \frac{-6q_H + 55q_L - 50c}{49}. \end{aligned}$$

Case 2. $\theta_1(p_L) < 0$. Firm L chooses not to defend its market share. To maximize the two firms' profits, we have the following functions:

$$\begin{aligned} & \operatorname{argmax}_{p_H} \left[(1 - \theta_N)(p_H - c) + \int_0^1 (p_H(\theta) - c)d\theta \right]. \\ & \operatorname{argmax}_{p_L} \theta_N p_L. \end{aligned}$$

We can easily prove that both π_H and π_L are concave functions. By the first-order derivatives, we have $\frac{\partial \pi_H}{\partial p_H} = \frac{c - 2p_H + p_L + c}{\Delta} = 0$, $\frac{\partial \pi_L}{\partial p_L} = \frac{p_H - 2p_L}{\Delta} = 0$. Solving the two equations simultaneously, we have $p_H = \frac{2(c + \Delta)}{3}$ and $p_L = \frac{c + \Delta}{3}$. To ensure the necessary condition $\theta_1(p_L) = \frac{c - p_L}{\Delta} < 0$, we have $c - \frac{c + \Delta}{3} < 0$ and $\mu = \frac{c}{\Delta} < \frac{1}{2}$. The firms' profits are expressed by $\pi_H = \frac{2c^2 - 20c\Delta + 23\Delta^2}{18\Delta}$ and $\pi_L = \frac{(c + \Delta)^2}{9\Delta}$, respectively. And the total consumer surplus is given by

$$CS = \int_0^{\theta_N} (V + \theta q_L - p_L) d\theta + \int_{\theta_N}^1 (V + \theta q_H - p_H) d\theta + \int_0^1 (V + \theta q_H - p_H(\theta)) d\theta$$

$$= 2V + \frac{c^2}{18(q_H - q_L)} + \frac{-8q_H + 26q_L - 16c}{18}.$$

From the discussion above, it follows that when $\mu < \frac{1}{4}$, only case 2 is available, and when $\mu \geq \frac{1}{2}$, only case 1 is available. When $\frac{1}{4} \leq \mu < \frac{1}{2}$, firm L may choose between the two cases. Let us compare them to decide which is preferable. We have

$$\begin{aligned} \pi_L^{case1} - \pi_L^{case2} &= \frac{2(3c + \Delta)^2}{49\Delta} - \frac{(c + \Delta)^2}{9\Delta}. \\ &= \frac{113c^2 + 10c\Delta - 31\Delta^2}{441\Delta}. \\ &= \frac{\Delta}{441} (113\mu^2 + 10\mu - 31). \end{aligned}$$

We solve the inequality $113\mu^2 + 10\mu - 31 \geq 0$ which is due to $\mu \in [0, 1]$. There exists a threshold $\mu_1 = \frac{1}{113}(-5 + 42\sqrt{2}) \approx 0.48$. When $\mu \geq \mu_1$, $\pi_L^{case1} \geq \pi_L^{case2}$ and when $\mu < \mu_1$, $\pi_L^{case1} < \pi_L^{case2}$. Ultimately, the preceding conditions can be expressed as follows: when $0 \leq \mu < \mu_1$, the equilibrium operates as case 2; when $\mu_1 \leq \mu \leq 1$, the equilibrium operates as case 1. Thus we obtain the announced result.

Proof of Proposition 3 When only firm L has information, firm L have the exclusive power to price-discriminate and set a tailored price $p_L(\theta) = p_H - \theta\Delta$. Correspondingly, firm H has to decide whether to set a new price so as to retain its market share.

Case 1. $\theta_2(p_H) \leq 1$. Firm H sets a relatively low price to defend its market share. In order to maximize the two firms' profits, we have the following functions:

$$\begin{aligned} &\operatorname{argmax}_{p_H} [(1 - \theta_N)(p_H - c) + (1 - \theta_2)(p_H - c)]. \\ &\operatorname{argmax}_{p_L} [\theta_N p_L + \int_0^{\theta_2} p_L(\theta) d\theta]. \end{aligned}$$

We can easily prove that both π_H and π_L are concave functions. By the first-order derivatives, we have $\frac{\partial \pi_H}{\partial p_H} = \frac{2\Delta - 4p_H + p_L + 2c}{\Delta} = 0$, $\frac{\partial \pi_L}{\partial p_L} = \frac{p_H - 2p_L}{\Delta} = 0$. Solving the two equations simultaneously, we have $p_H = \frac{4(c+\Delta)}{7}$ and $p_L = \frac{2(c+\Delta)}{7}$. To ensure the necessary condition $\theta_2(p_H) = \frac{p_H}{\Delta} \leq 1$, we have $\frac{4(c+\Delta)}{7} - \Delta \leq 0$ and then $\mu = \frac{c}{\Delta} \leq \frac{3}{4}$. The firms' profits are expressed by $\pi_H = \frac{2(3c-4\Delta)^2}{49\Delta}$ and $\pi_L = \frac{12(c+\Delta)^2}{49\Delta}$, respectively. And the total consumer surplus is given by

$$\begin{aligned} CS &= \int_0^{\theta_N} (V + \theta q_L - p_L) d\theta + \int_{\theta_N}^1 (V + \theta q_H - p_H) d\theta \\ &+ \int_0^{\theta_2} (V + \theta q_L - p_L(\theta)) d\theta + \int_{\theta_2}^1 (V + \theta q_H - p_H) d\theta \\ &= 2V + \frac{2c^2}{49(q_H - q_L)} + \frac{-5q_H + 54q_L - 52c}{49}. \end{aligned}$$

Case 2. $\theta_2(p_H) > 1$. Firm H chooses not to defend its market share. In order to maximize the two firms' profits, we have the following functions:

$$\begin{aligned} & \operatorname{argmax}_{p_H} (1 - \theta_N)(p_H - c). \\ & \operatorname{argmax}_{p_L} [\theta_N p_L + \int_0^1 p_L(\theta) d\theta]. \end{aligned}$$

We can easily prove that both π_H and π_L are concave functions. By the first-order derivatives, we have $\frac{\partial \pi_H}{\partial p_H} = \frac{c - 2p_H + p_L + c}{\Delta} = 0$, $\frac{\partial \pi_L}{\partial p_L} = \frac{p_H - 2p_L}{\Delta} = 0$. Solving the two equations simultaneously, we have $p_H = \frac{2(c + \Delta)}{3}$ and $p_L = \frac{c + \Delta}{3}$. To ensure the necessary condition $\theta_2(p_H) = \frac{p_H}{\Delta} > 1$, we have $\frac{2(c + \Delta)}{3} - \Delta > 0$ and then $\mu = \frac{c}{\Delta} > \frac{1}{2}$. The firms profits are expressed by $\pi_H = \frac{(c - 2\Delta)^2}{9\Delta}$ and $\pi_L = \frac{4c^2 + 20c\Delta + 7\Delta^2}{18\Delta}$. And the total consumer surplus is given by

$$\begin{aligned} CS &= \int_0^{\theta_N} (V + \theta q_L - p_L) d\theta + \int_{\theta_N}^1 (V + \theta q_H - p_H) d\theta + \int_0^1 (V + \theta q_L - p_L(\theta)) d\theta \\ &= 2V + \frac{c^2}{18(q_H - q_L)} + \frac{-5q_H + 23q_L - 22c}{18}. \end{aligned}$$

From the discussion above, we know that when $\mu \leq \frac{1}{2}$, only case 1 is available and when $\mu > \frac{3}{4}$, only case 2 is available. When $\frac{1}{2} \leq \mu < \frac{3}{4}$, firm H has a choice. Let us compare the two cases to decide which is preferable. We calculate

$$\begin{aligned} \pi_H^{case1} - \pi_H^{case2} &= \frac{2(3c - 4\Delta)^2}{49\Delta} - \frac{(c - 2\Delta)^2}{9\Delta} \\ &= \frac{113c^2 - 236c\Delta + 92\Delta^2}{441\Delta} \\ &= \frac{\Delta}{441} (113\mu^2 - 236\mu + 92). \end{aligned}$$

We solve the inequality $113\mu^2 - 236\mu + 92 \geq 0$ which is due to $\mu \in [0, 1]$. There exists a threshold $\mu_2 = \frac{2}{113} (59 - 21\sqrt{2}) \approx 0.52$. When $\mu \geq \mu_2$, $\pi_L^{case1} \leq \pi_L^{case2}$ and when $\mu \in [0, \mu_2)$, $\pi_L^{case1} > \pi_L^{case2}$. In summary, the preceding conditions can be expressed as follows: when $\mu \in [0, \mu_2)$, the equilibrium operates as case 1; when $\mu \in [\mu_2, 1]$, the equilibrium operates as case 2.

