**Selling through online marketplaces with consumer profiling**

**Abstract**

Retail platforms obtain consumers’ individual preferences by gathering vast amounts of data and can deliver such information to online retailers to support their pricing activities; this is called consumer-profiling services (CPS). We develop a game-theoretic model to study how a retail platform should provide CPS in light of retailers’ competition and consumers’ data-blocking activities. We show that exclusively providing data to high-quality retailers results in a net benefit for the platform and retailers. Low-quality retailers benefit from refusing the CPS provided by the platform to avoid head-to-head competition. In addition, we find that consumers’ data blocking can benefit both the platform and retailers when the data-blocking cost is moderate, which is counterintuitive. We also find that data blocking always hurts consumer surplus and social welfare. To test the robustness of the main model, three extensions are discussed: sequential pricing, asymmetric production costs, and positive service fees.

**Keywords**: Retail platforms; Consumer profiling services; Data blocking; Competition; Pricing

**1 Introduction**

The digital economy has transformed business practices and consumer choices (Chen et al., 2020) in the past decade, resulting in the rapid and strong growth of platform-based service providers such as Amazon, eBay, Alibaba, and Airbnb. Using Liu et al.’s (2021) definition, originally from Manville (2016), ‘the platform’ refers to ‘*the digitized, open, and participative business models creating commercially connected ecosystems of producers and consumers’*. In this sense, this digitised marketplace is not merely a distributional channel, but it also creates infrastructural support and coordinates commercial activities amongst both consumers and sellers. Powered by the advantage of platform economy, retail E-commerce has witnessed nearly a five-fold growth in sales from 2014 to 2022 and is projected to reach approximately USD 7.4 trillion in sales worldwide by 2025 (Statista, 2022).

Compared to brick-and-mortar retailing, online retail platforms can store and access richer data about consumers’ behaviours and engagements, allowing individual consumers to be thoroughly and diversely profiled. For example, platforms can gather clickstream data from customers’ browsing histories to record the number of searches or views for a particular product; the amount of time an individual customer has spent viewing products; and the view counts and click-through rates of advertisements, etc. (Liu et al., 2021). Such data are not directly accessible to sellers but can be essential for improving their operational decisions. Therefore, consumer data create a business opportunity for platform operators to sell consulting services based on this information. For instance, Alibaba offers a business-analytics tool called *Shengyicanmou* that allows sellers to obtain customer-behaviour analysis as well as information about competitors. Similar services exist on Amazon and Ebay as *Retail Analytics Premium* and *Terapeak Research*, respectively.

In business practice, consumer profiling influences retailers’ selling decisions in different ways. On the one hand, online retailers gain the ability to adopt behaviour-based pricing (BBP) to target more consumers and to extract more consumer surplus (CS). As such, consumer profiling benefits both the retailer and the platform, incentivising the platform to provide consumer profiling to its retailers. On the other hand, if customer-profiling service (CPS) is freely available to competing retailers, they will engage in fierce price competition, which may hurt both the retailers’ and the platform’s profits. Therefore, to maximise the overall earnings from the consumer base, platforms must control their retailers’ adoption of consumer profiling.

Although BBP (supported by consumer profiling) can benefit retailers, it might cause consumers' utility losses. Accordingly, online buyers are concerned that their digital footprint will be used against them and take precautions not to leak private information (Fortes and Rita, 2016). Such actions can include emptying their browsers’ viewing histories, blocking cookies, using the incognito mode or temporary e-mail addresses for new-accounts creation, maintaining ‘fake’ online identities, using different payment methods, and spreading purchases among unrelated vendors (Conitzer et al., 2012). Such actions are costly to consumers in terms of time, effort, and sometimes money. Consequently, this behaviour falls under the *hide and seek* game category proposed by Conitzer et al. (2012), which discounts consumers’ utility and distorts platforms’ records of consumer behaviour. Therefore, when shopping online, consumers must also decide whether to protect their private data. Such data-blocking activities affect platforms’ profiling outcomes as well as their service-provision strategies and pricing decisions.

Based on the above discussion, we aim to optimise CPS strategies for platform providers when customers’ data-blocking behaviour is considered and sellers are competing with substitutable products with different quality levels. Specifically, the research questions are as follows:

* For consumers, how does data-blocking behaviour interact with their purchasing behaviours?
* For a platform, how can it provide its profiling service to competing online retailers given consumers’ data blocking?
* For online retailers, how can they optimally determine their prices in response to consumers’ data blocking and platform’s profiling-service provision?

The reminder of our paper is organised as follows: Section 2 identifies the research gaps through an extensive review of relevant studies. Section 3 describes the problem in detail and formulates it based on a game-theoretic model. Section 4 solves the model and analyses the equilibrium results from three cases. Section 5 examines the impacts of customers’ data-blocking behaviour on profits, CS, and social welfare (SW). Section 6 provides two extensions to test the robustness of the main results. Section 7 concludes the paper.

**2 Literature Review**

We divide our literature review into three research streams: first, the marketplace; second, consumer profiling and firms’ pricing decisions; and third, BBP, where past purchases are used to learn about an individual consumer’s purchasing preference.

**2.1 Marketplace**

For the first research stream, a common question is whether an online retailer should (also) function as a marketplace for third-party sellers, as is the case with Amazon. Hagiu and Wright (2014) show that a reseller’s role is favoured by increasing marketing-spillover effects between products, and they investigate which type of products (long- or short-tail) the two roles should provide. Mantin et al. (2014) investigate a similar setting and find that the presence of a marketplace decreases profits for direct suppliers to retailers, and that in the presence of a marketplace, manufacturers should let retailers dictate contractual terms to minimise competition between the latter and third-party sellers. Tian et al. (2018) adopt the perspective of product heterogeneity and find that with increased product similarity, and thus competition among suppliers, a reseller’s role is more profitable for an online retailer, whereas less product similarity favours the marketplace selling.

Another approach is agency selling, whereby an online retailer provides manufactures with direct access to consumers for a fee. Abhishek et al. (2016) find that agency selling is more efficient than regular reselling. However, online marketplaces prefer agency selling, particularly when it has a negative impact on demand in physical stores. Yi et al. (2018) investigate a manufacturer’s choice between directly selling to customers or through an agent. Using a newsvendor model to maximise profit, they find that agent selling is preferable when customers place a high value on fairness, but that the opposite case favours direct selling.

Some research into online marketplaces considers information and customer-search costs. Among these studies, Dukes and Liu (2016) investigate how online marketplaces can reduce customer’s cost of searching for products. They identify an equilibrium search environment in which a limited number of suppliers is audited (breadth), while product features are completely audited (depth). Considering supplier heterogeneity, Zennyo (2020) investigates an online platform with agency and wholesale selling for low- and high-volume suppliers and finds that low-volume suppliers prefer agency selling at equilibrium. Kwark et al. (2017) consider wholesale versus marketplace selling on online platforms in the light of information availability. They find the wholesale scheme to be a preferred strategy if quality information is most important, while marketplace selling takes preference when product fit is most important.

Regarding online marketplaces and pricing, Geng et al. (2018) investigate product bundling. They find that wholesalers prefer bundling while marketplace providers prefer add-on pricing, and that agency selling rather than wholesale can lead to higher profits for both a platform and a reseller. In an empirical study, Zhang et al. (2018) report on a field experiment on the Alibaba online marketplace, where customers with items in their basket were offered discounts on those items. While this experiment increased customer engagement compared to a reference case, it also increased customers' strategic behaviour and reduced the overall price paid. Sun et al. (2020) investigate how traffic and sales to online retailers are affected by promotion either through sponsored search or through social-media endorsement. They find that sponsored search universally drives traffic but only benefits high-reputation sellers’ sales. Social-media endorsement repeats the same pattern to a lesser degree.

Furthermore, some research focuses on the role of information in online marketplaces. Hao and Tan (2019) investigate how a supplier and a retailer are affected by information disclosure. In particular, they find that the optimal degree of information disclosure for either party strongly relates to the revenue-sharing mechanism. Li et al. (2021) examine co-opetition between an online marketplace, a reseller, and a manufacturer under various settings of information sharing. They find equilibria at full information sharing or information sharing only with the manufacturer. Another approach to information is adopted by Morath and Munster (2018), who investigate an online platform that can decide whether to require registration before customers receive full product information. They demonstrate that the profit gain from requiring registration relates to the prospect of repeat purchases and the effect of offering store credit to encourage registration.

Our study falls into the research scope of marketplace-information management; however, it differs from previous studies in two respects: First, we consider consumers’ data-blocking activities in a platform’s information-management models. Second, we consider the platform’s consumer-profiling information-provision strategies for vertically differentiated online retailers.

**2.2 Consumer profiling**

Consumer profiling helps firms obtain valuable information, which helps them make better decisions on pricing. For example, Lei et al. (2018) consider an online retailer selling multiple products through fulfilment centres. They provide two heuristics for setting the sales price and for selecting the origin warehouse for shipping to customers.

However, to safeguard against being profiled, customers may engage in privacy protection. Acquisti et al. (2016) highlight the difficulty of establishing a unifying theory of privacy protection due to the wide context studied. They show empirical and theoretical results where privacy protection can both improve or worsen total and consumer welfare. As part of the theory on privacy protection, Conitzer et al. (2012) investigate a monopolist facing heterogenous consumers who may engage in identity management to prevent being recognised as past customers. They find that all customers engage in identity management when there is no cost attached, but that this leads to the highest possible profit for the monopolist. They also show that customer benefit may be maximised for a cost of anonymity greater than zero. Shy and Stenbacka (2016) investigate how switching costs relate to customer-privacy protection, firm profits, and welfare. Their results show that firm profits are maximal under weak but nonzero privacy protection, while total welfare and CS increase with privacy protection. Montes et al. (2019) investigate a duopolistic model in which companies can price discriminate and customers can guard their privacy at a cost. An increased cost of privacy is followed by increased customer welfare and decreased profits. For companies selling customer information, they find the optimal strategy to be to sell it to only one firm. Chen and Jiang (2019) investigate a duopoly where one company may record customer behaviour, e.g. through a black box in a car, and offer differential pricing. They find that it is optimal for exactly one company to record customer behaviour as the differential pricing reduces competition and increases both companies’ profits. Finally, Li and Li (2022) consider a situation in which consumers can manipulate their own data to confuse retailers, thus fighting against personalised pricing.

Other approaches to customer profiling and privacy protection include Taylor and Wagman (2014), who investigate the effect of privacy regulation in four oligopoly models and find that agents who stand to gain from privacy regulation depend both on the choice of model and on economic parameters. Another approach follows from Casadesus-Masanell and Hervas-Drane (2015). They investigate competing companies that gain revenue from selling to customers as well as from selling information about customers in a secondary market. They find that firms achieve maximum profits when they have one single revenue source targeted towards the broad market, and that competition drives the emergence of high-privacy services.

Our study is also closely related to the research on consumer profiling. However, none of the above studies considers consumer profiling as a business model between online platforms and their online retailers. In this study, we aim to determine how platforms can use consumer profiling as leverage to enhance the performance of the platform supply chain.

**2.3 BBP**

BBP concerns differentiated pricing based on past transactions with specific customers. Jing (2016) investigated how two competing firms fared with and without BBP. They found that the effects of implementing BBP depended on quality differentiation: when exogenous, a more efficient firm suffered reduced profits while a less efficient firm increased profits; when quality differentiation was endogenous, both firms’ profits increased at the cost of SW. Choe et al. (2017) investigate a similar two-firm model with two stages: one for behaviour exploration and one for exploitation. They identify two asymmetric equilibria favouring an aggressive strategy in terms of positioning or pricing. Nominally, both firms are worse off implementing BBP. Studies on BBP were extended to a supply-chain setting with serially linked manufacturers and retailers in Li (2018). Compared to a uniform pricing setting, Li (2018) finds that supply chains in which both manufacturers and retailers adopt BBP increase supply-chain profits while reducing CS, whereas BBP applied only at the retailer echelon reduces profits. Jullien et al. (2022) studies personalised pricing strategies for a dual-channel supply chain in which the retailer acts as the upstream supplier’s competitor.

Among studies investigating the economics of BBP, Elmachtoub et al. (2021) identify upper regret bounds for selecting feature-based BBP over uniform pricing and provide tighter bounds when additional information is known about the distribution of customers’ valuation. Xu and Dukes (2021) investigate a monopolist using BBP who, additionally, may display a list price next to the personalised price. They show that this strategy increases total SW at the expense of average CS. Hajihashemi et al. (2021) investigate the connection between BBP and network effects and identify conditions under which BBP, as compared to uniform pricing, leads to an overall reduction in demand, profit, and SW.

Additionally, two studies connect privacy protection to BBP: First, Chen et al. (2020) investigate the effects of customers using identity management to avoid being targeted by BBP. With multiple competing vendors, they find that active identity management may benefit an individual customer; however, as the proportion of such customers rises, firm profits increase and SW decreases. Second, Chen et al. (2021) investigate dynamic pricing when the demand function linking price and customer information is unknown. They develop a pricing policy for learning the demand function while simultaneously achieving a privacy guarantee and a regret-based performance guarantee. The privacy guarantee prevents information about customer transactions from being leaked to third parties. However, in the above literature, none of the scholars considers the interactions of BBP and consumers’ data-blocking activities in platform supply chains, which is a research gap that will be filled in this study.

**3 Model Formulation**

There are three types of agents: consumers, online retailers, and an online marketplace. There are two retailers, both selling substitutable products through the single marketplace to many end consumers. The retailers are vertically differentiated in terms of their product offering. Let H denote the retailer selling high-quality products and L denote the one selling low-quality products. Without loss of generality, we normalise Retailer H’s product quality to 1 and Retailer L’s to . There are two sets of consumers in the market. We refer to the first set as *new consumers*. For consumers in this set, private information cannot be obtained by the marketplace, and BBP is not applicable for them. These consumers can be those who have never purchased or browsed a product category on the sales platform (Valletti and Wu, 2020). We then refer to the second set as *old consumers*. For consumers in this set, their private information may be obtained by the marketplace and online retailers. If firms know their information, they will be offered behaviour-based prices. We set the number of consumers in the first set to and that in the second set to , so that the total market size is normalised to 1. Throughout the study, we use ‘n’ and ‘o’ to denote the new and old consumers’ decisions, respectively.

In online business practice, the marketplace often charges a fixed commission as a percentage of revenue, , to all online retailers. Therefore, the retailers’ revenue gain directly affects the platform’s profits. In our model, we assume that only the marketplace can collect and analyse consumer information. This is motivated by large selling platforms, such as Amazon, Taobao, and JD, who provide the sales infrastructure and therefore are the parties that can collect the most data, as well as because they have significant resources for data analytics. Having obtained consumers’ preference information, the marketplace has three options in dealing with it: no information offering (Case N), offering information to online Retailer H (Case H), and offering information to both online retailers (Case HL). The information-provision strategy affects both the consumers’ purchasing behaviour and the retailers’ pricing decisions and revenues, which eventually determine the platform’s profits.

For both sets of consumers in the market, the basic willingness to pay is modelled as , which is uniformly distributed on the line of [0,1]. Considering the online retailers’ pricing decisions in response to the information-provision strategies and product qualities, we model consumer-purchasing behaviours for the two sets of consumers in the three cases as follows:

(1) For new consumers, the marketplace’s information-provision strategy does not affect their purchasing behaviour. In the three cases, retailers can only provide uniform prices to all new consumers. Let and denote the uniform prices that are offered by online Retailers H and L, respectively; the gross utility for a type consumer who purchases Product H or L can respectively be expressed as and . A type consumer will then purchase Product H when or product L when .

(2) For old consumers, their purchasing behaviour will be different across the three cases, depending on the information structure. Specifically,

* In Case N, the marketplace does not provide consumer-preference information to either online retailer. Without such information, BBP is not applicable, and old consumers exhibit the same purchasing behaviours as new consumers.
* In Case H, the marketplace will provide consumer-preference information to Retailer H. In response to the marketplace’s potential information provision, we assume that old consumers can choose to block their data, i.e. hide their individual preference information, at a cost of , or not to block the data. We use to denote a threshold above which all the consumers block data, whereas no consumers block data while below . Consumers who choose to hide their information (,1]) will receive uniform prices from retailers H and L. However, for consumers who do not block data (]), Retailer H provides behaviour-based prices, while Retailer L provides a uniform price. We use the notation ‘~’ to denote the behaviour-based price in subsequent sections. The gross utility for a type consumer who purchases Product H or L can be expressed as follows:

and *,* respectively*.*

Similarly, a type consumer will purchase Product H when or Product L when . Here, is the decision of consumers and will be presented in the model analysis.

* In Case HL, the marketplace provides consumer-preference information to both retailers. As before, the consumers may choose whether to hide their individual information. We also assume that there exists a threshold of above which old consumers block data. For consumers who choose to block data (,1]), both retailers provide uniform prices. However, for consumers who do not block data (]), both firms can provide behaviour-based prices. The gross utility for a type consumer who purchases Product H or L can be expressed as

and *,* respectively*.*

A type consumer will then purchase Product H when or Product L when . Here, is a consumer-decision variable.

