https://pubsonline.informs.org/journal/msom
MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
Submitted to *Manufacturing & Service Operations Management*
ISSN 1523-4614, EISSN 1526-5498

Optimal Pricing Strategies of Information Goods with Data-enabled Learning

Fang Liu
Durham University Business School, Durham University, United Kingdom fang.liu@durham.ac.uk

Yu Tang
College of Business, Shanghai University of Finance and Economics, P.R.China cosmostang@stu.sufe.edu.cn

Lei Xie
College of Business, Shanghai University of Finance and Economics, P.R.China xie.lei@mail.shufe.edu.cn

Houmin Yan
College of Business, City University of Hong Kong, Hong Kong Houmin.Yan@cityu.edu.hk

Abstract. ***Problem definition*:** We consider a monopoly firm that sells information goods in a two-period model with data-enabled learning. Empowered by data-enabled learning, the users' product valuation may increase over time. Specifically, data-enabled learning consists of both across-user and within-user learning. Across-user learning enhances the users' product valuation through data collected from all users, meanwhile, within-user learning independently enhances each user's valuation by consolidating their individual usage data. The firm may offer pay-per-use and/or flat-rate payment options. We characterize optimal pricing strategies and analyze how data-enabled learning affects market dynamics. ***Methodology/results*:** We consider a game-theoretic setting between the firm and users. The firm selects a two-period pricing strategy to maximize its total profit. The users then choose their purchasing option in each period. We find that across-user learning enhances the degree of homogeneity within the user base, while within-user learning leads to a more heterogeneous user base. As a result, under across-user learning, the firm prefers the pay-per-use option to expand its user base in the first period and leans towards the flat-rate option in the second period to focus on maximizing profits. In contrast, the firm prefers flat-rate options in the first period under within-user learning because it can boost user engagement through the flat-rate bias. This gives the firm the flexibility to choose the most effective segmentation strategy in the second period to optimize profits. ***Managerial implications*:** Data-enabled learning plays a critical role in how firms should decide on their pricing strategies. The results offer insight into how businesses can leverage their data strategies to shape their user demographics and improve their competitiveness in the market.

Key words: Data-enabled learning; Data network effects; Behavior-based pricing

1. Introduction

1.1. Background and Motivation

Data and AI technologies are fundamentally reshaping the world, driving unprecedented transformations across all business sectors. A data flywheel describes a self-reinforcing cycle in which the accumulation and analysis of data enhance products and services, improve customer value, and ultimately lead to increased business growth. Trautman (2018) further emphasizes that AI-driven initiatives can harness data to promote continuous growth and optimization through what is known as the AI Flywheel effect. Therefore, as a critical part of the data flywheel, data-enabled learning provides substantial potential benefits in terms of efficiency, innovation, and competitive advantage. The concept of data-enabled learning was first introduced by Hagiu and Wright (2023), who suggests that effective use of data can significantly enhance a firm's market adaptability and product innovation. In general, data-enabled learning consists of across-user and within-user learning. Across-user learning enhances all users' experience by analyzing data accumulated from different users. For example, Tesla's Full Self-Driving (FSD) system gathers driving data from its users to enhance its autonomous driving capabilities. This system consolidates the collected data to improve service for all users. Dropbox and Slack show that continuous data-driven improvements can reduce customer acquisition costs while boosting user engagement and retention. While, within-user learning independently enhances each user's experience by consolidating their own usage data. For instance, Netflix uses sophisticated personalized recommendations powered by user data to significantly boost user engagement. TikTok employs sophisticated algorithms to cater to each user's unique preferences by analyzing their browsing and social network data. Data-enabled learning can also be effectively applied in category-based scenarios. In general, users are classified into distinct categories based on their characteristics. When a user generates a new data point, it can either enhance the valuation of users within the same category, referred to as within-category valuation improvement, or improve the valuation for the entire user population, known as across-category valuation improvement.

In a data flywheel, data-enabled learning is implemented through the *data network effect* (Gregory et al. 2021, Haftor et al. 2021), which leverages data accumulation and analysis to enhance the value of a product or service and user experience. In contrast to user-based network effects that generate the value of the user network by increasing the number of users or offering complementary goods and services, Levine and Jain (2023) highlights the distinct role of the data network effects in AI technologies. Data network effects enhance AI by leveraging extensive amounts of user-generated data to improve predictive capabilities and hence increase the user valuation of the product or service. This process involves a virtuous feedback loop, which not only increases algorithm value but also attracts more users and their engagement, thereby fostering continuous improvement in AI applications and business growth. Ideally, each dataset contributes in multiple ways to improving algorithm performance. This impact is also influenced by a specific user; data generated by the same user may have different effects compared to data produced by different users. Furthermore, the data network effect creates values based on different machine learning algorithms utilized in different scenarios. For example, self-driving algorithms are trained using driving data collected from various users, resulting in an improved self-driving system that benefits all users. In contrast, personalized recommendation algorithms rely on individual user data, enhancing the experiences of those who share their information.

To harness data-enabled learning to achieve market success, effective pricing strategies become essential. Two frequently employed pricing models are pay-per-use and flat-rate options. Under the pay-per-use option, the firm charges users a price each time they use the good or service. As a result, the more frequently a user engages with the offering, the higher the total charges incurred. In contrast, the flat-rate option involves a fixed price that users pay, regardless of how often they use the good or service. This option can incentivize users to use the good more frequently, leading to the *flat-rate bias*. Interestingly, as more users engage or engage more frequently in using the good, a larger volume of data is collected. The growing volume of data enhances the data network effect. This enables the firms to enhance their learning processes, refine personal experiences, and ultimately improve overall performance. However, little attention has been paid to exploring how firms can effectively leverage data-enabled learning to optimize their pricing strategies. In addition, it remains unclear how market structures defined by these pricing strategies evolve over time, both before and after the implementation of data-enabled learning.

Our research is motivated by the following research questions: What are the optimal pricing strategies for information goods in the presence of data-enabled learning? How does data-enabled learning affect market dynamics over time? What is the difference between across-user and within-user learning in terms of the impact on optimal pricing strategies and market dynamics?

In this paper, we consider a firm that sells information goods to its users over two periods. In the first period, during the firm's startup period, users' valuation of the good is uniformly distributed between 0 and 1. The firm offers both pay-per-use and flat-rate payment options. In the second period, empowered by data-enabled learning, users' valuation change. In the context of across-user learning, each user experiences the same increase in valuation, with the magnitude of this increase depending on the amount of user data collected in the first period. Conversely, in the context of within-user learning, each user's valuation increase varies; the more frequently a user uses the good in the first period, the more personal data is collected, resulting in a greater increase in her valuation. The good continues to support both pay-per-use and flat-rate payment options in the second period. In this user-firm interaction, users aim to choose the most suitable purchasing option from no-purchase, pay-per-use, and flat-rate options and consuming units in each period, while the firm decides an optimal two-period pricing strategy to maximize total profits.

1.2. Key Results and Contributions

Our key results are as follows. First, data-enabled learning profoundly impacts the user base, offering insights into how businesses can leverage their data strategies to shape their user demographics and enhance their competitiveness in the market. Across-user learning strengthens the data network effect by accumulating data from all paying users and provides inclusive benefits to all users, which enhances the degree of homogeneity of the user base. When the across-user learning capability is sufficiently high, the flat-rate payment option will cover the entire market. However, when the across-user learning capability is not sufficiently high, the firm may initially set lower prices to expand its user base and then increase prices in the second period to maximize profits. Interestingly, within-user learning enhances the data network effect by accumulating personalized data and providing exclusive benefits to each user, leading to a more heterogeneous user base. Consequently, fewer users are served following the implementation of within-user learning.

Second, the firm may effectively use pricing strategies to influence market dynamics over the two periods under data-enabled learning. This suggests that pricing strategies play a critical role in enhancing data strategies to achieve market success. Alternatively, effective pricing not only influences user valuation but also directly impacts how well a business can leverage its data-driven insights. We find that the flat-rate bias plays different roles under across-user and within-user learning. Specifically, with across-user learning, the firm prefers the pay-per-use option to expand its user base in the first period, while opting for the flat-rate option to maximize profits in the second period due to the flat-rate bias and the higher degree of homogeneity in the user base. However, under within-user learning, the flat-rate bias generated by the flat-rate option can further amplify the data network effect and enhance the degree of heterogeneity in the user base. Consequently, the firm opts for the flat-rate option in the first period to boost user engagement and leverages the advantages of frequent interactions in the second period by segmenting users when the within-user learning capability is either sufficiently low or high.

1.3. The Organization of the Paper

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature. Section 3 introduces the basic model, outlines some preliminary concepts, and formulates the model. Sections 4 and 5 investigate the impacts of across-user and within-user learning, respectively. Section 6 revisits the general with both across-user and within-user learning. Finally, Section 7 presents the concluding remarks.

2. Literature Review

Our research is closely related to the literature on network effects. Katz and Shapiro (1985) first reveals that consumer demand depends on the number of customers in addition to price and product quality, which lays a groundwork for analyzing how network externalities affect market outcomes. In particular, the data network effect is first introduced by Gregory et al. (2021) and Haftor et al. (2021). Hagiu and Wright (2023) categorizes these two value creation mechanisms as across-user learning and within-user learning, respectively, and collectively refers to them as data-enabled learning. The presence of network effects also has a significant impact on pricing strategies. A substantial body of literature has investigated pricing strategies under user-based network effects (Chien and Chu 2008, Jiang et al. 2015, Hajihashemi et al. 2022, Liu and Luo 2023). Chien and Chu (2008) highlights that a seller can attract a larger customer base by setting lower initial prices and raising them as the network grows when consumers are more cautious about the prospect of the durable-good's success. Jiang et al. (2015) indicates that larger networks can support higher prices due to the increased utility that a larger user base provides to each consumer. Gurkan and de Véricourt (2022) explores the dynamics among machine learning (ML), data acquisition, pricing, and contracting within a two-period moral hazard framework. They find that a firm's decisions are significantly influenced by the interaction between the amount of data used to train the machine and the provider's effort to improve the algorithm. However, little attention has been paid to connecting data-based network effects to pricing strategies for information goods. To the best of our knowledge, we are the first to leverage the concept of data-enabled learning to characterize the generation process of data network effects and further investigate the pricing strategies under data-enabled learning.

Another closely related area of research is the literature on pricing strategies for information goods, which have been extensively studied. Information goods have their unique characteristics, such as low marginal costs and high replicability. Research in this area has explored various pricing mechanisms, including sales, pay-per-use, and subscription models. Essegaier et al. (2002) examine multi-version pricing and personalized pricing strategies, emphasizing how differentiated pricing strategies tailored to various consumer segments can maximize profits. Their model reveals that price discrimination and product versioning are effective tools for profit maximization in the information goods market. Sundararajan (2004) investigates nonlinear pricing strategies, particularly focusing on optimizing pricing through a mix of fixed fees (flat-rate charges) and usage-based pricing under conditions of incomplete information. His findings suggest that incorporating a fixed-fee pricing element can enhance a seller's profits by reducing transaction costs associated with usage-based pricing and increasing consumer surplus, even in the presence of such costs. Balasubramanian et al.

