Cyclic Pricing When Customers Queue with Rating Information

Fengfeng Huang

School of Management and Economics, University of Electronic Science and Technology of China, China, ff.huang@connect.polyu.hk

Pengfei Guo

Faculty of Business, The Hong Kong Polytechnic University, Hong Kong, pengfei.guo@polyu.edu.hk

Yulan Wang*[1]

Faculty of Business, The Hong Kong Polytechnic University, Hong Kong, yulan.wang@polyu.edu.hk

Consider a situation where a service provider serves two types of customers, sophisticated and naive. Sophisticated customers are well-informed of service-related information and make their joining-or-balking decisions strategically, whereas naive customers do not have such information and rely on online rating information to make such decisions. We demonstrate that under certain conditions a service provider can increase its profitability by simply 'dancing' its price, that is, replacing the static pricing strategy with a high-low cyclic pricing strategy. The success of this strategy relies on two key conditions: the potential market size is large enough so that congestion is a key concern in the service system, and the rating provides the average price and average utility information. Finally, we show that the cyclic pricing strategy is not socially optimal.

Keywords: Queueing strategy; unobservable queue; pricing; customer rating; game theory

[1]For correspondence.

1 Introduction

When customers seek service, they make their joining-or-balking decisions based on the information they have access to. Customers have varying accesses to information; some are aware of service-related information such as service quality and service rate, whereas others are not. Instead of making blind decisions, customers who are not aware of service-related information often rely on buyer-generated information, such as customer ratings and reviews online, to make their decisions. In fact, buyer-generated information has reshaped customers' habits. Checking ratings/reviews before making consumption decisions has become a ritual for many of today's customers. According to the latest global report on www.pwc.com,[2] 78% of respondents say that social media has influenced their purchase decisions.

The cyclic pricing strategy is often adopted by service providers. For example, in the United States, many theme parks offer discounts from March to October, the high tourist season.[3] Similarly, many tourist cities in China offer discounts on entrance fees for various tourist destinations from July to September. For example, in 2018, the Guizhou Provincial Tourism Bureau published advertisement on several portal websites announcing that a 50% discount of entrance fees would be offered to tourists in several provincial popular tourist destinations.[4]

In this work, we will demonstrate that, the cyclic pricing strategy, together with rating-dependent customers, can achieve a higher profit than the static pricing strategy. To illustrate the mechanism behind this, we consider a monopoly service provider serving two types of customers who are heterogeneous in information access, namely, sophisticated and naive customers. Sophisticated customers are aware of service-related information, such as service rate and service reward, and they make strategic decisions by taking into consideration the joining-and-balking decisions made by others. Naive customers are those one-time shoppers, and they do not have such service-related information. They rely on buyer-generated information to make the decisions; they join if price is not higher than the average rating and balk otherwise. For example, local residents often know the food quality (service reward) and the probable waiting times of their nearby restaurants. On the contrary, tourists generally do not know such information, and they have to rely on reviews/ratings posted on websites

[2]The information can be accessed at `https://www.pwc.com/gx/en/industries/retail-consumer/global-total-retail/global-key-findings.html`.

[3]See `https://www.moneysavingexpert.com/deals/cheap-theme-parks/#longleat`.

[4]See, for example `http://www.sohu.com/a/236603031_395859`.

such as Dianping.com and Yelp.com to decide whether to join or to balk.

The proportion of each type of customer, that is, the customer type composition parameter, is known to the service provider. For example, according to a report on Variety.com,[5] local visitors account for 39% of the total attendance at Hong Kong Disneyland, while tourists (mainland Chinese and international) comprise the other 61% (41% and 20%, respectively, for mainland Chinese and international visitors). The interaction between the service provider and customers is modeled as an $M/M/1$ queue. The queue length is unobservable. We assume that a customer, after obtaining the service, posts a rating equal to her *consumption utility* (service reward minus the waiting cost). The *average rating* on the consumption utility is then advertised to future arriving customers and we assume that incoming customers do not observe the whole history of rating information. We also extend our study to the other information scenario where the service provider reveals the average ratings on both the *net utility* (consumption utility minus price) and the *average price* to incoming customers. We illustrate that the cyclic pricing mechanism still works in this information scenario.

We show that the optimal pricing strategy, if cyclic, must be of high-low type: sophisticated customers join during the high-price phase and obtain an expected net utility of zero, while naive customers join during the low-price phase and obtain a negative expected utility. Recall that a customer's consumption utility is determined by the congestion level of the system. The system is less congested during the high-price phase; hence, the corresponding ratings on the consumption utility are higher. Ratings from both sophisticated and naive customers are averaged, and the average rating is then revealed to the incoming customers. The service provider just needs to charge a price equal to the average rating to lure naive customers into joining. This price, however, turns out to be higher than the naive customers' expected consumption utility. In short, the high-price phase uplifts the average rating, which then allows the service provider to overcharge naive customers (i.e., charging naive customers a price higher than their expected consumption utility) during the low-price phase. Numerical studies show that the profit increment brought by cyclic pricing is around 5%. This incremental amount may not be regarded as high, but it could be particularly important for entertainment parks like Hong Kong Disneyland that are struggling to break even financially.[6]

[5] The details can be found at `http://variety.com/2016/biz/asia/hong-kong-disneyland-in-loss-1201706466/`.

[6] Please refer to http://www.scmp.com/news/hong-kong/economy/article/2133962/hong-kong-

The intriguing aspect about this mechanism of improving profitability with cyclic pricing is that it assumes *honest feedbacks* from customers who have experienced the service, without any distortion of their ratings. This feature makes it easy to be implemented without worrying about the ethical issues. It is worth mentioning that the adoption of the cyclic pricing strategy requires *the potential market size to be above a certain threshold value* so that congestion is a significant factor affecting customers' patronizing decision. A customer's consumption utility is affected by the system's congestion, which, in turn, is determined by the effective arrival rates controlled by the prices. When the potential arrival rate is small or the service capacity is very large, the congestion effect does not play a profound role. Thus, the aforementioned cyclic pricing mechanism does not work; instead, static pricing is preferred. One extreme case is the goods market, in which delay is not an issue and how much a customer enjoys a goods item is determined merely by her idiosyncratic features. In such a case, our "dancing price" strategy does not work.

Furthermore, we consider an alternative information scenario wherein customers post ratings on their net utility. Under such scenario, an extra condition is required for the cyclic pricing strategy to work, that is, the average price shall be provided to the incoming customers. With the knowledge about the average price, the incoming customers can still nail down the rating on the consumption utility, and hence the above cyclic pricing mechanism still works.

Finally, we examine whether the cyclic pricing strategy is socially desired; that is, whether it can maximize the sum of the service provider's profit and customers' surplus. In the classic queueing literature with identical customers and unobservable queues, all customer surplus is internalized through pricing and goes to the service provider. Therefore, welfare- and profit-maximization are equivalent; see Hassin and Haviv (2003). However, in our model, customers differ on their information access. We show that the welfare-maximizing pricing strategy is always static. This indicates that a high-low cyclic pricing strategy is never socially optimal. The key rationale behind is that the cyclic pricing strategy allows the service provider to overcharge naive customers during the low-price phase. This is because when naive customers are lured to join the system by the rating information, they do not consider the negative externality of their joining behavior on others. The system becomes too congested, causing welfare loss. In fact, we show that when the cyclic pricing strategy is adopted, naive customers always obtain a negative expected utility during the low-price

disneyland-falls-further-red-losses-double-2017-hit for the details.

phase, which renders that the service provider's profit is higher than the social welfare it offers. Welfare-maximization prefers even workloads across the time periods, whereas high-low pricing leads to uneven workloads. In most cases, the system is heavily congested during the low-price phase and underutilized during the high-price phase. Interestingly, in certain situations wherein the market size is not large and sophisticated customers comprise no less than half of the population, welfare loss is caused by the underutilization in both phases.