Throughout the study, we normalise the retailers’ selling and replenishment costs to zero to focus on consumer-profiling issues. The impacts of cost differentiation will be discussed in the extensions. The game in this model proceeds in several steps as follows:

* **Step 1:** Each consumer realises their willingness to pay as and decides whether to block their individual data.
* **Step 2:** The marketplace conducts consumer profiling and chooses from the three CPS-provision strategies: N, H or HL.
* **Step 3:** The two competing retailers decide whether to accept the marketplace’s offers of CPS provision (if available) and then simultaneously determine their sales prices.
* **Step 4:** Consumers make purchase decisions and the market clears.

**4 Model analysis**

In this section, we analyse the equilibrium results for the three cases. Additionally, we derive the marketplace’s optimal information-sharing strategies.

**4.1 Case N: Marketplace does not provide CPS**

In Case N, no CPS is provided to the retailers. Therefore, the retailers cannot obtain access to consumers’ individual preferences, and prices to both new and old consumers are uniform. Retailer H’s market demand is and Retailer L’s demand is . The retailers’ and marketplace’s profit functions can be expressed as follows:

*,*

*,*

*.*

Following the decision sequence described in Section 3, we derive the equilibrium results as follows:

**Proposition 1. *(Equilibrium results for Case N)*** *When the marketplace does not provide CPS to online Retailers H and L,*

*(1) the equilibrium selling prices are, respectively,* and

*(2) the profits for the retailers and marketplace are, respectively, ,* , and *.*

Proposition 1 provides a benchmark case in which the marketplace does not provide CPS. Therefore, neither retailer uses BBP and consumers never hide their information. From the results, the selling price and profit for online Retailer H decrease in Parameter , while those for online Retailer L are concave in . Additionally, the marketplace’s total profit is decreasing in . The above results indicate that intensified competition (higher ) negatively impacts the platform’s and powerful retailer’s total profits. However, it benefits the strong retailer when is small and hurts the weak retailer when is large. We use Case N as a benchmark in this study to investigate the impact of the marketplace’s CPS decisions.

**4.2 Case H: Marketplace provides CPS to Retailer H**

Turning to case H, in which the marketplace provides CPS only to the powerful retailer. In this case, the demands for new consumers can be obtained with respect to the (uniform) price decisions, and . The demands from old consumers is more complex. We present consumers’ data blocking and purchasing behaviours in response to the online retailers’ pricing decisions as follows.

**Lemma 1. (Old consumers’ data blocking and purchasing activities for case H)** *Let*  *and ,*

*(1) When , old consumers with the type of will block data; otherwise, they will not block data. All of them will purchase from Retailer H, and pay the prices conditional on their individual type, i.e.,*

*.*

*(2) When , old consumers with the type of will block data; otherwise, they will not block data. All old customers will purchase from Retailer H, and pay the prices conditional on their individual type, i.e.,*

*.*

In Lemma 1, we show the old consumers’ data blocking and purchasing decisions with respect to the two retailers’ pricing decisions. First, we show that when or , and the powerful retailer (H) obtains the consumers’ preference information, it can drive Retailer L out of the market and extract as much consumer surplus as possible. For example, when and , Retailer H will extract all the consumer surplus and leave . When , it becomes more difficult to extract consumer surplus because of the existence of Retailer L. In this condition, Retailer H will set a behaviour-based price of such that , is never below , which deters Retailer L’s entry and extracts the most possible consumer surplus. Second, we show that consumers’ data blocking decisions are dependent on the selling prices, the product quality and the data blocking cost. Specifically, when the selling price of Retailer H (L) becomes higher (lower), Retailer L’s product quality or consumers’ data blocking cost becomes higher, consumers are less willing to block their private information.

As we have assumed case H, the marketplace cannot obtain private information from new consumers. Therefore, new consumers behave the same in case H as in case N. Based on the consumers’ purchasing and data blocking responses, we formulate the profit functions for the retailers and the marketplace, and derive their equilibrium results.

*,*

*,*

*.*

Following the decision sequence described in Section 3, we derive the equilibrium results for Case H as follows (the equilibrium results are presented in Table 1 in Appendix B):

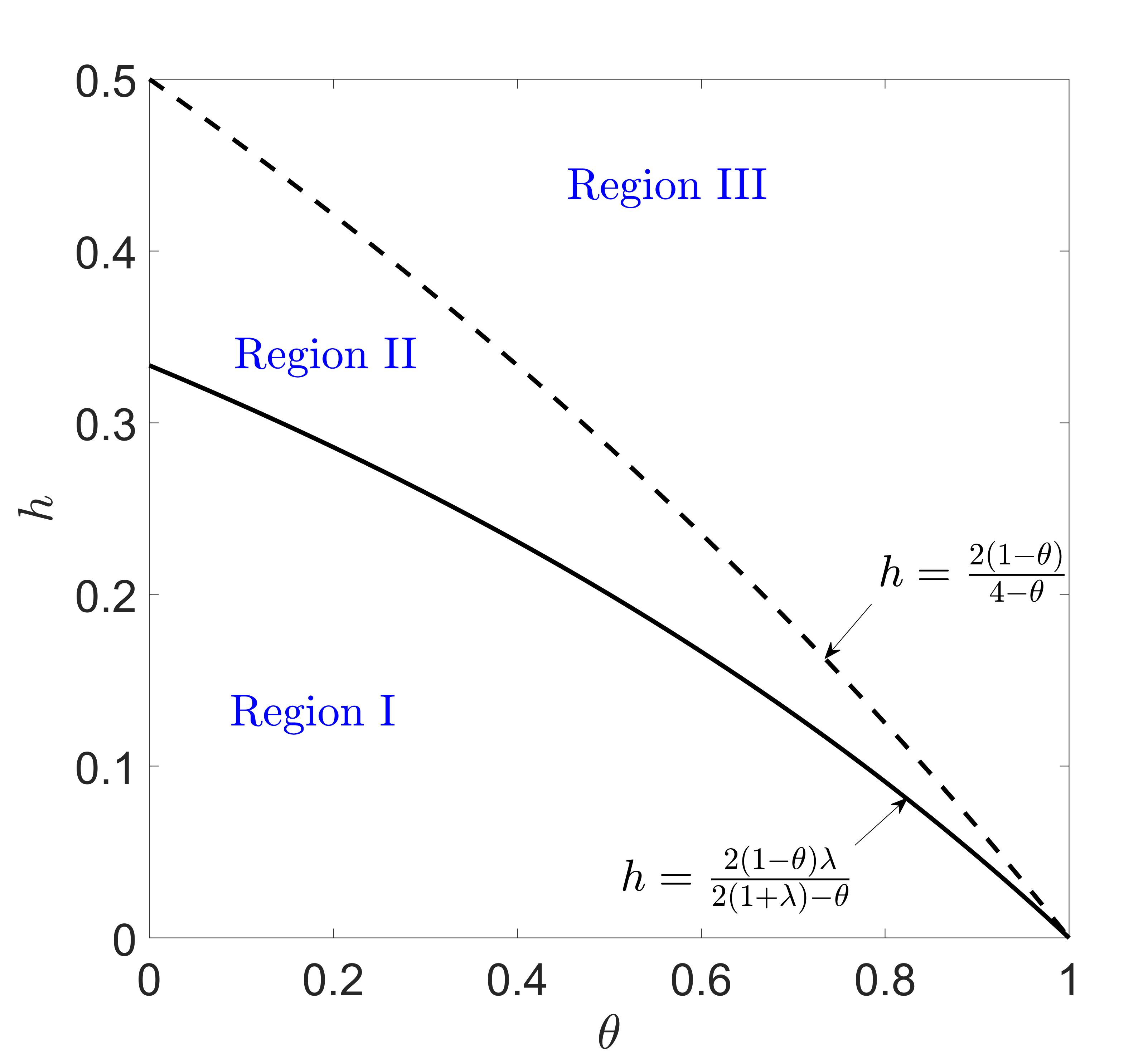
**Proposition 2. (Equilibrium results for Case H)** *Let*

*(1) When , old consumers with the type of will block their data;*

*(2) When , old consumers will never block their data. In this condition,*

*(i) if , they have incentives to block data but will be deterred by Retailer H;*

*(ii) if , they have no incentives to block data.*



**Figure 1.** Decision region for case H. Note: .

In Proposition 2 and in Figure 1, we show the equilibrium results for Case H. First, we show that when the consumers’ data-blocking cost is relatively small, i.e. , the powerful retailer will tolerate some of the data blocking (see Region I). In this situation, stopping consumers’ data-blocking actions is uneconomical because it requires Retailer H to set a sufficiently high price to the anonymous consumers, which reduces the utility of data blocking. Second, for a moderate data-blocking cost, , high-value consumers still have incentives to hide their information; however, online Retailer H will stop the consumers’ data blocking by raising their selling price to the anonymous consumers (including all the new consumers and some of the old ones; see Region II). Therefore, consumers’ data blocking is deterred by online Retailer H. Last, when the data-blocking cost is sufficiently high, , consumers have no incentives to hide their data because of its negative impact on consumer utility (see Region III). Therefore, in this region, Retailer H can ignore data blocking and provide behaviour-based prices to all old consumers.

In addition to the data-blocking cost, we observe in Figure 1 that the two thresholds are decreasing in . This implies that the sizes of Regions I and II shrink while that of Region III rises in . In practice, it means that when the quality of Retailer L’s product is higher, the intensified competition reduces consumers’ incentives to hide their data. This happens because intensified competition causes both retailers to offer lower selling prices and consumers enjoy higher utility without hiding their information.

**4.3 Case HL: Marketplace provides CPS to Retailers H and L**

We turn to Case HL in which the marketplace provides CPS to both retailers. In this case, the demand for new consumers remains the same: and . We present consumers’ data-blocking and purchasing behaviour in response to the retailers’ pricing decisions as follows (see Appendix A for the proof of Lemma 2):

**Lemma 2. (Old consumers’ data blocking activities in case HL)**

*Let , old consumers with the type of will block data; otherwise, they will not block data. All of them will purchase from Retailer H, and pay the prices conditional on their individual type, i.e.,*

*.*

In Lemma 2, we show the old consumers’ purchasing decisions with respect to the two retailers’ pricing decisions. First, we show that there exists a threshold, , above which a consumer will hide private information to avoid BBP. Otherwise, the consumer will not hide the private information as doing so would result in lower utility. Second, when both retailers receive private consumer information from the platform, the powerful retailer will dominate the old consumer market and will set aggressive prices to drive Retailer L out of that market.

Notice that the data-blocking threshold is determined by the selling price, , quality of product, L (), and consumers’ data-blocking cost (). Based on the consumers’ purchasing and data-blocking decisions, we formulate the profit functions for the retailers and marketplace and derive their equilibrium results:

*,*

*,*

*.*

Following the decision sequence described in Section 3, we derive the equilibrium results for Case H as follows (the equilibrium results are presented in Table 2 in Appendix B):

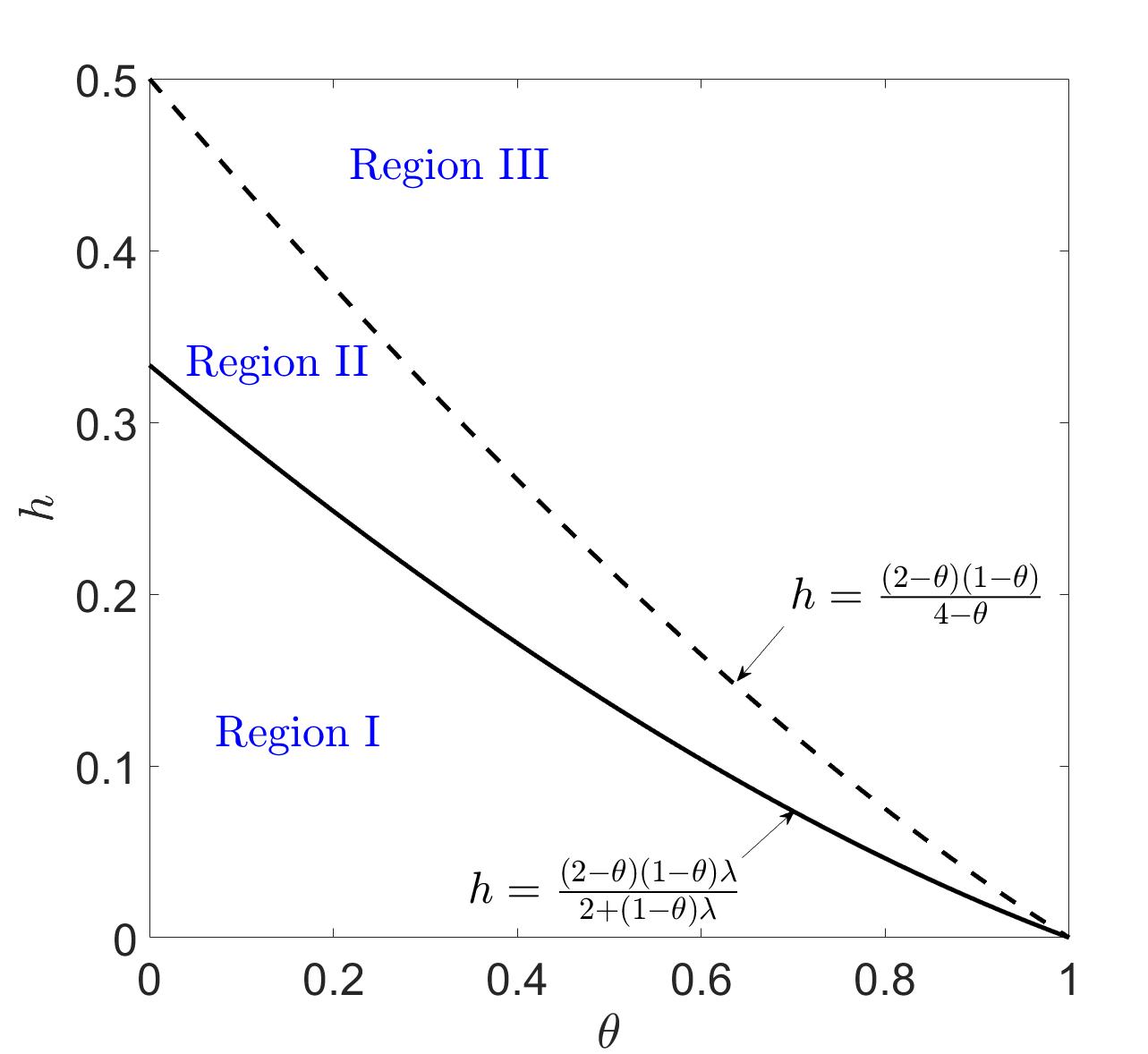
**Proposition 3. (Equilibrium results for Case HL)** *Let :*

*(1) When , old consumers with the type of will block data.*

*(2) When , old consumers will never block data. In this condition,*

*(i) if , they have incentives to block data but will be deterred by Retailer H; and*