(2015) offer a comprehensive analysis, demonstrating that pay-per-use can yield higher profits in a monopoly scenario, provided that the psychological cost associated with per-use pricing is minimal. In addition, as mentioned in Balasubramanian et al. (2015), "When consumers are faced with an expected pattern of use for an information good, they tend to favor a flat rate (consistent with selling) over pay-per-use, even when the expected payments under the mechanisms do not differ. This preference is referred to as the flat-rate bias." Lambrecht and Skiera (2006) detail the multiple effects that can lead to the flat-rate bias, including the insurance effect, the taxi meter (or ticking meter) effect, the convenience effect, and overestimation effect. However, much of the existing literature on information goods pricing strategies focuses on single-period models. In contrast, our research extends this by developing a two-period monopolistic model to study optimal pricing strategies under data-enabled learning. Unlike the models presented by Essegaier et al. (2002) and Balasubramanian et al. (2015), we analyze how data-enabled learning would affect the pricing strategies by explicitly considering the impact of pricing on both market share and consumer usage frequency.

Our research also falls into the literature of Behavior-Based Pricing (BBP) which refers to the practice that the firm differentiates and charges new customers a different price based on the data collected from the customer's purchasing history (Li et al. 2020). Fudenberg and Villas-Boas (2006) provides a comprehensive literature review. BBP leverages detailed customer data, such as customers' purchase history or response to previous pricing offers, to maximize the firm's profit by tailoring prices to individual customers. Varian (1989) discusses the potential of using customer data to implement personalized pricing and capture more consumer surplus. Fudenberg and Tirole (2000) investigates how firms could use consumer lock-in by changing prices based on past purchase behavior. Ma et al. (2021) finds that online service providers can significantly influence users' perceived quality of service through social learning mechanisms, which in turn allows these providers to adjust their dynamic pricing strategies in response to user feedback. Furthermore, Jing (2011) analyzes the effects of market learning on the pricing and timing of durable goods releases, revealing how consumer learning shapes corporate market strategies. The pricing strategy we study in this paper is a kind of BBP because the firm can perform price discrimination after observing which pricing option the customers have selected. Moreover, our research emphasizes data-enabled learning as a function of both market share and customer usage frequency, both of which are endogenous variables directly influenced by pricing strategies.

We also contribute to the literature on customer loyalty programs (Liu 2007, Kuksov and Zia 2021, Chun and Ovchinnikov 2019). Liu (2007) conducts a longitudinal analysis over two years, revealing that loyalty programs did not significantly change the spending levels or store-exclusive loyalty of heavy buyers who were already frequent customers. However, the programs had a notable positive impact on light and moderate buyers, increasing their purchase frequency and transaction sizes, thus enhancing their loyalty to the store. Similarly, Chun and Ovchinnikov (2019) corroborates these findings, showing that the effectiveness of loyalty programs is not uniform across all customer segments. Kuksov and Zia (2021) incorporates endogenous search costs into the analysis of customer loyalty and its impact on market dynamics and pricing strategies. In this paper, as within-user learning is based on individual user's personal data, particularly their historic usage frequency, users who have frequently used a product or service in the past derive greater benefits and become more loyal due to lock-in effects. It can be considered as one type of loyalty program.

3. Data-enabled Learning and User-firm Interactions

In this section, we first introduce the basic setting and show some preliminaries. Then, we introduce a general data-enabled learning model that includes two commonly observed models: across-user learning and within-user learning. Finally, we describe the interactions between the firm and its users and formulate the firm's problem.

3.1. Basic Settings and Preliminaries

We consider a monopolist firm that sells a good over two periods. The good can be repeatedly consumed by users in a market in both periods. Without loss of generality, assume that the unit production cost of the good is zero. Users are heterogeneous and their valuations are uniformly located over interval $[0,1]$ at the beginning of the horizon. Denote by v a user's valuation, which represents her utility of consuming one unit of the good. Due to data-enabled learning, user valuation v may change over the two periods, which will be elaborated in Section 3.2. Similar to Sundararajan (2004), let $U(q,v,p)$ be the net utility of a user with valuation v consuming q units of the good, i.e.,

$$U(q,v,p) = qv - \gamma q^2 - p(q), \tag{1}$$

where the first term qv accounts for the total utility of consuming q units, the second term is the dis-utility due to the decreasing marginal utility of consuming the good, and the third term $p(q)$ is the price of consuming q units paid by the user. The parameter γ indicates how quickly the marginal utility decreases as more goods are consumed.

Pricing options. We consider two prevalent pricing options that the firm offers to users in practice: *flat-rate* option that targets users who use the most number of times and *pay-per-use* option that targets users who use less.

1. *Flat-rate* option. A user pays a flat rate p_f to the firm regardless of the number of times she uses the good. The flat rate p_f can be interpreted as a membership fee for a user to have unlimited access to the good. The net utility of a user with valuation v consuming q units of the good then is given by $U_f(q, v, p_f) = qv - \gamma q^2 - p_f$.

2. *Pay-per-use* option. Users pay a price p_0 for each purchase or usage of the good. The price p_0 can be interpreted as a unit price paid by a user each time she uses the good. The net utility of a user with valuation v consuming q units of the good is $U_p(q, v, p_0) = q(v - p_0) - \gamma q^2$.

The firm offers both flat-rate and pay-per-use options to all users. We assume each user is rational. If a user with valuation v selects the flat-rate option, then the optimal amount of good she would consume is $q_f^*(v) = v/(2\gamma)$. Alternatively, if a user with valuation v selects the pay-per-use option, then the optimal amount of good she would consume is $q_p^*(v) = (v - p_0)/(2\gamma)$. Obviously, a user with a higher valuation consumes more under either option. In addition, users consume more under the flat rate option than the pay-per-use option due to the flat rate bias (Balasubramanian et al. 2015, Lambrecht and Skiera 2006). However, users may not always favor the flat rate option in terms of net utility. Despite the two options offered by the firm, a user may also choose not to consume. In this case, she receives a zero net utility. That is, a user would choose the no-purchase option only if both pricing options yield negative net utilities yet otherwise select the one that yields a larger net utility.

Note that if the firm offers its product or service for free, i.e., $p_0 = 0$ or $p_f = 0$, then all users consume $q_f^* = v/(2\gamma)$. This scenario can arise when a firm prioritizes increasing consumption at the expense of profits in the first period, ultimately aiming to maximize total profit.

Given a pair of flat rate p_f and pay-per-use price p_0, the users can be segmented into at most three groups: those who purchase under the flat-rate option, those who purchase under the pay-per-use option, and those who do not purchase. Because those who choose the flat-rate option always have a higher valuation than those who chose the pay-per-use option, we define critical valuation v_f as the user's valuation that is indifferent between the flat rate and pay-per-use price, i.e., $U_f\left(q_f^*(v_f), v_f, p_f\right) = U_p\left(q_p^*(v_f), v_f, p_0\right) \geq 0$. Similarly, we define critical valuation v_p as the user's valuation that is indifferent between the pay-per-use price and no-purchase option, i.e., $U_p\left(q_p^*(v_p), v_p, p_0\right) = 0$. In general, we refer to the *flat-rate users* as those who choose the flat-rate option and the *pay-per-use users* as those who choose the pay-per-use option.

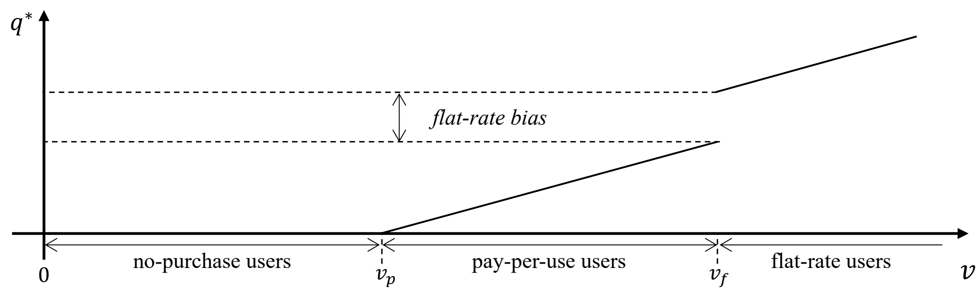
We characterize the two critical valuations v_f and v_p in the following proposition.

Proposition 1. (USER SEGMENTATION) *Given the flat-rate price $p_f > 0$ and pay-per-use price $p_0 > 0$, we have $v_p = p_0$ and $v_f = 2\gamma p_f/p_0 + p_0/2$.*

(a) When $p_f \geq p_0^2/(4\gamma)$, a user with valuation v would choose the flat-rate option with consumption quantity $q_f^* = v/(2\gamma)$ if $v \geq v_f$; she would choose the pay-per-use option with consumption quantity $q_p^* = (v - p_0)/(2\gamma)$ if $v_p \leq v < v_f$; and otherwise she would choose not to consume.

(b) When $p_f < p_0^2/(4\gamma)$, a user with valuation v would choose the flat-rate option with consumption quantity $q_f^* = v/(2\gamma)$ if $v \geq 2\sqrt{\gamma p_f}$, and otherwise would choose not to consume.

Figure 1 User Segmentation



For illustrative purposes and simplicity, we use a two-letter system to represent the structure of a market. The first letter indicates the extent of market coverage, while the second letter denotes whether the market serves multiple user segments. Specifically, 'P' and 'F' respectively represent partially and fully covered markets. For the second letter, we use 'T' to represent a market with multiple user segments, 'F' to represent a market with only flat-rate users, and 'P' to represent a market with only pay-per-use users. This notation allows us to concisely describe the various market structures under consideration.

Part (a) of Proposition 1 suggests that when the flat-rate price is relatively high, i.e., $p_f \geq p_0^2/(4\gamma)$, users fall into three segments in the market referred as a *Partially Covered Tiered Market* (PT market), as shown in Figure 1. In this PT market, flat-rate users ($v \geq v_f$) choose the flat-rate option, pay-per-use users ($v_p \leq v < v_f$) select the pay-per-use option, and the rest of the users ($v < v_p$) leave without a purchase. In addition, as pay-per-use price p_0 increases, the amount of flat-rate users becomes more and the amount of pay-per-use users becomes less. The higher the flat-rate price is, the fewer users would choose the flat-rate option. Note that, when the flat-rate price is relatively low, i.e., $p_f < p_0^2/(4\gamma)$, the flat rate option would dominate the pay-per-use option due to flat rate bias. This reduces to Part (b) of Proposition 1, which can be referred as a *Partially Covered Flat-rate Market* (PF market). Interestingly, the firm's pricing strategy determines the number of user segments to serve in the market, which can be seen as a reflection of the market structure.

In addition, Proposition 1 implies that, a pair of critical valuations corresponds to a pair of flat-rate and pay-per-use prices in each period and vice versa. Particularly, for a given pair of prices (p_0, p_f), the critical valuations are $v_p = p_0$ and $v_f = 2\gamma p_f / p_0 + p_0/2$. Conversely, if we have a pair of critical valuations (v_p, v_f), the pay-per-use price is $p_0 = v_p$ and the flat-rate price is $p_f = v_p(2v_f - v_p)/(4\gamma)$. For ease of analysis, we can use the critical valuations as decision variables in place of the prices.

We assume that the firm can distinguish old and new users in the second period. For convenience, we denote the flat-rate and pay-per-use prices at period $t \in \{1, 2\}$ as $p_{t,f}$ and $p_{t,0}$, respectively. The corresponding critical valuations at period t are $v_{t,p}$ and $v_{t,f}$ ($t \in \{1, 2\}$) and the corresponding market structure in period t is denoted by $s_t \in$ {PT, PF, PP, FT, FF, FP} ($t \in \{1, 2\}$).