The remainder of the paper is organized as follows. In Section 2, we review the related literature. We introduce basic assumptions and the model setup in Section 3. In Section 4, we analyze the provider's optimal profit- and welfare-maximizing pricing strategies by first considering that customers rate over the consumption utility. We then consider another information scenario where customers rate over the net utility. We compare the pricing strategies and the corresponding equilibrium outcomes in Section 5. Section 6 concludes the paper. All the proofs are relegated to the online Appendix.

2 Literature Review

Our work belongs to the stream of study on customers' strategic queueing behavior and the provider's information disclosure. Pioneering work on this can be traced back to Naor (1969), who studies the equilibrium joining strategies of customers under observable queues and proposes regulating the demand rate by imposing fees or tolls. Edleson and Hildebrand (1975) extend the analysis to an unobservable $M/M/1$ queue. Many studies have been conducted along these lines since then; see Hassin and Haviv (2003) and Hassin (2016) for the comprehensive surveys of studies in this field. In the following, we shall review the most closely related works, which we classify into the following three streams of research: information disclosure in queues, customers' bounded rationality in making queueing decisions, and service pricing.

Studies concerning information disclosure in queues can be further classified according to information content: waiting time, service reward, and service rate. A number of papers consider whether queue length information should be released for free, including Hassin (1986), Whitt (1999), Armony and Maglaras (2004), Dobson and Pinker (2006), Guo and Zipkin (2007), Allon et al. (2011) and some others. Another group of studies consider offering customers the choice of paying a fee to inspect queue length, including Hassin and Haviv (1994), Hassin and Roet-Green (2011, 2018). All these papers assume that customers are strategic and they have the same access to the service information. In our work, however, only a proportion of customers are aware of such information; others simply follow the online rating information.

Our paper is closely related to the work of Hu et al. (2017), who consider the effect of customer information heterogeneity on waiting time. In their model, some customers are informed of a real-time delay (informed customers) and others (uninformed customers) make their joining/balking decisions based on past experiences. They find that a certain level of information heterogeneity leads to more efficient outcomes. Similar to Hu et al. (2017), we consider customers' information heterogeneity. In our model, a proportion of customers (sophisticated customers) are aware of service-related information such as service reward and service rate, whereas others (naive customers) do not have such information. Different from Hu et al. (2017), the queue length under our setting is unobservable to both types of customers. We find that a profit-maximizing provider can make higher profits by taking advantage of information heterogeneity, but it never leads to higher social welfare.

A second stream of literature on information disclosure in queues considers imperfect service reward information; these studies include Veeraraghavan and Debo (2009, 2011), Debo et al. (2012), and Guo et al. (2015). The study by Veeraraghavan and Debo (2009) shows how information externalities due to congestion affect customers' choice between two servers. In their model, customers have private information about service quality and queue length. They find that information externalities lead to cycles during which one server is thriving and the other is not. Veeraraghavan and Debo (2011) study the herding behavior of customers choosing between two congested services with unknown service qualities in which customers observe an imperfect private signal on service qualities and queue lengths before making their choices. They characterize the equilibrium joining behavior and the extent of customer learning from the queue information. Debo et al. (2012) examine customers' queueing strategy when the service reward is known only by a proportion of customers, and find that uninformed customers adopt a "hole" strategy, i.e., they balk when the queue is at a certain length (called the "hole") and behave the same as informed customers otherwise. Guo et al. (2015) examine two competing servers with unknown qualities. They show that under certain conditions, the low-quality server has a higher incentive to reveal queue length. In contrast to these models in which customers are always strategic, only a portion of customers in our model are strategic, with others being naive.

A third stream of work on information disclosure in queues focuses on the information about service rate and arrival rate, including Guo et al. (2011); Debo and Veeraraghavan (2014); Cui and Veeraraghavan (2014), and Afeche and Ata (2011). Guo et al. (2011) and Debo and Veeraraghavan (2014) assume that customers do not know the service rate but have information on its distribution. Cui and Veeraraghavan (2014) consider "blind queues", in which customers only know some vague information about service rate. Afeche and Ata (2011) propose the "learning-and-earning" problem where customers are classified into patient and impatient customers but the proportion of each type is unknown.

Our work is related to studies on the bounded rationality of customers, including Huang et al. (2013), Huang and Chen (2015), and Li et al. (2016). In these studies, customers make their decisions without fully assessing service quality and waiting time due to limited cognitive ability or a lack of opportunities. Huang et al. (2013) demonstrate that strategically taking advantage of the customer bounded rationality may lead to significant increases in both revenue and social welfare. Li et al. (2016) consider customer-intensive services and find that revenue-maximizing firms do not always exploit customers' bounded rationality and may leave positive net utility to customers under certain circumstances. Huang and Chen (2015) consider anecdotal reasoning customers, who rely on a limited sample to make queueing decisions. They find that with anecdotal reasoning, customers are less price-sensitive, and revenue and welfare maximization lead to different pricing strategies. Similar to these models, naive customers in our model are bounded-rational. However, customers in our system consist of both sophisticated and naive customers, whereas most of the above studies just consider a single type of customers.

Our paper is also related to the pricing strategy for service facilities. Dewan and Mendelson (1990) first propose a joint optimization of capacity and pricing. Stidham (1992) shows that even for a simple $M/M/1$ system, the joint pricing and capacity problems have multiple local optima. Numerous studies on capacity investment and admission control have been conducted; see Stidham (2009) for a comprehensive review.

In addition, much research has been conducted on vacation queueing systems, in which a service shuts down when no customers are present and resumes when the queue reaches a critical length (Guo and Hassin, 2011, 2012; Guo and Li, 2013; Guo and Zhang, 2013; Wang and Li, 2008; Zhang et al., 2013; Economou et al., 2011). Similar to these vacation queueing systems, the system in our study oscillates between high- and low-arrival states when the provider adopts a cyclic pricing strategy.

Lastly, our paper relates to the research on consumer ratings/reviews. Numerous researches have explored how rating and review information can affect a firm's pricing and information disclosure strategies in the retailing business. Two of these works are closely relevant to our own work. The first is Crapis et al. (2016), which analyzes the social learning mechanism and its effect on a seller's pricing decision. In their model, customers, after consumption, rate the product as either "like" or "dislike"; later-arriving customers observe the rating profile (the proportion of "like" and "dislike") and infer how much they are likely to enjoy the product. They compare two pricing strategies– a static price and one with a single price change– and suggest that pricing policies that account for social learning may increase revenues. The second is Shin and Zeevi (2017), which studies a fluid model and investigates a monopolist's optimal dynamic pricing strategy over a finite horizon. In both models, customers have private information about their preferences, and the demand function evolves in conjunction with the review profile/dynamics. As far as we know, our work is one of the first papers on customer ratings in a service/queueing context.

3 Model Setup and Preliminaries

Consider a monopoly service provider (he) whose service times are i.i.d and exponentially distributed with rate μ. Potential customers arrive according to a Poisson process with rate Λ. Once served, a customer (she) receives a service reward R, and incurs a price p and a waiting cost that is proportional to her waiting time in the system (the sum of the waiting time in the queue and the service time) with a unit-time cost c. The queue length is unobservable. Those customers who have joined the queue form an arrival process with rate $\lambda(p)$, which is called the effective arrival rate. The service system can hence be modeled as an $M/M/1$ queue, and the waiting time W is an exponential random variable with parameter $\mu - \lambda(p)$, i.e., $W \sim \exp(\mu - \lambda(p))$. Clearly, the expected waiting time $E(W) = 1/(\mu - \lambda(p))$.