*(ii) if , they have no incentives to block data.*



**Figure 2**. Decision region for Case HL. **Note**: .

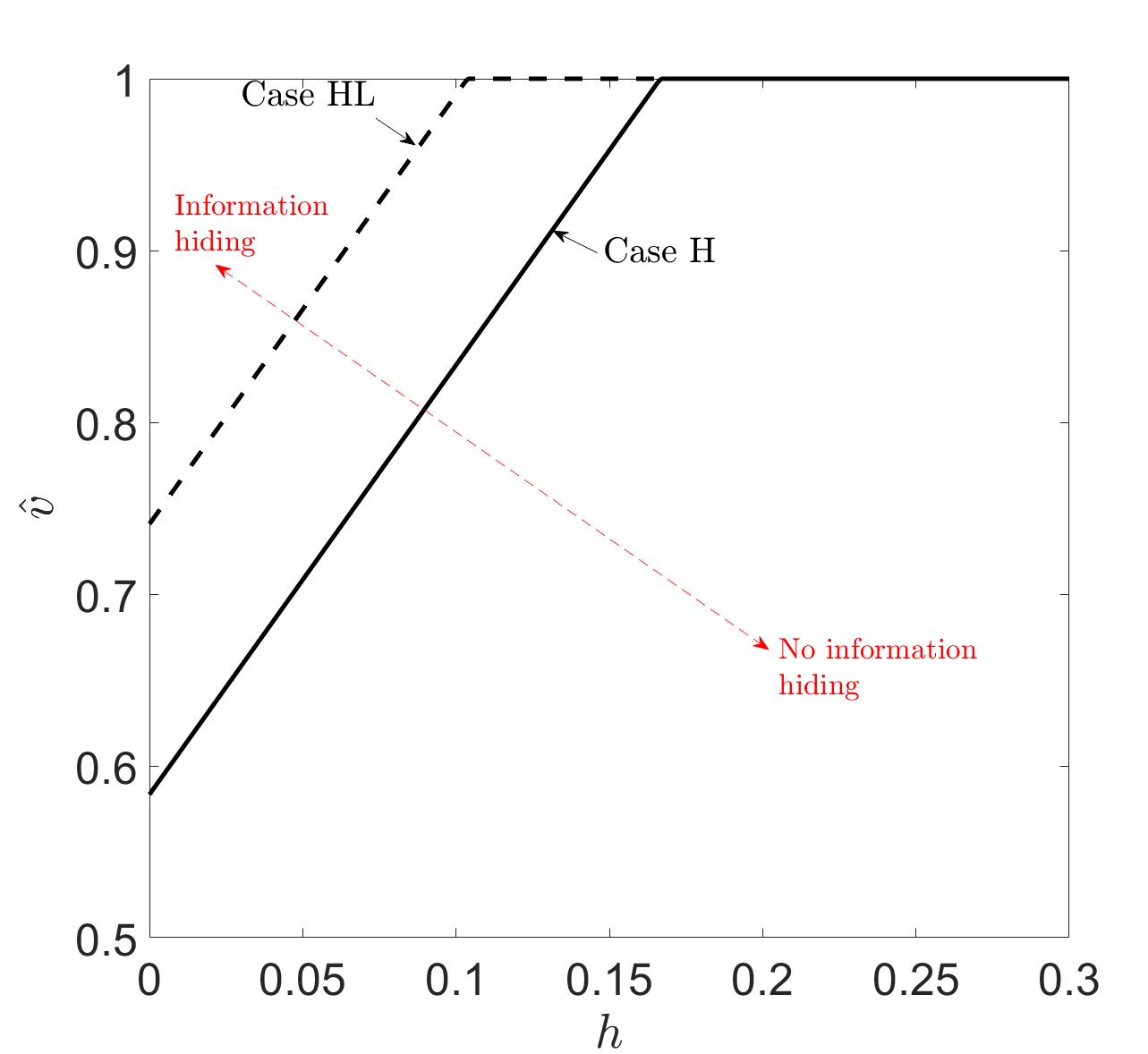
In Proposition 3 and in Figure 2, we show the equilibrium results for Case HL. The decision regions appear in Figure 2. First, we show that when the data-blocking cost is relatively small, i.e. , the two online retailers will tolerate that some consumers hide their information (see Region I in Figure 2). In this situation, as explained in Proposition 2, stopping consumers’ data-blocking actions is uneconomical because the retailers must set high prices to make consumers switch from data blocking to not hiding. Second, for a moderate data-blocking cost, i.e. , there exists a region in which data-blocking behaviours are deterred. Last, when the data-blocking cost is high, i.e. , retailers can adopt BBP without any obstacles because no consumer will experience positive utility from blocking.

Moreover, we observe in Figure 2 that the sizes of Regions I and II are decreasing in , while that for Region III is increasing in . Therefore, in Case HL, the result that intensified competition reduces consumers’ data-blocking incentives also obtains.

**4.4 Comparison of strategies N, H and HL**

In this section, we compare the equilibrium results for the three cases.

**Proposition 4. (Comparison of consumers’ data blocking behaviours)** *The threshold value above which old consumers will block data in the three cases satisfies* .



**Figure 3.** The segmentation of anonymous consumers *w.r.t.* . **Note**: , .

In Proposition 4 and in Figure 3, we compare the segmentation of anonymous old consumers in the three cases. First, we show that in each case, the sizes of the anonymous-consumer segments () in Cases H and HL are weakly increasing in consumers’ data-blocking cost, , which agrees with the intuition the data-blocking cost reduces consumers’ incentives to be anonymous. Second, we find that in Case H, a higher number of consumers hide their information than in Case HL. One might think that when consumers’ private information was obtainable by more retailers, consumers would be more concerned about privacy issues and would be more likely to hide their information to protect their interests. However, in our model, this intuition fails. When information is obtained by the two competing retailers (Case HL), fewer consumers choose to hide their information. This follows because the intensified competition weakens the retailers’ pricing power in the non-anonymous consumer segment, thereby protecting these consumers from BBP.

Next, we compare equilibrium profits for online Retailers H and L and the marketplace to derive the equilibrium strategy for the three entities.

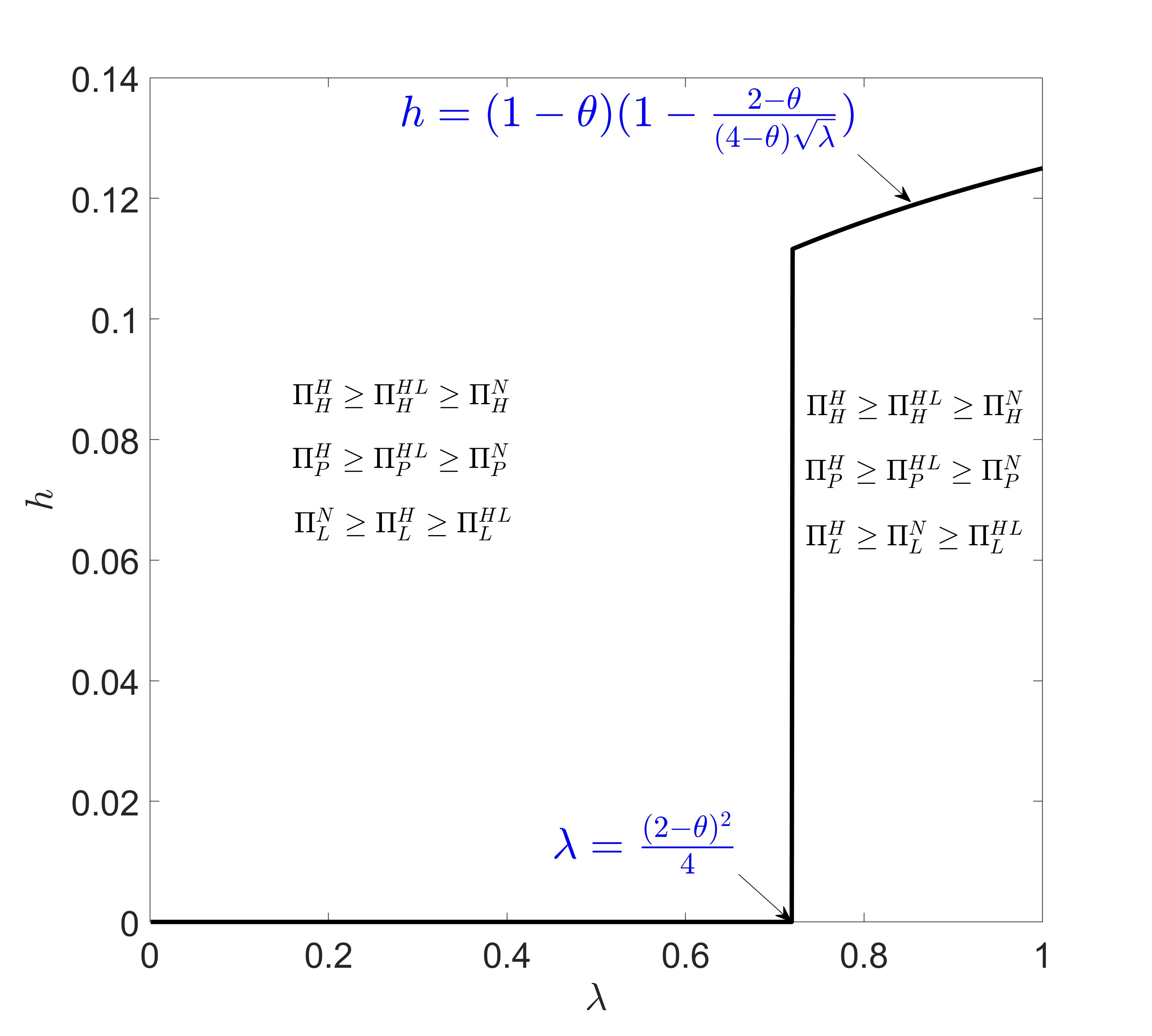
**Proposition 5. (Comparison of profits)** *Comparing the equilibrium profits for the three cases, we have the following results.*

*(1) , ;*

*(2) When , or and , ;*

*When and , .*

*(3) The equilibrium strategy is H.*



**Figure 4.** Comparison of profits in the three cases.

In Proposition 4 and in Figure 4, we show the comparative-profit results for the three cases. First, we show that when the marketplace provides information exclusively to the powerful retailer, both Retailer H and the marketplace obtain the highest profits. However, when no information is provided to either retailer, both Retailer H and the marketplace obtain the lowest profit. Interestingly, providing information to both retailers does not result in the highest profit for the marketplace, which is counterintuitive.

Second, investigating the weak retailer’s profit, we find that Strategy HL always results in the lowest profit. When the hiding cost is high or new-consumer segmentation is low, Strategy N results in the highest profit; when the hiding cost is low and the new-consumer segmentation is high, Strategy H dominates Strategy N and results in higher profit. In conventional wisdom, obtaining information helps firms to provide behaviour-based prices and to extract more surplus from consumers. However, in a vertically competing market, where a weak retailer can obtain information and use BBP, competition in the market will be intensified and both retailers’ profitabilities will suffer if information is obtained by all the retailers in the market. Therefore, a weak competitor in the market is better off avoiding head-to-head competition. This also explains why Strategy HL is neither preferred by the powerful retailer nor by the marketplace. Investigating the comparative results, we find the equilibrium strategy to be H. This implies that in a vertically differentiated online market, the marketplace should only provide information to the powerful retailer, which could result in a *universally beneficial* outcome for all entities in the online marketplace.

**5 Impacts of consumers’ data blocking**

In this section, we examine the impact of consumers’ data blocking on profits, CS, and SW. We compare the model in Case H (which is the equilibrium strategy) with an extreme case in which data blocking is infeasible (i.e. the blocking cost is prohibitively high). We first show the impact on profit as follows (see the expressions for and in Table 4, Appendix B):

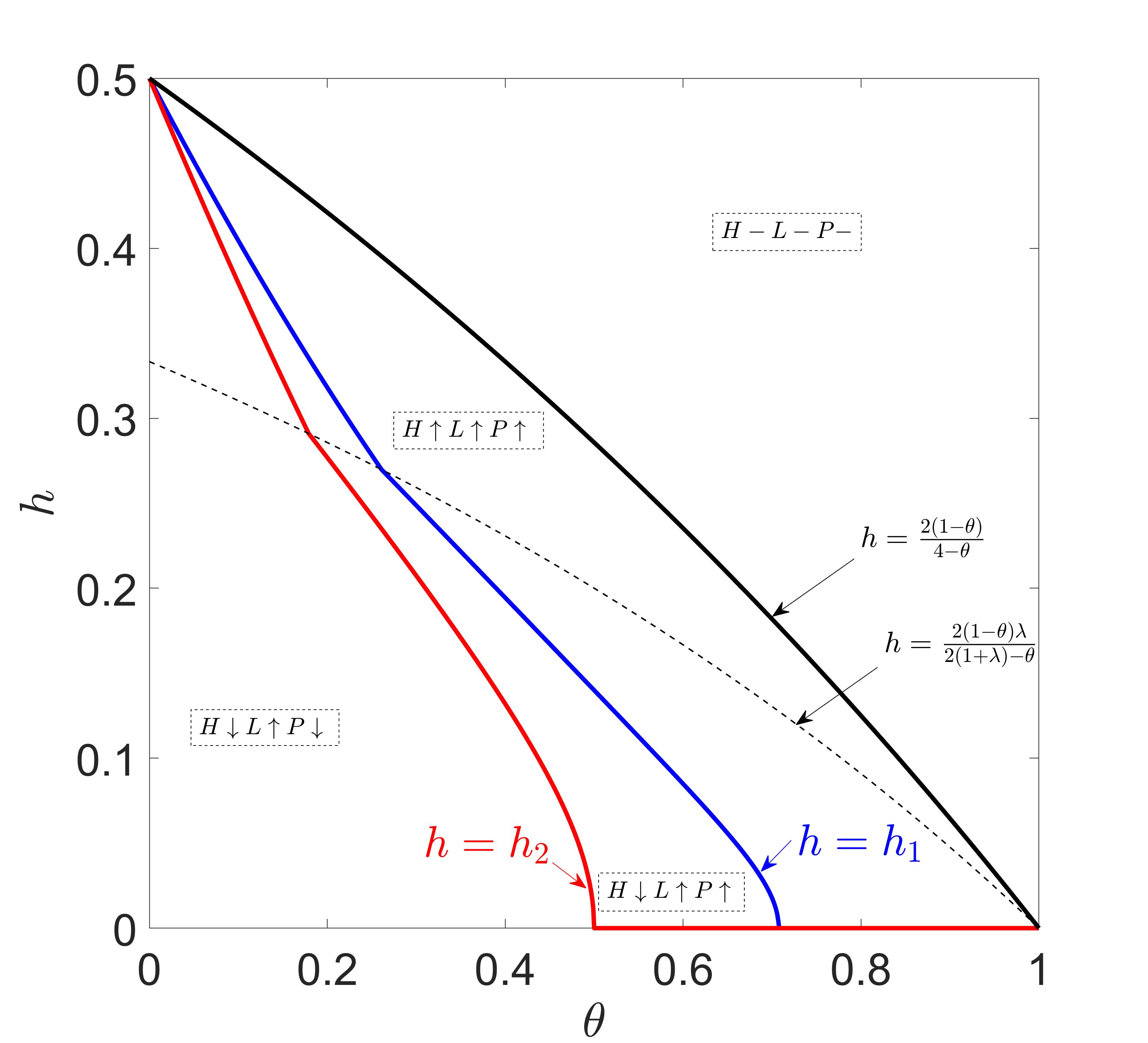
**Proposition 6. (Impacts of data blocking on profit)** *In the presence of consumers’* *data blocking,*