Proof of Proposition 4 In the preceding discussion, we determined the profits of firm H and firm L in the different cases, as shown in the following chart.

μ	$0 \leq \mu < \mu_1$	$\mu_1 \leq \mu < \mu_2$	$\mu_2 \leq \mu \leq 1$
π_H^{BN}	$\frac{2c^2 - 20c\Delta + 23\Delta^2}{18\Delta}$	$\frac{12(c-2\Delta)^2}{49\Delta}$	
π_L^{BN}	$\frac{(c+\Delta)^2}{9\Delta}$	$\frac{2(3c+\Delta)^2}{49\Delta}$	
π_H^{NB}	$\frac{2(3c-4\Delta)^2}{49\Delta}$		$\frac{(c-2\Delta)^2}{9\Delta}$
π_L^{NB}	$\frac{12(c+\Delta)^2}{49\Delta}$		$\frac{4c^2 + 20c\Delta + 7\Delta^2}{18\Delta}$
π_H^{BB}	$\frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta}$		
π_L^{BB}	$\frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta}$		

Now we consider the price of consumer information

For $0 \leq \mu < \mu_1$,

$$K_H = \pi_H^{BN} - \pi_H^{NB} = \frac{2c^2 - 20c\Delta + 23\Delta^2}{18\Delta} - \frac{2(3c - 4\Delta)^2}{49\Delta} = -\frac{226c^2 + 116c\Delta - 551\Delta^2}{882\Delta}.$$

$$K_L = \pi_L^{NB} - \pi_L^{BN} = \frac{12(c + \Delta)^2}{49\Delta} - \frac{(c + \Delta)^2}{9\Delta} = \frac{59(c + \Delta)^2}{441\Delta}.$$

$$\begin{aligned} K_{HL} &= (\pi_H^{BB} - \pi_H^{NB}) + (\pi_L^{BB} - \pi_L^{BN}) \\ &= \frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta} - \frac{2(3c - 4\Delta)^2}{49\Delta} + \frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta} - \frac{(c + \Delta)^2}{9\Delta} \\ &= \frac{656c^2 - 410c\Delta + 257\Delta^2}{882\Delta}. \end{aligned}$$

We can easily prove that $K_H > K_{HL} > K_L$ when $0 \leq \mu < \mu_1$. So the information price here is $K = K_H = -\frac{226c^2 + 116c\Delta - 551\Delta^2}{882\Delta}$.

For $\mu_1 \leq \mu < \mu_2$,

$$K_H = \pi_H^{BN} - \pi_H^{NB} = \frac{12(c - 2\Delta)^2}{49\Delta} - \frac{2(3c - 4\Delta)^2}{49\Delta} = \frac{16\Delta^2 - 6c^2}{49\Delta}.$$

$$K_L = \pi_L^{NB} - \pi_L^{BN} = \frac{12(c + \Delta)^2}{49\Delta} - \frac{2(3c + \Delta)^2}{49\Delta} = \frac{-6c^2 + 12c\Delta + 10\Delta^2}{49\Delta}.$$

$$\begin{aligned} K_{HL} &= (\pi_H^{BB} - \pi_H^{NB}) + (\pi_L^{BB} - \pi_L^{BN}) \\ &= \frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta} - \frac{2(3c - 4\Delta)^2}{49\Delta} + \frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta} - \frac{2(3c + \Delta)^2}{49\Delta} \\ &= \frac{430c^2 - 430c\Delta + 319\Delta^2}{882\Delta}. \end{aligned}$$

We can easily prove that $K_H > K_{HL}$ and $K_L > K_{HL}$ when $\mu_1 \leq \mu < \mu_2$. Then we have

$$K_H - K_L = \frac{16\Delta^2 - 6c^2}{49\Delta} - \frac{2(-3c^2 + 6c\Delta + 5\Delta^2)}{49\Delta} = -\frac{6}{49}(2c - \Delta).$$

Thus we have $K_H \geq K_L$ when $\mu_1 \leq \mu \leq \frac{1}{2}$ and $K_H < K_L$ when $\frac{1}{2} < \mu < \mu_2$. The respective information prices are $K = K_H = \frac{16\Delta^2 - 6c^2}{49\Delta}$ when $\mu_1 \leq \mu \leq \frac{1}{2}$ and $K = K_L = \frac{-6c^2 + 12c\Delta + 10\Delta^2}{49\Delta}$ when $\frac{1}{2} < \mu < \mu_2$.

For $\mu_2 \leq \mu \leq 1$,

$$\begin{aligned}
 K_H &= \pi_H^{BN} - \pi_H^{NB} = \frac{12(c - 2\Delta)^2}{49\Delta} - \frac{(c - 2\Delta)^2}{9\Delta} = \frac{59(c - 2\Delta)^2}{441\Delta}. \\
 K_L &= \pi_L^{NB} - \pi_L^{BN} = \frac{4c^2 + 20c\Delta + 7\Delta^2}{18\Delta} - \frac{2(3c + \Delta)^2}{49\Delta} = \frac{-128c^2 + 764c\Delta + 307\Delta^2}{882\Delta}. \\
 K_{HL} &= \left(\pi_H^{BB} - \pi_H^{NB}\right) + \left(\pi_L^{BB} - \pi_L^{BN}\right) \\
 &= \frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta} - \frac{(c - 2\Delta)^2}{9\Delta} + \frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta} - \frac{2(3c + \Delta)^2}{49\Delta}. \\
 &= \frac{656c^2 - 902c\Delta + 503\Delta^2}{882\Delta}.
 \end{aligned}$$

We can easily prove that $K_H < K_L$ and $K_{HL} < K_L$ when $\mu_2 \leq \mu \leq 1$. Thus the information price here is $K = K_L = \frac{-128c^2 + 764c\Delta + 307\Delta^2}{882\Delta}$. Summarizing the preceding discussion, we obtain the announced result.

Proof of Lemma 1 Suppose that a strictly positive mass of old consumers pay for privacy. We assume that consumers $\theta \in [0, \theta_L]$ pay for privacy and purchase from firm L while consumers $\theta \in [\theta_H, 1]$ pay for privacy and purchase from firm H, $0 \leq \theta_L \leq \mu$ and $\mu \leq \theta_H \leq 1$. The functions for the maximization of the two firms’ profits are

$$\begin{aligned}
 &\operatorname{argmax}_{p_H} \left[(1 - \theta_N)(p_H - c) + \int_{\mu}^{\theta_H} (p_H(\theta) - c)d\theta + (1 - \theta_H)(p_H - c) \right]. \\
 &\operatorname{argmax}_{p_L} \left[\theta_N p_L + \int_0^{\theta_L} p_L d\theta + \int_{\theta_L}^{\mu} p_L(\theta)d\theta \right].
 \end{aligned}$$

We can easily prove that both π_H and π_L are concave functions. By the first-order derivatives, we have $\frac{\partial \pi_H}{\partial p_H} = \frac{2\Delta - 2p_H + p_L + c - \Delta\theta_H}{\Delta} = 0$, $\frac{\partial \pi_L}{\partial p_L} = \frac{p_H - 2p_L + \Delta\theta_L}{\Delta} = 0$. Solving the two equations simultaneously, we have $p_H = \frac{2c + 4\Delta - 2\theta_H\Delta + \theta_L\Delta}{3}$ and $p_L = \frac{c + 2\Delta - 2\theta_H\Delta + 2\theta_L\Delta}{3}$.

Consumers choose to pay for privacy only when they receive a higher utility, which means

$$V + \theta q_H - p_H - c_0 \geq V + \theta q_H - p_H(\theta)$$

and

$$V + \theta q_L - p_L - c_0 \geq V + \theta q_L - p_L(\theta).$$

Then we have $\frac{2c + 4\Delta - 2\theta_H\Delta + \theta_L\Delta}{3} + c_0 \leq \theta\Delta$ and $\frac{c + 2\Delta - 2\theta_H\Delta + 2\theta_L\Delta}{3} + c_0 \leq c - \theta\Delta$. The result can be expressed as $\frac{2\mu + 4 - 2\theta_H + \theta_L}{3} + \mu_0 \leq \theta \leq \mu - \frac{\mu + 2 - 2\theta_H + 2\theta_L}{3} - \mu_0$. Comparing the lower and upper bounds of θ , we find that for any $c_0 > 0$, there does not exist any θ_H and θ_L , $0 \leq \theta_L \leq \mu$ and $\mu \leq \theta_H \leq 1$, to make the inequity available. So when both firms have information and the privacy cost $c_0 > 0$, no consumer will pay for privacy.