There are two types of customers: θ proportion of them are *sophisticated* and have the knowledge about the service rate μ and the service reward R; the rest $1 - \theta$ proportion are *naive* and do not have such information, representing those one-time customers. For example, local residents generally are aware of service-related information such as the probable waiting times and food quality of a restaurant whereas tourists have no such information. We assume that the customer type composition parameter θ is common knowledge. Denote the potential market size, i.e., the potential arrival rates, of sophisticated and naive customers by Λ_s and

Λ_n, respectively. Then,

$$\Lambda_s = \theta\Lambda \text{ and } \Lambda_n = (1-\theta)\Lambda.$$

For the pricing strategy, we consider a general cyclic pricing strategy. The time horizon is divided into periods with length T. Each period is divided into N phases indexed by i. The price in phase i is denoted as p_i and the corresponding phase length is denoted as T_{p_i}. The general cyclic pricing strategy $\boldsymbol{PT}$ can be captured by a finite sequence of N combinations of *prices* p_i and the *phase length* T_{p_i} that price p_i continues for, i.e., $\boldsymbol{PT} = \{(p_i, T_{p_i})|i = 1, 2, ...N\}$. Clearly, $T = \sum_{i=1}^{i=N} T_{p_i}$. When $N = 1$, the cyclic pricing strategy degenerates into a static one. We assume that the cycle length T is long enough compared to the expected waiting time such that the time of the transient process to the steady states is negligible in each pricing cycle. For example, for typical service providers on yelp.com such as restaurants, dentists and optometrists, the waiting time per service episode is normally no more than 1 hour, whereas the cyclic length is one month (Yelp.com displays a "Monthly Trend", which shows the average rating of each month). Hence, the transient process within each pricing cycle is negligible.

The service provider decides on his pricing policy $\boldsymbol{PT}$ to maximize his long-run average profit per unit of time as follows:

$$\max_{\boldsymbol{PT}} \quad \Pi = \frac{1}{T}\left[\sum_{i=1}^{N} p_i\lambda(p_i)T_{p_i}\right],$$

where $\lambda(p_i)$ is the effective arrival rate of customers when the service provider charges a price p_i, which includes both naive and sophisticated customers. This objective function is in principle equivalent to maximizing the "profit rate" as in the queueing literature (see, for example, the comprehensive review of Hassin and Haviv (2003)), where the provider's objective is generally to maximize $p\lambda(p)$, where $\lambda(p)$ is an arrival rate. Define L_{p_i} as follows:

$$L_{p_i} = \frac{T_{p_i}}{T}.$$

In other words, L_{p_i} represents the proportion of time the price p_i is charged by the service provider in a cycle. Then, we can rewrite the provider's long-run average profit as

$$\max_{\boldsymbol{PT}} \quad \Pi = \sum_{i=1}^{N} p_i\lambda(p_i)L_{p_i}.$$

3.1 Average Rating and Equilibrium Arrival Rates

After the service, each customer, regardless of her type, *honestly rates* her *consumption utility*, $R - cw$, where w is the actual waiting time she has experienced. Ratings accumulate over time and the *average rating on the consumption utility* is advertised to the incoming customers by the service provider. Customers are also informed about the current-period price upon arrival. By simply comparing the average rating on the consumption utility with the posted current-period price, naive customers decide whether to join or not.

As our goal is to illustrate the effect of cyclic pricing in the long run, we directly consider the system performance *during the stable states* and ignore the transient process leading up to them. Given a pricing strategy $\boldsymbol{PT} = \{(p_i, T_{p_i})|i = 1, 2, ...N\}$, the average rating converges to a constant number in the long run, which is a function of $\boldsymbol{PT}$ denoted by $\eta(\boldsymbol{PT})$. The online Appendix B illustrates such a convergence process if customers adopt exponential smoothing to aggregate all ratings. Here, we shall not expand our discussion on this as it is not our focus. Instead, we move on to solve the long-run performance; that is, for a given pricing strategy $\boldsymbol{PT}$, the average rating $\eta(\boldsymbol{PT})$ and equilibrium arrival rates in each phase $\lambda(p_i)$'s, $i = 1, 2, ...N$, can be obtained through solving multiple equations. Given arrival rates $\lambda(p_i)$'s, we can determine the average rating $\eta(\boldsymbol{PT})$; given the average rating $\eta(\boldsymbol{PT})$, we can derive the equilibrium arrival rates in each phase. Below, we illustrate these two steps in detail.

Step 1: Determine the average rating $\eta(\boldsymbol{PT})$ given arrival rates.

$\eta(\boldsymbol{PT})$ is determined by customers' queueing behavior through the whole pricing cycle, and can be written as

$$\eta(\boldsymbol{PT}) := \frac{\sum_{i=1}^{i=N} v(p_i)\lambda(p_i)T_{p_i}}{\sum_{i=1}^{i=N} \lambda(p_i)T_{p_i}} = \frac{\sum_{i=1}^{i=N} v(p_i)\lambda(p_i)L_{p_i}}{\sum_{i=1}^{i=N} \lambda(p_i)L_{p_i}}, \tag{1}$$

where

$$v(p_i) := R - cE[W(p_i)] = R - \frac{c}{\mu - \lambda(p_i)} \tag{2}$$

is the expected consumption utility of a customer who joins at price p_i.

Step 2: Determine arrival rates given average rating $\eta(\boldsymbol{PT})$.

A customer arriving during phase i observes the current price p_i and the average rating $\eta(\boldsymbol{PT})$. She then makes the joining-or-balking decision to maximize her utility. Denote $\delta_n(p_i)$ and $\delta_s(p_i)$ as the joining decisions of the naive and sophisticated customers, respectively, when faced with the price p_i.

As naive customers do not have information about service-related parameters R and μ, they are incapable of anticipating their expected utility. Instead, they rely on the online rating information to decide whether to join or to balk. A naive customer joins if $p_i \leq \eta(\boldsymbol{PT})$ and balks otherwise. Note that the assumption of customer joining when $p_i = \eta(\boldsymbol{PT})$ (the break-even case) is not critical, as the service provider can always reduce the price by an infinitesimal amount to lure naive customers to join in practice. The joining decision of a naive customer is hence captured by a binary variable:

$$\delta_n(p_i) = \begin{cases} 1, & \text{if } p_i \leq \eta(\boldsymbol{PT}) \\ 0, & \text{if } p_i > \eta(\boldsymbol{PT}) \end{cases}. \tag{3}$$

As all naive customers have the same information access, they all join ($\delta_n(p_i) = 1$) or all balk ($\delta_n(p_i) = 0$). Note that this scenario is different from the no-information case referred to in some queueing strategy literature (see Guo and Zipkin (2007)). There, the queue is unobservable but customers still know other information about the system such as the potential arrival and service rates; thus, uninformed customers can still make a strategic queueing decision, that is, by adopting a mixed strategy in joining. Here, the herding behavior of naive customers results from lacking information. The system capacity is assumed to be large enough such that when naive customers join collectively, their expected consumption utility is non-negative i.e., $R - \frac{c}{\mu-(1-\theta)\Lambda} \geq 0$.

Sophisticated customers are strategic in making their queueing decisions. They choose a joining probability $\delta_s(p_i)$ to maximize their expected utility $R - cE[W(p_i)] - p_i$, in which they also take naive customers' joining decision $\delta_n(p_i)$ into consideration:

$$\mathcal{U}(p_i) = \max\{R - cE[W(p_i)] - p_i, 0\} = \max\left\{R - \frac{c}{\mu - \lambda(p_i)} - p_i, 0\right\}, \tag{4}$$

where the effective arrival rate is

$$\lambda(p_i) = \delta_s(p_i)\Lambda_s + \delta_n(p_i)\Lambda_n. \tag{5}$$

Clearly, if the potential market size is large enough, in equilibrium, $\delta_s(p_i)$ simply solves the equation

$$R - \frac{c}{\mu - \lambda(p_i)} - p_i = 0.$$

For the purpose of easy notations, let

$$V_s := R - \frac{c}{\mu - \Lambda_s} \text{ and } V_n := R - \frac{c}{\mu - \Lambda_n}.$$

Note that V_s (V_n, respectively) represents the expected consumption utility when *only* the sophisticated (naive, respectively) customers join the service system. We then have the following proposition regarding the joining/balking decisions of the two types of customers.