*(1) Retailer H benefits when , and is hurt when ;*

*(2) The platform benefits when , while is hurt when ;*

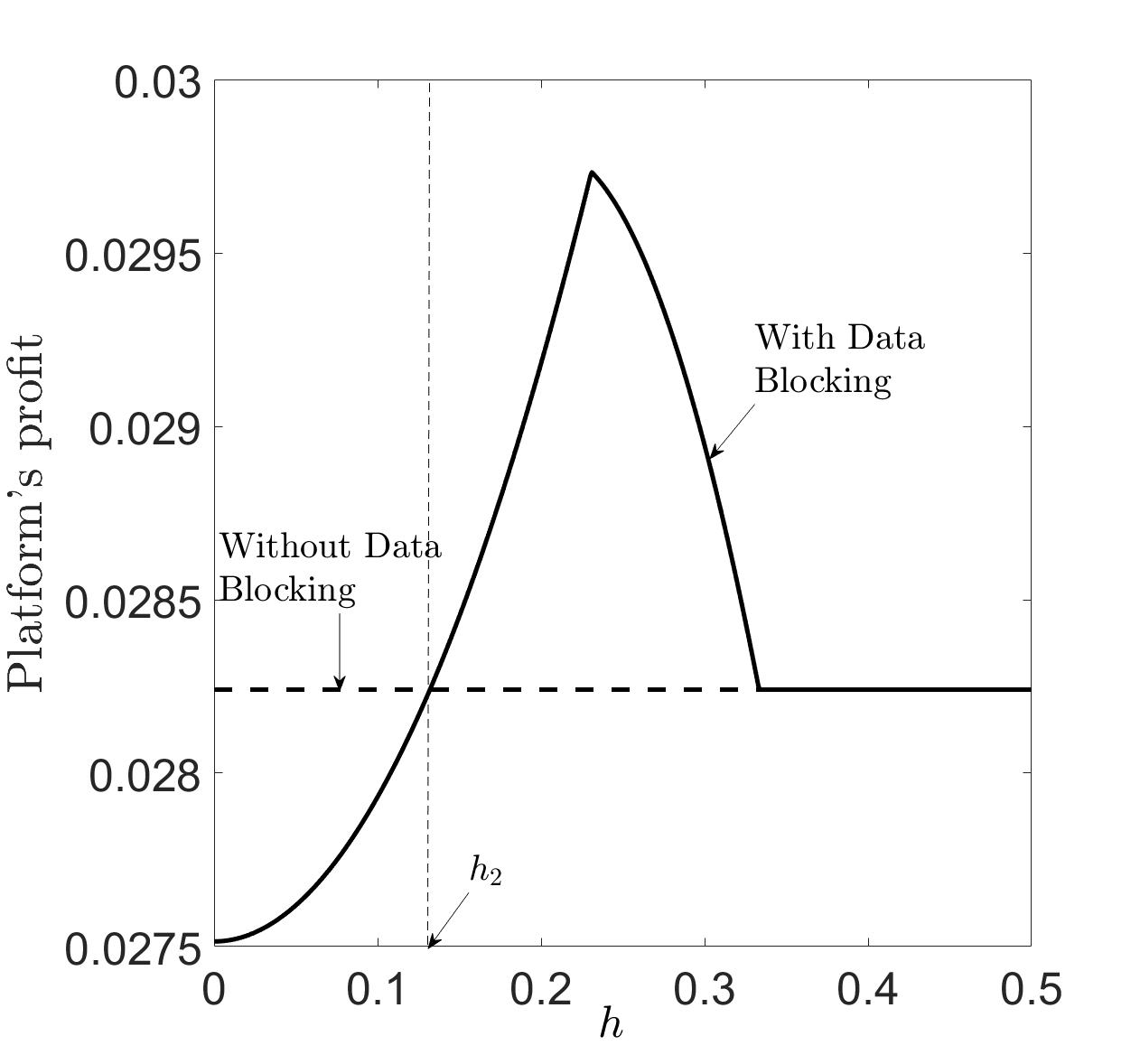
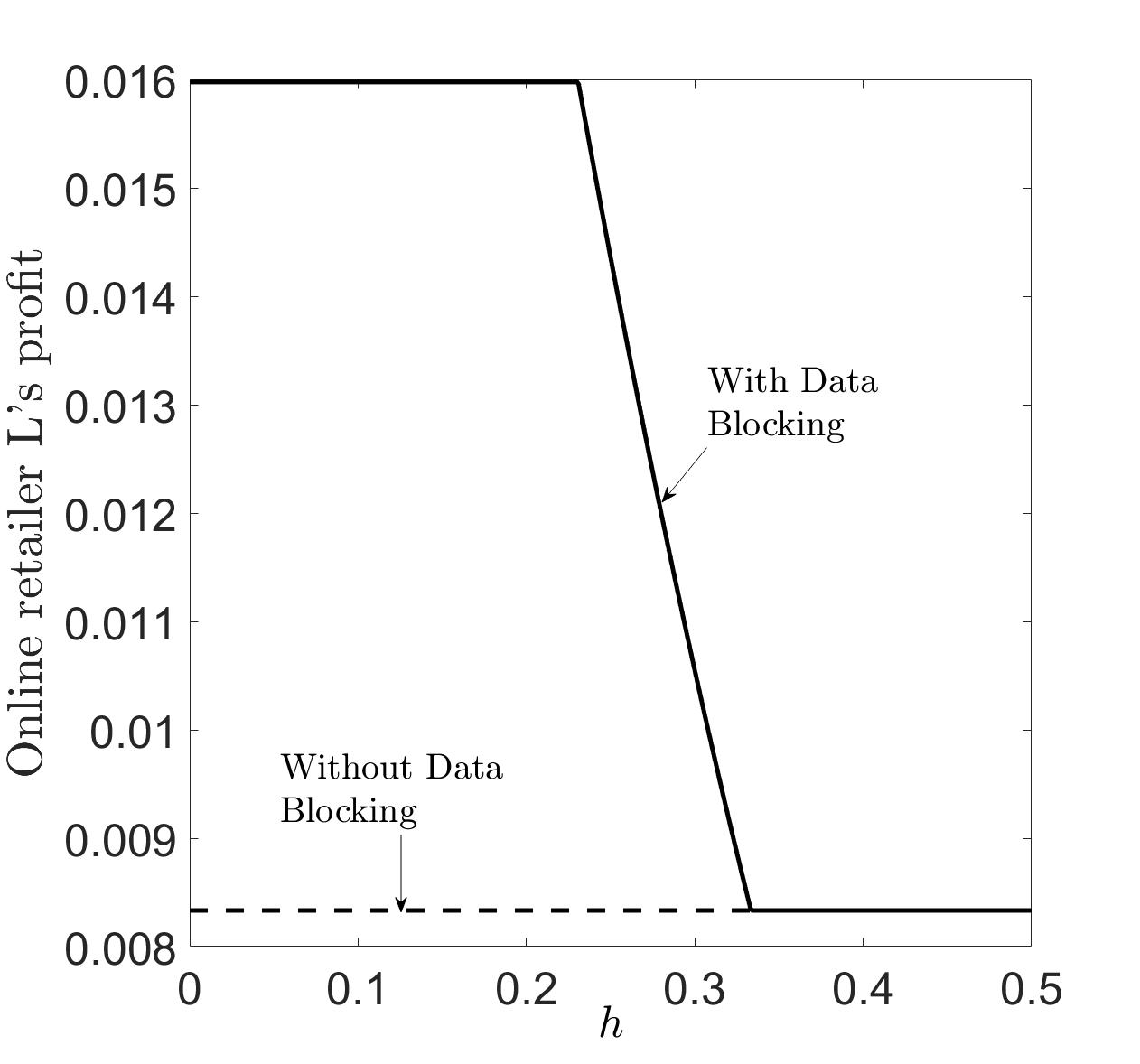
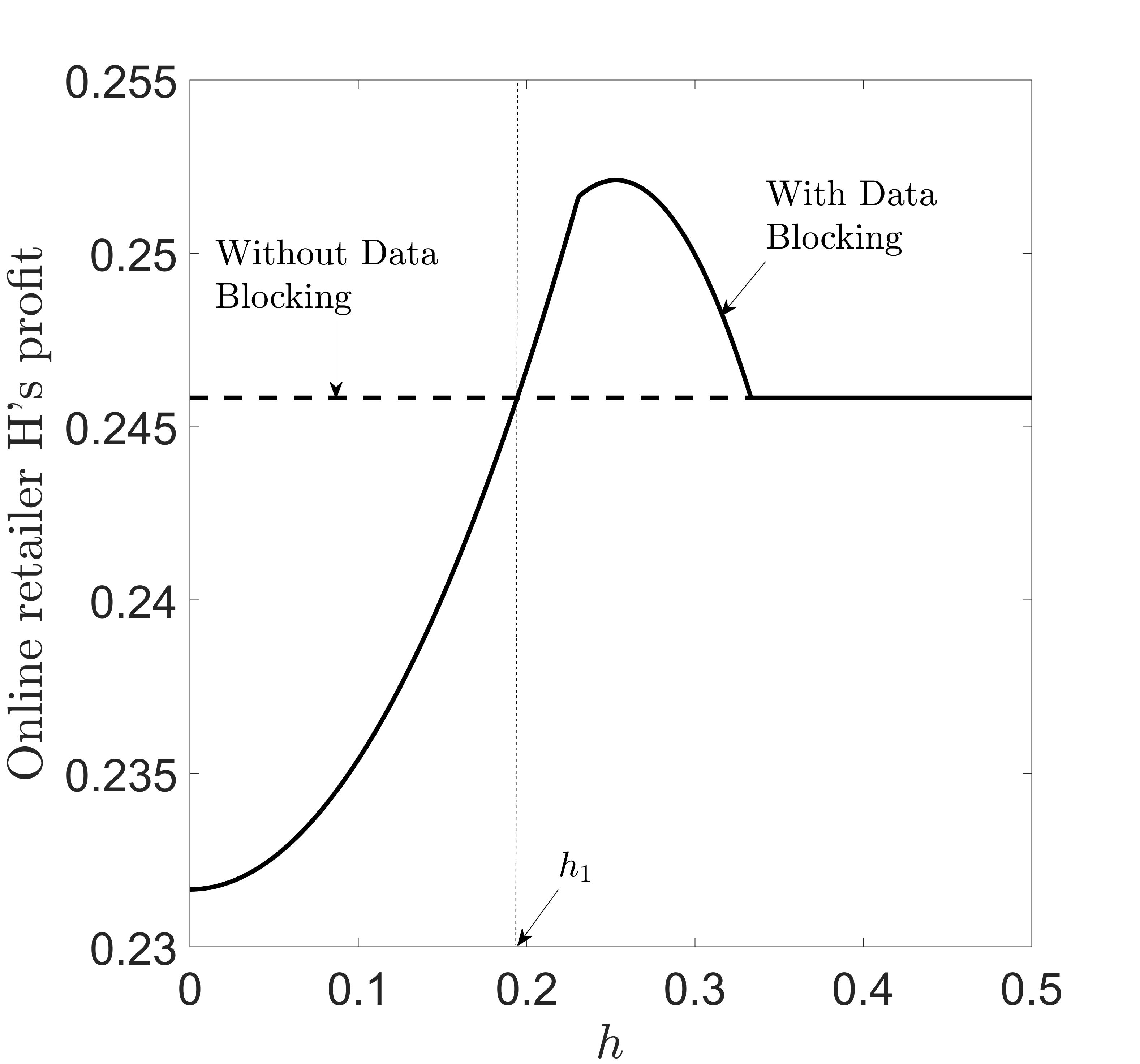
*(3) Retailer L benefits when ;*

*(4) All the entities’ profits do not change when .*

**

**Figure 5.** Impacts of consumers’ data blocking on profit. **Note:** H, L and P represent the two online retailers and the platform, respectively. The marks of “↑”, “↓” and “−” represent “positive”, “negative” and “no” impacts on profits, respectively.

In Proposition 6, Figures 5 and 6 show the impact of consumers’ data blocking on the three entities’ profits. There exists a threshold data blocking cost, , below which Retailer H will suffer when consumers can block their data (see Figures 5 and 6(a)). This result is consistent with our intuition that when the data-blocking cost is low, consumers are more willing to block their data and BBP becomes less applicable. Thus, Retailer H will suffer. In other words, when the data-blocking cost rises to a moderate level, i.e. , an interesting result arises: consumers’ data blocking then benefits Retailer H (see Figure 5 and 6(b)). In this region, Retailer H offers a behaviour-based price to low-end consumers (whose willingness to pay is low) and a uniform price to high-end consumers. When the data-blocking cost is not high, Retailer H can maintain a high uniform price to the high-end consumers. This helps Retailer H to earn the highest profit. However, when the data-blocking cost continues to rise, consumers will forego data blocking and the impact vanishes.

**

(a) Online Retailer H (b) Online Retailer L (c) Platform

**Figure 6.** Profits changing with for the cases with and without data blocking. **Note**: ,, .

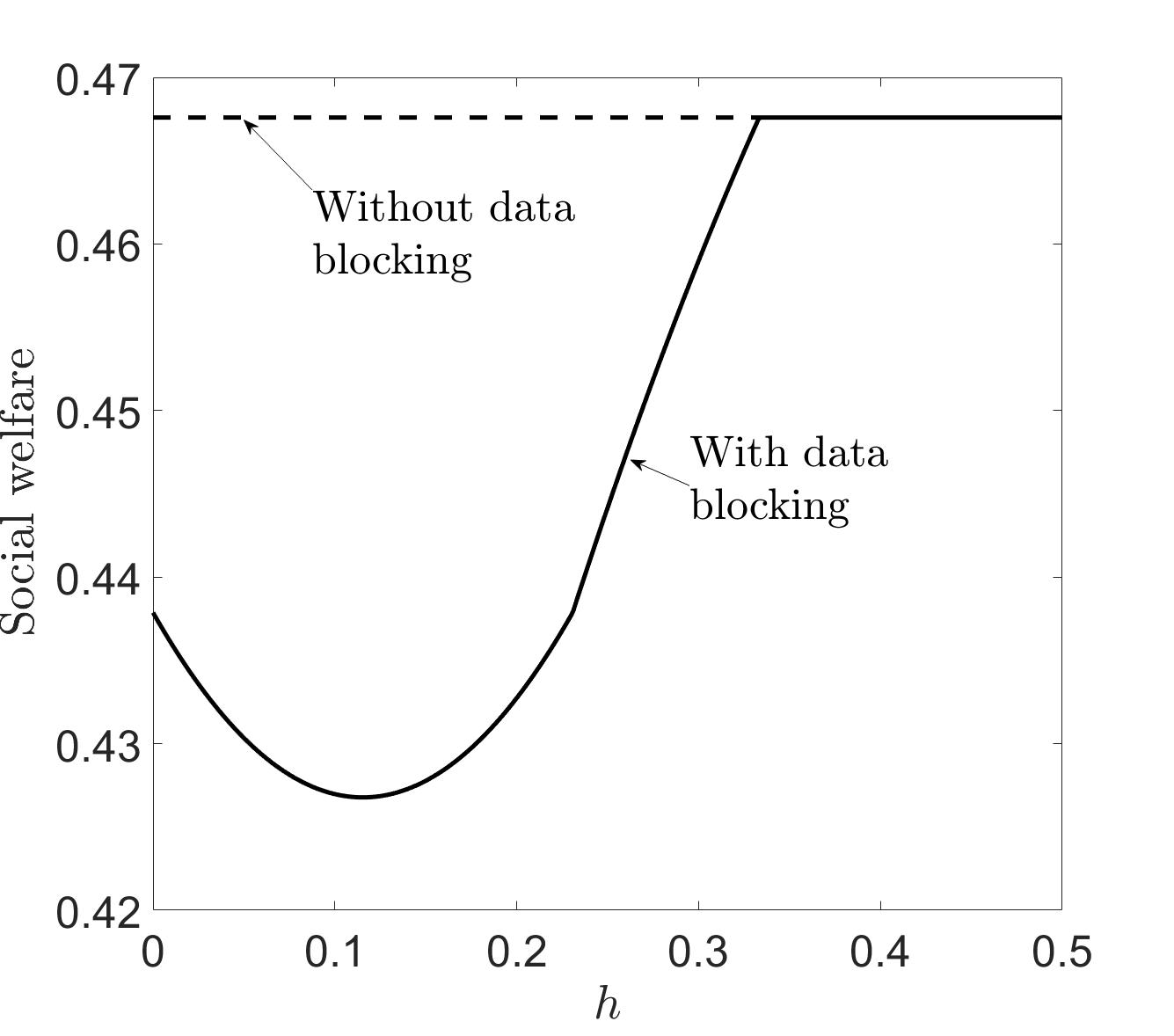
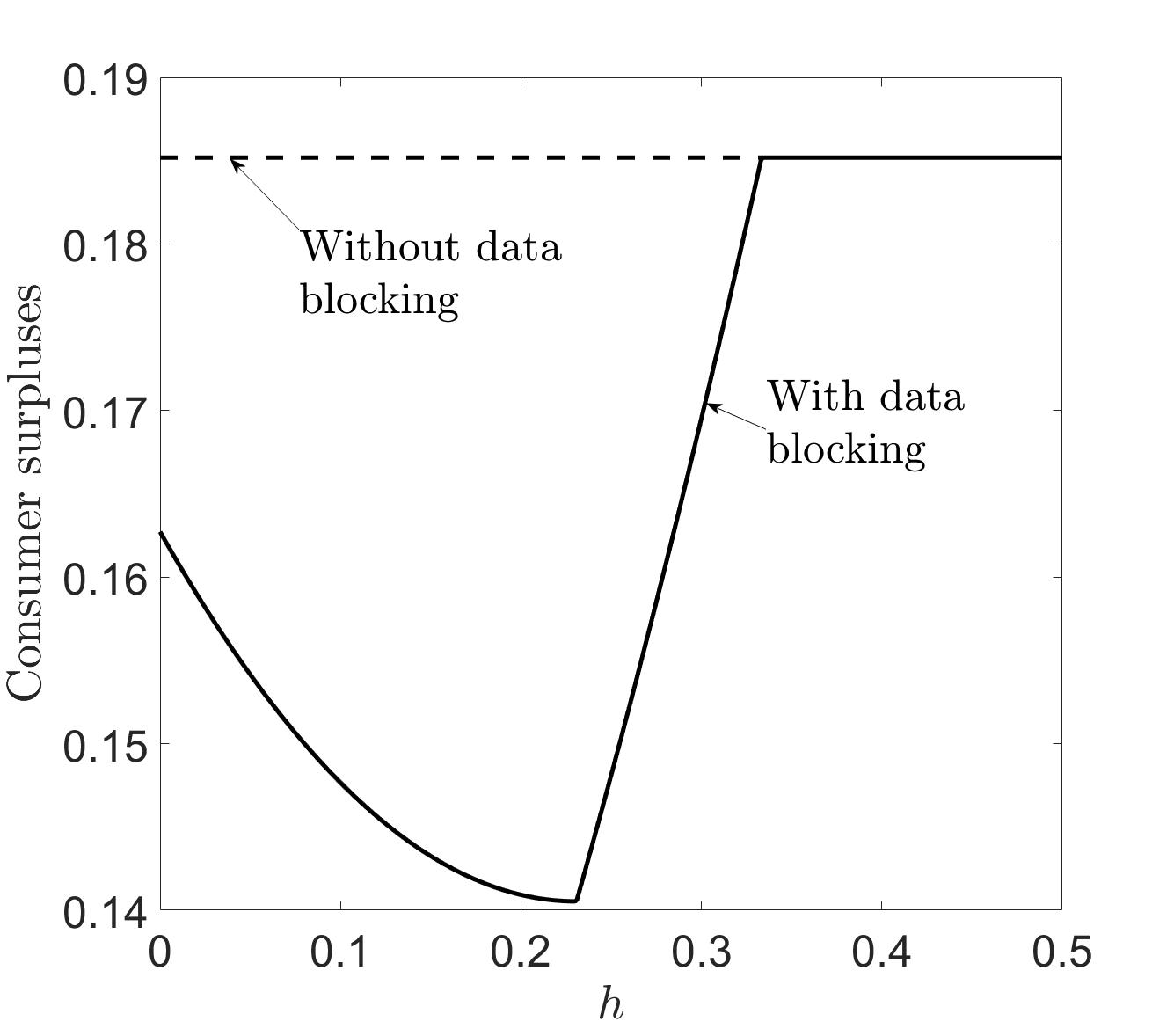
A similar impact of data blocking on the platform’s profit is depicted in Figures 5 and 6(c). Specifically, when the data-blocking cost is small, , the platform will suffer from data-blocking actions. When the data-blocking cost is moderate, , the platform will benefit from consumers’ data blocking. From the platform’s viewpoint, in addition to the price-enhancement effect, consumers’ data-blocking acts as a tool to ease competition between the two online retailers. A higher results in a less competitive market and benefits the platform.