Proof of Proposition 5 When only firm H has consumer data, it alone has the power to price–discriminate, while consumers can pay for privacy and firm L must decide whether to set a price to retain its share of the old market.

Case 1. $\theta_1(p_L) \geq 0$. In this case, firm L sets a relatively low price to keep its previous market share. The function for the maximization of the two firms' profits is

$$\begin{aligned} & \operatorname{argmax}_{p_H} \left[(1 - \theta_N)(p_H - c) + \int_{\theta_1}^{\theta_H^c} (p_H(\theta) - c) d\theta + (1 - \theta_H^c)(p_H - c) \right]. \\ & \operatorname{argmax}_{p_L} [\theta_N p_L + \int_0^{\theta_1} p_L(\theta) d\theta]. \end{aligned}$$

We can easily prove that both π_H and π_L are concave functions. By the first-order derivatives, we have $\frac{\partial \pi_H}{\partial p_H} = \frac{c+2\Delta-3p_H+2p_L}{\Delta} = 0$, $\frac{\partial \pi_L}{\partial p_L} = \frac{c+p_H-4p_L}{\Delta} = 0$. Solving the two equations simultaneously, we have $p_H = \frac{3c+4\Delta}{5}$ and $p_L = \frac{2c+\Delta}{5}$. To ensure the necessary condition $\theta_1(p_L) = \frac{c-p_L}{\Delta} \geq 0$, we have $c - \frac{2c+\Delta}{5} \geq 0$ and then $\mu = \frac{c}{\Delta} \geq \frac{1}{3}$. Another constraint $\theta_H^c = \frac{p_H-p_L+c_0}{\Delta} = \frac{c+3\Delta+5c_0}{5\Delta} \leq 1$ and then we have $\mu_0 = \frac{c_0}{\Delta} \leq \frac{2}{5} - \frac{1}{5}\mu$. The firms' profits are expressed by $\pi_H = \frac{12(c-2\Delta)^2+25c_0^2}{50\Delta}$ and $\pi_L = \frac{2(2c+\Delta)^2}{25\Delta}$, respectively, and the total consumer surplus is given by

$$\begin{aligned} CS &= \int_0^{\theta_N} (V + \theta q_L - p_L) d\theta + \int_{\theta_N}^1 (V + \theta q_H - p_H) d\theta + \int_0^{\theta_1} (V + \theta q_L - p_L) d\theta \\ &+ \int_{\theta_1}^{\theta_H^c} (V + \theta q_H - p_H(\theta)) d\theta + \int_{\theta_H^c}^1 (V + \theta q_H - p_H) d\theta \\ &= \frac{2c^2 + 25c_0^2 + 10c_0(c - 2q_H + 2q_L)}{50(q_H - q_L)} - \frac{1}{25}(24c + 6q_H - 31q_L) + 2V. \end{aligned}$$

Case 2. $\theta_1(p_L) < 0$. In this case, firm L chooses to relinquish its share of the old market. The profit maximization functions are

$$\begin{aligned} & \operatorname{argmax}_{p_H} \left[(1 - \theta_N)(p_H - c) + \int_0^{\theta_H^c} (p_H(\theta) - c) d\theta + (1 - \theta_H^c)(p_H - c) \right]. \\ & \operatorname{argmax}_{p_L} \theta_N p_L. \end{aligned}$$

It is obvious that both π_H and π_L are concave functions. By the first-order derivatives, we have $\frac{\partial \pi_H}{\partial p_H} = \frac{c+2\Delta-3p_H+2p_L}{\Delta} = 0$, $\frac{\partial \pi_L}{\partial p_L} = \frac{p_H-2p_L}{\Delta} = 0$. Solving the two equations simultaneously, we have $p_H = \frac{c+2\Delta}{2}$ and $p_L = \frac{c+2\Delta}{4}$. To ensure the necessary condition $\theta_1(p_L) = \frac{c-p_L}{\Delta} < 0$, we have $c - \frac{c+2\Delta}{4} < 0$ and then $\mu = \frac{c}{\Delta} < \frac{2}{3}$. Another constraint $\theta_H^c = \frac{p_H-p_L+c_0}{\Delta} = \frac{c+2\Delta+4c_0}{4\Delta} \leq 1$ and then we have $\mu_0 = \frac{c_0}{\Delta} \leq \frac{1}{2} - \frac{1}{4}\mu$. The firms' profits

are expressed respectively by $\pi_H = \frac{3c^2 - 36c\Delta + 44\Delta^2 + 16c_0^2}{32\Delta}$ and $\pi_L = \frac{(c+2\Delta)^2}{16\Delta}$, and the total consumer surplus is given by

$$\begin{aligned}
 CS &= \int_0^{\theta_N} (V + \theta q_L - p_L) d\theta + \int_{\theta_N}^1 (V + \theta q_H - p_H) d\theta \\
 &+ \int_0^{\theta_H^c} (V + \theta q_H - p_H(\theta)) d\theta + \int_{\theta_H^c}^1 (V + \theta q_H - p_H) d\theta \\
 &= \frac{c^2 + 8c_0^2 + 4c_0(c - 2q_H + 2q_L)}{16(q_H - q_L)} + \frac{1}{4}(-3c - 3q_H + 7q_L) + 2V.
 \end{aligned}$$

From the preceding discussion, we know that when $\mu \geq \frac{2}{3}$, only case 1 is available and when $\mu < \frac{1}{3}$, only case 2 is available. When $\frac{1}{3} \leq \mu < \frac{2}{3}$, firm L can choose between the two cases. Comparing them, we calculate

$$\begin{aligned}
 \pi_L^{case1} - \pi_L^{case2} &= \frac{2(2c + \Delta)^2}{25\Delta} - \frac{(c + 2\Delta)^2}{16\Delta} = \frac{103c^2 + 28c\Delta - 68\Delta^2}{400\Delta} \\
 &= \frac{\Delta}{400}(103\mu^2 + 28\mu - 68).
 \end{aligned}$$

We solve the inequality $103\mu^2 + 28\mu - 68 \geq 0$ which is due to $\mu \in [\frac{1}{3}, \frac{2}{3}]$. We find that $\pi_L^{case1} < \pi_L^{case2}$ is always true when $\frac{1}{3} \leq \mu < \frac{2}{3}$. At the same time, the upper bound of μ_0 also satisfies $\frac{1}{2} - \frac{1}{4}\mu > \frac{2}{5} - \frac{1}{5}\mu$. Ultimately, the preceding conditions can be expressed as follows: when $0 \leq \mu < \frac{2}{3}$, the equilibrium operates as case 2; when $\frac{2}{3} \leq \mu \leq 1$, the equilibrium operates as case 1. Thus we obtain the announced result.

Proof of Proposition 6 When only firm L has information, this firm have the exclusive power to price-discriminate, while consumers can pay for privacy and firm H must decide whether to set a price to retain a share of the old market.