Proposition 1. *Given that the service provider charges a price $p_i \in \boldsymbol{PT}$, the equilibrium arrival rates are as follows.*

1. *If $p_i > \eta(\boldsymbol{PT})$, naive customers all balk, i.e., $\delta_n(p_i) = 0$. Only sophisticated customers join, and they join with probability $\delta_s(p_i) = \min\left\{\frac{1}{\Lambda_s}\left(\mu - \frac{c}{R-p_i}\right), 1\right\}$. The effective arrival rate is $\lambda(p_i) = \min\left\{\mu - \frac{c}{R-p_i}, \Lambda_s\right\}$.*

2. *If $p_i \leq \eta(\boldsymbol{PT})$, all the naive customers join, i.e., $\delta_n(p_i) = 1$, whereas sophisticated customers' joining probability is determined by the magnitude of $\eta(\boldsymbol{PT})$ and V_n:*

    I. *if $\eta(\boldsymbol{PT}) \geq V_n$ and $V_n \leq p_i \leq \eta(\boldsymbol{PT})$, all the sophisticated customers balk, i.e., $\delta_s(p_i) = 0$. The effective arrival rate is $\lambda(p_i) = \Lambda_n = (1-\theta)\Lambda$;*

    II. *otherwise, sophisticated customers join with probability*

    $$\delta_s(p_i) = \min\left\{\frac{1}{\Lambda_s}\left(\mu - \Lambda_n - \frac{c}{R-p_i}\right), 1\right\}.$$

    The effective arrival rate is $\lambda(p_i) = \min\left\{\mu - \frac{c}{R-p_i}, \Lambda\right\}$.

Proposition 1 implies that naive customers join only when the price is not high ($p_i \leq \eta(\boldsymbol{PT})$). Sophisticated customers, however, always choose a positive joining probability except that the price falls into an intermediate range of $V_n \leq p_i \leq \eta(\boldsymbol{PT})$ provided that $\eta(\boldsymbol{PT}) \geq V_n$. This is because in this case, naive customers all join but they gain a non-positive expected utility as the price charged is higher than their expected consumption utility V_n.

3.2 A Benchmark Case: $\theta = 1$

Before proceeding with the detailed analysis, we first introduce a benchmark case where all the customers are sophisticated, i.e., $\theta = 1$. This case has been widely studied in the literature on customers' strategic queueing behavior; see Chapter 3 of Hassin and Haviv (2003). Here, we briefly review the result of this traditional setting where the queue length is unobservable.

Consider a monopoly service provider, modeled as an $M/M/1$ queue, who decides on his profit-maximizing price p. As customers are all strategic, the effective arrival rate λ must solve $R - p - c/(\mu - \lambda) = 0$. The profit-maximizing provider's problem is to maximize $p\lambda$, subject to $0 \leq \lambda \leq \Lambda$ and $R - p - c/(\mu - \lambda) = 0$. It can be easily shown that the optimal price

$$p^* = \begin{cases} p_b := R - \sqrt{\frac{cR}{\mu}}, & \text{if } \Lambda > \mu - \sqrt{\frac{c\mu}{R}}, \\ R - \frac{c}{\mu - \Lambda}, & \text{otherwise.} \end{cases} \tag{6}$$

For ease of notation, denote

$$\lambda_b := \mu - \sqrt{\frac{c\mu}{R}}. \tag{7}$$

That is, when all the customers are strategic, if $\Lambda \leq \lambda_b$, it is optimal to admit all the customers into the system; if $\Lambda > \lambda_b$, some customers balk, and the equilibrium effective arrival rate is λ_b. Note that under this benchmark setting, the optimal decision of a profit-maximizing provider is the same as that of a welfare-maximizing provider.

4 Optimal Pricing Decision

In this section, we first investigate the service provider's optimal pricing strategies under both welfare maximization and profit maximization when customers rate over the consumption utility. We then consider another information scenario where customers rate over the net utility.

4.1 Welfare Maximization

Anticipating the customers' joining decisions stated in Proposition 1, the service provider decides his optimal pricing strategy $\boldsymbol{PT} = \{(p_i, T_{p_i}) | i = 1, 2, ..., N\}$. For a welfare-maximizing service provider, the price transfer between the customer and the provider is internalized. Maximizing the long-run average social welfare is equivalent to maximizing the expected total consumption utility given as follows:

$$\max_{\boldsymbol{PT}} \mathcal{SW} = \sum_{p_i \in \boldsymbol{PT}} v(p_i) \lambda(p_i) L_{p_i} \tag{8}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} L_{p_i} = 1, L_{p_i} \geq 0,$$

where $\lambda(p_i)$ is obtained from Proposition 1, and $v(p_i)$ is given in (2).

For ease of notation, let

$$\underline{\Lambda} = \frac{\lambda_b}{\max\{(1-\theta), \theta\}}, \quad \overline{\Lambda} = \frac{\lambda_b}{\min\{(1-\theta), \theta\}},$$

and

$$\widehat{\Lambda} = \frac{1}{2\theta(1-\theta)} \left(\mu - \sqrt{\mu \left[\mu - 4\theta(1-\theta) \left(\mu - \frac{c}{R} \right) \right]} \right).$$

It can be easily shown that $\underline{\Lambda}$ and $\widehat{\Lambda}$ increase (decrease, respectively) in θ while $\overline{\Lambda}$ decreases (increases, respectively) in θ when $\theta < \frac{1}{2}$ ($\theta \geq \frac{1}{2}$, respectively). Then, we have the following proposition.

Proposition 2. *The welfare-maximizing pricing strategy is static, i.e., $N = 1$.*

1. *When $\underline{\Lambda} < \Lambda < \overline{\Lambda}$ and $\theta < \frac{1}{2}$, the optimal price is as follows:*
   - *(a) If $\underline{\Lambda} < \Lambda \leq \widehat{\Lambda}$, $p^*_{sw} = V_n = R - \frac{c}{\mu-(1-\theta)\Lambda}$. All naive customers join while all sophisticated customers balk, that is, $\delta_s = 0$ and $\delta_n = 1$;*
   - *(b) If $\widehat{\Lambda} < \Lambda < \overline{\Lambda}$, $p^*_{sw} = V_s = R - \frac{c}{\mu-\theta\Lambda}$. All sophisticated customers join while all naive customers balk, i.e., $\delta_s = 1$ and $\delta_n = 0$.*

2. *Otherwise, the service provider always sets the price equal to that under the benchmark case in which all customers are sophisticated, that is, $p^*_{sw} = p^*$, where p^* is given by (6). In particular,*
   - *(a) If $\Lambda \leq \lambda_b = \mu - \sqrt{\frac{c\mu}{R}}$, $p^*_{sw} = R - \frac{c}{\mu-\Lambda}$. Both naive and sophisticated customers join, i.e., $\delta_s = \delta_n = 1$;*
   - *(b) If $\lambda_b < \Lambda \leq \underline{\Lambda}$, or if $\underline{\Lambda} < \Lambda < \overline{\Lambda}$ and $\frac{1}{2} \leq \theta < 1$, $p^*_{sw} = R - \sqrt{\frac{cR}{\mu}}$. Naive customers all join, i.e., $\delta_n = 1$, whereas sophisticated customers join with probability $\delta_s = \frac{\lambda_b - \Lambda_n}{\Lambda_s}$;*
   - *(c) If $\Lambda \geq \overline{\Lambda}$, $p^*_{sw} = R - \sqrt{\frac{cR}{\mu}}$. Naive customers all balk, i.e., $\delta_n = 0$, whereas sophisticated customers join with probability $\delta_s = \frac{\lambda_b}{\Lambda_s}$.*

Proposition 2 shows how market conditions affect the market equilibrium under welfare maximization; see Figure 1 for the illustration. If the potential market size is either small ($\Lambda \leq \underline{\Lambda}$) or large ($\Lambda \geq \overline{\Lambda}$), the welfare-maximizing provider charges a price that is the same as when all customers are sophisticated; naive customers are served when $\Lambda \leq \underline{\Lambda}$, while