This also shows that Retailer L always benefits from data blocking (Figures 5 and 6(c)). Although they earn zero profit from the segment of old consumers, their data-blocking actions ease the competition in the new consumers’ market segment, thus causing higher profits. If the data-blocking cost is extremely high, i.e. , consumers lack an incentive to block their data, eliminating the effect of data blocking on the three entities’ profits.

From Figures 6(a) and 6(b), we also observe that Retailer H and the platform’s profits are first increasing and then decreasing in the data-blocking cost. This result tells us that in a competitive business environment, it is unwise to eliminate consumers’ data-blocking actions by raising their data-blocking costs. It is sometimes optimal to maintain a moderate level of data-blocking cost and allow some consumers to block their data.

In addition to the analysis of profits, we continue to analyse CS and SW in the entire market. The expressions for CS and SW in Case H are presented in Table 3.

**Proposition 7. (Impacts of data blocking on CS and SW)** *The existence of consumers’ data blocking weakly hurts CS and SW in the entire market.*

**

**Figure 7.** CS and SW changing with for the cases with and without data blocking. **Note**: .

In Proposition 7 and in Figure 7, we show that when consumers can block their data, both CS and SW are negatively impacted. First, for CS, when consumers can block their data, old consumers benefit because BBP is prevented by high-end consumers and their utility is protected. However, new consumers suffer from the increased uniform selling price. Consequently, the total CS declines when data blocking is feasible for the consumers. Second, Figure 7 reveals that the total SW also decreases when data blocking is feasible. Although, under some conditions, the total market profit may increase in the presence of data blocking, the increase does not compensate for the sudden drop in CS, thus resulting in a decrease in total SW.

**6 Extended discussions**

**6.1 Extension 1: Retailer H as the price leader**

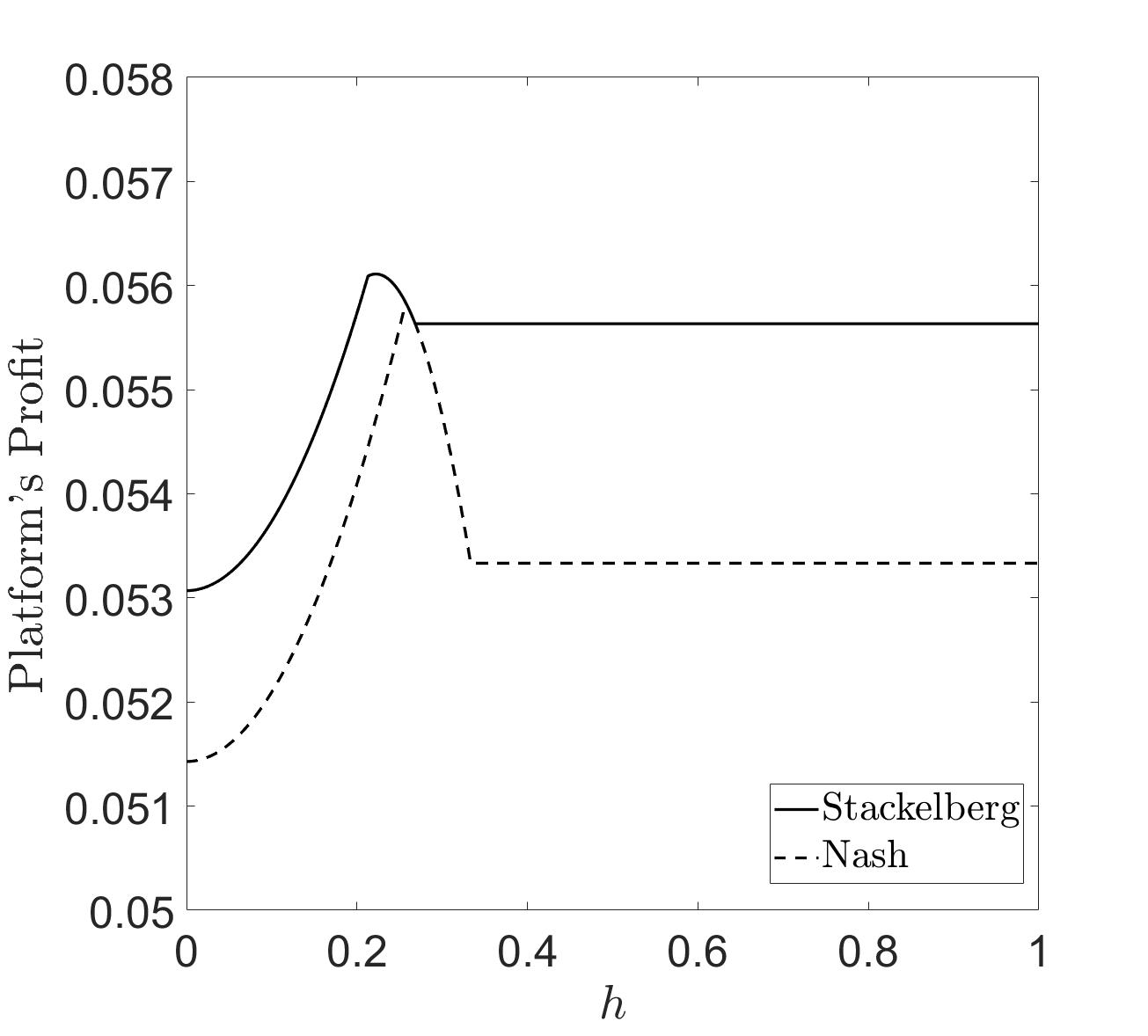
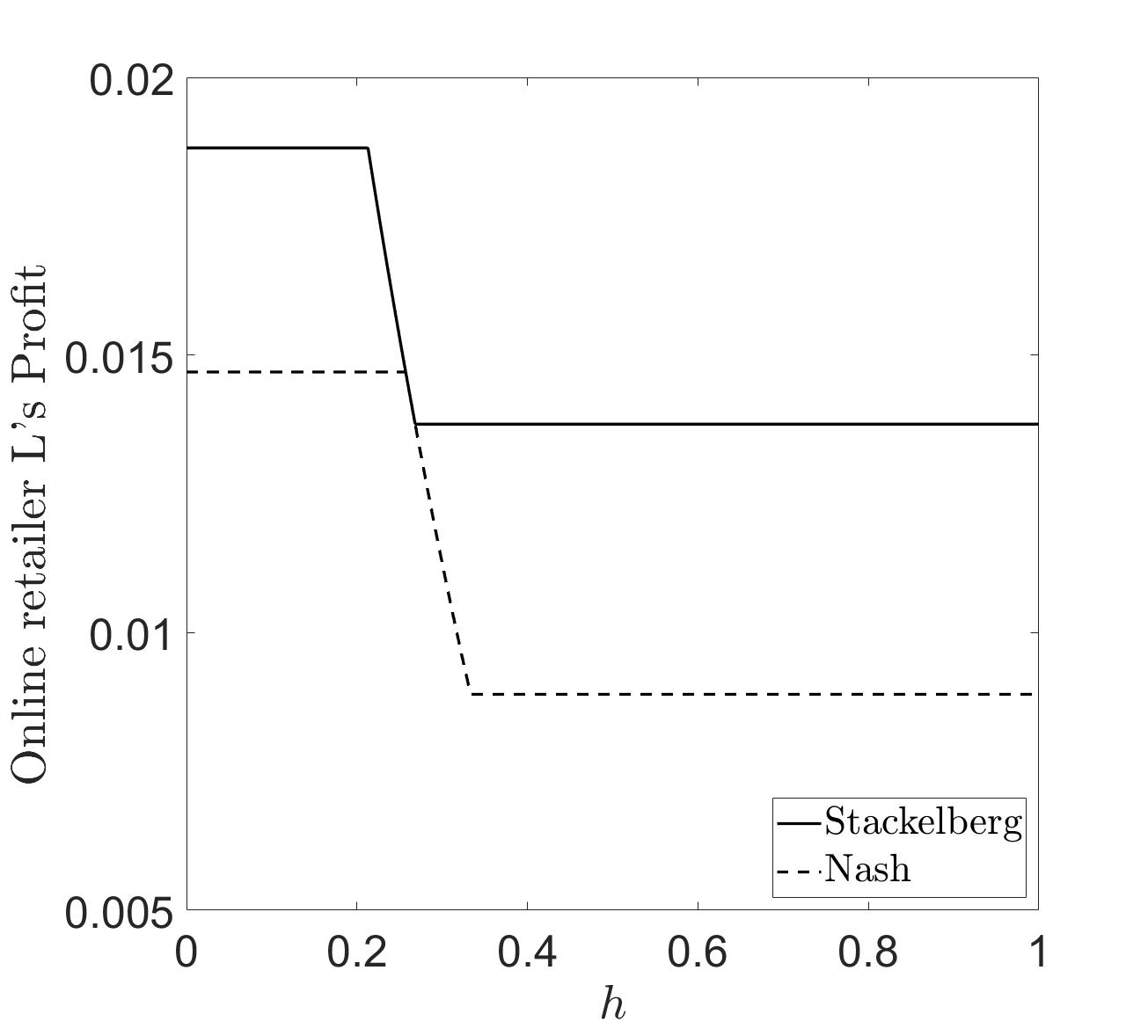
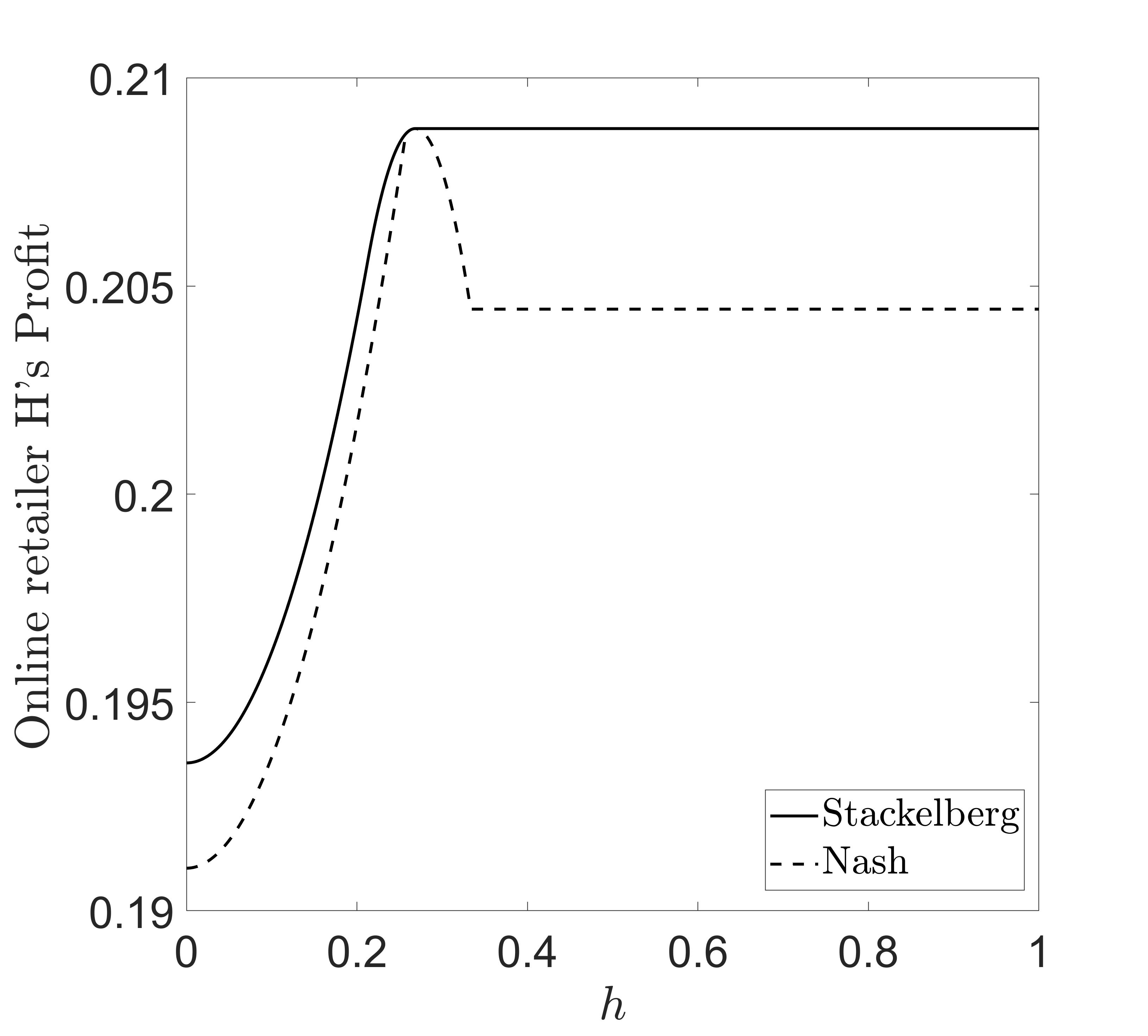
In the main model, we assumed that the two retailers simultaneously made price decisions. In this extension, we discuss models in which Retailer H is the Stackelberg leader when determining the selling price. We present the analytical results for the three cases (N, H, and HL) in Appendix C1. Relative to the main model, we summarise the impacts of the decision-sequence change as follows:

***Proposition 8.*** *When Retailer H becomes the price leader,*

*(1) the equilibrium CPS strategy does not change and remains H;*

*(2) all the entities (Retailers H and L and the platform) will benefit; and*

*(3) consumers will be less willing to block data.*



(a) Retailer H (b) Retailer L (c) Platform

**Figure 8.** Profit comparison for the Stackelberg and Nash game. **Note**: ,, .

In Proposition 8, we show that when Retailer H acts as the Stackelberg leader, in comparison to the main model, the equilibrium result does not change (i.e. Strategy H). Second, we compare the profits for the three entities in this model with those in the main model. Figure 8 reveals that all three entities benefit. This is consistent with the intuition that Retailer H benefits from their first-mover advantage and gains higher profits than in the main model; however, it is interesting that Retailer L also benefits. This results in eased competition in the Stackelberg game in comparison to the Nash game. When both retailers benefit, the platform also benefits. Third, comparing the decision regions, we find that the threshold under which old consumers block data is lower in the Stackelberg game. This implies that consumers are less willing to block data in the Stackelberg game than in the Nash game.