3.2. Data-enabled Learning

Based on the data collected in the first period, the firm can improve the value of the information goods by learning from users' usage, i.e., data-enabled learning (Hagiu and Wright 2023), and hence improve the users' valuation in the second period. Specifically, we model the enhancement of user valuation under two commonly observed data-enabled learning processes: across-user learning and within-user learning.

Denote by v_o the valuation of the good for a user in the first period where the subscript "o" represents the original valuation.

The data is generated each time a user uses the product or service. Furthermore, each individual use of the product or service by a user generates one unit of data. We assume that the firm can collect every unit of data generated by the users. Then $q_p^*(v_o)$ represents the number of data units collected by the firm from a pay-per-use user with original valuation v_o, while $q_f^*(v_o)$ represents the number of data units collected by the firm from a flat-rate user with original valuation v_o. Consequently, the total number of data units collected by the firm in the first period is given by:

$$n_1 = \int_{v_{1,p}}^{v_{1,f}} q_p^*(v_o) dv_o + \int_{v_{1,f}}^{1} q_f^*(v_o) dv_o, \tag{2}$$

where the first term implies the total number of data units collected from the pay-per-use users and the second term implies that collected from the flat-rate users.

1. *Across-user* learning. Across-user learning implies that a firm can enhance product value via data gained by gathering insights and knowledge from the experiences, behaviors, and feedback of multiple users in the past. The more users the firm serves, the greater the amount of user data it collects, which in turn increases the value each user can derive from their consumption. Therefore, across-user learning can be regarded as a type of network effect that provides inclusive benefits to each user. We denote $\delta_a(n_1, \alpha)$ as the across-user learning function, which represents the enhanced value to each user's valuation in the second

period due to across-user learning. Note that, the across-user learning function is independent of the user's purchasing choice in the first period. In particular, $\alpha(>0)$ represents the firm's capability of implementing across-user learning. A larger α implies a greater capability of implementing across-user learning, which leads to a larger added value to each user's valuation in the second period. We further impose the following assumption on the across-user learning function.

ASSUMPTION 1. *Assume that across-user learning function $\delta_a(n_1,\alpha)$ is twice differentiable and satisfies the following conditions within its domain:1) $\delta_a(0,\alpha)=\delta_a(n_1,0)=0$; 2) $\delta_a(n_1,\alpha)$ is increasingly convex with respect to n_1; and 3) $\delta_a(n_1,\alpha)$ increases in α and is supermodular in (n_1,α).*

Part 1) implies that the user valuation in the second period does not increase if the number of data units collected in the first period is zero or the firm is unable to implement across-user learning ($\alpha=0$). Part 2) indicates a positive network effect of across-user learning on the first period's demand. In part 3), a better capability of implementing across-user learning results in a greater increase in user valuation in the second period. In addition, the number of data units collected in the first period n_1 and the capability of implementing across-user learning α are complementary.

2. *Within-user* learning. Within-user learning refers to the process of gathering and analyzing data from individual users to personalize and improve their unique experiences over time. Usually, the more a user consumes a firm's product, the more value the user can gain from within-user learning by better understanding and responding to her specific needs. In contrast to across-user learning, within-user learning offers exclusive benefits for specific users. As flat-rate users consume more than pay-per-use users, the primary beneficiaries of within-user learning are the flat-rate users. For ease of analysis, we assume that within-user learning only enhances user valuations of the flat-rate users in the first period. Denote by $\delta_w(q_f^*(v_o),\beta)$ within-user learning function representing the enhanced value in the second period to the valuation of the user with original valuation v_o due to within-user learning. Similarly, parameter $\beta(>0)$ captures the firm's capability of implementing within-user learning.

Summing the enhanced values of across-user and within-user learning, the updated valuation in the second period can be expressed as

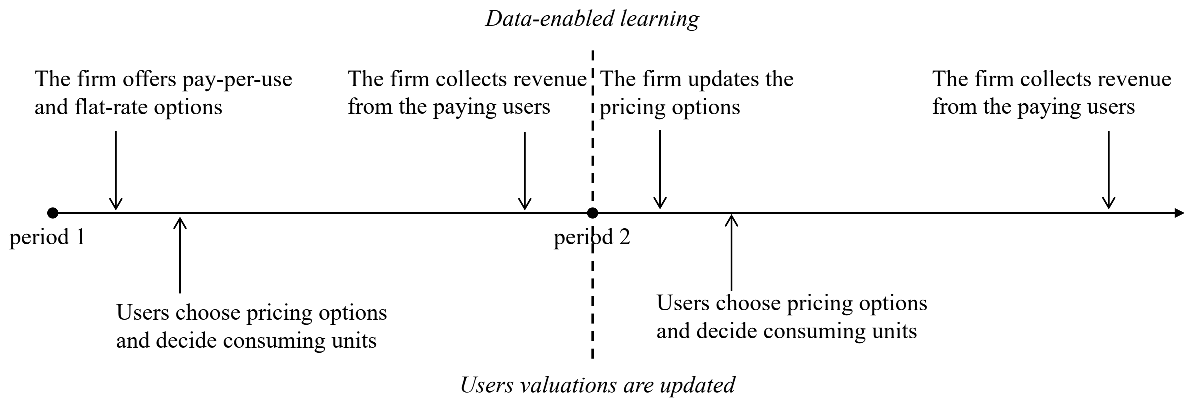
$$v_u(v_o)=v_o+\delta_a(n_1,\alpha)+\delta_w\left(q_f^*(v_o),\beta\right)\mathbf{I}_{v_o\geq v_{1,f}}, \tag{3}$$

where the subscript "u" represents the updated user valuation after implementing data-enabled learning.

3.3. Interactions between the Firm and Users

We next discuss the interactions between the monopolist firm and users in this two-period model. At the beginning of the first period, the firm offers the two pricing options with the flat rate $p_{1,f}$ and pay-per-use price $p_{1,0}$. Each user chooses a pricing option that maximizes her net utility. If she chooses the flat-rate option, then she pays the firm a flat rate $p_{1,f}$ regardless of the amount of information goods consumed. If she chooses the pay-per-use option, then she pays the firm $p_{1,0}$ for each purchase. Based on users' choices in the first period, the firm implements data-enabled learning to improve users' valuations. At the beginning of the second period, the firm updates the flat rate to $p_{2,f}$ and the pay-per-use price to $p_{2,0}$, respectively. Users then select a pricing option and optimize their consuming units based on their updated valuations. The firm then collects revenue from the paying users. Figure 2 illustrates the sequence of events and the interactions between the firm and users.

Figure 2 The Sequence of Events and the Interactions between the Firm and Users



The firm's problem. We formulate the firm's problem in a backward manner. Given period t's flat rate $p_{t,f}$ and pay-per-use price $p_{t,0}$, let $d_{t,p}$ represent the number of pay-per-use users and $d_{t,f}$ represent the number of flat-rate users at period t, respectively. In addition, denote by $d_t = d_{t,p} + d_{t,f}$ the total number of paying users at period t. Clearly, $d_{t,p}$ and $d_{t,f}$ are functions of flat rate $p_{t,f}$ and pay-per-use price $p_{t,0}$. The firm maximizes its revenue in the second period by choosing a flat rate $p_{2,f}$ and pay-per-use price $p_{2,0}$, i.e.,

$$
\begin{aligned}
\Pi_2^*(p_{1,0}, p_{1,f}) &= \max_{p_{2,0}, p_{2,f}} \Pi_2(p_{1,0}, p_{1,f}, p_{2,0}, p_{2,f}) \\
&= \max_{p_{2,0}, p_{2,f}} \Bigg\{ p_{2,f} \cdot d_{2,f} \\
&\qquad + p_{2,0} \int_{v_{2,p}}^{v_{2,f}} q_p^* \left(v_u(v_o) \right) dv_o \Bigg\}.
\end{aligned}
$$

In the first period, the firm maximizes its total discounted profit over the two periods by choosing the flat rate $p_{1,f}$ and pay-per-use price $p_{1,0}$,

$$\begin{aligned}\Pi_1^* &= \max_{p_{1,0},p_{1,f}} \Pi_1(p_{1,0},p_{1,f}) \\ &= \max_{p_{1,0},p_{1,f}} \Bigg\{ p_{1,f}\cdot d_{1,f} \\ &\qquad + p_{1,0}\int_{v_{1,p}}^{v_{1,f}} q_p^*(v_o)dv_o + \lambda\Pi_2^*(p_{1,0},p_{1,f}) \Bigg\},\end{aligned}$$

where $\lambda \in [0,1]$ is the discount factor. We first present the optimal pricing strategies without data-enabled learning in the following lemma.

Lemma 1. *In the absence of data-enabled learning, i.e., $v_u(v_o) = v_o$,*

1. *The optimal critical valuations satisfy $v_{1,p}^* = v_{2,p}^* = 2/5$ and $v_{1,f}^* = v_{2,f}^* = 4/5$.*
2. *The number of paying users is $d_1^* = d_2^* = 3/5$ and the number of flat-rate users is $d_{1,f}^* = d_{2,f}^* = 1/5$.*

Lemma 1 implies that without data-enabled learning, user valuation remains the same across both periods, resulting in identical optimal pricing strategies. Moreover, the firm serves $3/5$ and $1/5$ of the total users as the paying and flat-rate users, respectively.

Next, we separately explore the impacts of across-user learning and within-user learning and move back to the general model in Section 6.

3.4. Data Network Effect and It's Relationship with Flat-rate Bias

Unlike conventional network effects, data network effects stem from the positive externalities generated by the accumulation of data from users who consume information goods. Specifically, data network effects can be categorized into two types based on their source: effects generated from different users and effects generated from repeated use by individual users. The first type occurs through across-user learning, where the value of the information good for all paying users in the second period is enhanced by the consumption behavior of users in the first period. The second type occurs through within-user learning, where only the value for flat-rate users in the first period increases as they continue to the second period. This second source of the data network effect is further amplified by flat-rate bias, which leads to more frequent usage of the information good by flat-rate users compared to pay-per-use users in the first period.

According to Proposition 1, the flat-rate bias can be quantified by Δq, which represents the incremental amount in usage when a user switches from a pay-per-use option to a flat-rate option. Specifically, Δq is given by

$$\Delta q = \frac{\sqrt{(4\gamma p_f) \wedge p_0^2}}{2\gamma}.$$

The flat-rate bias increases with higher values of p_f and p_0, but decreases as γ increases. A larger γ implies that the marginal utility diminishes more rapidly with increased consumption, resulting in reduced usage and a smaller flat-rate bias. Conversely, higher prices tend to discourage frequent usage, leading to a larger flat-rate bias.

4. The Impact of Across-user Learning

In this section, we investigate how across-user learning affects market dynamics by optimizing the firm's pricing strategies in both periods when the user's valuation in the second period is updated with only the across-user learning function. Specifically, we assume that the updated user valuation v_u in (3) consists of only the across-user learning function, i.e.,

$$v_u(v_o) = v_o + \delta_a(n_1, \alpha). \tag{4}$$

We first characterize the optimal critical valuations in the second period in the following lemma. Let $\Delta_a = \delta_a(n_1, \alpha)$ be the inclusive enhanced value on user valuation in the second period due to across-user learning.