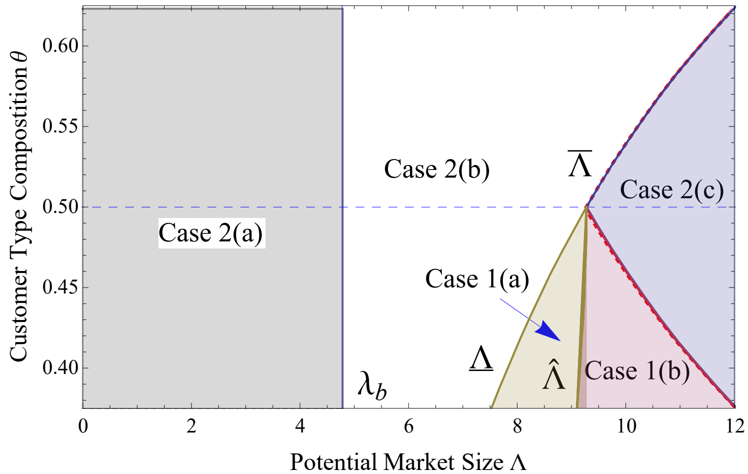

Figure 1: Equilibrium Pricing and Customer Joining Behavior under the Welfare Maximization: $R = 40$, $c = 180$, $\mu = 12$

only sophisticated customers are served when $\Lambda \geq \overline{\Lambda}$. When the potential market size is in an intermediate range ($\underline{\Lambda} < \Lambda < \overline{\Lambda}$), the customer type composition parameter θ plays a critical role in the provider's admission strategy. If $\theta \geq 0.5$, the provider admits all naive customers into the system since the potential market size of naive customers is small, i.e., $\Lambda_n < \lambda_b$. However, when $\theta < 0.5$, the potential market size of naive customers is larger than (whereas that of sophisticated customers is smaller than) what the welfare-maximizing provider desires; that is, $\Lambda_s < \lambda_b < \Lambda_n$. As naive customers join or balk collectively, the welfare-maximizing provider can serve only one type of customers. Specifically, the provider serves only naive customers when $\underline{\Lambda} < \Lambda < \widehat{\Lambda}$ and only sophisticated customers when $\widehat{\Lambda} < \Lambda < \overline{\Lambda}$.

To summarize, Proposition 2 indicates that the service provider serves all the customers when the potential market size is small ($\Lambda \leq \lambda_b$). Proposition 2 also implies that the service provider would ideally serve λ_b customers, if possible, by taking into consideration naive customers' herding behavior.

Figure 2 shows how the potential market size Λ and the customer type composition parameter θ affect both the social welfare and the optimal price. The solid line and the dashed line correspond to the cases $\theta = 0.4$ and $\theta = 0.5$, respectively. Figure 2(a) confirms that social welfare increases in the potential market size Λ when Λ is small, and reaches its peak at $\Lambda = \lambda_b$, the socially-desired arrival rate. A further increase of Λ harms social welfare when $\theta < 0.5$ and $\underline{\Lambda} < \Lambda < \overline{\Lambda}$ due to naive customers' herding behavior, which makes the effective arrival rate $\lambda_b = 4.65$ unachievable. Figure 2(b) shows that the optimal price

(a) Social Welfare

(b) The Optimal Price

Figure 2: Social Welfare and Optimal Price under the Welfare Maximization: $R = 40$, $c = 180$, $\mu = 12$

decreases in the potential market size Λ when Λ is smaller than what is socially desired (i.e, $\Lambda \leq \lambda_b = 4.65$), and remains unchanged when it is beyond the socially desired λ_b as long as λ_b is achievable. Note that when $\theta = 0.4$ and $\underline{\Lambda} = 7.7 < \Lambda < 11.6 = \overline{\Lambda}$, $\lambda_b = 4.65$ cannot be achieved. Figure 2(b) shows that in this range, the optimal price first decreases, then jumps up at $\widehat{\Lambda} = 9.2$, and decreases again as the market size Λ increases. This is due to the switch from serving naive customers only to serving sophisticated customers only.

4.2 Profit Maximization

The optimization problem of the profit-maximizing provider is very similar to that of the welfare-maximizing provider, except that $v(p_i)$'s are replaced by p_i's in the objective function. Thus, his optimization problem is given as follows:

$$\max_{\boldsymbol{PT}} \Pi = \sum_{p_i \in \boldsymbol{PT}} p_i \lambda(p_i) L_{p_i}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} L_{p_i} = 1, L_{p_i} \geq 0,$$

where $\lambda(p_i)$ is stated in Proposition 1. We then obtain the following result:

Proposition 3. *The profit-maximizing pricing strategy satisfies the following properties.*

1. *There exists a threshold $\widetilde{\Lambda}$ on the potential market size, above which the cyclic pricing strategy is preferred by the service provider. Otherwise, the service provider prefers*

the static pricing strategy and behaves exactly like that under welfare maximization as stated in Proposition 2.

2. *The optimal cyclic pricing strategy is high-low cyclic, which can be simply denoted as $\boldsymbol{PT} = \{(p_h, L), (p_l, 1-L)\}$, where p_h and p_l represent the high and low price, respectively, and L represents the proportion of time the high price remains. Under the optimal cyclic pricing strategy, sophisticated customers join during the high-price phase and naive ones join during the low-price phase.*

3. *The optimal low price satisfies $p_l^* = \eta(\boldsymbol{PT})$, and the other two decision variables, (p_h^*, L^*), are given as follows.*

   I. *If $\theta < \frac{1}{2}$ and $\widetilde{\Lambda} < \Lambda < \ddot{\Lambda}$ ($\ddot{\Lambda}$ is given by (44)), $p_h^* = V_s$ and $L^* = L_b$, where L_b solves (43). Here, $\delta_s = \delta_n = 1$.*

   II. *If $\theta < \frac{1}{2}$ and $\Lambda \geq \ddot{\Lambda}$, or if $\theta \geq \frac{1}{2}$ and $\Lambda > \widetilde{\Lambda}$, $p_h^* = p_h^0 > V_s$ and $L^* = L^0$, where (p_h^0, L^0) solves (37) and (45) simultaneously. Moreover, $\delta_s < 1$ and $\delta_n = 1$.*

The detailed expressions of (37), (43), (44) and (45) can be found in the online Appendix. Moreover, $p_h^* > R - \sqrt{\frac{cR}{\mu}}$.

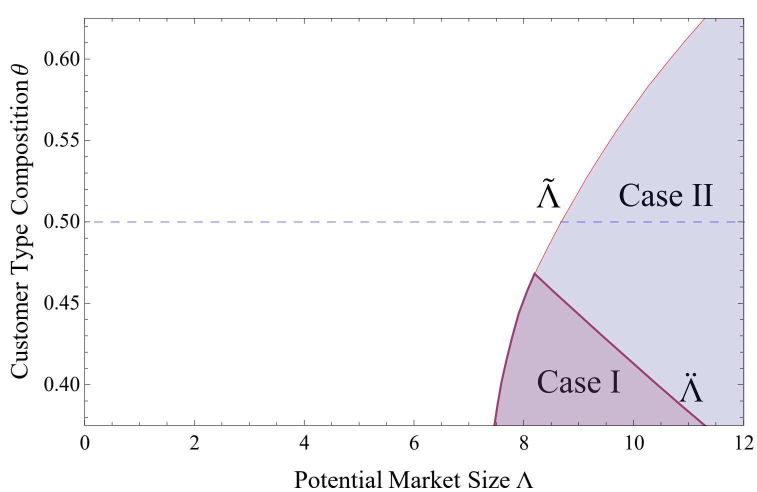

Figure 3: Equilibrium Pricing and Customer Joining Behavior under the Profit Maximization: $R = 40$, $c = 180$, $\mu = 12$

Proposition 3 shows that only when the potential market size Λ is above a certain value will the cyclic pricing strategy be preferred by the service provider over the static pricing strategy; see Figure 3 for the illustration. As illustrated in Figure 3, at the left side of the red line $\Lambda = \widetilde{\Lambda}$, the profit-maximizing provider adopts exactly the same static pricing strategy as the welfare-maximizing provider does (see Figure 1); while at the right side of the red line, the cyclic pricing strategy is adopted. Under the optimal cyclic pricing, sophisticated customers join only within the high-price phase and they obtain an expected net utility of zero; naive customers join only within the low-price phase and their expected consumption utility is V_n. Since the optimal low price p_l^* equals the long-run average rating and is higher than V_n, naive customers' expected net utility, $V_n - p_l^*$, is always negative. Thus, although naive customers pay a lower price, the price is not low enough to guarantee a non-negative utility. This overcharging is possible due to the higher ratings generated from sophisticated customers, which boost up the average rating. Figure 3 also shows that $\widetilde{\Lambda}$ increases with θ, the proportion of sophisticated customers in the market. That is, when the market comprises a larger proportion of sophisticated customers, static pricing is more likely to be favored.