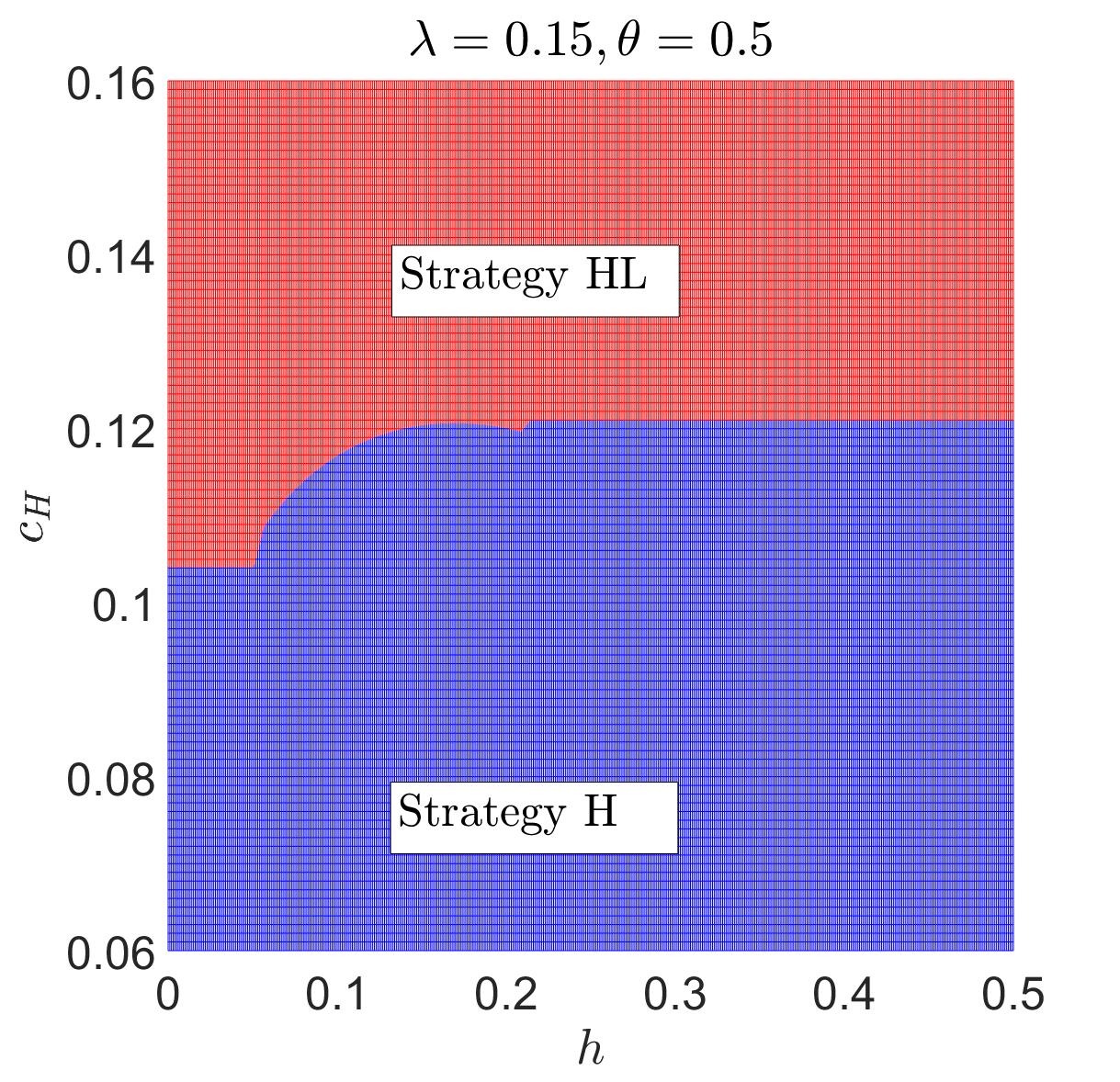
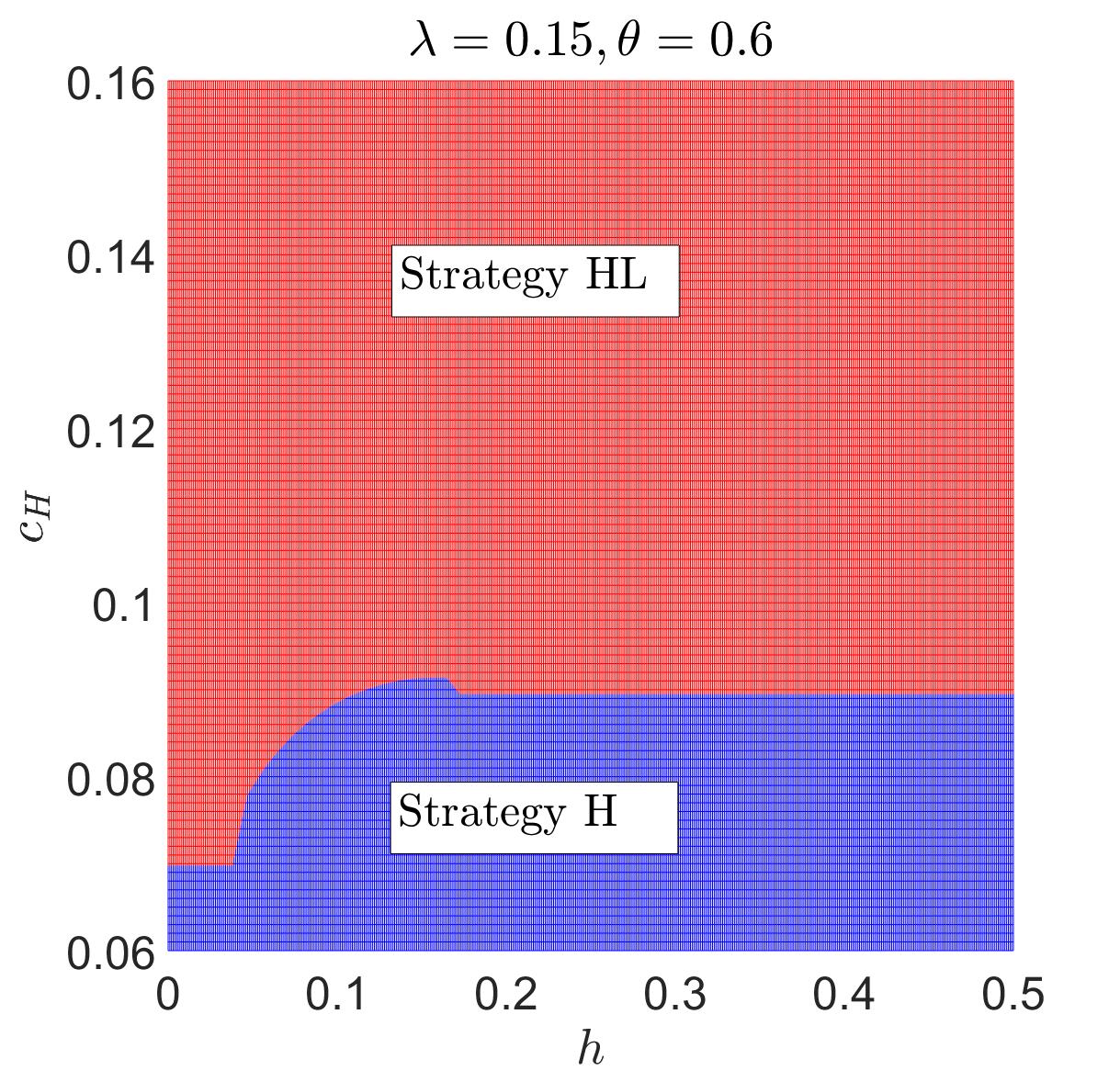
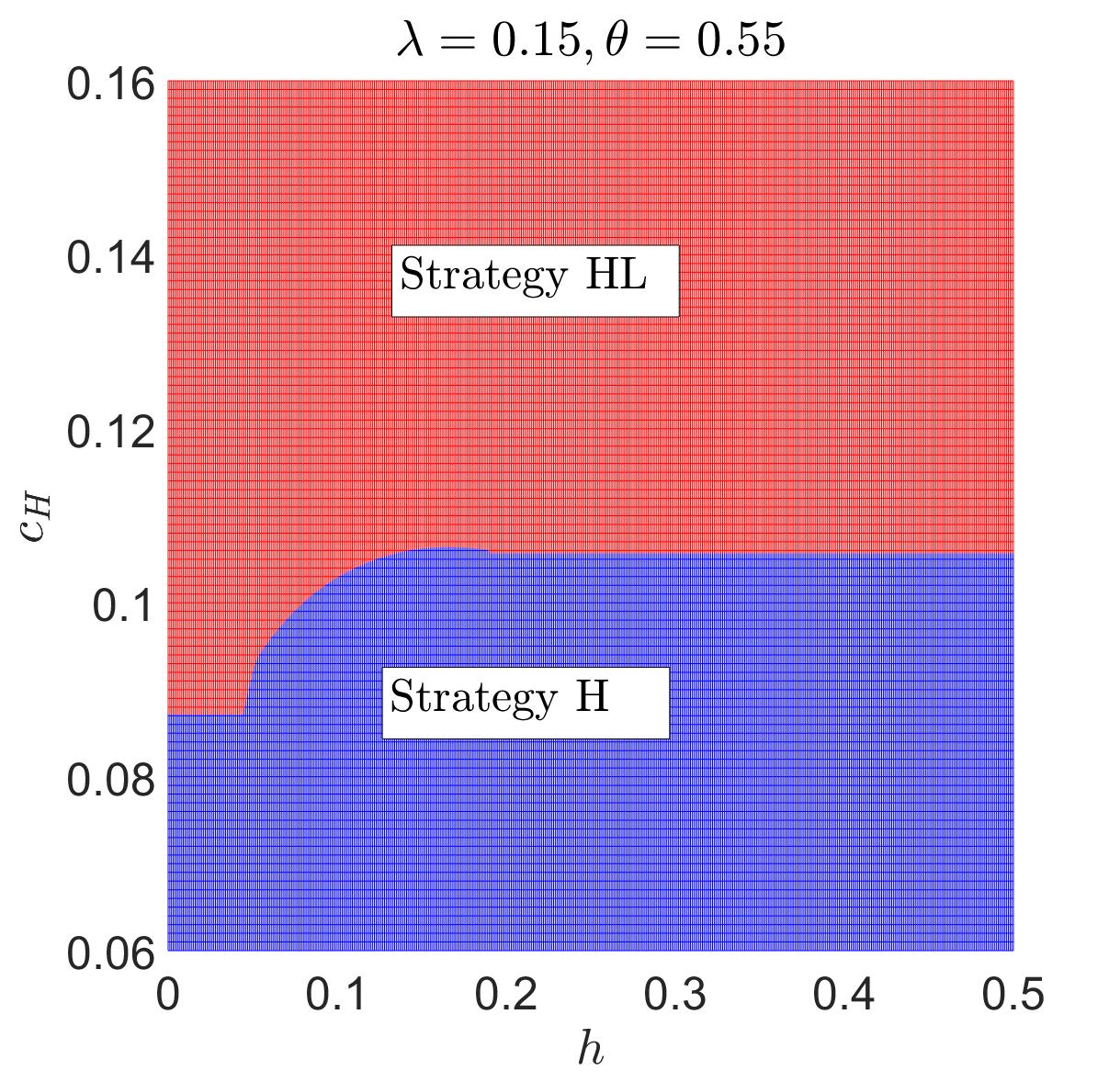
**6.****2 Extension 2: Asymmetric production costs for Retailers H and L**

In the main model, we assumed that both retailers’ productions costs were the same (zero). In this extension, we investigate the impact of positive production costs on the equilibrium result. For simplicity, we also set Retailer L’s production cost to zero, while we set Retailer H’s production cost to , which is positive. The derivations of the analytical results are presented in Appendix C2. Due to the complexity of the expressions, we only show the numerical results in this extension. The parameter settings are as follows: , , , , .

***Observation 1.*** *In the case with asymmetric production costs,*

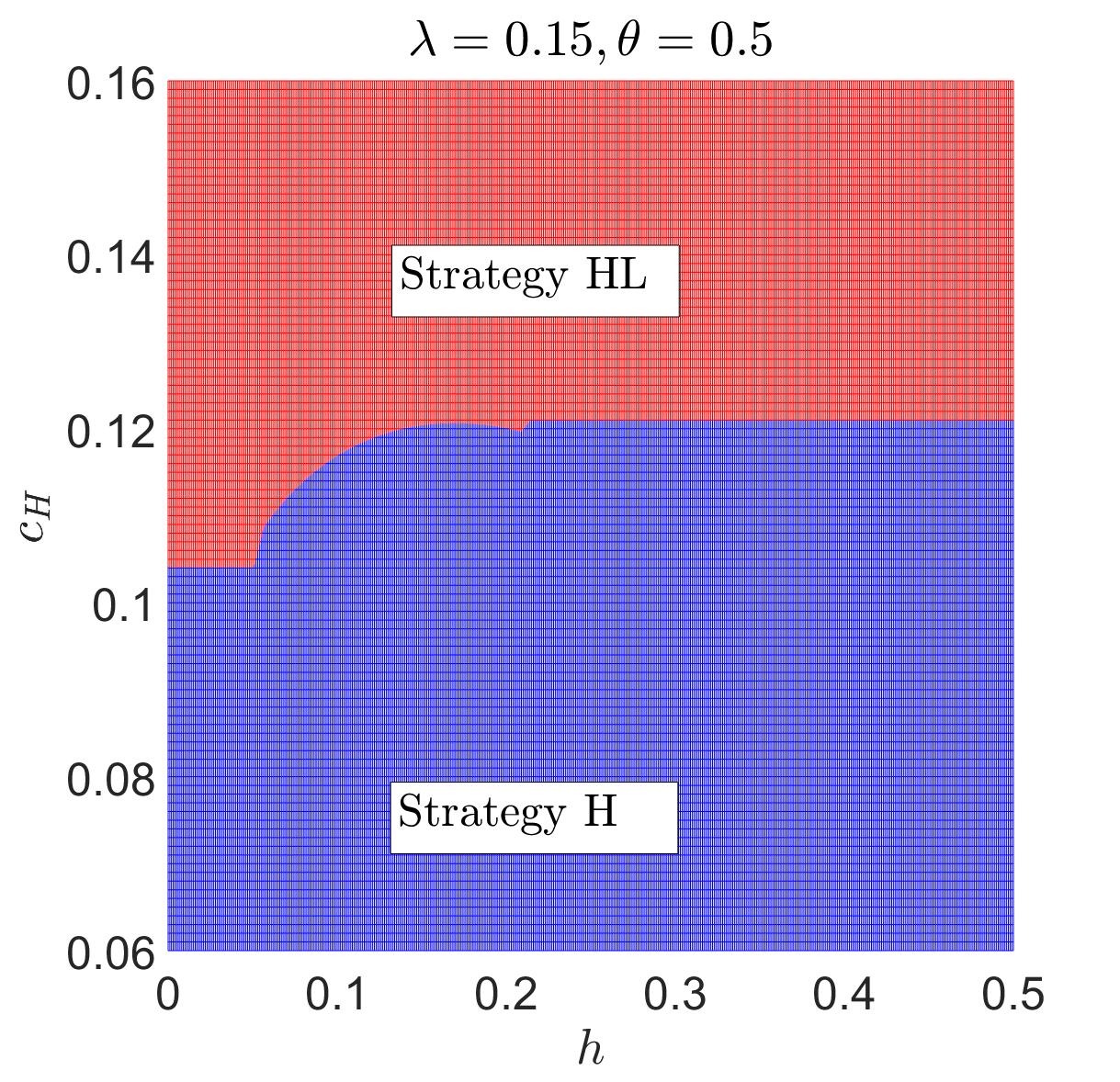
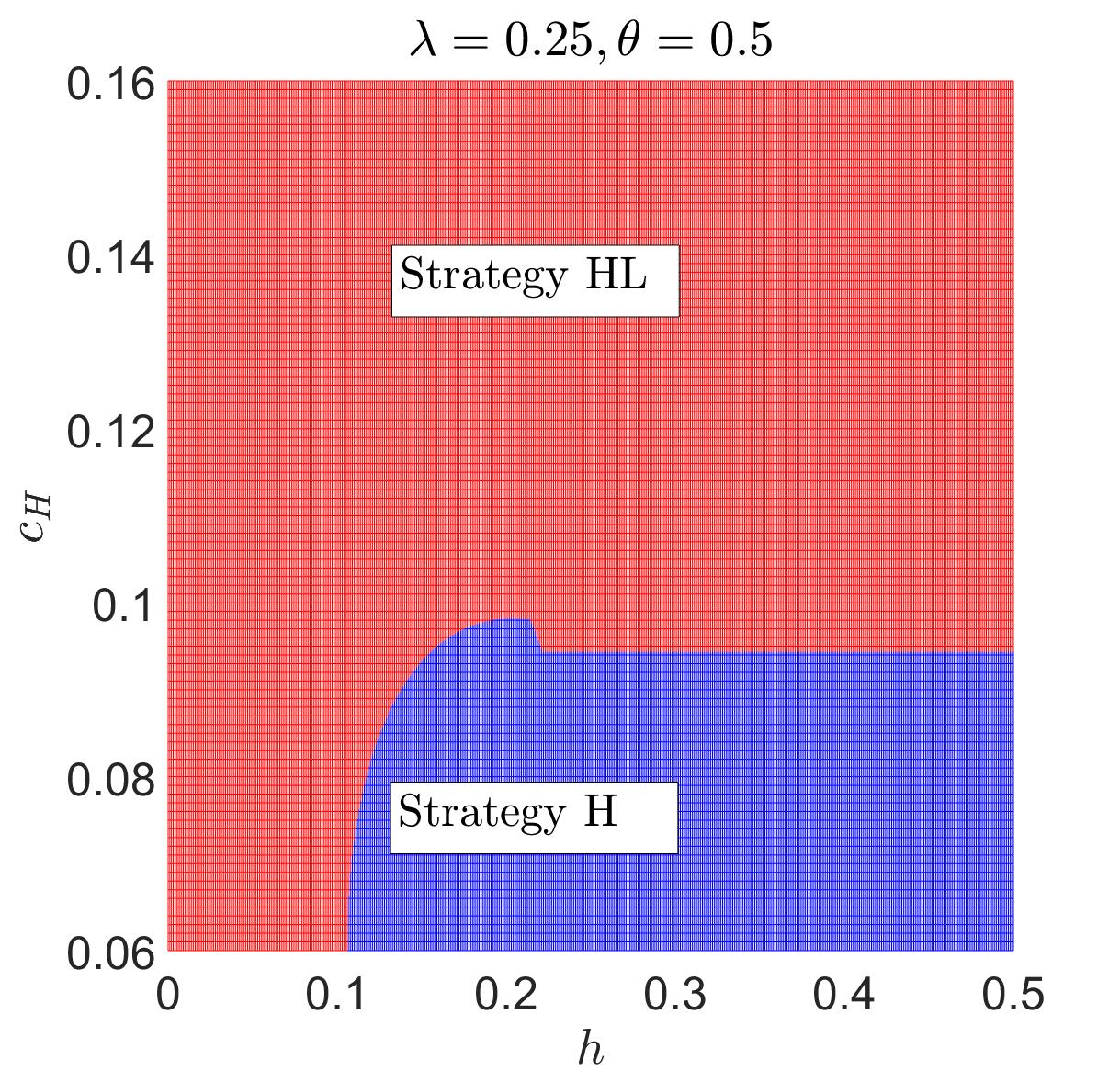
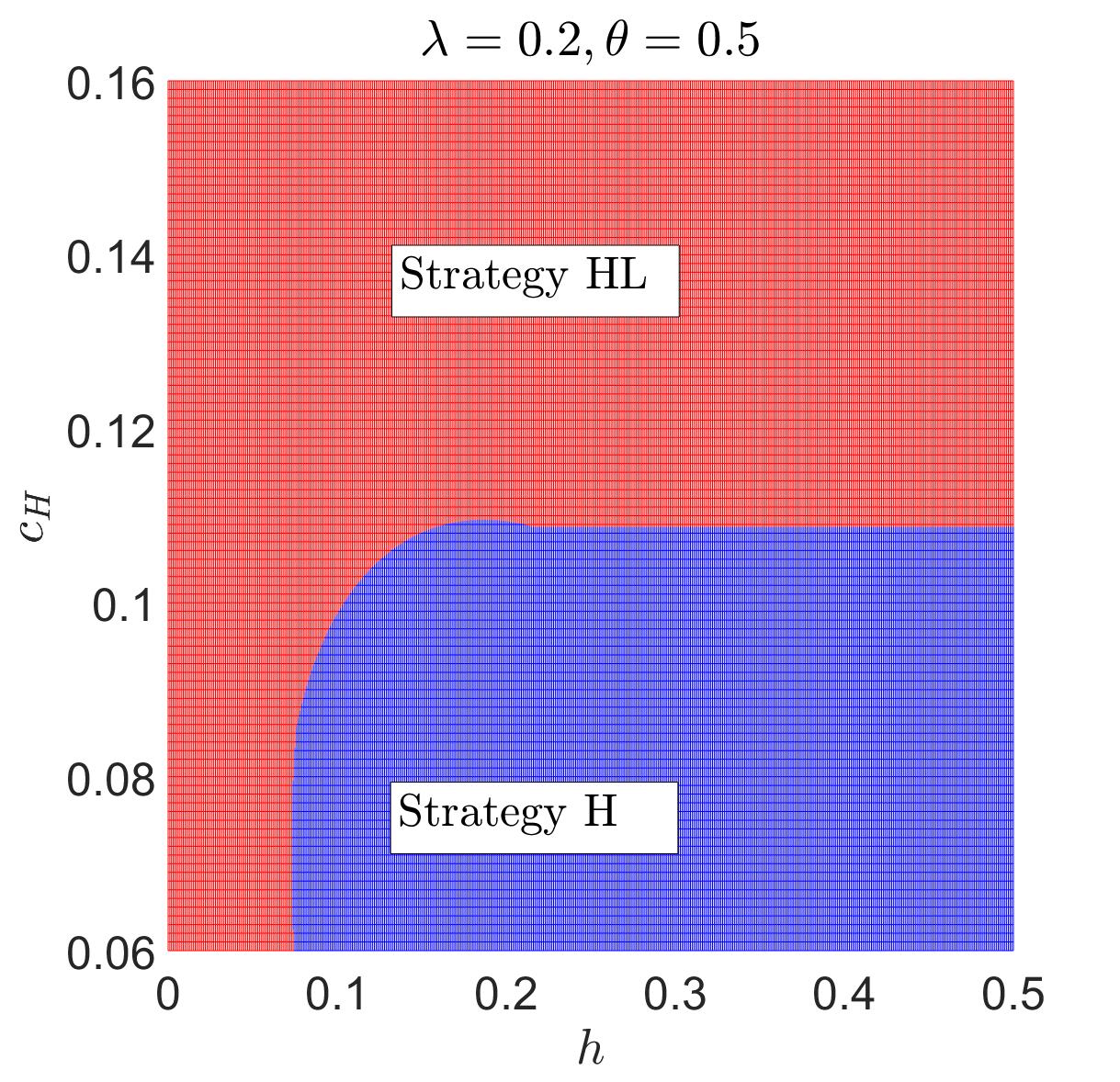
*(1) the equilibrium CPS-provision strategy switches from H to HL only when the cost difference is sufficiently high; and*