Case 1. $\theta_2(p_H) \leq 1$. In this case, firm H sets a relatively low price to hold on to its market share. The profit maximization functions are

$$\begin{aligned}
 &\operatorname{argmax}_{p_H} [(1 - \theta_N)(p_H - c) + (1 - \theta_2)(p_H - c)]. \\
 &\operatorname{argmax}_{p_L} [\theta_N p_L + \int_0^{\theta_L^c} p_L d\theta + \int_{\theta_L^c}^{\theta_2} p_L(\theta) d\theta].
 \end{aligned}$$

It can readily be proved that both π_H and π_L are concave functions. By the first-order derivatives, we have $\frac{\partial \pi_H}{\partial p_H} = \frac{2(c+\Delta)-4p_H+p_L}{\Delta} = 0$, $\frac{\partial \pi_L}{\partial p_L} = \frac{2p_H-3p_L}{\Delta} = 0$. Solving the two equations simultaneously, we have $p_H = \frac{3(c+\Delta)}{5}$ and $p_L = \frac{3(c+\Delta)}{5}$. To ensure the necessary condition $\theta_2(p_H) = \frac{p_H}{\Delta} \leq 1$, we have $\frac{3(c+\Delta)}{5} - \Delta \leq 0$ and then $\mu = \frac{c}{\Delta} \leq \frac{2}{3}$. Another constraint $\theta_L^c = \frac{p_H-p_L-c_0}{\Delta} = \frac{c+\Delta-5c_0}{5\Delta} \geq 0$ and then we have $\mu_0 = \frac{c_0}{\Delta} \leq \frac{1}{5} + \frac{1}{5}\mu$. The firms

profits are expressed by $\pi_H = \frac{2(2c-3\Delta)^2}{25\Delta}$ and $\pi_L = \frac{12(c+\Delta)^2+25c_0^2}{50\Delta}$, respectively, and the total consumer surplus is given by

$$\begin{aligned}
 CS &= \int_0^{\theta_N} (V + \theta q_L - p_L) d\theta + \int_{\theta_N}^1 (V + \theta q_H - p_H) d\theta + \int_0^{\theta_L^c} (V + \theta q_L - p_L) d\theta \\
 &+ \int_{\theta_L^c}^{\theta_2} (V + \theta q_L - p_L(\theta)) d\theta + \int_{\theta_2}^1 (V + \theta q_H - p_H) d\theta \\
 &= \frac{2c^2 + 25c_0^2 - 10c_0(c + q_H - q_L)}{50(q_H - q_L)} - \frac{1}{25}(28c + 4q_H - 29q_L) + 2V.
 \end{aligned}$$

Case 2. $\theta_2(p_H) > 1$. In this case, firm H chooses to relinquish its previous share of the old market. The profit maximization functions are

$$\begin{aligned}
 &\operatorname{argmax}_{p_H} (1 - \theta_N)(p_H - c). \\
 &\operatorname{argmax}_{p_L} [\theta_N p_L + \int_0^{\theta_L^c} p_L d\theta + \int_{\theta_L^c}^1 p_L(\theta) d\theta].
 \end{aligned}$$

We can easily prove that both π_H and π_L are concave functions. By the first-order derivatives, we have $\frac{\partial \pi_H}{\partial p_H} = \frac{c+\Delta-2p_H+p_L}{\Delta} = \frac{\partial \pi_L}{\partial p_L} = \frac{2p_H-3p_L}{\Delta} = 0$. Solving the two equations simultaneously, we have $p_H = \frac{3(c+\Delta)}{4}$ and $p_L = \frac{c+\Delta}{2}$. To ensure the necessary condition $\theta_2(p_H) = \frac{p_H}{\Delta} > 1$, we have $\frac{3(c+\Delta)}{4\Delta} - 1 > 0$ and then $\mu = \frac{c}{\Delta} > \frac{1}{3}$. Another constraint $\theta_L^c = \frac{p_H-p_L-c_0}{\Delta} = \frac{c+\Delta-4c_0}{4\Delta} \geq 0$ and then we have $\mu_0 = \frac{c_0}{\Delta} \leq \frac{1}{4} + \frac{1}{4}\mu$. The firms' profits are expressed by $\pi_H = \frac{(c-3\Delta)^2}{16\Delta}$ and $\pi_L = \frac{3c^2+30c\Delta+11\Delta^2+16c_0^2}{32\Delta}$, and the total consumer surplus is given by

$$\begin{aligned}
 CS &= \int_0^{\theta_N} (V + \theta q_L - p_L) d\theta + \int_{\theta_N}^1 (V + \theta q_H - p_H) d\theta \\
 &+ \int_0^{\theta_L^c} (V + \theta q_L - p_L) d\theta + \int_{\theta_L^c}^1 (V + \theta q_L - p_L(\theta)) d\theta \\
 &= \frac{c^2 + 8c_0^2 - 4c_0(c + q_H - q_L)}{16(q_H - q_L)} - \frac{1}{16}(22c + 7q_H - 23q_L) + 2V.
 \end{aligned}$$

From the preceding discussion, we know that when $\mu \leq \frac{1}{3}$, only case 1 is available; and when $\mu > \frac{2}{3}$, only case 2 is available. When $\frac{1}{3} < \mu \leq \frac{2}{3}$, firm H can choose between the two cases. Comparing them, we calculate

$$\pi_H^{case1} - \pi_H^{case2} = \frac{2(2c - 3\Delta)^2}{25\Delta} - \frac{(c - 3\Delta)^2}{16\Delta} = \frac{103c^2 - 234c\Delta + 63\Delta^2}{400\Delta}.$$

We solve the inequality $113\mu^2 + 234\mu + 63 \geq 0$ which is due to $\mu \in [\frac{1}{3}, \frac{2}{3}]$. The result shows that $\pi_H^{case1} - \pi_H^{case2} < 0$ is always true for $\frac{1}{3} < \mu \leq \frac{2}{3}$. At the same time, the upper bound of μ_0 also satisfies $\frac{1}{4} + \frac{1}{4}\mu > \frac{1}{3} + \frac{1}{5}\mu$. Ultimately, the preceding conditions can be

expressed as follows: when $0 \leq \mu \leq \frac{1}{3}$, the equilibrium operates as case 1; when $\frac{1}{3} < \mu \leq 1$, the equilibrium operates as case 2. Thus we obtain the announced result.

Proof of Proposition 7 Through the preceding analysis, we have determined the profits of firm H and firm L profits for the different cases in which consumers can pay for privacy. These are shown in the following charts.

(1) In region 1, the customers of both firms can pay for privacy, and we have

μ	$0 \leq \mu < \frac{1}{3}$	$\frac{1}{3} \leq \mu < \frac{2}{3}$	$\frac{2}{3} \leq \mu \leq 1$
π_H^{BN}	$\frac{3c^2 - 36c\Delta + 44\Delta^2 + 16c_0^2}{32\Delta}$		$\frac{12(c-2\Delta)^2 + 25c_0^2}{50\Delta}$
π_L^{BN}	$\frac{(c+2\Delta)^2}{16\Delta}$		$\frac{2(2c+\Delta)^2}{25\Delta}$
π_H^{NB}	$\frac{2(2c-3\Delta)^2}{25\Delta}$	$\frac{(c-3\Delta)^2}{16\Delta}$	
π_L^{NB}	$\frac{12(c+\Delta)^2 + 25c_0^2}{50\Delta}$	$\frac{3c^2 + 30c\Delta + 11\Delta^2 + 16c_0^2}{32\Delta}$	
π_H^{BB}	$\frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta}$		
π_L^{BB}	$\frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta}$		

Now we consider the price of consumer information.

For $0 \leq \mu < \frac{1}{3}$

$$\begin{aligned}
 K_H &= \pi_H^{BN} - \pi_H^{NB} = \frac{3c^2 - 36c\Delta + 44\Delta^2 + 16c_0^2}{32\Delta} - \frac{2(2c - 3\Delta)^2}{25\Delta} \\
 &= \frac{-181c^2 - 132c\Delta + 524\Delta^2 + 400c_0^2}{800\Delta} \\
 K_L &= \pi_L^{NB} - \pi_L^{BN} = \frac{12(c + \Delta)^2 + 25c_0^2}{50\Delta} - \frac{(c + 2\Delta)^2}{16\Delta} = \frac{c^2 + 22c\Delta + 3\Delta^2 + 16c_0^2}{32\Delta} \\
 K_{HL} &= (\pi_H^{BB} - \pi_H^{NB}) + (\pi_L^{BB} - \pi_L^{BN}) \\
 &= \frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta} - \frac{2(2c - 3\Delta)^2}{25\Delta} + \frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta} - \frac{(c + 2\Delta)^2}{16\Delta} \\
 &= \frac{3023c^2 - 1844c\Delta + 308\Delta^2}{3600\Delta}
 \end{aligned}$$

We can easily prove that $K_H > K_{HL}$ and $K_H > K_L$ for any $c_0 > 0$ when $0 \leq \mu < \frac{1}{3}$. So the information price here is $K = K_H = \frac{-181c^2 - 132c\Delta + 524\Delta^2 + 400c_0^2}{800\Delta}$.