To thoroughly understand the mechanism underlying the profit gain, we now turn to the comparison between the static and cyclic pricing strategies. Under a static pricing strategy, all customer surplus is internalized through pricing and goes towards the service provider's profit; thus, maximizing profit is the same as maximizing welfare. The classic queueing literature on pricing problems has shown that the objective function of a profit-maximizing service provider is concave in the effective arrival rate, and that λ_b is the best demand level that a monopoly provider should maintain whenever it is achievable. According to Proposition 2, the optimal arrival rate λ_b can be achieved under the static pricing strategy when more than half of the customers are sophisticated ($\theta \geq 0.5$). One might believe that the average arrival rate under the cyclic pricing strategy should be equal to this one. That is, if the demand rate during the high-price phase is smaller than λ_b, the one during the low-price phase should be larger than λ_b. Is this belief true? The following corollary provides a different answer.

Corollary 1. *When $\theta \geq 0.5$ and $\widetilde{\Lambda} < \Lambda < \overline{\Lambda}$, the effective arrival rates under the optimal cyclic prices are always less than that under the static one, specifically, $\lambda(p_h^*) < \lambda(p_l^*) < \lambda(p_{sw}^*)$.*

Corollary 1 shows that, under the cyclic pricing strategy, a service provider may deliberately serve less than λ_b customers during both low- and high-price phases and achieve a higher profit. Why not lower both high and low prices to attract more customers? Why does the original price-demand tradeoff not work here? A closer investigation of the cyclic pricing strategy reveals that the service provider cannot do that. During the low-price phase, naive customers all join and obtain a negative expected utility. Sophisticated customers know that, and hence they will never join then. Therefore, demand cannot be increased by lowering the price unless naive customers can obtain a non-negative expected utility. During the high-price phase, demand can be increased by lowering the price; however, this results in more congestion, which will reduce the ratings on the consumption utility and jeopardize the strategy of overcharging naive customers during the low-price phase. Consequently, due to a higher profit margin from overcharging naive customers, the original price-demand tradeoff is twisted, and fewer customers are served in both phases.

(a) Optimal Profit

(b) Profit Increment

Figure 4: Optimal Profit and Performance Comparsion between Cyclic Pricing and Static Pricing: $R = 40$, $c = 180$, $\mu = 12$, $\widetilde{\Lambda} = 7.62$ when $\theta = 0.4$ and $\widetilde{\Lambda} = 8.66$ when $\theta = 0.5$

Next, we numerically examine the benefit brought by the cyclic pricing strategy. In our numerical study, we change the value of the customer type composition parameter θ from 0.3875 to 0.6125 with a step length 0.0125. For each given θ, we then vary the potential market size Λ to find the highest profit increment brought by the cyclic pricing strategy over the static one. Our aggregated numerical results show that the highest increment amount is around 5.19% on average but can reach as high as 11.22%; see Figure 4 for an illustration of the results when $\theta = 0.4$ and $\theta = 0.5$, respectively.

We also numerically investigate how market size affects the optimal cyclic pricing strategy. Figures 5(a) and 5(b) depict the optimal high and low prices when the customer type composition parameter θ equals 0.4 and 0.5, respectively. They show that for a given customer type composition parameter, both optimal prices decrease as the potential market size Λ increases. For the sake of comparison, we also plot the optimal static price (that is,

(a) Optimal Prices: $\theta = 0.4$

(b) Optimal Prices: $\theta = 0.5$

(c) Proportion of Time Staying at the High Price (L)

Figure 5: Optimal Pricing Strategy under Profit Maximization: $R = 40$, $c = 180$, $\mu = 12$, $\widetilde{\Lambda} = 7.62$ when $\theta = 0.4$ and $\widetilde{\Lambda} = 8.66$ when $\theta = 0.5$

the welfare-maximizing price) with a black dashed line in Figures 5(a) and 5(b). We find that under both cases, when the market size is not large, both high and low prices adopted in the cyclic pricing strategy are larger than the static one. This observation implies that the conclusion stated in Corollary 1 might be generally true for a small-size market and not necessarily just restricted to the case of $\theta \geq 0.5$. Figure 5(c) shows that L^*, the proportion of time that the high price shall be charged, increases with the potential market size Λ. That is, as Λ increases, the service provider charges the high price for a longer time.

4.3 An Alternative Information Scenario: Rating on the Net Utility

Here, we consider another information scenario where customers take price into consideration in their rating; that is, customers rate over the net utility $R - cw - p_i$ rather than the consumption utility $R - cw$. In practice, some online review sites, such as dianping.com and openrice.com, allow a customer to rate on multiple *classified items*, such as experience, dining, and cost-effectiveness (i.e., price). This allows the incoming customers to infer the consumption utility. Other review sites such as yelp.com, however, provide only an overall rating score. In the information scenario where only the average rating on the net utility is available, naive customers join only if the average rating is non-negative. Consequently, a static pricing strategy is enough as the service provider cannot change naive customers' joining behavior by cyclically changing the price. In other words, the validity of the cyclic pricing strategy requires rating over both the net utility and the price. Below, we examine this information scenario in detail.

Recall that with the rating on the consumption utility, a naive customer joins if and only if the average rating is higher than the price. In such situation, a negative rating will drive away the naive customer. However, with the ratings on both the price and the net utility, a naive customer may find it worthwhile to join even when the average rating on the net utility is negative. Let us use the following simple numerical example to illustrate naive customers' joining decision when customers rate over both the price and the net utility. At the beginning of a certain period, naive customers come and they see a rating of average price \$100 and an average rating of the net utility -10. Suppose that the current-period price is \$80, a 20% discount of the average one. Naive customers then would join because the \$20 gain due to the price discount overwhelms the negative part of the net utility. From this example, we can see that, with the knowledge about the average ratings on both the net utility and the price, incoming customers can still make their joining/balking decision by comparing the amount of price discount with the average rating over the net utility. Alternatively, they can simply add up the average price (denoted by $\bar{p}$) and the average rating of the net utility to infer the average consumption utility. They can then make their joining/balking decisions by comparing this inferred consumption utility with the current-period price. To see this equivalence, let $\eta'(\boldsymbol{PT})$ be the average rating of the net utility. Then, we have

$$\begin{aligned}\eta'(\boldsymbol{PT}) &\equiv \frac{\sum_{i=1}^{i=N}(v(p_i)-p_i)\lambda(p_i)L_{p_i}}{\sum_{i=1}^{i=N}\lambda(p_i)L_{p_i}} \\ &= \frac{\sum_{i=1}^{i=N}v(p_i)\lambda(p_i)L_{p_i}}{\sum_{i=1}^{i=N}\lambda(p_i)L_{p_i}} - \frac{\sum_{i=1}^{i=N}p_i\lambda(p_i)L_{p_i}}{\sum_{i=1}^{i=N}\lambda(p_i)L_{p_i}} = \eta(\boldsymbol{PT}) - \bar{p}.\end{aligned}$$

The above equation exactly shows that the incoming customers can infer the average rating on the consumption utility $\eta(\boldsymbol{PT})$ by adding up the average price $\bar{p}$ and the average rating on the net utility $\eta'(\boldsymbol{PT})$. Regarding customers' joining/balking decisions, we obtain the following result.

Proposition 4. *Give a pricing strategy* $\boldsymbol{PT}$*, the joining-or-balking decisions of both types of customers remain the same under the two rating information scenarios, rating over the consumption utility only and rating over both the price and the net utility.*

Proposition 4 shows that customers' responses towards joining or balking for a given pricing strategy remain the same in the two rating information scenarios. Note that the service provider makes his optimal pricing decisions in anticipation of the customers' joining-or-balking decisions. Proposition 4 then implies that the service provider's optimal pricing strategy shall remain the same no matter whether the customers rate over the consumption utility only or rate over both the price and the net utility. That is, all the results obtained in §4.1 and §4.2 continue to hold with the rating over both the price and the net utility. For example, given a cyclic pricing strategy, sophisticated customers join only during the high-price phase and naive customers join only during the low-price phase under both rating information scenarios.