Lemma 2. *Given the critical valuations $v_{1,p}$ and $v_{1,f}$ of the first period, 1) the optimal critical valuations of the second period $v^*_{2,p}$ and $v^*_{2,f}$ can be characterized as follows,*

$$(v^*_{2,p}, v^*_{2,f}) = \begin{cases} \left(\frac{2}{5}(1+\Delta_a), \frac{4}{5}(1+\Delta_a)\right), & 0 \le \Delta_a < \frac{2}{3}, \\ \left(\Delta_a, 1+\frac{1}{2}\Delta_a\right), & \frac{2}{3} \le \Delta_a < 2, \\ (\Delta_a, \Delta_a), & \Delta_a \ge 2. \end{cases}$$

2) Furthermore, the number of pay-per-use users $d^*_{2,p}$, flat-rate users $d^*_{2,f}$, and paying users d^*_2 can be explicitly characterized as,

$$(d^*_{2,p}, d^*_{2,f}, d^*_2) = \begin{cases} \left(\frac{2}{5}(1+\Delta_a), \frac{1}{5}(1+\Delta_a), \frac{3}{5}(1+\Delta_a)\right), & 0 \le \Delta_a < \frac{2}{3}, \\ \left(1-\frac{1}{2}\Delta_a, \frac{1}{2}\Delta_a, 1\right), & \frac{2}{3} \le \Delta_a < 2, \\ (0, 1, 1), & \Delta_a \ge 2. \end{cases}$$

Lemma 2 part 1) indicates that the critical valuations in the second period increase as across-user learning provides more inclusive enhanced value to all users. Alternatively, when across-user learning provides more inclusive enhanced value to all users, more no-purchase users would switch to the pay-per-use option, and more pay-per-use users switch to the flat-rate option. Lemma 2 part 2) indicates that in a FT market, where each user consumes the good, as across-user learning provides more inclusive enhanced value to all users, the number of pay-per-use users decreases because some switch to the flat-rate option. Eventually, the flat-rate option may dominate the pay-per-use option, turning the market into a FF market if the inclusive enhanced value is sufficiently large.

We next analyze the market dynamics by comparing the critical valuations of the two periods. To link the segments delineated by $v_{2,p}$ and $v_{2,f}$ in the second period to those delineated by $v_{1,p}$ and $v_{1,f}$ in the first period, we define the inverse function $v_u^{-1}(\cdot)$. Specifically, for any critical valuation v_2 in the second period,

$$v_u^{-1}(v_2) = v_1 \in [0,1] \text{ such that } v_u(v_1) = v_2.$$

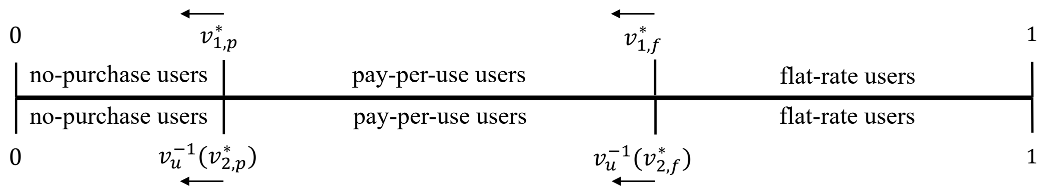
That is, $v_u^{-1}(v_{2,f})$ and $v_u^{-1}(v_{2,p})$ represent the user's valuations in the first period whose valuations in the second period are $v_{2,f}$ and $v_{2,p}$, respectively, allowing for a coherent analysis of user choices and market shifts over time. Proposition 2 characterizes the properties of the optimal critical valuations of the two periods.

Proposition 2. 1. *The optimal critical valuations in the first period $v^*_{1,p}(\alpha,\lambda,\gamma)$ and $v^*_{1,f}(\alpha,\lambda,\gamma)$ decrease with respect to α and λ but increase with respect to γ. In addition, for any $\lambda \in (0,1]$ and $\gamma > 0$, if δ_a is strictly increasing with respect to n_1 and α, there exists $\alpha_1(\lambda,\gamma)$ such that $v^*_{1,p}(\alpha,\lambda,\gamma) = 0$ and $v^*_{1,f}(\alpha,\lambda,\gamma) = 0$ for all $\alpha \geq \alpha_1(\lambda,\gamma)$.*

2. *Both $v_u^{-1}(v^*_{2,p}(\alpha,\lambda,\gamma))$ and $v_u^{-1}(v^*_{2,f}(\alpha,\lambda,\gamma))$ decrease with respect to α and λ. However, they both increase with respect to γ. In addition, for any $\lambda \in (0,1]$ and $\gamma > 0$, if δ_a is strictly increasing with respect to α, there exists $\alpha_{2,p}(\lambda,\gamma)$ and $\alpha_{2,f}(\lambda,\gamma)$ ($\alpha_{2,p}(\lambda,\gamma) \leq \alpha_{2,f}(\lambda,\gamma)$) such that $v_u^{-1}(v^*_{2,p}(\alpha,\lambda,\gamma)) = 0$ when $\alpha \geq \alpha_{2,p}(\lambda,\gamma)$ and $v_u^{-1}(v^*_{2,f}(\alpha,\lambda,\gamma)) = 0$ when $\alpha \geq \alpha_{2,f}(\lambda,\gamma)$.*

Figure 3 provides an illustration of Proposition 2. Part (1) of Proposition 2 implies that, as the firm becomes more capable of implementing across-user learning (i.e., a larger α) or the discount factor increases (i.e., a larger λ), it would act more aggressively in the first period. Specifically, the firm would set lower pay-per-use and flat-rate prices to increase the number of paying users and flat-rate users, respectively, with the goal of acquiring more user data. In contrast, as the marginal utility of users decreases more rapidly with increasing consumption (that is, a larger γ), the firm would act less aggressively in the first period due to the reduced consumption resulting in a reduced amount of data generated. Furthermore, if the capability to implement across-user learning is sufficiently large, that is, $\alpha \geq \alpha_1(\lambda,\gamma)$, the firm covers the entire market and adopts only the flat-rate option in the first period. Part (2) of Proposition 2 implies that, as α or λ increases, the number of flat-rate and paying users in the second period increase. In addition, the firm favors the flat-rate option in the second period when across-user learning provides a significant inclusive benefit, i.e., $\alpha \geq \alpha_{2,f}(\lambda, \gamma)$. Part (2) of Proposition 2 also provides a necessary and sufficient condition for the market to be fully covered in the second period. Specifically, when the capability of using across-user learning is low, i.e., $\alpha \leq \alpha_{2,p}(\lambda, \gamma)$, the market is partially covered and both pay-per-use and flat-rate users increase in the second period. Conversely, when the capability of using across-user learning is high, i.e., $\alpha > \alpha_{2,p}(\lambda, \gamma)$, all users at the second period become paying users and choose the flat-rate option as the capability of using across-user learning is sufficiently large, i.e., $\alpha \geq \alpha_{2,f}(\lambda, \gamma)$.

Figure 3 The Changes of the Optimal Critical Valuations as Across-user Learning Capability α or λ Increases



Note. The figure illustrates the changes of the optimal critical valuations at period 1 (above) and period 2 (below).

Therefore, due to across-user learning, the firm prefers the flat-rate option in the first period as it increases user engagement and thus expands the data pool, and favors the flat-rate option in the second period to maximize its profits. This is because flat-rate bias can lead to a larger revenue.

We further characterize how the capability of implementing across-user learning and the discount factor affect the market structure and optimal demand sizes of pay-per-use users and flat-rate users.

Corollary 1. 1. *The total number of paying users in both periods d_1^* and d_2^* increase in α. Furthermore, $d_1^* = 1$ when $\alpha \geq \alpha_1$ and $d_2^* = 1$ when $\alpha \geq \alpha_{2,p}$.*

2. *The number of pay-per-user users in the first period $d_{1,p}^*$ decreases in α. The number of pay-per-user users in the second period $d_{2,p}^*$ first increases ($\alpha \leq \alpha_{2,p}$) and then decreases ($\alpha > \alpha_{2,p}$) in α. Furthermore, $d_{1,p}^* = 0$ when $\alpha \geq \alpha_1$ and $d_{2,p}^* = 0$ when $\alpha \geq \alpha_{2,f}$.*

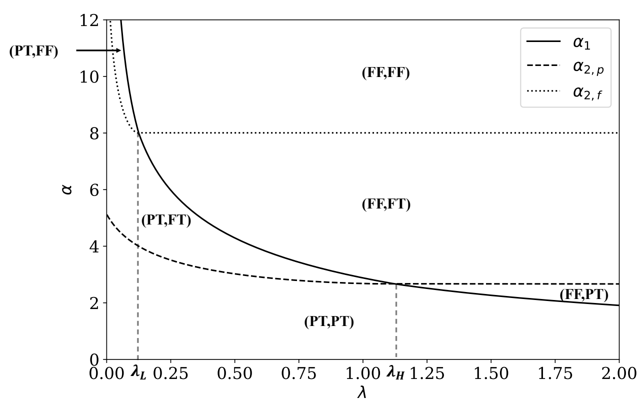
3. *The number of flat-rate users in both period $d_{1,f}^*$ and $d_{2,f}^*$ increases in α. Furthermore, $d_{1,f}^* = 1$ when $\alpha \geq \alpha_1$ and $d_{2,f}^* = 1$ when $\alpha \geq \alpha_{2,f}$.*

In addition, for any $\gamma > 0$, $\alpha_1(\lambda, \gamma), \alpha_{2,p}(\lambda, \gamma)$, and $\alpha_{2,f}(\lambda, \gamma)$ decrease with respect to λ. Furthermore, there exist $\lambda_L(\gamma)$ and $\lambda_H(\gamma)$ such that

$$\begin{cases} \alpha_{2,p}(\lambda, \gamma) < \alpha_{2,f}(\lambda, \gamma) < \alpha_1(\lambda, \gamma), & \lambda < \lambda_L(\gamma), \\ \alpha_{2,p}(\lambda, \gamma) \leq \alpha_1(\lambda, \gamma) \leq \alpha_{2,f}(\lambda, \gamma), & \lambda_L(\gamma) \leq \lambda \leq \lambda_H(\gamma), \\ \alpha_1(\lambda, \gamma) < \alpha_{2,p}(\lambda, \gamma) < \alpha_{2,f}(\lambda, \gamma), & \lambda > \lambda_H(\gamma). \end{cases}$$

Corollary 1 indicates that the discount factor λ determines the order of $\alpha_1, \alpha_{2,p}$, and $\alpha_{2,f}$, which in turn determines how the market segmentation changes over the two periods. Figure 4 uses a linear across-user learning function and sets $\gamma = 1$ as an example to illustrate the results in Corollary 1. For simplicity, we represent the market structures of these two periods using a vector. The first element indicates the market structure of the first period, while the second element represents the market structure of the second period. Specifically, when the discount factor is low, i.e., $\lambda < \lambda_L$, as α increases, the market in the second period initially becomes fully covered, then transitions to a FF market. Eventually, the market in the first period also becomes a FF market as the inclusive enhanced value offered by across-user learning increases. In other words, across-user learning has a more significant effect on the market in the second period compared to its impact in the first period. However, when the discount factor is high, i.e., $\lambda > \lambda_H$, as α increases, the market in the first period initially becomes a FF market. Then, the market in the second period transitions to a FT market and eventually becomes a FF market. Interestingly, when the discount factor is medium, i.e., $\lambda_L \leq \lambda \leq \lambda_H$, as α increases, the market in the second period first becomes fully covered. Then, the market in the first period transitions to a FF market. Later, the market in the second period becomes a FF market.