For $\frac{1}{3} \leq \mu < \frac{2}{3}$

$$\begin{aligned}
 K_H &= \pi_H^{BN} - \pi_H^{NB} = \frac{3c^2 - 36c\Delta + 44\Delta^2 + 16c_0^2}{32\Delta} - \frac{(c - 3\Delta)^2}{16\Delta} \\
 &= \frac{c^2 - 24c\Delta + 26\Delta^2 + 16c_0^2}{32\Delta} \\
 K_L &= \pi_L^{NB} - \pi_L^{BN} = \frac{3c^2 + 30c\Delta + 11\Delta^2 + 16c_0^2}{32\Delta} - \frac{(c + 2\Delta)^2}{16\Delta} \\
 &= \frac{c^2 + 22c\Delta + 3\Delta^2 + 16c_0^2}{32\Delta}
 \end{aligned}$$

$$\begin{aligned}
 K_{HL} &= (\pi_H^{BB} - \pi_H^{NB}) + (\pi_L^{BB} - \pi_L^{BN}) \\
 &= \frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta} - \frac{(c - 3\Delta)^2}{16\Delta} + \frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta} - \frac{(c + 2\Delta)^2}{16\Delta} \\
 &= \frac{1}{144} \left(-158c + \frac{158c^2}{\Delta} + 35\Delta \right).
 \end{aligned}$$

We can easily prove that $K_H > K_{HL}$ and $K_L > K_{HL}$ for any $c_0 > 0$ when $\frac{1}{3} \leq \mu < \frac{2}{3}$. And we have

$$K_H - K_L = \frac{23}{32}(-2c + \Delta) = \frac{23\Delta}{32}(-2\mu + 1).$$

Thus we have $K_H \geq K_L$ when $\frac{1}{3} \leq \mu \leq \frac{1}{2}$ and $K_H < K_L$ when $\frac{1}{2} < \mu < \frac{2}{3}$. The information prices here are $K = K_H = \frac{c^2 - 24c\Delta + 26\Delta^2 + 16c_0^2}{32\Delta}$ when $\frac{1}{3} \leq \mu \leq \frac{1}{2}$ and $K = K_L = \frac{71c^2 + 92c\Delta - 4\Delta^2 + 200c_0^2}{400\Delta}$ when $\frac{1}{2} < \mu \leq 1$.

For $\frac{2}{3} \leq \mu \leq 1$

$$\begin{aligned}
 K_H &= \pi_H^{BN} - \pi_H^{NB} = \frac{12(c - 2\Delta)^2 + 25c_0^2}{50\Delta} - \frac{(c - 3\Delta)^2}{16\Delta} \\
 &= \frac{71c^2 - 234c\Delta + 159\Delta^2 + 200c_0^2}{400\Delta}. \\
 K_L &= \pi_L^{NB} - \pi_L^{BN} = \frac{3c^2 + 30c\Delta + 11\Delta^2 + 16c_0^2}{32\Delta} - \frac{2(2c + \Delta)^2}{25\Delta} \\
 &= \frac{-181c^2 + 494c\Delta + 211\Delta^2 + 400c_0^2}{800\Delta}. \\
 K_{HL} &= (\pi_H^{BB} - \pi_H^{NB}) + (\pi_L^{BB} - \pi_L^{BN}) \\
 &= \frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta} - \frac{2(2c + \Delta)^2}{25\Delta} + \frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta} - \frac{(c - 3\Delta)^2}{16\Delta} \\
 &= \frac{3023c^2 - 4202c\Delta + 1487\Delta^2}{3600\Delta}.
 \end{aligned}$$

We can easily prove that $K_H < K_L$ and $K_{HL} < K_L$ for any $c_0 > 0$ when $\frac{2}{3} \leq \mu \leq 1$. Thus the information price here is $K = K_L = \frac{-181c^2 + 494c\Delta + 211\Delta^2 + 400c_0^2}{800\Delta}$.

(2) In regions 2 and 3, only exclusive firm’s consumers may purchase product and pay for privacy, then we have

μ	$0 \leq \mu < \frac{1}{2}$	$\frac{1}{2} < \mu \leq 1$
π_H^{BN}	$\frac{3c^2 - 36c\Delta + 44\Delta^2 + 16c_0^2}{32\Delta}$	$\frac{12(c - 2\Delta)^2}{49\Delta}$
π_L^{BN}	$\frac{(c + 2\Delta)^2}{16\Delta}$	$\frac{2(3c + \Delta)^2}{49\Delta}$
π_H^{NB}	$\frac{2(3c - 4\Delta)^2}{49\Delta}$	$\frac{(c - 3\Delta)^2}{16\Delta}$
π_L^{NB}	$\frac{12(c + \Delta)^2}{49\Delta}$	$\frac{3c^2 + 30c\Delta + 11\Delta^2 + 16c_0^2}{32\Delta}$
π_H^{BB}	$\frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta}$	
π_L^{BB}	$\frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta}$	

Now we consider the price of consumer information.

For $0 \leq \mu < \frac{1}{2}$,

$$\begin{aligned}
 K_H &= \pi_H^{BN} - \pi_H^{NB} = \frac{3c^2 - 36c\Delta + 44\Delta^2 + 16c_0^2}{32\Delta} - \frac{2(3c - 4\Delta)^2}{49\Delta} \\
 &= \frac{-429c^2 - 228c\Delta + 1132\Delta^2 + 784c_0^2}{1568\Delta} \\
 K_L &= \pi_L^{NB} - \pi_L^{BN} = \frac{12(c + \Delta)^2}{49\Delta} - \frac{(c + 2\Delta)^2}{16\Delta} = \frac{143c^2 + 188c\Delta - 4\Delta^2}{784\Delta} \\
 K_{HL} &= \left(\pi_H^{BB} - \pi_H^{NB}\right) + \left(\pi_L^{BB} - \pi_L^{BN}\right) \\
 &= \frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta} - \frac{2(3c - 4\Delta)^2}{49\Delta} + \frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta} - \frac{(c + 2\Delta)^2}{16\Delta} \\
 &= \frac{5591c^2 - 3476c\Delta + 1076\Delta^2}{7056\Delta}
 \end{aligned}$$

We can easily prove that $K_H > K_L$ and $K_H > K_{HL}$ for any $c_0 > 0$ when $0 \leq \mu \leq \frac{1}{2}$.

Thus the information price here is $K = K_H = \frac{-429c^2 - 228c\Delta + 1132\Delta^2 + 784c_0^2}{1568\Delta}$.

For $\frac{1}{2} < \mu \leq 1$,

$$\begin{aligned}
 K_H &= \pi_H^{BN} - \pi_H^{NB} = \frac{12(c - 2\Delta)^2}{49\Delta} - \frac{(c - 3\Delta)^2}{16\Delta} = \frac{143c^2 - 474c\Delta + 327\Delta^2}{784\Delta} \\
 K_L &= \pi_L^{NB} - \pi_L^{BN} = \frac{3c^2 + 30c\Delta + 11\Delta^2 + 16c_0^2}{32\Delta} - \frac{2(3c + \Delta)^2}{49\Delta} \\
 &= \frac{-429c^2 + 1086c\Delta + 475\Delta^2 + 784c_0^2}{1568\Delta} \\
 K_{HL} &= \left(\pi_H^{BB} - \pi_H^{NB}\right) + \left(\pi_L^{BB} - \pi_L^{BN}\right) \\
 &= \frac{11c^2 - 26c\Delta + 17\Delta^2}{18\Delta} - \frac{(c - 3\Delta)^2}{16\Delta} + \frac{11c^2 + 4c\Delta + 2\Delta^2}{18\Delta} - \frac{2(3c + \Delta)^2}{49\Delta} \\
 &= \frac{5591c^2 - 7706c\Delta + 3191\Delta^2}{7056\Delta}
 \end{aligned}$$

We can easily prove that $K_H < K_L$ and $K_L > K_{HL}$ for any $c_0 > 0$ when $\frac{1}{2} < \mu \leq 1$.

Thus the information price here is $K = K_L = \frac{-429c^2 + 1086c\Delta + 475\Delta^2 + 784c_0^2}{1568\Delta}$. Summarizing the discussion above, we obtain the announced Proposition 7.