Note that naive customers make their ex-ante joining/balking decisions fully based on the rating information. In the aforementioned numerical example, they had thought that they would make a $10 gain in the net utility if they join. However, they ignore that other naive customers are thinking exactly the same, and thus all the naive customers join. As a result, the system eventually becomes overcrowded, and all customers obtain a negative ex-post net utility. In other words, what makes the naive customers' ex-ante and ex-post payoffs different from each other is that they ignore their peer-customers' joining behavior. This is actually the key difference between a goods market and a service market: quality in terms of congestion in a service system is *endogenously* formed by customers themselves while quality in a goods market is exogenous. Indeed, an ex-ante "correct" decision may actually result in ex-post disappointing outcomes. Clearly, the pricing mechanism requires an underlying assumption that these naive customers are one-time customers. The rationale for the cyclic pricing strategy to yield a higher profit for the service provider is that, even though the long-run average rating is low and negative, naive customers derive an additional gain when they observe a lower-than-average price. Despite their dissatisfaction and low ratings after consumption, the overall ratings can still be uplifted by the high ratings posted by the sophisticated customers, which will be used to attract another batch of new-arriving naive customers in the next pricing phase.

5 Comparison between Profit Maximization and Welfare Maximization

So far, we have derived the service provider's optimal pricing strategies under both welfare and profit maximization. Recall that the profit-maximizing provider behaves exactly as a welfare-maximizing provider when the potential market size Λ is below the threshold $\widetilde{\Lambda}$ (see Proposition 3). Thus, the system performances under these two objectives are the same when $\Lambda \leq \widetilde{\Lambda}$. Hereafter, we focus on comparing the system performance under these two objectives when $\Lambda > \widetilde{\Lambda}$, that is, when the service provider adopts static pricing under welfare maximization but cyclic pricing under profit maximization. We examine how the cyclic pricing strategy affects the system's welfare. We define *welfare loss* of the system under profit maximization as

$$\frac{\mathcal{SW}^* - \mathcal{SW}_{pm}}{\mathcal{SW}^*},$$

where $\mathcal{SW}_{pm}$ denotes the social welfare under the optimal cyclic pricing strategy. $\mathcal{SW}_{pm}$ is given as follows:

$$\mathcal{SW}_{pm} = v(p_h^*)\lambda(p_h^*)L^* + V_n\Lambda_n(1 - L^*) = p_h^*\lambda(p_h^*)L^* + V_n\Lambda_n(1 - L^*),$$

where p_h^* and L^* are the optimal high price and the corresponding proportion of time the high price remains under the cyclic pricing strategy, respectively.

Figure 6: Optimal Cyclic Pricing Strategy's Social Efficiency: $R = 40$, $c = 180$, $\mu = 12$, $\widetilde{\Lambda} = 7.62$ when $\theta = 0.4$ and $\widetilde{\Lambda} = 8.66$ when $\theta = 0.5$

Figure 6 illustrates the impact of market conditions on welfare loss when the customer type composition parameter $\theta = 0.4, 0.5$. An interesting observation is that welfare loss is not monotone in the market size. For the case of $\theta = 0.4$, the optimal pricing strategy is cyclic only when the potential market size $\Lambda > \widetilde{\Lambda} = 7.62$. Starting from $\Lambda = 7.62$, the welfare loss first increases, then decreases and finally increases again in Λ. Specifically, when Λ falls into the range between $\widetilde{\Lambda} = 7.62$ and $\widehat{\Lambda} = 9.19$, the welfare loss first increases, then decreases and reaches 0 at $\Lambda = \widehat{\Lambda} = 9.19$. In this range, the welfare maximization requires that only naive customers be served, and the socially optimal price is $p^*_{sw} = V_n$. Thus, the maximal social welfare is $\mathcal{SW}^* = V_n \Lambda_n$. However, under the cyclic pricing strategy, only sophisticated customers are served at the high price $p^*_h = V_s$ (see Case I of Proposition 3), and all naive customers are served at the low price. Therefore, the difference in welfare only occurs during the high-price phase. In short, when $\Lambda \in [7.62, 9.19)$, the welfare loss is caused by the underutilization of the system during the high-price phase, due to the fact that the number of sophisticated customers is less than that of naive customers. Such welfare loss can be expressed as

$$\frac{\mathcal{SW}^* - \mathcal{SW}_{pm}}{\mathcal{SW}^*} = \frac{(V_n \Lambda_n - V_s \Lambda_s)}{V_n \Lambda_n} \cdot L^*. \tag{9}$$

We can see that welfare loss is the product of two terms, $\frac{V_n \Lambda_n - V_s \Lambda_s}{V_n \Lambda_n}$ and L^*. We now consider the impact of marker size on these two terms. One can check that

$$\frac{V_n \Lambda_n - V_s \Lambda_s}{V_n \Lambda_n} = 1 - \frac{\theta}{1-\theta} \cdot \frac{V_s}{V_n}$$

is decreasing in Λ. As the market size increases, the number of sophisticated customers increases proportionally and the underutilization effect is reduced. Meanwhile, L^* is increasing in Λ, because the proportion of high-price phase must be lengthened to balance the negative reviews from an increased number of naive customers. When Λ is very small, the second effect (the increasing monotonicity of L^*) dominates, but as Λ further increases, the first effect (the diminishing effect of underutilization) dominates. These two effects jointly drive the product term in (9) to first increase and then decrease in Λ. Note that when $\Lambda = \widehat{\Lambda} = 9.19$, $V_n\Lambda_n = V_s\Lambda_s$. Hence, the welfare loss stated in (9) becomes 0. For the range $\Lambda \geq \widehat{\Lambda} = 9.19$, welfare maximization requires that only sophisticated customers should be served, whereas cyclic pricing admits all naive customers during the low-price phase. This indicates that when $\Lambda > 9.19$, the welfare loss is mainly caused by the overutilization of the system during the low-price phase. As the potential market size increases, welfare loss becomes larger due to the increased over-crowdedness of the system.

When $\theta = 0.5$, the cyclic pricing strategy is adopted by the profit-maximizing provider when the potential market size $\Lambda \geq \widetilde{\Lambda} = 8.66$. Figure 6 shows that when $\Lambda \geq 8.66$, welfare loss first decreases and then increases, reaching the minimum at $\Lambda = \overline{\Lambda} = 9.30$. This observation can be explained as follows. According to Proposition 2, when $\theta \geq 0.5$, the socially optimal price is $p^*_{sw} = p(\lambda_b)$, where λ_b is the socially desirable effective arrival rate defined in (7). Also, according to Case II of Proposition 3, the optimal high price satisfies $p^*_h > p(\lambda_b)$. Hence, under profit maximization, compared with what is socially desirable, the system is always underutilized during the high-price phase. In addition, the provider always serves all the naive customers during the low-price phase. When Λ is small such that $\Lambda_n < \lambda_b$, the system is also underutilized during the low-price phase. This observation is also consistent with the conclusion in Corollary 1: cyclic pricing could cause underutilization in both phases. It contradicts our conventional belief that welfare loss caused by cyclic pricing is due to the uneven workloads in the two phases– the underutilization in the high-price phase and overutilization in the low-price phase. Thus, when $\Lambda_n < \lambda_b$, increasing Λ reduces the effect of underutilization and thus, mitigates welfare loss. Welfare loss reaches its minimum when $\Lambda_n = \lambda_b$, i.e., $\Lambda = \overline{\Lambda} = 9.30$. When $\Lambda_n > \lambda_b$, the system is overutilized during the low-price phase. In this range, increasing Λ further increases the level of crowdedness and thus enlarges the welfare loss.