Figure 4 The Impact of Across-User Learning on Market Dynamics



Note. Across-user learning function $\delta_a = \alpha n_1$ and $\gamma = 1$.

According to Proposition 2, as α increases, the critical valuations $v^*_{1,p}, v^*_{2,p}, v^*_{1,f}$, and $v^*_{2,f}$ all decrease. In addition, Corollary 1 shows that if $\alpha \geq \alpha_1$, then $d^*_1 = 1 \geq d^*_2$ and $d^*_{1,f} = 1 \geq d^*_{2,f}$; if $\alpha \geq \alpha_{2,p}$, then $d^*_2 = 1 \geq d^*_1$, and if $\alpha \geq \alpha_{2,f}$, then $d^*_{2,f} = 1 \geq d^*_{1,f}$. However, when both periods have PT markets, i.e, $\alpha < \min\{\alpha_1, \alpha_{2,p}\}$, the impact of across-user learning on the number of paying users and flat-rate users in these two periods remains unclear. Proposition 3 further analyzes the impact of across-user learning on price and demand across the two periods.

PROPOSITION 3. *Given the presence of across-user learning, i.e., $\alpha > 0$, the following properties hold under equilibrium:*

1. $p^*_{1,0} \leq p^*_{2,0}$ *and* $p^*_{1,f} \leq p^*_{2,f}$.
2. *When* $\alpha < \min\{\alpha_1, \alpha_{2,p}\}$, $d^*_1 \geq d^*_2$ *if and only if*

$$\frac{\lambda}{\gamma}\frac{\partial \delta_a(n^*_1, \alpha)}{\partial n_1} \geq f\left(\delta_a(n^*_1, \alpha)\right), \tag{5}$$

$d^*_{1,p} \leq d^*_{2,p}$, *and* $d^*_{1,f} \geq d^*_{2,f}$ *if and only if*

$$\frac{\lambda}{\gamma}\frac{\partial \delta_a(n^*_1, \alpha)}{\partial n_1} \geq \frac{1}{6} f\left(\delta_a(n^*_1, \alpha)\right) \tag{6}$$

where

$$f\left(\delta_a(n^*_1, \alpha)\right) = \frac{25}{2}\frac{\delta_a(n^*_1, \alpha)}{\left(1+\delta_a(n^*_1, \alpha)\right)^2}$$

increases with respect to $\delta_a(n^*_1, \alpha)$.

Part 1 of Proposition 3 suggests that across-user learning enables the firm to charge higher pay-per-use and flat-rate prices in the second period. Consequently, the firm would increase its number of flat-rate and paying users in the first period by setting lower pay-per-use and flat-rate prices. Part 2 of Proposition 3 implies that across-user learning could result in more paying users (flat-rate users) in the first period than those in the second period when its marginal benefit on the first-period data, $\partial\delta_a/\partial n_1$, is sufficiently large, i.e., Equation (5) (Equation (6)) holds. In other words, the firm offers a low price to gain sufficient data from paying users in the first period and then switches to high pricing in the second period.

We take Tesla's Full Self-Drive (FSD) service as an example. Tesla has implemented a "Shadow Mode" to collect real-world data. Under the Shadow Mode, while the vehicle is being manually driven, the autonomous driving system and sensors remain active but do not engage in vehicle control. Instead, the system algorithms make simulated decisions that are compared with the driver's actions. Discrepancies trigger data feedback. Thus, The more users who use the FSD service and the more frequently they use it, the more data Tesla collects, forming the dataset for training Tesla's FSD model to further enhance the technology. In 2021, Tesla introduced a new subscription pricing for the FSD service at 199 or 99 per month, in addition to the existing flat-rate fee of $10,000$[1]. This monthly payment model aims to attract users who are less inclined to pay the high upfront cost, generating more data for Tesla to improve the service. As the FSD technology matured, Tesla increased the flat-rate price of the FSD to $15,000$ in September 2022, establishing a more appropriate pricing structure for the advanced service[2]. Elon Musk, Tesla's CEO, explained the rationale behind the pricing strategy, stating, “Well, we just wanted to make it more affordable as more people try it. Yes, I think, over time, the price of FSD will increase proportionate to its value. So, with regard [to] the current price as a kind of a temporary low.”[3]

5. The Impact of Within-user Learning

Next, we examine how within-user learning affects the firm’s optimal pricing strategies. In particular, the updated user valuation in the second period v_u in Equation (3) consists of only the within-user learning function, i.e.,

$$v_u(v_o) = v_o + \delta_w\left(q_f^*(v_o), \beta\right)\mathbf{I}_{v_o \geq v_{1,f}}. \tag{7}$$

Recall that, in contrast to across-user learning, within-user learning increases a user’s valuation based on the number of her consumption in the first period. The more a user uses the product in the first period, the greater the exclusive benefits they can derive from within-user learning in the second period. It is important to note that first-period high valuation users tend to consume more units compared to the low valuation users, thereby gaining additional user value in the second period. Without loss of generality, we restrict our attention to a linear within-user learning function and introduce the following assumption.

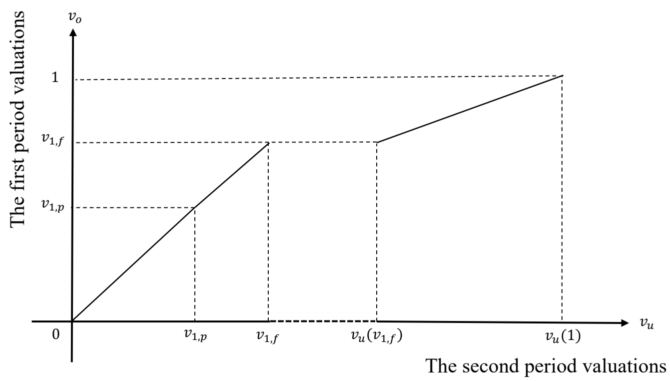
Assumption 2. *The within-user learning function $\delta_w(q,\beta)$ has the following format, i.e., $\delta_w(q,\beta) = \beta q, \beta \geq 0$.*

Assumption 2 indicates that no enhanced value in user valuation if a user does not use the good, $\delta_w(0,\beta) = 0$ or the firm lacks within-user learning capability $\delta_w(q,0) = 0$. In addition, $\delta_w(q,\beta)$ is increasing in q and β, suggesting that the more a user uses the good or the greater capability of within-user learning, the more exclusive benefits within-user learning can provide.

Since within-user learning applies only to the flat-rate users in the first period, the updated user valuation in the second period $v_u(v_o)$ becomes discontinuous at the critical valuation $v_{1,f}$ shown in Figure 5. In particular, the no-purchase users and pay-per-use users in the first period gain no benefit from within-user learning, while the flat-rate users gain an additional exclusive benefit from within-user learning depending on how often they use the good. Lemma 3 characterizes the optimal pricing strategies of the firm in both periods determined by the critical valuations under within-user learning.

Lemma 3. *The critical valuations $(v_{1,p}^*, v_{1,f}^*)$ in the first period exist and $v_{1,p}^*, v_{1,f}^* \in [\frac{2}{5}, \frac{4}{5}]$. In the second period, the critical valuations $(v_{2,p}^*, v_{2,f}^*)$ also exist. In addition, $v_u^{-1}(v_{2,p}^*) \in [v_{1,p}^*, v_{1,f}^*]$ and $v_u^{-1}(v_{2,f}^*) \in [v_{1,f}^*, 1]$.*

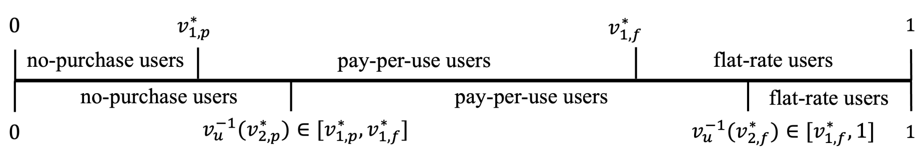
Figure 5 The Valuation Change over Both Periods under Within-user Learning



Note. $v_u(v_o)$ is discontinuous at the critical valuation $v_{1,f}$.

The findings of Lemma 3 are shown in Figure 6. Interestingly, on the one hand, the number of flat-rate users becomes smaller in the second period compared to that of the first period, i.e., $v_u^{-1}(v_{2,f}^*) \in [v_{1,f}^*, 1]$. This suggests that the flat-rate users from the first period are further segmented into pay-per-use and flat-rate users in the second period. This segmentation arises due to the uneven impact of within-user learning on flat-rate users: those who consume the good less frequently in the first period experience a smaller increase in valuation in the second period. Consequently, significant differences in valuation emerge among flat-rate users as they transit into the second period. Segmenting these users in the second period proves more profitable, as a single flat-rate option would fail to accommodate the substantial valuation discrepancies present in the market. Conversely, $v_u^{-1}(v_{2,p}^*)$ consistently falls between $v_{1,p}^*$ and $v_{1,f}^*$, indicating that pay-per-use users from the first period are consistently divided into no-purchase users and pay-per-use users in the second period.

Figure 6 Critical Valuations and Market Dynamics over Both Periods



Lemma 3 implies that compared to across-user learning, firms are incentivized to serve more flat-rate users in the first period under within-user learning, capitalizing on the benefits of frequent engagement. The combination of increased user engagement and the cumulative data generated contributes to a stronger data network effect. Specifically, the flat-rate bias under within-user learning can significantly amplify the data network effect by encouraging users to engage more frequently. This increased engagement not only enhances the data collection process but also drives higher profitability by maximizing user interaction.

Recall that, parameter γ indicates the speed of the marginal utility decreases with increased usage and β represents unit enhanced value from within-user learning. We denote by $\tilde{\beta} = \beta/\gamma$ the relative capability of within-user learning. The following proposition shows how market dynamics varies with the relative capability of implementing within-user learning $\tilde{\beta}$ and discount factor γ.

PROPOSITION 4. *Given within-user learning function $\delta_w(q,\beta)$ satisfying Assumption 2, the optimal critical valuations $v^*_{1,p}(\tilde{\beta},\lambda), v^*_{1,f}(\tilde{\beta},\lambda)$, $v_u^{-1}(v^*_{2,p}(\tilde{\beta},\lambda))$, and $v_u^{-1}(v^*_{2,f}(\tilde{\beta},\lambda))$ exist and satisfy the following properties.*

1. $v^*_{1,p}(\tilde{\beta},\lambda)$ is increasing in $\tilde{\beta}$ and λ; $v^*_{1,f}(\tilde{\beta},\lambda)$ is decreasing in $\tilde{\beta}$ and λ. In addition, for any λ, $\lim_{\tilde{\beta}\to\infty} v^*_{1,p}(\tilde{\beta},\lambda) > \lim_{\tilde{\beta}\to\infty} v^*_{1,f}(\tilde{\beta},\lambda) = 2/5$.