Proof of Corollary 2 In the BN case, only firm H has the information and consumers pay for privacy. We know firm H’s profits

$$\pi_H = \begin{cases} \frac{3c^2 - 36c\Delta + 44\Delta^2 + 16c_0^2}{32\Delta}, & 0 \leq \mu < \frac{2}{3} \text{ and } 0 < \mu_0 \leq \frac{1}{2} - \frac{1}{4}\mu \\ \frac{12(c - 2\Delta)^2 + 25c_0^2}{50\Delta}, & \frac{2}{3} \leq \mu \leq 1 \text{ and } 0 < \mu_0 \leq \frac{2}{3} - \frac{1}{5}\mu \end{cases}$$

While firm L’s profits are not affected by the value of c_0 . For any $c_0 > 0$, we can easily determine that $\pi'_H(c_0) > 0$ which is due to $0 \leq \mu \leq 1$. That means π_H always increases with c_0 . When c_0 exceeds the upper bound, the profits are the same as in the no-privacy case (i.e., they are linear).

In the NB case, only firm L has the information and consumers can pay for privacy. We know firm L’s profits

$$\pi_H = \begin{cases} \frac{12(c+\Delta)^2+25c_0^2}{50\Delta}, & 0 \leq \mu \leq \frac{1}{3} \text{ and } 0 < \mu_0 \leq \frac{1}{5} + \frac{1}{5}\mu, \\ \frac{3c^2+30c\Delta+11\Delta^2+16c_0^2}{32\Delta}, & \frac{1}{3} < \mu \leq 1 \text{ and } 0 < \mu_0 \leq \frac{1}{4} + \frac{1}{4}\mu, \end{cases}$$

while firm H’s profits are not related to c_0 . For any $c_0 > 0$, we can easily determine that $\pi'_L(c_0) > 0$ which is due to $0 \leq \mu \leq 1$. That means π_L also increases with c_0 . As μ_0 exceeds the upper bound, the profits are the same as in the no-privacy case. Thus we obtain the announced Corollary 2.

Proof of Corollary 3 In the BN case, when $0 \leq \mu < \frac{2}{3}$, we have $CS'(c_0) = \frac{16c_0+4(c-2q_H+2q_L)}{16(q_H-q_L)} = \mu_0 + \frac{1}{4}\mu - \frac{1}{2} < 0$ for any $\mu_0 \in (0, \frac{1}{2} - \frac{1}{4}\mu)$, which means that CS decreases as c_0 rises. When $\frac{2}{3} \leq \mu \leq 1$, $CS'(c_0) = \frac{50c_0+10(c-2q_H+2q_L)}{50(q_H-q_L)} = \mu_0 + \frac{1}{5}\mu - \frac{2}{5} < 0$ for any $\mu_0 \in (0, \frac{2}{5} - \frac{1}{5}\mu)$, which means that here too CS decreases as c_0 increases.

In the NB case, when $0 \leq \mu \leq \frac{1}{3}$, we have $CS'(c_0) = \frac{50c_0-10(c+q_H-q_L)}{50(q_H-q_L)} = \mu_0 - \frac{1}{5}\mu - \frac{1}{5} < 0$ for any $\mu_0 \in (0, \frac{1}{5} + \frac{1}{5}\mu)$, which means that CS decreases with rising c_0 . Likewise, when $\frac{1}{3} < \mu \leq 1$, $CS'(c_0) = \frac{16c_0-4(c+q_H-q_L)}{16(q_H-q_L)} = \mu_0 - \frac{1}{4}\mu - \frac{1}{4} < 0$ for any $\mu_0 \in (0, \frac{1}{4} + \frac{1}{4}\mu)$, which means that once again CS falls as c_0 rises.

Proof of Corollary 4 From the discussion in Proposition 7, we have the information prices in different regions, as summarized in Tables 1 and 2.

$K'(c_0) > 0$ is always available for any $c_0 > 0$ in whichever region. At the same time, the information prices in region II and III are always larger than the corresponding prices in region I. So we can say that the information prices offered by the DS always increase with c_0 , and we arrive at the announced results.

Table 1 The information price in different regions

Region μ	I
$0 \leq \mu < \frac{1}{3}$	$\frac{-181c^2-132c\Delta+524\Delta^2+400c_0^2}{800\Delta}$
$\frac{1}{3} \leq \mu < \frac{1}{2}$	$\frac{c^2-24c\Delta+26\Delta^2+16c_0^2}{32\Delta}$
$\frac{1}{2} \leq \mu < \frac{2}{3}$	$\frac{71c^2+92c\Delta-4\Delta^2+200c_0^2}{400\Delta}$
$\frac{2}{3} \leq \mu \leq 1$	$\frac{-181c^2+494c\Delta+211\Delta^2+400c_0^2}{800\Delta}$

Table 2 The information price in different regions

Region μ	II	III
$0 \leq \mu < \frac{1}{2}$	$\frac{-429c^2-228c\Delta+1132\Delta^2+784c_0^2}{1568\Delta}$	
$\frac{1}{2} \leq \mu \leq 1$		$\frac{-429c^2+1086c\Delta+475\Delta^2+784c_0^2}{1568\Delta}$

Proof of Proposition 8 (1) when $0 < \mu \leq \mu_1, 0 < \mu_0 < \frac{1}{2} - \frac{1}{4}\mu$, the difference in industry profits with and without transparency is given by

$$\begin{aligned} \pi^D &= \pi^{with} - \pi^{without} \\ &= \frac{3c^2 - 36c\Delta + 44\Delta^2 + 16c_0^2}{32\Delta} + \frac{(c + 2\Delta)^2}{16\Delta} - \frac{1}{18} \left(-20c + \frac{2c^2}{\Delta} + 23\Delta \right) - \frac{(c + \Delta)^2}{9\Delta} \\ &= \frac{-19c^2 + 4c\Delta + 68\Delta^2 + 144c_0^2}{288\Delta} > 0. \end{aligned}$$

The consumer surplus difference is

$$\begin{aligned} CS^D &= CS^{with} - CS^{without} \\ &= \frac{72c_0^2 + 36c_0(c - 2\Delta) + (c + 22\Delta)(c - 2\Delta)}{144\Delta} \\ &= \frac{\Delta}{144} [72\mu_0^2 + (36\mu_0 + \mu + 22)(\mu - 2)] < 0. \end{aligned}$$

The social welfare difference is

$$\begin{aligned} SW^D &= \pi^D + CS^D \\ &= \frac{-17c^2 + 44c\Delta - 20\Delta^2 + 72cc_0 - 144\Delta c_0 + 288c_0^2}{288\Delta} \\ &= \frac{\Delta}{288} (-17\mu^2 + 44\mu - 20) + \Delta \left[\mu_0^2 - \left(\frac{1}{2} - \frac{1}{4}\mu \right) \mu_0 \right]. \end{aligned}$$

We can easily find that $-17\mu^2 + 44\mu - 20 < 0$ and $\mu_0^2 - \left(\frac{1}{2} - \frac{1}{4}\mu\right)\mu_0 \leq 0$ for any $0 < \mu \leq \mu_1, 0 < \mu_0 < \frac{1}{2} - \frac{1}{4}\mu$, so we have $SW^D < 0$.

(2) When $\mu_1 < \mu \leq \frac{2}{3}, 0 < \mu_0 < \frac{1}{2} - \frac{1}{4}\mu$, the difference in industry profits with and without transparency is given by

$$\begin{aligned} \pi^D &= \pi^{with} - \pi^{without} \\ &= \frac{3c^2 - 36c\Delta + 44\Delta^2 + 16c_0^2}{32\Delta} + \frac{(c + 2\Delta)^2}{16\Delta} - \frac{12(c - 2\Delta)^2}{49\Delta} - \frac{2(3c + \Delta)^2}{49\Delta} \\ &= \frac{-715c^2 - 220c\Delta + 948\Delta^2 + 784c_0^2}{1568\Delta} > 0. \end{aligned}$$

The consumer surplus difference is

$$\begin{aligned} CS^D &= CS^{with} - CS^{without} \\ &= \frac{392c_0^2 + 196c_0(c - 2\Delta) + (17c + 246\Delta)(c - 2\Delta)}{784\Delta} \\ &= \frac{\Delta}{784} [392\mu_0^2 + (196\mu_0 + 17\mu + 246)(\mu - 2)] < 0. \end{aligned}$$

The social welfare difference is

$$\begin{aligned} SW^D &= \pi^D + CS^D \\ &= \frac{-681c^2 + 204c\Delta - 36\Delta^2 + 392cc_0 - 784\Delta c_0 + 1568c_0^2}{1568\Delta} \end{aligned}$$

$$= \frac{\Delta}{1568}(-681\mu^2 + 204\mu - 36) + \Delta \left[\mu_0^2 - \left(\frac{1}{2} - \frac{1}{4}\mu \right) \mu_0 \right].$$

We can easily find that $-681\mu^2 + 204\mu - 36 < 0$ and $\mu_0^2 - (\frac{1}{2} - \frac{1}{4}\mu)\mu_0 \leq 0$ for any $\mu_1 < \mu \leq \frac{2}{3}, 0 < \mu_0 < \frac{1}{2} - \frac{1}{4}\mu$, so we have $SW^D < 0$.