*(2) the marketplace will be less willing to adopt the HL strategy with lower or .*

****

(a) (b) (c)

**Figure 9.** Impact of and on strategy choice w.r.t. .

****

(a) (b) (c)

**Figure 10.** Impact of and on strategy choice w.r.t. .

In Figures 9 and 10, we show the impact of on the equilibrium CPS-provision strategy for the platform. From this extension, we see that when the production-cost differentiation is below a threshold, the cost asymmetry does not affect the equilibrium result in the base model. Only when the cost differentiation is above the threshold does the marketplace change its CPS strategy.

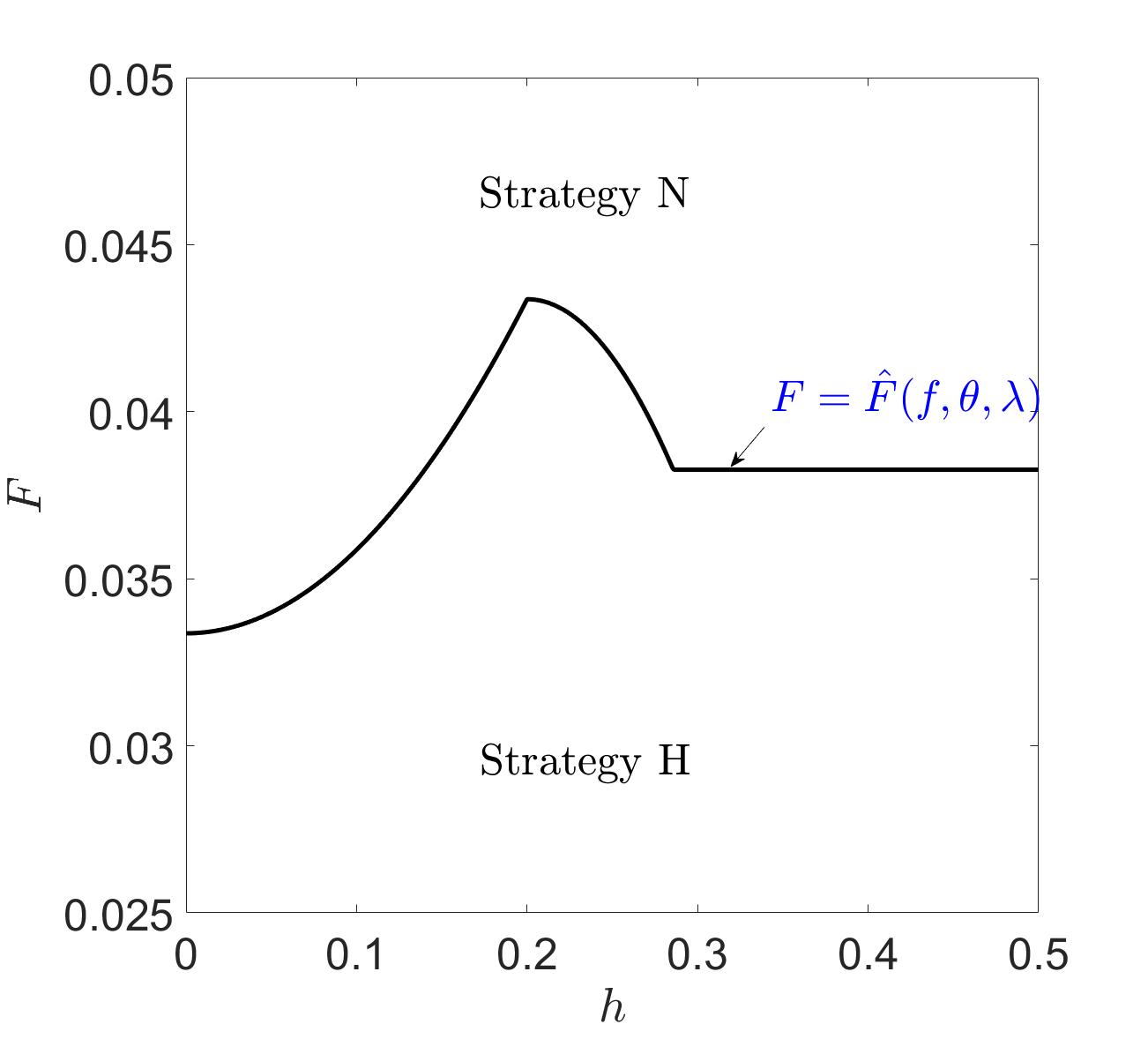
First, we show in Observation 1(a) that for fixed values of and , when is sufficiently high, the platform provides CPS to both retailers. This is because, when considering the impact of , Retailer H’s behaviour-based price will be bounded (see Lemmas D1 and D2 in Appendix D). Accordingly, their market coverage declines and they cannot serve some of the consumers. Under this condition, the platform should rely on the low-end retailer to target the consumers that cannot be targeted by Retailer H. Therefore, when is sufficiently high, the platform must provide CPS to Retailer L to enhance the entire retailer base’s total profit.

Second, we show the impacts of and on the equilibrium strategy in Proposition 1(2). When the two products are increasingly differentiated, (i.e. decreases), Figure 9 shows that the platform will be less willing to provide CPS to the low-end retailer (the HL-strategy region becomes smaller). The reason is that the drop in weakens Retailer L’s sales margin from adopting BBP. Therefore, when the product quality drops, the marketplace is less willing to provide CPS to Retailer L. Additionally, Figure 10 shows that when there are more old consumers in the market ( increases), the platform is more willing to adopt Strategy HL. This is because when the number of old consumers increases, providing CPS to both retailers helps them to target more consumers with BBP, which enhances the retailer base’s total profit.

**6.3 Extension 3: Positive CPS fee charged by the marketplace**

In the main model, we assumed that the platform provided CPS to the retailers for free. In this extension, we assume that the platform charges a service fee, , when providing CPS to the retailers. The impact of on the equilibrium strategy is as follows:

***Proposition 9.*** *There exists a threshold for , (1) when , the equilibrium strategy is H; (2) when , the equilibrium strategy is N.*



**Figure 11.** Impacts of service fee on CPS strategy choice. **Note**: ,, .

Proposition 9 and Figure 11 show the impact of a service fee on the equilibrium strategy. The expression for is presented in Appendix E. For any value of , HL cannot be an equilibrium strategy. According to Proposition 5, Retailer L will never accept CPS even when the service fee is zero. Therefore, when the service fee is positive, HL cannot be the equilibrium strategy. Investigating Retailer H’s preference for the strategies, we find that only if the service fee paid to the platform is small, , will Retailer H benefit from CPS. Otherwise, if , the profit increase generated by the adoption of CPS cannot compensate for the service fee.

**7 Conclusions**

Equipped by big-data technologies, retail platforms engage in consumer profiling by gathering and analysing vast amounts of private data from consumers to learn their individual preferences. Such preference information enables online retailers to provide behaviour-based prices to consumers, and they can then extract more CS, gain competitiveness, and earn higher profits. Therefore, many retail platforms (including Taobao, JD, and Amazon) are providing consumer-profiling information to their online retailers so that they can achieve a win-win outcome with the retailers. However, being aware of the losses incurred by private-data leakage, consumers can block their data, which in turn affects the platforms’ consumer-profiling results. Based on the above interactions between platforms, consumers, and online retailers, this study examines a platform’s CPS-provision strategies in the presence of online retailers’ competition and consumers’ data-blocking behaviours.

**7.1 Research findings and managerial implications**

In the main model, we find that the platform’s optimal strategy is to exclusively provide CPS to the powerful online retailer and to encourage their BBP decisions. For the powerful retailer, exclusively obtaining CPS also generates the highest profit, which results in a win-win outcome for the platform and powerful retailer. From the weak retailer’s perspective, interestingly, they will voluntarily give up the right to obtain CPS even if provided by the platform for free. The reason lies in the strategic impact of consumers’ preference data on market competition. When both retailers obtain consumers’ private information and both can adopt BBP, the competition in the entire market will be intensified, and all the entities’ profits will suffer. This provides important implications for the platform and retailers. First, for the platform, providing CPS for all the online retailers would be unwise; instead, it should provide CPS only to the high-end retailers and encourage their BBP to extract surplus from high-value consumers. Second, for the retailers, if they have disadvantages in product quality, adopting BBP will be harmful because it will result in fierce competition and a lose-lose outcome for both firms in the market.

In addition, we have also investigated the impacts of consumer-data blocking on the firms’ profits, CS, and SW. Counterintuitively, we show that data blocking can sometimes result in a win-win-win outcome for the two retailers and platform. The reason lies in the effects of competition alleviation and the uniform price increment for moderate values of the data-blocking cost. This provides implications for the firms that they should sometimes welcome data blocking when consumers’ data-blocking disutility is relatively high. Continuously, we show that the existence of data blocking is detrimental to the total CS and SW. For the CS, although data blocking protects high-value consumers’ benefit, it creates more harm to the low-value consumers’ benefit. Therefore, the total CS declines when data blocking is feasible. With the decreased CS follows a drop in total SW.

We have also investigated the impact of sequential price decisions, asymmetric production costs, and a service cost on the equilibrium results in the extended discussions. We find that the platform and Retailer H benefit from the latter’s first-mover advantage; interestingly, Retailer L also benefits when they act as a second mover. Additionally, we show that the equilibrium Strategy H is stable unless the cost differentiation is relatively high, which then requires a switch to Strategy HL. Then, we also find that when the platform’s positive service fee exceeds a threshold, the optimal strategy switches from H to N.

**7.2 Future research directions**

There are several promising directions for future research. First, we can consider a more general supply-chain structure with both individual and platform-owned retailers and examine its impacts on CPS strategies. Second, in the main model, we have considered the case in which all consumers face the same data-blocking cost. In the future, we can consider a more interesting case in which consumers are heterogeneous in the data-blocking cost. Third, we considered only two competing retailers on the platform. In the future, we can consider competition between multiple retailers and investigate the platform’s information-provision strategies in a highly competitive market environment. Fourth, in this study, we consider retailers’ personalised pricing decisions. In practice, many retailers also provide personalised advertisements to different consumers. Therefore, the combined personalised pricing and advertising problems would be interesting. Last, in this study, we have assumed that the marketplace’s consumer-profiling is accurate. However, in real operations, consumer profiling may be inaccurate. For such cases, we may consider information accuracy and investigate its impact on the CPS strategies and retailers’ operations.

**Conflict of Interest:** The authors declare that they have no conflict of interest.

**References**

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

Casadesus-Masanell, R., & Hervas-Drane, A. (2015). Competing with privacy. Management Science, 61(1), 229-246.

Chen, Y., & Jiang, B. (2019). Effects of monitoring technology on the insurance market. Production and Operations Management, 28(8), 1957-1971.

Chen, Y.J., Dai, T., Korpeoglu, C.G., Korpeoglu, E., Sahin, O., Tang, C.S., & Xiao, S. (2020). Innovative online platforms: research opportunities. Manufacturing and Service Operations Management, 22(3), 429-643.

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

Choe, C., King, S., & Matsushima, N. (2017). Pricing with cookies: behavior-based price discrimination and spatial competition. Management Science, 64(12), 5669-5687.

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.

Dukes, A., & Liu, L. (2016). Online shopping intermediaries: the strategic design of search environments. Management Science, 62 (4), 1064-1077.

Elmachtoub, A.N., Gupta, V., & Hamilton, M.L. (2021). The value of personalized pricing. Management Science, 67(10), 6055-6070.

Fortes, N., & Rita, P. (2016). Privacy concerns and online purchasing behaviour: towards an integrated model. European Research on Management and Business Economics, 22(3), 167-176.

Geng, X., Tan, Y., & Wei, L. (2018). How add-on pricing interacts with distribution contracts. Production and Operations Management, 27(4), 605-623.

Hagiu, A., & Wright, J. (2014). Marketplace or reseller? Management Science, 61(1), 184-203.

Hajihashemi, B., Sayedi, A., & Shulman, J.D. (2021). The perils of personalized pricing with network effects. Marketing Science, 41(3), 477-500.

Hao, L., & Tan, Y. (2019). Who wants consumers to be informed? facilitating information disclosure in a distribution channel. Information Systems Research, 30(1), 34-49.

Jing, B. (2016). Behavior-based pricing, production efficiency, and quality differentiation. Management Science, 63(7), 2365-2376.