2. $v_u^{-1}(v^*_{2,p}(\tilde{\beta},\lambda))$ and $v_u^{-1}(v^*_{2,f}(\tilde{\beta},\lambda))$ are discontinuous in $\tilde{\beta}$ and λ. In addition, for any λ, $\lim_{\tilde{\beta}\to\infty} v_u^{-1}(v^*_{2,p}(\tilde{\beta},\lambda)) = 2/5$ and $\lim_{\tilde{\beta}\to\infty} v_u^{-1}(v^*_{2,f}(\tilde{\beta},\lambda)) = 4/5$. Specifically, for any λ, there exist $\tilde{\beta}_1(\lambda)$ and $\tilde{\beta}_2(\lambda)$ that are decreasing in λ and present three possible market structures, described as follows:

i Partially covered tiered market ($0 \le \tilde{\beta} < \tilde{\beta}_1$): $v_u^{-1}(v^*_{2,p})$ is increasing in $\tilde{\beta}$ and $v_u^{-1}(v^*_{2,f}) = 4/5$.

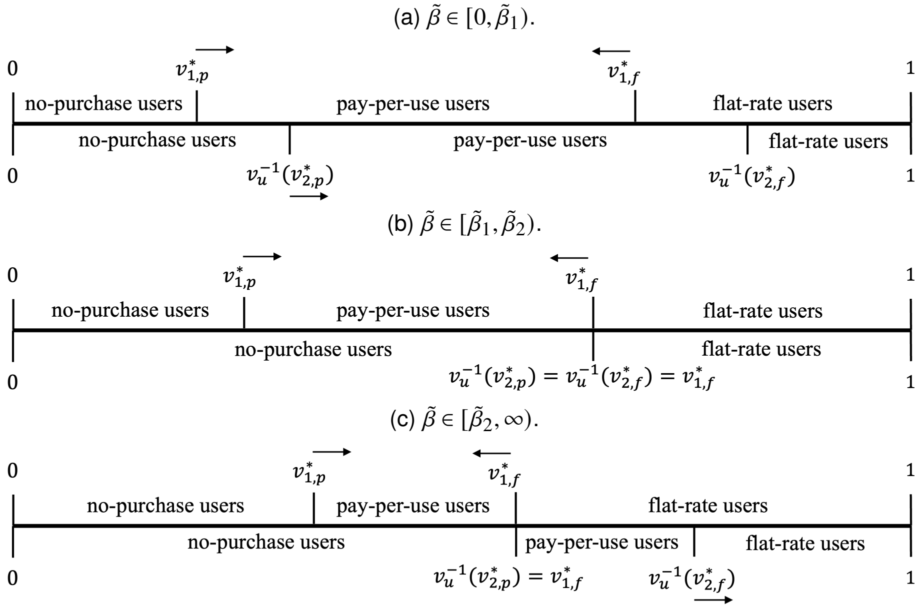
ii Partially covered flat-rate market ($\tilde{\beta}_1 \le \tilde{\beta} < \tilde{\beta}_2$): $v_u^{-1}(v^*_{2,p}) = v_u^{-1}(v^*_{2,f}) = v^*_{1,f}$ and they are decreasing in $\tilde{\beta}$.

iii Partially covered tiered market ($\tilde{\beta} \ge \tilde{\beta}_2$): $v_u^{-1}(v^*_{2,p}) = v^*_{1,f}$ is decreasing in $\tilde{\beta}$ and $v_u^{-1}(v^*_{2,f}) = 1 - v_u^{-1}(v^*_{1,f})/2$ is increasing in $\tilde{\beta}$.

For illustrative purpose, Figures 7 summarizes the results presented in Proposition 4.

First, Proposition 4 indicates that the optimal critical valuations depend on the relative within-user learning capability $\tilde{\beta}$, and the discount factor λ. Similar to across-user learning, a larger discount factor or a higher within-user learning capability increases the significance of the second period in the overall profit, thereby affecting the firm's pricing strategy in a similar way. Part 1) of Proposition 4 suggests that as the relative within-user learning capability improves, the critical valuation $v^*_{1,f}$ decreases, indicating a lower flat-rate price. This occurs because within-user learning enhances the valuation of flat-rate users, prompting the firm to offer a more attractive flat-rate option in the first period to encourage more users to choose it in the second period. This strategy would lead to higher user consumption in the second period and consequently increase profits. The higher the within-user learning capability, the more significant the enhancement in user valuation for the second period, resulting in a greater reduction in $v^*_{1,f}$. The increase in $v^*_{1,p}$ is a consequence of the decrease in $v^*_{1,f}$. This is because the firm can attract more users to select the flat-rate option in the first period (i.e., lower $v^*_{1,f}$) by offering a lower flat-rate price $p^*_{1,f}$ or a higher per-unit price $p^*_{1,0}$ (a higher $p^*_{1,0}$

Figure 7 The Changes of Optimal Critical Valuations as Within-user Learning Capability $\tilde{\beta}$ Increases



makes the pay-per-use option more expensive, converting some users into flat-rate customers). However, a higher $p_{1,0}^*$ leads to an increase in $v_{1,p}^*$.

Interestingly, Part 2) of Proposition 4 presents a cyclic change in market structure in the second period as the relative capability of within-user learning increases. This cyclic change is caused by the increment in user valuation of the first period flat-rate users due to within-user learning, and the increment in user valuation is triggered by the flat-rate bias, which strengthens the data network effect under within-user learning. Specifically, the market initially transitions from a PT market to a PF market, reflecting a greater impact of the flat-rate option in efficiently improving profits. However, as the relative capability of within-user learning continues to increase, the role of flat-rate bias in amplifying the data network effect becomes increasingly significant, leading to a greater number of flat-rate users being served in the first period. Consequently, this enhanced effect prompts a shift in the market structure during the second period, causing a reversion back to a PT market. This cyclical transition highlights the complex interplay between within-user learning and pricing strategies.

When the relative within-user learning capability is sufficiently low ($\tilde{\beta} \in [0, \tilde{\beta}_1)$), the initial second-period PT market is shown in Figure 7a. The enhancement in learning capability increases the valuation difference between the pay-per-use and flat-rate users in the first period due to flat-rate bias. However, the enhanced value is not significant and the firm can only monetize a subset of the first-period flat-rate users by the flat-rate option. Consequently, $v_u^{-1}(v_{2,f}^*)$ is always greater than $v_{1,f}^*$, which indicates a partially covered tiered market in the second period. As within-user learning enhances valuations on the first-period flat-rate users and a subset of them choose the pay-per-use option, this leads to a larger pay-per-use price in the second period, i.e., critical valuation $v_u^{-1}(v_{2,p}^*)$ increases with $\tilde{\beta}$.

Market structure undergoes a shift as $\tilde{\beta}$ increases from $\tilde{\beta}_1$ to $\tilde{\beta}_2$, entering a PF market (Figure 7b), i.e., $v_u^{-1}(v_{2,p}^*) = v_{1,f}^*$. In the first period, the firm serves a partially covered tiered market where pay-per-use users contribute to improved profits. Moderate within-user learning enhances user valuations of first-period flat-rate users, allowing the firm to focus exclusively on serving these users in the second period and transition to a flat-rate market.

If the relative capability of within-user learning $\tilde{\beta}$ is sufficiently high, i.e, $\tilde{\beta} \geq \tilde{\beta}_2$, the second-period market turns back to a PT market shown in Figure 7c. In the second period, the firm serves only the first-period flat-rate users but segments them into pay-per-use and flat-rate users. As the relative capability of within-user learning becomes higher, the firm serves fewer paying users but a greater proportion of flat-rate users in the first period, which encourages more data accumulation by converting more pay-per-use users to flat-rate users. As a result, individual data accumulation through flat-rate bias decreases, leading to an expansion of first-period flat-rate users. This strategy allows the firm to maximize the profits in the second period by segmenting these users into pay-per-use and flat-rate users.

Interestingly, as $\tilde{\beta}$ approaches infinity, $v_{1,p}^* > v_{1,f}^* = v_u^{-1}(v_{2,p}^*)$ converges to $2/5$, and $v_u^{-1}(v_{2,f}^*)$ converges to $4/5$. This indicates that even with sufficiently high learning capability, it is impossible to serve the entire market in both periods. This is because within-user learning focuses on individual data accumulation and amplifies customer heterogeneity through data network effects. This contrasts fundamentally with across-user learning, which provides inclusive benefits to all users.

Based on Proposition 4, we can further deduce the changes in the optimal market structures over the two periods as follows.

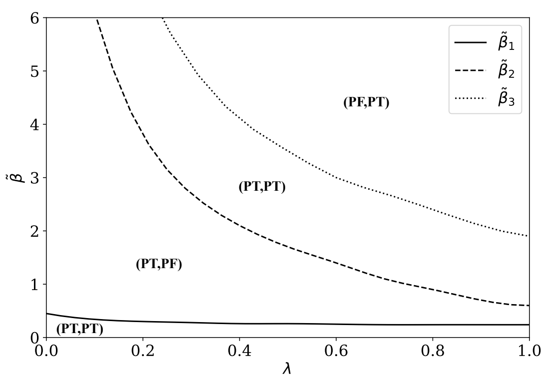
COROLLARY 2. *Given within-user learning function $\delta_w(q,\beta)$ satisfying Assumption 2, the optimal market structures in both periods can be explicitly characterized as,*

$$s_1^* = \begin{cases} PT, & 0 \leq \tilde{\beta} < \tilde{\beta}_3, \\ PF, & \tilde{\beta} \geq \tilde{\beta}_3. \end{cases} \quad s_2^* = \begin{cases} PT, & 0 \leq \tilde{\beta} < \tilde{\beta}_1, \\ PF, & \tilde{\beta}_1 \leq \tilde{\beta} < \tilde{\beta}_2, \\ PT, & \tilde{\beta} \geq \tilde{\beta}_2. \end{cases}$$

In addition, thresholds $\tilde{\beta}_1(\lambda), \tilde{\beta}_2(\lambda)$ and $\tilde{\beta}_3(\lambda)$ decrease with respect to λ and satisfies $\tilde{\beta}_1(\lambda) \leq \tilde{\beta}_2(\lambda) \leq \tilde{\beta}_3(\lambda)$.

As shown in Figure 8, Corollary 2 illustrates the evolution of market dynamics in response to changes in within-user learning capability or the discount factor. When both the within-user learning capability and the discount factor are high, within-user learning can substantially influence overall profits through data network effects and second-period profits. In this scenario, the firm initially uses a single flat-rate pricing strategy for the first-period market, leveraging flat-rate bias to accumulate data and enhance valuation. In the second period, the firm maximizes overall profits by segmenting paying users and offering both pay-per-use and flat-rate options. Conversely, when both within-user learning capability and the discount factor are low, the impact of within-user learning is limited, requiring the firm to utilize both pricing options in both periods to enhance overall profits. In cases where these factors are moderate, or when one is high and the other low, the flat-rate option tends to dominate the pay-per-use option in the second period, as it generates higher profits from high-valuation users. Additionally, as either within-user learning capability or the discount factor increases, the first-period market shifts from a PT to a PF market. Similarly, the second-period market transitions from a PT to a PF market, before reverting back to a partially covered tiered market.

Figure 8 The Impact of Within-User Learning on Market Dynamics



Finally, we summarize the impact of within-user learning in the following Proposition.

PROPOSITION 5. *Given the presence of within-user learning, i.e., $\beta > 0$, the following properties hold under equilibrium.*

1. $p^*_{1,0} \le p^*_{2,0}$ *and* $p^*_{1,f} \le p^*_{2,f}$.
2. $d^*_1 \ge d^*_2$ *and* $d^*_{1,f} \ge d^*_{2,f}$.