(3) When $\frac{2}{3} < \mu \leq 1, 0 < \mu_0 < \frac{2}{5} - \frac{1}{5}\mu$, the difference in industry profits with and without transparency is given by

$$\begin{aligned} \pi^D &= \pi^{with} - \pi^{without} \\ &= \frac{12(c - 2\Delta)^2 + 25c_0^2}{50\Delta} + \frac{2(2c + \Delta)^2}{25\Delta} - \frac{12(c - 2\Delta)^2}{49\Delta} - \frac{2(3c + \Delta)^2}{49\Delta} \\ &= \frac{8(-16c^2 + 29c\Delta + 6\Delta^2) + 1225c_0^2}{2450\Delta} > 0. \end{aligned}$$

The consumer surplus difference is

$$\begin{aligned} CS^D &= CS^{with} - CS^{without} \\ &= \frac{1225c_0^2 + 490c_0(c - 2\Delta) - 2(c - 2\Delta)(c - 72\Delta)}{2450\Delta} \\ &= \frac{\Delta}{2450} [1225\mu_0^2 + [490\mu_0 - 2\mu + 144\Delta](\mu - 2)] < 0. \end{aligned}$$

The social welfare difference is

$$\begin{aligned} SW^D &= \pi^D + CS^D \\ &= \frac{-13c^2 + 38c\Delta - 24\Delta^2 + 49cc_0 - 98\Delta c_0 + 245c_0^2}{245\Delta} \\ &= \frac{\Delta}{245} (-13\mu^2 + 38\mu - 24) + \Delta \left[\mu_0^2 - \left(\frac{2}{5} - \frac{1}{5}\mu \right) \mu_0 \right]. \end{aligned}$$

We can easily determine that $\mu_0^2 - (\frac{1}{2} - \frac{1}{4}\mu)\mu_0 \leq 0$ for any $\mu_1 < \mu \leq \frac{2}{3}, 0 < \mu_0 < \frac{1}{2} - \frac{1}{4}\mu$. When $\mu \leq \frac{12}{13}, -13\mu^2 + 38\mu - 24 \leq 0$; and when $\mu > \frac{12}{13}, -13\mu^2 + 38\mu - 24 > 0$. So we have $SW^D < 0$ when $\mu \leq \frac{12}{13}$. Thus we obtain the announced Proposition 8.

Proof of Proposition 9 (1) when $0 < \mu \leq \frac{1}{3}, 0 < \mu_0 < \frac{1}{5} + \frac{1}{5}\mu$, the difference in industry profits with and without transparency is given by

$$\begin{aligned} \pi^D &= \pi^{with} - \pi^{without} \\ &= \frac{2(2c - 3\Delta)^2}{25\Delta} + \frac{12(c + \Delta)^2 + 25c_0^2}{50\Delta} - \frac{2(3c - 4\Delta)^2}{49\Delta} - \frac{12(c + \Delta)^2}{49\Delta} \\ &= \frac{-8(16c - 19\Delta)(c + \Delta) + 1225c_0^2}{2450\Delta} > 0. \end{aligned}$$

The consumer surplus difference is

$$\begin{aligned} CS^D &= S^{with} - S^{without} \\ &= \frac{-2c^2 - 144c\Delta - 142\Delta^2 - 490cc_0 - 490\Delta c_0 + 1225c_0^2}{2450\Delta} \end{aligned}$$

$$= \frac{\Delta}{2450}(-2\mu^2 - 144\mu - 142) + \Delta\left[\frac{1}{2}\mu_0^2 - \left(\frac{1}{5} + \frac{1}{5}\mu\right)\mu_0\right].$$

We can easily determine that $-2\mu^2 - 144\mu - 142 < 0$ and $\frac{1}{2}\mu_0^2 - \left(\frac{1}{5} + \frac{1}{5}\mu\right)\mu_0 \leq 0$ when $0 < \mu \leq \frac{1}{3}, 0 < \mu_0 < \frac{1}{5} + \frac{1}{5}\mu$. Hence we have $CS^D < 0$.

The social welfare difference is

$$\begin{aligned} SW^D &= \pi^D + S^D \\ &= \frac{-13c^2 - 12c\Delta + \Delta^2 - 49cc_0 - 49\Delta c_0 + 245c_0^2}{245\Delta} \\ &= \frac{\Delta}{245}(-13\mu^2 - 12\mu + 1) + \Delta\left[\mu_0^2 - \left(\frac{1}{5} + \frac{1}{5}\mu\right)\mu_0\right]. \end{aligned}$$

It is easily to verify that $\mu_0^2 - \left(\frac{1}{5} + \frac{1}{5}\mu\right)\mu_0 \leq 0$ for any $0 < \mu \leq \frac{1}{3}, 0 < \mu_0 < \frac{1}{5} + \frac{1}{5}\mu$. When $\frac{1}{13} \leq \mu \leq \frac{1}{3}, -13\mu^2 - 12\mu + 1 \leq 0$; and when $\mu < \frac{1}{13}, -13\mu^2 - 12\mu + 1 > 0$. So we have $SW^D < 0$ when $\frac{1}{13} \leq \mu \leq \frac{1}{3}$.

(2) When $\frac{1}{3} < \mu \leq \mu_2, 0 < \mu_0 < \frac{1}{4} + \frac{1}{4}\mu$, the difference in industry profits with and without transparency is expressed by

$$\begin{aligned} \pi^D &= \pi^{with} - \pi^{without} \\ &= \frac{(c - 3\Delta)^2}{16\Delta} + \frac{3c^2 + 30c\Delta + 11\Delta^2 + 16c_0^2}{32\Delta} - \frac{2(3c - 4\Delta)^2}{49\Delta} - \frac{12(c + \Delta)^2}{49\Delta} \\ &= \frac{-715c^2 + 1650c\Delta + 13\Delta^2 + 784c_0^2}{1568\Delta} > 0. \end{aligned}$$

The consumer surplus difference is

$$\begin{aligned} CS^D &= S^{with} - S^{without} \\ &= \frac{392c_0^2 - 196c_0(c + \Delta) + (c + \Delta)(17c - 263\Delta)}{784\Delta} \\ &= \frac{\Delta}{784}(17\mu^2 - 246\mu - 263) + \Delta\left[\frac{1}{2}\mu_0^2 - \left(\frac{1}{4} + \frac{1}{4}\mu\right)\mu_0\right]. \end{aligned}$$

We can easily determine that $17\mu^2 - 246\mu - 263 < 0$ and $\frac{1}{2}\mu_0^2 - \left(\frac{1}{4} + \frac{1}{4}\mu\right)\mu_0 \leq 0$ when $\frac{1}{3} < \mu \leq \mu_2, 0 < \mu_0 < \frac{1}{4} + \frac{1}{4}\mu$. So we have $CS^D < 0$.

The social welfare difference is

$$\begin{aligned} SW^D &= \pi^D + S^D \\ &= \frac{-681c^2 + 1158c\Delta - 513\Delta^2 - 392(c + \Delta)c_0 + 1568c_0^2}{1568\Delta} \\ &= \frac{\Delta}{1568}(-681\mu^2 + 1158\mu - 513) + \Delta\left[\mu_0^2 - \left(\frac{1}{4} + \frac{1}{4}\mu\right)\mu_0\right]. \end{aligned}$$

We can easily determine that $-681\mu^2 + 1158\mu - 513 < 0$ and $\mu_0^2 - \left(\frac{1}{4} + \frac{1}{4}\mu\right)\mu_0 \leq 0$ when $\frac{1}{3} < \mu \leq \mu_2, 0 < \mu_0 < \frac{1}{4} + \frac{1}{4}\mu$. So we have $SW^D < 0$.