Kwark, S., Chen, J., & Raghunathan, S. (2017). Platform or wholesale? a strategic tool for online retailers to benefit from third-party information. MIS Quaterly, 41(3), 763-785.

Lei, Y., Jasin, S., & Sinha, A. (2018). Joint dynamic pricing and order fulfillment for e-commerce retailers. Manufacturing and Service Operations Management, 20(2), 269-284.

Li, K.J. (2018). Behavior-based pricing in marketing channels. Marketing Science, 37(2), 310-326.

Li, G., Tian, L., & Zheng, H. (2021). Information sharing in an online marketplace with co-opetitive sellers. Production and Operations Management, 30(10), 3713-3734.

Mantin, B., Krishnan, H., & Dhar, T. (2014). The strategic role of third-party marketplaces in retailing. Production and Operations Management, 23(11), 1937-1949.

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

Morath, F., & Munster, J. (2018). Online shopping and platform design with ex ante registration requirements. Management Science, 64(1), 360-380.

Shy, O., & Stenbacka, R. (2016). Customer privacy and competition. Journal of Economics & Management Strategy, 25(3), 539-562.

Statista.com. (2022). Retail e-commerce sales worldwide from 2014 to 2025. Accessed from: <https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/>.

Sun, H., Fan, M., & Tan, Y. (2020). An empirical analysis of seller advertising strategies in an online marketplace. Information Systems Research, 31(1), 37-56.

Taylor, C., & Wagman, L. (2014). Consumer privacy in oligopolistic markets: winners, losers, and welfare. International Journal of Industrial Organization, 34(1), 80-84.

Tian, L., Vakharia, A. J. Tan, Y. & Xu, Y. (2018). Marketplace, reseller, or hybrid: strategic analysis of an emerging e-commerce model. Production and Operations Management, 27(8), 1595-1610.

Valletti, T., & Wu, J. (2020). Consumer profiling with data requirements: structure and policy implications. Production and Operations Management, 29(8), 309-329.

Xu Z, & Dukes, A. (2021). Personalization from customer data aggregation using list price. Management Science, 68(2), 960-980.

Yi, Z., Wang, Y., Liu, Y., & Chen, Y. J. (2018). The impact of consumer fairness seeking on distribution channel selection: direct selling vs. agent selling. Production and Operations Management, 27(6), 1148-1167.

Zennyo, Y. (2020). Strategic contracting and hybrid use of agency and wholesale contracts in e-commerce platforms. European Journal of Operational Research, 281(1), 231-239.

Zhang, D. J., Dai, H., Dong, L., Qi, F., Zhang, N., Liu, X., Liu, Z., & Yang, J. (2018). How do price promotions affect customer behavior on retailing platforms? Evidence from a large randomized experiment on Alibaba. Production and Operations Management, 27(12), 2343-2345.

**Appendix A: proofs for the main results**

**Proof for Proposition 1**

Solving the first order derivatives of and *w.r.t.* and ,

, ,

we can easily obtain the equilibrium selling prices for the two online retailers as

and *.*

Substitute the prices into the profit functions for the three entities, we have

*, , .*

**Proof for Lemma 1**

To derive old consumers’ data blocking decisions, we compare two scenarios, i.e., data blocking and no data blocking.

(1) Data blocking

Suppose that consumers that locate within the region of [0,1] are all block their data to the platform. Then, behaviour-based pricing is infeasible for all consumers. A type consumer’s utility that obtained from product H and L are and .

If , consumers who locates in the region of will purchase from Retailer H; those who locates in the region of will purchase from Retailer L; those who locates in the region of will not purchase. Then, for known , and , their final utility can be expressed as

If .

If , consumers who locates in the region of will purchase from Retailer H; those who locates in the region of will not purchase. Then, for known , and , their final utility can be expressed as

If .

(2) No data blocking

Suppose that all the consumers within the region of [0,1] do not block their data and Retailer H can use the data to provide behaviour-based prices to consumers. Since Retailer L cannot provide behaviour-based price, a type consumer’s utility that obtained from product L is .

Then, based on the utility for product L, Retailer H sets its behaviour-based price as follow. In the region of , consumers can only purchase from Retailer H, because purchasing from Retailer L obtains negative utility. Therefore, in this region, Retailer H will set a price of to extract all the surplus from consumers. In the region of , consumers can purchase from Retailer H or Retailer L. In this region, Retailer H will drive Retailer L out of the market by setting the behaviour-based price of . In this case, consumers will never purchase from Retailer L. Then, for known and , their final utility can be expressed as

(3) Comparison of and

Then, we obtain consumers’ decision of blocking or not by comparing the utility of and .

* If , for , is satisfied and consumers will block their data; for , is satisfied and consumers will not block their data.
* If , for , is satisfied and consumers will block their data; for , is satisfied and consumers will not block their data.

Summarizing the above results, we obtain consumers’ responses to retailers’ prices in Lemma 1.

**Proof for Proposition 2**

We first consider Retailer L’s pricing problem. Solving the first order condition of w.r.t. , we have the responsive pricing function of .

Then, we turn to Retailer H’s problem. Consider Retailer L’s responsive decision, we find that the condition of is always satisfied. This denotes that the Nash equilibrium never exists in the region of . Therefore, Retailer H’s piecewise profit function can be simplified to

.

where .

There are two scenarios.

First, we consider the condition that which means that is always satisfied. Then, we solve the optimization problem of

Max: .

s.t. .

Then, the Lagrangian and Karush-Kuhn-Tucker (KKT) optimality conditions for Retailer H’s optimization problem are as follows.

,

,

,

where is the non-negative Lagrangian multiplier.

According to the Complementary Relaxation Theorem, and with algebra analysis, we find that the optimal depends on the value of . When , there exists an interior solution of ; otherwise, when , there exists a boundary solution of .

Second, we consider the condition that , which means that . Then, we solve the optimization problem of

Max: .

s.t. , .

Following the same fashion, we find that when , there exists an interior solution of ; otherwise, when , there exists a boundary solution of .

Combining the above two scenarios, we finally obtain the optimal response of as

.

The Nash equilibrium exists only if the two online retailers’ response functions are satisfied at the same time. Solving the equations of the two retailers’ responsive prices, we finally obtain the equilibrium prices of

In the region of , some of the old consumers will block their data (); in the region of , old consumers’ data blocking incentive will be deterred (); in the region of , old consumers has no incentive to block data ().

Substituting the results into the corresponding profit functions, we obtain the results in Proposition 2 and Table 1.

**Proof for Lemma 2**

The proof for Lemma 2 is similar to that of Lemma 1, which is omitted here.

**Proof for Proposition 3**

The proof for Proposition 3 is similar to that of Proposition 2, which is omitted here.

**Proof for Proposition 4**

We compare the values of and as follows.

.

.

Therefore, we obtain the results in Proposition 4.

**Proof for Proposition 5**

We first consider online Retailer L’s problem. The results can be obtained by comparing the equilibrium profits in the three cases.

(1) The comparison of case N and case H.

* In the region of , . Solving the condition of , we obtain . Otherwise, .
* In the region of , . Solving the condition of , we obtain and . Otherwise, .
* In the region of , .

Summarizing the above results, we obtain that (1) when , or and , ; (2) When  and , .

(2) Then, we turn to the platform and Retailer H’s problem. Following the above fashion, the results in Proposition 5(1) can be proved.

With the results in Proposition 5(1) and (2), we continue to analysis the equilibrium consumer profiling service. For the platform, it will be more willing to provide CPS to Retailer H exclusively. For Retailer H, if CPS is only obtained by itself, comparing to strategy HL, it obtains higher profit. Both the platform and Retailer H prefers strategy H, and Retailer L has no way to own CPS. Therefore, the equilibrium result is H.

**Proof for Proposition 6**

To obtain the results in Proposition 6, we need to compare the retailers’ and the platform’s profit in case H with the one in which data blocking is infeasible (case NB). The equilibrium results for case NB equal to the results in region III where consumers give up data blocking because of the high costs.

(1) we first consider Retailer H’s profit

* For , we let

.

When , it can be solved that .

* For , we let

**.**

When , it can be solved that .

Combining the above results, we find that when

,

data blocking hurts Retailer H’s profit. Otherwise, for , data blocking benefits Retailer H’s profit. In the region of , data blocking has no impact on Retailer H’s profit. Following the same fashion, the results for Proposition 6(2), (3) and (4) can be obtained.

**Proof for Proposition 7**

To obtain the results in Proposition 7, we need to compare the CS and SW in case H with the case in which data blocking is infeasible (case NB). The equilibrium CS and SW for case NB is equal to the results in region III where consumers give up data blocking because of the high costs.

We show the comparison of CS as follows.

* For , we let

.

In this region, is always satisfied.

* For , we let

.

In this region, it can be verified that is increasing in parameter . In addition, when , is satisfied. Therefore, in the while region, is satisfied.

* For , is satisfied.

Therefore, we obtain the result that data blocking weakly hurts CS. Follow the same fashion, we can also obtain the result that data blocking also weakly hurts SW.

**Appendix B: Equilibrium results and expressions for the main model**

**Table 1.** The equilibrium results for case H

|  |
| --- |
| **Region I**: |
| ,, |
| , |
| *,* |
| , |
| . |
| **Region II**: |
| ,, |
| , |
| ， |
| , |
| , |
| **Region III**: |
| , , |
| , |
| , |
| , |
| . |

**Table 2.** The equilibrium results for case HL

|  |
| --- |
| **Region I**: |
| ,, |
| , |
| , |
| , |
| . |
| **Region II**: |
| ,, |
| , |
| , |
| , |
| , |
| **Region III**: |
| , , |
| , |
| , |
| , |
| , |

**Table 3.** The equilibrium results of CS and SW for case H

|  |
| --- |
| **Region I**: |
| , |
| . |
| **Region II**: |
| , |
| . |
| **Region III**: |
| , |
| . |

**Table 4.** Expressions for the thresholds

|  |
| --- |
| . |
| . |
| , |
| . |

**Appendix C. Results for the extensions**

**Appendix C1: Retailer H as the price leader.**

In this extension, the three entities’ profit functions in case N, case H and case HL are the same as those in the base model, respectively. Next, we discuss the optimal decisions for the three cases.

(1) **Case N:** Using backward induction, we first solve Retailer L’s pricing problem, which obtains . Then, we substitute the response function in to Retailer H’s target function, and solve the first order condition w.r.t. , we obtain Substitute it into Retailer L’s response price and the profit functions, we obtain , , , .

(2) **Case H:** Using backward induction, we first solve Retailer L’s pricing problem, which obtains . With the response function, we simplify Retailer H’s profit function to

,

where .

The constraint that should be satisfied in the above function. Therefore, we discuss Retailer H’s problem in two aspects. Firstly, when , is always satisfied. The optimization problem becomes

Max: ,

s.t. .

Solving the problem, we can obtain the optimal price decision as

The solution procedure is straightforward, which is omitted here. Substituting the results into the target function, we obtain the corresponding result as.

Secondly, when , is always satisfied. The optimization problem becomes

Max:.

s.t. .

The solution procedure is straightforward, which is omitted here. Solving the problem, we can obtain the optimal price decision as

Comparing the two scenarios, we find that

* when , Retailer H should set a low price, such that part of the consumers will block their data;
* when , Retailer H should set a price of such that all the consumers’ data blocking will be deterred;
* when , Retailer H will set a price of such that all the consumers will give up blocking data.

Substituting the optimal price into the profit functions, we obtain the results in Table 5.

(3) **Case HL:** Using backward induction, we first solve Retailer L’s pricing problem, which obtains . With the response function, we simplify Retailer H’s profit function to

*,*

where .

The constraint that should be satisfied in the above function. Therefore, we discuss Retailer H’s problem in two aspects. Firstly, when , is always satisfied. The optimization problem becomes

Max: *,*

s.t. .

Solving the problem, we can obtain the optimal price decision as

The solution procedure is straightforward, which is omitted here. Substituting the results into the target function, we obtain the corresponding result as.

Secondly, when , is always satisfied. The optimization problem becomes

Max:*,*

s.t. .

The solution procedure is straightforward, which is omitted here. Solving the problem, we can obtain the optimal price decision as

Comparing the two scenarios, we find that

* when , Retailer H should set a low price, such that part of the consumers will block their data;
* when , Retailer H should set a price of such that all the consumers’ data blocking will be deterred;
* when , Retailer H will set a price of such that all the consumers will give up blocking data.

Substituting the optimal price into the profit functions, we obtain the results in Table 6.

**Appendix C2: Asymmetric production costs for Retailers H and L**

(1) **Case N**: In this case, the two retailers cannot use behaviour-based pricing because the platform provides no CPS to both firms. Therefore, the corresponding profit functions can be obtained as

*, ,*

*.*

Solving the Nash game of Retailer H and Retailer L, we obtain that , . The corresponding profits for the two retailers and the platform can be expressed as , , . We use case N as a benchmark comparing to the cases H and HL. Note that, in this case it is necessary to assume that to ensure the positivity of Retailer H’s demands.

(2) **Case H:** In this case, the platform only provides CPS to Retailer H. Therefore, Retailer L has no ability to provide behaviour-based prices. For both types of consumers, its selling price is . However, Retailer H can provide behaviour-based prices to the old consumers after obtaining the CPS from the platform. Note that, in the extended model, when considering the production cost , an important constraint of should be satisfied, which ensures positive sales margin for Retailer H. Also, we only consider the case of , which guarantees positive demands in the new consumer group.

**Lemma C1. (Consumers’ Response in Case H)** *Let ,*

*When , old consumers within the region of will choose to block data; otherwise, they will not block data.*

*(1) Old consumers within the region of will purchase from H. The prices are conditional in their individual type, i.e.,*

*.*

*(2) Old consumers within the region of will purchase from L, and pay the price of ;*

*(3) Old consumers within the region of will not purchase anything.*

From Lemma C1, when firm L’s price drops to a certain level or Retailer H’s production cost is relatively high (), Retailer L can have positive demand in the old consumer group. Therefore, in this extension, we concentrate on an interesting case that is relatively high.

The two retailers set their selling prices simultaneously to maximize their individual profits.

*,*

*.*

*s.t. .*

Follow the same fashion as the base model, we obtain the equilibrium results are presented in Table 7.

(2) **Case HL:** In this case, the platform provides CPS to both retailers. Therefore, both retailers can provide behaviour-based prices to consumers if they do not block their individual data. We show the consumers’ choice of data blocking w.r.t. retailers’ prices as follows.

**Lemma C2. (Consumers’ Response in Case HL)** *Let , old consumers within the region of will choose to block data; otherwise, they will not block data.*

*They will purchase from H if and pay the prices conditional on their individual type, i.e.,*

*.*

*They will purchase from L if and the price is .*

Next, based on the consumers’ purchasing and data blocking decisions, we formulate the profit functions for the retailers and the marketplace, and derive their equilibrium results.

*,*

*s.t.*

*,*

*.*

Follow the same fashion as the base model, we obtain the equilibrium results are presented in Table 8.

**Appendix C3: Positive CPS fee charged by the marketplace**

We analyse the impacts of CPS fee based on the results of Proposition 5 in the main model.

First, we notice that for Retailer L, the adoption of HL is always unbeneficial comparing to the case of N and H. Therefore, when considering positive CPS fee, Retailer L is still not willing to accept CPS from the platform.

Second, we turn to Retailer H. Since HL is never reached, we only need to compare the outcomes of case N and case H. Considering service fee, its profits can be expressed as

,

.

Then, if , i.e.,

,

Retailer H will accept the platform’s CPS. Otherwise, it will not accept CPS from the platform.

Last, we turn to the platform. We find that any positive value of service fee will result in its preference for strategy H.

Therefore, we conclude that

(1) the equilibrium result is H when ;

(2) the equilibrium result is N when

**Appendix D: Equilibrium results and expressions for the extended models**

**Table 5.** The equilibrium results for case H in Extension 1.

|  |
| --- |
| **Region I**: |
| ,, |
| , |
| , |
| , |
| . |
| **Region II**: |
| ,, |
| , |
| , |
| , |
| , |
| **Region III**: |
| , , |
| , |
| , |
| , |
| . |

**Table 6.** The equilibrium results for case HL in Extension 1.

|  |
| --- |
| **Region I**: |
| ,, |
| , |
| , |
| , |
| . |
| **Region II**: |
| ,, |
| , |
| , |
| , |
| , |
| **Region III**: |
| , , |
| , |
| , |
| , |
| , |

**Table 7.** The equilibrium results for case H in Extension 2.

|  |
| --- |
| **Region I**: |
| ,, |
| , |
| , |
| , |
| . |
| **Region II**: |
| ,, |
| , |
| , |
| , |
| , |
| **Region III**: |
| ,, |
| , |
| , |
| , |
| , |

**Table 8.** The equilibrium results for case HL in Extension 2.

|  |
| --- |
| **Region I**: |
| ,, |
| , |
| , |
| , |
| . |
| **Region II**: |
| ,, |
| , |
| , |
| , |
| , |
| **Region III**: |
| , , |
| , |
| , |
| , |
| , |