Similar to across-user learning, Part 1 of Proposition 5 suggests that within-user learning also enables the firm to gain greater market power, allowing it to charge higher pay-per-use and flat-rate prices in the second period. Interestingly, Part 2 of Proposition 5 implies that within-user learning may lead to a reduction in the number of paying users and pay-per-use users in the second period compared to the first period. This occurs because the personalized valuation enhancements brought about by within-user learning tend to increase the valuation differences among users. As a result, the firm needs to further differentiate users in the first period through appropriate pricing strategies.

When the firm applies within-user learning, it prefers the flat-rate option in the first period. This is because, compared to the pay-per-use option, the flat-rate option is better at boosting user engagement in the first period and hence increases product valuation in the second period. However, whether to use the flat-rate option or the pay-per-use option in the second period depends highly on the firm's within-user learning capability. In particular, Proposition 4 implies that if the within-user learning capability is moderate, the firm would only serve flat-rate users in the second period. For example, Peloton, a firm that combines exercise equipment with streaming technology and offers a subscription service that grants access to a wide range of classes, adopts a within-user learning strategy that tailors class recommendations and progress reports to better match the users' specific fitness goals. Peloton's within-user learning capability is moderate, and hence it charges a fixed monthly fee after a one-month free trial, reflecting a flat-rate pricing strategy. Specifically, it offers a one-month trial of its "App One" plan, allowing users a limited number of classes. This trial encourages user engagement and facilitates data collection. Following the trial period, users are charged a fixed monthly fee. In contrast, if the firm's within-user learning capability is high, the heterogeneity of the customer base is significant. In this case, introducing the pay-per-use option in the second period can better accommodate heterogeneous preferences. For example, Cresta, a generative AI startup that helps firms enhance customer experiences, adopts a within-user learning strategy that collects input from its customers over time and refines its service to better fit the context of each of its customers. Cresta's within-user learning capability is high, and hence it initially charges a flat-rate fee per user, but later experiments with shifting to a per-interaction pricing model—charging for each conversation supported by its AI agent[4].

6. The Impact of Data-enabled Learning

In this section, we revisit the general model with both across-user and within-user learning where the updated valuation $v_u(\cdot)$ is updated according to Equation (3). We compare the impact of across-user learning and within-user learning on the optimal pricing strategies in Section 6.1, and then examine the effect of data-enabled learning on the optimal pricing strategies in Section 6.2.

6.1. Comparison between Across-user Learning and Within-user Learning

Table 1 summarizes how the optimal pricing strategies, critical valuations, and market dynamics change in across- and within-user learning capabilities. Note that across-user learning enhances each user's valuation

Table 1 Summary of Results: Enhancing Across-user Learning Capability α and Within-user Learning Capability β

	Period	Across-user Learning		Within-user Learning	
Prices	1	$p^*_{1,0}\downarrow$	$p^*_{1,f}\downarrow$	$p^*_{1,p}\uparrow$	$p^*_{1,f}\downarrow$
	2	$p^*_{2,0}\uparrow$	$p^*_{2,f}\uparrow$	$p^*_{2,0}\uparrow$	$p^*_{2,f}\uparrow$
Critical Valuations	1	$v^*_{1,p}\downarrow$	$v^*_{1,f}\downarrow$	$v^*_{1,p}\uparrow$	$v^*_{1,f}\downarrow$
	2	$v_u^{-1}(v^*_{2,p})\downarrow$	$v_u^{-1}(v^*_{2,f})\downarrow$	$v_u^{-1}(v^*_{2,p})\uparrow,\downarrow$	$v_u^{-1}(v^*_{2,f})\downarrow,\uparrow$
Market Dynamics	1	$PT\to FF$		$PT\to PF$	
	2	$PT\to FT\to FF$		$PT\to PF\to PT$	

Note. $\uparrow$: increase; $\downarrow$: decrease; $\rightarrow$: transition from left to right.

via the aggregated data of all users. As the across-user learning capability increases, the firm serves more paying and flat-rate users in the first period by lowering the first period prices, which strengthens the data network effect. The increased user's value in the second period due to the network effect allows the firm to increase the second period prices and hence increase its profit in the second period. In this case, across-user learning leverages the valuation of each user and hence reduces the users' heterogeneity. Among a group of users with low heterogeneity, the flat-rate option would result in more profit. However, within-user learning provides exclusive benefits to each individual user. The firm serves fewer paying users but more flat-rate users in the first period, which strengthens the data network effect. Therefore, within-user learning strengthens the heterogeneity of the customer base and the firm can only partially cover the market in both periods. The market structure remains tiered when the within-user learning capability is sufficiently weak or strong.

6.2. Optimal Pricing Strategies with Both Across-user and Within-user Learning

In what follows, we first show the optimal pricing strategies when the across-user learning capability is sufficiently large, i.e., $\alpha \geq \alpha_1(\lambda,\gamma)$, and then we conduct a numerical analysis when $\alpha < \alpha_1(\lambda,\gamma)$.

PROPOSITION 6. *For any $\lambda \in (0,1]$ and $\gamma > 0$, given an across-user learning function $\delta_a(n_1,\alpha)$ with $\alpha \geq \alpha_1(\lambda,\gamma)$ that satisfies Assumption 1, and a within-user learning function $\delta_w(q,\beta)$ with $\beta \geq 0$ that satisfies Assumption 2, the optimal critical valuations $v^*_{1,p}$, $v^*_{1,f}$, $v^*_{2,p}$, and $v^*_{2,f}$ exist and can be characterized as follows: $(v^*_{1,p}, v^*_{1,f}) = (0,0)$ and*

$$\left(v_u^{-1}(v^*_{2,p}), v_u^{-1}(v^*_{2,f})\right) = \begin{cases} (\frac{2}{5}-\frac{3h}{5}, \frac{4}{5}-\frac{h}{5}), & 0 \leq h < \frac{2}{3}, \\ (0, 1-\frac{h}{2}), & \frac{2}{3} \leq h < 2, \\ (0,0), & h \geq 2, \end{cases}$$

where $\Delta_a = \delta_a(1/4\gamma,\alpha)$, $\Delta_w = \delta_w(1/2\gamma,\beta)$, and $h = \Delta_a/(1+\Delta_w)$. Furthermore, the optimal prices are given by, $(p^*_{1,0}, p^*_{1,f}) = (0,0)$ and

$$(p_{2,0}^*, p_{2,f}^*) = \begin{cases} (\frac{2}{5}(1+\Delta_a+\Delta_w), \frac{3}{25\gamma}(1+\Delta_a+\Delta_w)^2), & 0 \le h < \frac{2}{3}, \\ (\Delta_a, \frac{1}{2\gamma}\Delta_a(1+\Delta_w)), & \frac{2}{3} \le h < 2, \\ (\Delta_a, \frac{1}{4\gamma}\Delta_a^2), & h \ge 2. \end{cases}$$

And the firm's optimal total discounted profit is

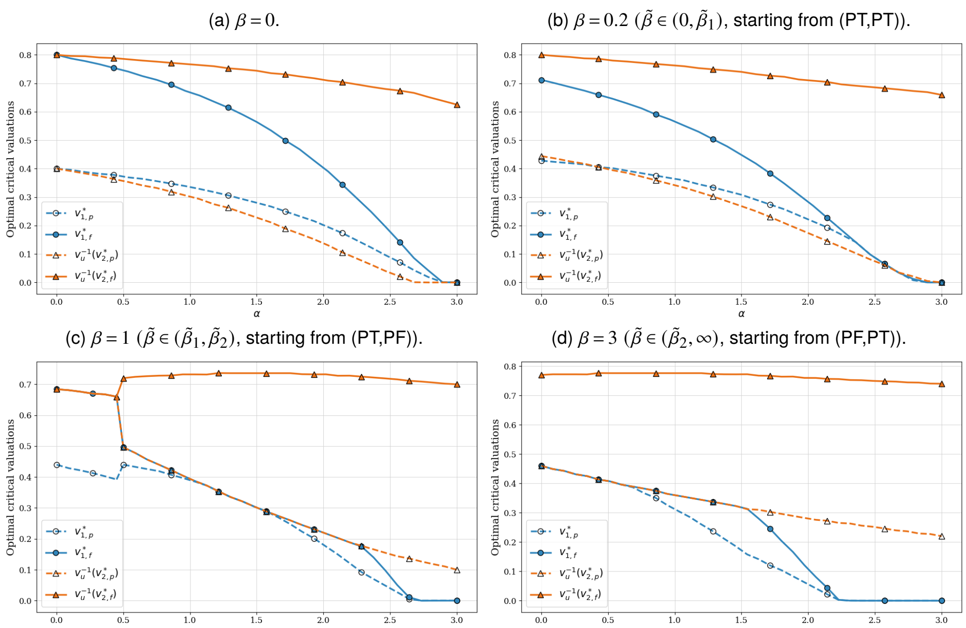
$$\Pi_1^* = \begin{cases} \frac{\lambda}{25\gamma}\frac{(1+\Delta_a+\Delta_w)^3}{1+\Delta_w}, & 0 \le h < \frac{2}{3}, \\ \frac{\lambda}{16\gamma}\Delta_a[4(1+\Delta_w) + \frac{\Delta_a^2}{1+\Delta_w}], & \frac{2}{3} \le h < 2, \\ \frac{\lambda}{4\gamma}\Delta_a^2, & h \ge 2. \end{cases}$$

Proposition 6 reveals that, as the across-user learning capability is sufficiently large, i.e., $\alpha > \alpha_1$, it is optimal for the firm to earn zero profit in the first period ($v_{1,p}^* = v_{1,f}^* = 0$) by serving a FF market. The optimal critical valuations in the second period is closely related to the ratio of the across-user learning capability to the within-user learning capability measured by $h = \Delta_a/(1+\Delta_w)$. Note that, as the across-user learning capability α increases, h becomes larger and both $v_u^{-1}(v_{2,p}^*)$ and $v_u^{-1}(v_{2,f}^*)$ decrease, eventually fully covering a flat-rate market in the second period. Conversely, as the within-user learning capability β increases, h becomes smaller and both $v_u^{-1}(v_{2,p}^*)$ and $v_u^{-1}(v_{2,f}^*)$ increase, eventually partially covering a tiered market. These findings are consistent with the results in Sections 4 and 5. Across-user learning enhances valuation in a user-independent manner. Therefore, a larger α leads to a more homogeneous customer base in the second period, making the market more suitable for a uniform flat-rate pricing scheme. In contrast, within-user learning enhances valuation in a user-dependent manner. As a result, a larger β leads to a more heterogeneous customer base in the second period, making the market more suitable for segmentation, i.e., offering different payment options. Interestingly, $h = v_u(0)/(v_u(1) - v_u(0))$ also indicates the degree of market heterogeneity.

When the across-user learning capability is low, i.e., $\alpha < \alpha_1$, it is optimal for the firm to achieve a positive profit in the first period. As it is challenging to derive the analytical results of the optimal critical valuations, we conduct a numerical analysis to investigate the effects of data-enabled learning capabilities α and β on the optimal critical valuations by assuming $\lambda = \gamma = 1$ and $\delta_a = \alpha n_1$ for simplicity.