(3) When $\mu_2 < \mu \leq 1, 0 < \mu_0 < \frac{1}{4} + \frac{1}{4}\mu$, the difference in industry profits with and without transparency is given by

$$\begin{aligned} \pi^D &= \pi^{with} - \pi^{without} \\ &= \frac{(c - 3\Delta)^2}{16\Delta} + \frac{3c^2 + 30c\Delta + 11\Delta^2 + 16c_0^2}{32\Delta} - \frac{(c - 2\Delta)^2}{9\Delta} - \frac{1}{18} \left(20c + \frac{4c^2}{\Delta} + 7\Delta \right) \\ &= \frac{-17c^2 - 10c\Delta + 7\Delta^2 + 48c_0^2}{96\Delta} < 0. \end{aligned}$$

The consumer surplus difference is

$$\begin{aligned} S^D &= S^{with} - S^{without} \\ &= \frac{c^2 - 22c\Delta - 23\Delta^2 - 36cc_0 - 36\Delta c_0 + 72c_0^2}{144\Delta} \\ &= \frac{\Delta}{144} (\mu^2 - 22\mu - 23) + \Delta \left[\frac{1}{2}\mu_0^2 - \left(\frac{1}{4} + \frac{1}{4}\mu \right) \mu_0 \right]. \end{aligned}$$

It is easily determined that $\mu^2 - 22\mu - 23 < 0$ and $\frac{1}{2}\mu_0^2 - \left(\frac{1}{4} + \frac{1}{4}\mu \right) \mu_0 \leq 0$ when $\mu_2 < \mu \leq 1, 0 < \mu_0 < \frac{1}{4} + \frac{1}{4}\mu$. So we have $CS^D < 0$.

The social welfare difference is

$$\begin{aligned} SW^D &= \pi^D + S^D \\ &= \frac{-719c^2 + 746c\Delta - 299\Delta^2 - 294(c + \Delta)c_0 + 1764c_0^2}{1176\Delta} \\ &= \frac{\Delta}{1176} (-719\mu^2 + 746\mu - 299) + \Delta \left[\mu_0^2 - \left(\frac{1}{4} + \frac{1}{4}\mu \right) \mu_0 \right]. \end{aligned}$$

It is easily determined that $-719\mu^2 + 746\mu - 299 < 0$ and $\mu_0^2 - \left(\frac{1}{4} + \frac{1}{4}\mu \right) \mu_0 \leq 0$ when $\mu_2 < \mu \leq 1, 0 < \mu_0 < \frac{1}{4} + \frac{1}{4}\mu$. So we have $SW^D < 0$. Thus we arrive at the announced Proposition 9.

References

- Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature*, 54(2), 442–492.
- Acquisti, A., & Varian, H. R. (2005). Conditioning prices on purchase history. *Marketing Science*, 24(3), 367–381.
- Amaldoss, W., & He, C. (2019). The charm of behavior-based pricing: When consumers’ taste is diverse and the consideration set is limited. *Journal of Marketing Research*, 56(5), 767–790.
- Casadesus-Masanell, R., & Hervas-Drane, A. (2015). Competing with privacy. *Management Science*, 61(1), 229–246.
- Choe, C., King, S., & Matsushima, N. (2018). Pricing with cookies: Behavior-based price discrimination and spatial competition. *Management Science*, 64(12), 5669–5687.
- Choi, J. P., Jeon, D. S., & Kim, B. C. (2019). Privacy and personal data collection with information externalities. *Journal of Public Economics*, 173, 113–124.
- Choudhary, V., Ghose, A., & Mukhopadhyay, T. (2005). Personalized pricing and quality differentiation. *Management Science*, 51(7), 1120–1130.
- Colombo, S. (2016). Imperfect behavior-based price discrimination. *Journal of Economics & Management Strategy*, 25(3), 563–583.
- Conitzer, V., Taylor, C. R., & Wagman, L. (2012). Hide and seek: Costly consumer privacy in a market with repeat purchases. *Marketing Science*, 31(2), 277–292.
- De Nijs, R. (2017). Behavior-based price discrimination and customer information sharing. *International Journal of Industrial Organization*, 50, 319–334.

- Esteves, R. B. (2009). A survey on the economics of behaviour-based price discrimination. *NIPE-Universidade do Minho*.
- Esteves, R.-B. (2010). Pricing with customer recognition. *International Journal of Industrial Organization*, 28(6), 669–681.
- Esteves, R.-B. (2014). Behavior-based price discrimination with retention offers. *Information Economics and Policy*, 27, 39–51.
- Esteves, R.-B., & Cerqueira, S. (2017). Behavior-based pricing under imperfectly informed consumers. *Information Economics and Policy*, 40, 60–70.
- Esteves, R.-B., & Reggiani, C. (2014). Elasticity of demand and behaviour-based price discrimination. *International Journal of Industrial Organization*, 32, 46–56.
- Federal Trade Commission of United States. (2014). Data brokers: A call for transparency and accountability. Report, FTC, Washington, DC.
- Fudenberg, D., & Villas-Boas, J. M. (2006). Behavior-based price discrimination and customer recognition. *Handbook on Economics and Information Systems*, 1, 377–436.
- Goldfarb, A., & Tucker, C. E. (2011). Privacy regulation and online advertising. *Management Science*, 57(1), 57–71.
- Hann, I.-H., Hui, K.-L., & Lee, S.-Y. T. (2008). Consumer privacy and marketing avoidance: A static model. *Management Science*, 54(6), 1094–1103.
- Jing, B. (2017). Behavior-based pricing, production efficiency, and quality differentiation. *Management Science*, 63(7), 2365–2376.
- Koh, B., Raghunathan, S., & Nault, B. R. (2017). Is voluntary profiling welfare enhancing? *MIS Quarterly*, 41(1), 23.
- Li, G., Lim, M. K., & Wang, Z. (2020a). Stakeholders, green manufacturing, and practice performance: empirical evidence from Chinese fashion businesses. *Annals of Operations Research*, 290, 961–982.
- Li, G., Liu, M., Bian, Y., & Sethi, S. P. (2020b). Guarding against disruption risk by contracting under information asymmetry. *Decision Sciences*. <https://doi.org/10.1111/dec.12437>.
- Li, G., Zheng, H., Sethi, S. P., & Guan, X. (2020c). Inducing downstream information sharing via manufacturer information acquisition and retailer subsidy. *Decision Sciences*, 51(3), 691–719.
- Li, K. J. (2018). Behavior-based pricing in marketing channels. *Marketing Science*, 37(2), 310–326.
- Li, K. J., & Jain, S. (2016). Behavior-based pricing: An analysis of the impact of peer-induced fairness. *Management Science*, 62(9), 2705–2721.
- Li, X., Li, K. J., & Wang, X. (2020d). Transparency of behavior-based pricing. *Journal of Marketing Research*, 57(1), 78–99.
- Montes, R., et al. (2019). The value of personal information in online markets with endogenous privacy. *Management Science*, 65(3), 1342–1362.
- Pazgal, A., & Soberman, D. (2008). Behavior-based discrimination: Is it a winning play, and if so, when? *Marketing Science*, 27(6), 977–994.
- Posner, R. A. (1981). The economics of privacy. *American Economic Review*, 71(2), 405–409.
- Rhee, K. E., & Thomadsen, R. (2017). Behavior-based pricing in vertically differentiated industries. *Management Science*, 63(8), 2729–2740.
- Shin, J., & Sudhir, K. (2010). A Customer management dilemma: When is it profitable to reward one's own customers. *Marketing Science*, 29(4), 671–689.
- Shy, O., & Stenbacka, R. (2016). Customer privacy and competition. *Journal of Economics & Management Strategy*, 25(3), 539–562.
- Taylor, C. R. (2004). Consumer privacy and the market for customer information. *Rand Journal of Economics*, 35(4), 631–650.
- Taylor, C., & Wagman, L. (2014). Consumer privacy in oligopolistic markets: Winners, losers, and welfare. *International Journal of Industrial Organization*, 34, 80–84.
- Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. *Journal of Marketing Research*, 51(5), 546–562.
- Valletti, T., & Wu, J. (2019). Consumer profiling with data requirements: Structure and policy implications. *Production and Operations Management*, 29(2), 309–329.
- Villas-Boas, J. M. (1999). Dynamic competition with customer recognition. *Rand Journal of Economics*, 30(4), 604–631.
- Villas-Boas, J. M. (2004). Price cycles in markets with customer recognition. *Rand Journal of Economics*, 35(3), 486–501.