Figure 9 shows how the optimal critical valuations change when the across-user learning capability α increases given different values of the within-user learning capability β. Figures 9a and 9b reveal that, when the within-user learning capability is low, the impact of across-user learning dominates that of within-user learning, leading to smaller critical valuations as across-user learning capability increases. More paying users and pay-per-use users are served in the second period than those in the first period. While, more flat-rate users are served in the first period than that in the second period. However, Figures 9c and 9d show that when the within-user learning capability is greater than a certain value, the first-period market is fully covered by the flat-rate users but the second-period market is partially covered with multiple consumer segments when the across-user learning capability is sufficiently high. Interestingly, as the across-user learning capability α increases, both $v_{1,p}^*$ and $v_u^{-1}(v_{2,f}^*)$ are not monotone. In addition, the number of flat-rate users in the first period is more than that of paying users in the second period as the impact of the within-user learning becomes significant.

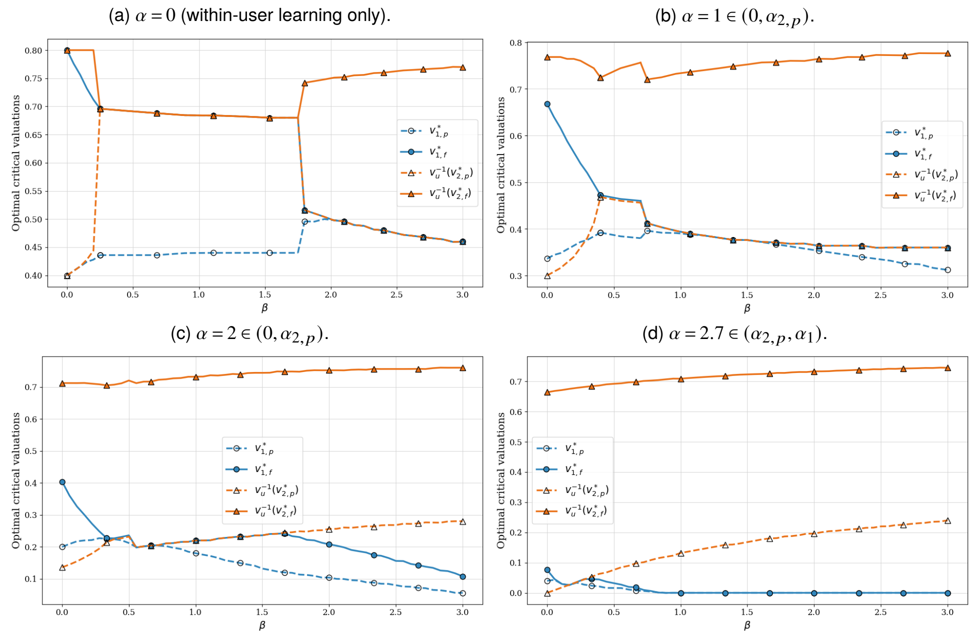
Figure 9 **The Changes of Optimal Critical Valuations as α Increases**



Note. $\lambda = \gamma = 1$ and $\delta_a = \alpha n_1$.

We next examine how the optimal critical valuations change when the within-user learning capability β increases under different values of the across-user learning capability α. Figure 10 shows that, in the first period, there is a larger number of paying and flat-rate users due to the enhanced across-user learning capability. However, optimal critical valuations $v_{1,p}^*$ and $v_{1,f}^*$ do not exhibit a monotonic trend, such as the market structure in the first period may even shift back to a PT market, as the within-user learning capability improves, particularly when the across-user learning capability is high. In the second period, as the across-user learning capability improves, the value for each user increases, leading to the disappearance

Figure 10 The Changes of Optimal Critical Valuations as β Increases



Note. $\lambda = \gamma = 1$.

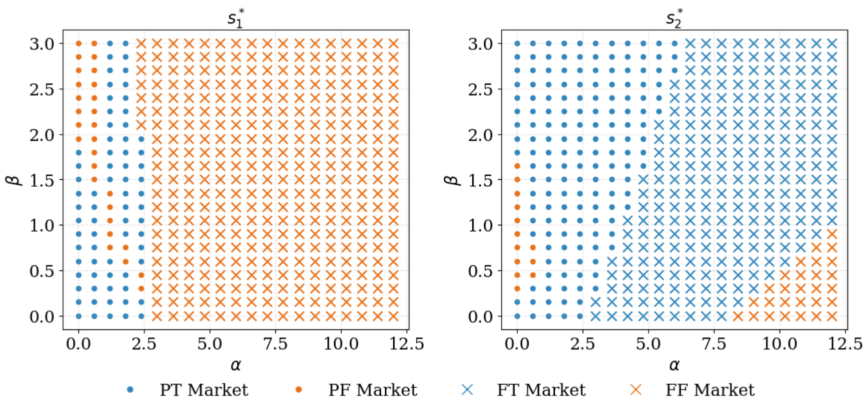
of the PF market. Interestingly, when the across-user learning capability reaches a sufficiently high level, the number of paying and flat-rate users declines because higher prices can be charged.

Figure 11 further illustrates the market structures in both periods across different levels of data-enabled learning capabilities. The influences of both across-user and within-user learning significantly enhance the changes in market structure, contrasting with the effects observed under single learning scenarios. In the first period, if the across-user learning capability is sufficiently high, the market operates as a FF market. Conversely, when the across-user learning capability is lower, the market structure starts at a PT market, shifts to a PF market, and then returns to a PT market as the within-user learning capability increases. In the second period, when both within-user and across-user learning are substantial, within-user learning contributes to greater market heterogeneity, while across-user learning enhances the overall benefits for all users. As a result, the market takes on a tiered structure with multiple user segments. The determination of whether the market is fully covered hinges on whether the across-user learning is relatively greater than the within-user learning. When the ratio of within-user learning to across-user learning falls below a certain threshold, the market becomes a FF market due to its high level of homogeneity.

7. Conclusions

In this paper, we analyze a monopolistic firm that sells a good over two periods. The firm provides both pay-per-use and flat-rate payment options during each period. In the first period, the firm gathers and accumulates

Figure 11 Optimal Market Structures in Both Periods for Various Data-enabled Learning Capabilities



user data through its pricing strategies. This data-enabled learning process allows the firm to leverage the accumulated user data to enhance user valuation in the second period. Under across-user learning, each user has the same increase in valuation, with the magnitude of this increase depending on the amount of user data collected in the first period. Conversely, in within-user learning, the more frequently a user uses the good in the first period, the more personal data is collected, resulting in a greater increase in her valuation. Users aim to choose the best purchasing option from no-purchase, pay-per-use, and flat-rate options and consuming units in each period. The firm aims to optimize pricing strategies.

We investigate the differences between the learning between users and within users and explore how optimal pricing strategies can enhance data-enabled learning, a critical component of data strategy, to influence the user base and market dynamics over time. In particular, across-user learning increases the homogeneity of the user base by providing inclusive benefits to all users. This underscores the importance of data collection in the first period, as it allows the firm to collect valuable insights that can be utilized to enhance user experiences. In the second period, this accumulated data facilitates profit maximization strategies, enabling the firm to leverage the insights gained to optimize pricing and offerings for a more cohesive user experience. As a result, the firm prefers the flat-rate option in the first period as it increases user engagement and thus expands the data pool, and favors the flat-rate option in the second period to maximize its profits. However, within-user learning provides exclusive benefits to each individual user. The firm serves fewer paying users but more flat-rate users in the first period, capitalizing on the benefits of frequent engagement further strengthened by the flat-rate bias. The complex interplay between within-user learning and pricing strategies leads to a cyclic change in market structure in the second period as the relative capability of within-user learning increases.

Notes

[1]https://electrek.co/2021/07/16/tesla-launches-full-self-driving-subscription-package-199-per-\
\month/.

[2]https://www.notateslaapp.com/tesla-reference/958/tesla-fsd-price-increase-history.

[3]https://www.teslarati.com/tesla-ceo-elon-musk-explains-strategy-behind-full-self-driving-price-drop/.

[4]https://www.theinformation.com/articles/ai-startups-are-trying-to-make-usage-based-pricing-happen?rc=6cz8j7.

References

Balasubramanian S, Bhattacharya S, Krishnan VV (2015) Pricing information goods: A strategic analysis of the selling and pay-per-use mechanisms. *Marketing Sci.* 34(2):218–234.

Chien HK, Chu CYC (2008) Sale or lease? Durable-goods monopoly with network effects. *Marketing Sci.* 27(6):1012–1019.

Chun SY, Ovchinnikov A (2019) Strategic consumers, revenue management, and the design of loyalty programs. *Management Sci.* 65(9):3969–3987.

Essegaier S, Gupta S, Zhang ZJ (2002) Pricing access services. *Marketing Sci.* 21(2):139–159.

Fudenberg D, Tirole J (2000) Customer poaching and brand switching. *RAND J. Econom.* 31(4):634–657.

Fudenberg D, Villas-Boas JM (2006) Behavior-based price discrimination and customer recognition. *Handbook of Economics and Information Systems* (Elsevier), 377-436.

Gregory RW, Henfridsson O, Kaganer E, Kyriakou H (2021) The role of artificial intelligence and data network effects for creating user value. *Acad. Management Rev.* 46(3):534–551.

Gurkan H, de Véricourt F (2022) Contracting, pricing, and data collection under the AI flywheel effect. *Management Sci.* 68(12):8791–8808.

Haftor DM, Costa Climent R, Lundström JE (2021) How machine learning activates data network effects in business models: Theory advancement through an industrial case of promoting ecological sustainability. *J. Bus. Res.* 131:196–205.

Hagiu A, Wright J (2023) Data-enabled learning, network effects, and competitive advantage. *RAND J. Econom.* 54(4):638–667.

Hajihashemi B, Sayedi A, Shulman JD (2022) The perils of personalized pricing with network effects. *Marketing Sci.* 41(3):477–500.

Jiang B, Chen PY, Mukhopadhyay T (2015) Software licensing: Pay-per-use versus perpetual. Preprint, submitted April 21, https://ssrn.com/abstract=1088570.

Jing B (2011) Exogenous learning, seller-induced learning, and marketing of durable goods. *Management Sci.* 57(10):1788–1801.

Katz M, Shapiro C (1985) Network externalities, competition, and compatibility. *Amer. Econom. Rev.* 75(3):424–440.

Kuksov D, Zia M (2021) Benefits of customer loyalty in markets with endogenous search costs. *Management Sci.* 67(4):2171–2190.

Lambrecht A, Skiera B (2006) Paying too much and being happy about it: Existence, causes, and consequences of tariff-choice biases. *J. Marketing Res.* 43(2):212–223.

Levine SS, Jain D (2023) How network effects make AI smarter. *Harvard Bus. Rev.* (March 14), https://hbr.org/2023/03/how-network-effects-make-ai-smarter.

Li X, Li KJ, Wang XS (2020) Transparency of behavior-based pricing. *J. Marketing Res.* 57(1):78–99.

Liu Y (2007) The long-term impact of loyalty programs on consumer purchase behavior and loyalty. *J. Marketing* 71(4):19–35.

Liu Y, Luo R (2023) Network effects and multinetwork sellers' dynamic pricing in the U.S. smartphone market. *Management Sci.* 69(6):3297–3318.

Ma Q, Shou B, Huang J, Başar T (2021) Monopoly pricing with participation-dependent social learning about quality of service. *Production Oper. Management* 30(11):4004–4022.

Sundararajan A (2004) Nonlinear pricing of information goods. *Management Sci.* 50(12):1660–1673.

Trautman E (2018) The virtuous cycle of AI products. Accessed January 27, 2020, https://www.eriktrautman.com/posts/\\the-virtuous-cycle-of-ai-products.

Varian HR (1989) Price discrimination. *Handbook of Industrial Organization* (Elsevier), 597-654.