Recall that welfare is the sum of the service provider's profit and all customers' utilities. The price becomes the internal transfer between the two parties, and hence does not appear in the welfare expression. Thus, only the number of customers served and each customer's consumption utility matter when determining the social welfare. The tradeoff between the number of customers served and consumption utility per customer determines an optimal congestion level. The foregoing analysis shows that welfare loss is mainly caused by the uneven workloads between the two phases — over-congestion in the low-price phase and underutilization in the high-price phase. However, in a situation where the market size is not large and sophisticated customers comprise not less than half of the population, cyclic pricing can cause underutilization in both phases, also resulting in welfare loss. In the latter case, too few customers are served, which is also socially undesirable.

6 Conclusion

In this paper, we consider a typical service situation where customers are heterogeneous in information accesses: some customers know the service-related information, whereas others do not; the latter relies on the buyer-generated information to make their queueing decisions. We demonstrate that a cyclic pricing strategy can be used to improve the profitability of a service provider without distorting customers' ratings. Under the optimal high-low cyclic pricing strategy, sophisticated customers join at the high price, and naive customers join at the low price. During the high-price phase, the system is less congested, and thus the ratings are relatively high, which boosts the average rating and allows the provider to charge a price higher than naive customers' expected consumption utility during the low-price phase.

The validity of the cyclic pricing strategy requires the potential market size to be above a threshold value such that congestion is a significant factor in affecting customers' joining decision. The interesting part is that under the optimal cyclic pricing strategy, even though naive customers feel unsatisfied and post low ratings after consumption, the average rating can still be maintained by getting high ratings in the high-price phase. This, in turn, allows the service provider to obtain a higher profit than that under a static pricing strategy. We also find that although the cyclic pricing strategy can improve profitability, it harms social welfare. Welfare maximization prefers even workloads across periods, and hence prefers the static pricing strategy. However, the system can be either over- or under-utilized under the cyclic pricing strategy, leading to welfare loss.

Our study's main takeaway is to show that even without manipulating customer ratings, a service provider can still take advantage of uninformed customers by implementing a cyclic pricing strategy and advertising the average rating to the incoming customers. From the viewpoint of naive customers, they shall take in the aggregated rating information with *caution*: when the products are service goods, an average rating in the past does not mean that they can get the same consumption utility if they join. How to protect naive customers' welfare remains to be an interesting research question.

There are other issues calling for further study. First, our paper considers the stable states of the system, but the transient process to reach such stable states is left out. Studying such transient process is an interesting future research question and particularly important for short-life-cycle service products. Second, in our work, naive customers rely on the average rating information. In practice, they might use anecdotal reasoning; that is, they draw on a sample of ratings and reviews and then make their joining-or-balking decisions based on this limited sample. Thus, a pricing strategy based on customer anecdotal reasoning could be an interesting research question. Third, learning over the service rate is not critical in our study as incoming customers only care about the rating on the consumption utility. However, it would be an interesting research question to extend our current study to a health care setting where the service quality is related with the service rate. In such setting, learning over the service rate would be critical. We leave it as a future research question.

Acknowledgments

We are grateful to the departmental editor (Professor Michael Pinedo), an anonymous senior editor, and two anonymous referees for very helpful comments and suggestions. The second author Pengfei Guo acknowledges the financial support by the Research Grants Council of Hong Kong (GRF grant number: PolyU15506417). The third author Yulan Wang's work was supported by the Research Grants Council of Hong Kong (RGC Reference Number: PolyU 15502917).

References

Afeche, P., B. Ata. 2013. Bayesian dynamic pricing in queueing systems with unknown delay cost characteristics. *Manufacturing & Service Operations Management*, 15(2), 292-304.

Allon G., A. Bassamboo, I. Gurvich. 2011. "We will be right with you": Managing customer expectations with vague promises and cheap talk. *Operations Research*, 59, 1382-1394.

Armony M., C. Maglaras. 2004. Contact center with a call-back option and real-time delay information. *Operations Research*, 52, 527–545.

Besbes, O., Scarsini, M. 2018. On information distortions in online ratings. Operations Research.

Branco, F., Sun, M., Villas-Boas, J. M. 2012. Optimal search for product information. Management Science, 58(11), 2037-2056.

Chen, Y., Xie, J. 2008. Online consumer review: Word-of-mouth as a new element of marketing communication mix. Management science, 54(3), 477-491.

Chen, H., Wan, Y. W. 2003. Price competition of make-to-order firms. IIE Transactions, 35(9), 817-832.

Crapis, D., Ifrach, B., Maglaras, C., Scarsini, M. 2016. Monopoly pricing in the presence of social learning. *Management Science*, 63(11), 3586-3608.

Cui, S., S. K. Veeraraghavan. 2014. Blind queues: The impact of consumer beliefs on revenues and congestion. Available at SSRN 2196817.

Debo, L., C. Parlur, U. Rajan. 2012. Signaling quality via queues. *Management Science*, 58(5), 876-891.

Debo, L., S. Veeraraghavan. 2014. Equilibrium in queues under unknown service times and service value. *Operations Research*, 62(1), 38-57.

Dellarocas, C. 2006. Strategic manipulation of internet opinion forums: Implications for consumers and firms. Management science, 52(10), 1577-1593.

Dewan, S., H. Mendelson. 1990. User delay costs and internal pricing for a service facility. *Management Science*, 36(12), 1502-1517.

Dobson G., E. Pinker. 2006. The value of sharing lead time information. *IIE Transactions*,38, 171-183.

Economou, A., A. Gomez-Corral, S. Kanta. 2011. Optimal balking strategies in single-server queues with general service and vacation times. *Performance Evaluation*. 68(10), 967-982.

Edleson, N., D. Hildebrand. 1975. Congestion tolls for Poisson queueing process. *Econometrica*, 43, 81–92.

Guo, P., P. Zipkin. 2007. Analysis and comparison of queues with different levels of delay information. *Management Science*, 53, 962-970.

Guo, P., R. Hassin. 2011. Strategic behavior and social optimization in Markovian vacation queues. *Operations Research*, 59(4), 986-997.

Guo, P., R. Hassin. 2012. Strategic behavior and social optimization in Markovian vacation queues: the case of heterogeneous customers. *European Journal of Operational Research*, 222(2), 278-286.

Guo, P., Q. Li. 2013. Strategic behavior and social optimization in partially-observable Markovian vacation queues. *Operations Research Letters*, 41(3), 277-284.

Guo, P., W. Sun, Y. Wang. 2011. Equilibrium and optimal strategies to join a queue with partial information on service times. *European Journal of Operational Research*, 214(2), 284-297.

Guo, P., Z. Zhang. 2013. Strategic queueing behavior and its impact on system performance in service systems with the congestion-based staffing policy. *Manufacturing & Service Operations Management*, 15(1), 118-131.

Guo, P., M. Haviv, Y. Wang. 2015. Equilibrium queueing strategies when service quality is unknown to some customers. Working paper. The Hong Kong Polytechnic University, Hong Kong.

Hassin, R. 1986. Consumer information in markets with random product quality: The case of queues and balking. *Econometrica*, 54, 1185-1195.

Hassin, R. 2016. *Rational Queueing*. CRC Press.

Hassin, R., M. Haviv. 1994. Equilibrium strategies and the value of information in a two line queueing system with threshold jockeying. *Stochastic Models*, 10, 415–435.

Hassin, R., Haviv, M. 2003. *To queue or not to queue: Equilibrium behavior in queueing systems*. Kulwer Academic Publisher, Nowerll, MA.

Hassin, R., R. Roet-Green. 2011. Equilibrium in a two dimensional queueing game: When inspecting the queue is costly. Working paper. Tel Aviv University, Israel.

Hassin, R., Roet-Green, R. 2018. Cascade equilibrium strategies in a two-server queueing system with inspection cost. European Journal of Operational Research, 267(3), 1014-1026.

Heidhues, P., Koszegi, B. 2014. Regular prices and sales. Theoretical Economics, 9(1), 217-251.

Hu, M., Li, Y., Wang, J. 2017. Efficient ignorance: Information heterogeneity in a queue. *Management Science*, 64(6), 2650-2671.

Huang, T., G. Allon, A. Bassamboo. 2013. Bounded rationality in service systems. *Manufacturing & Service Operations Management*, 15(2), 263-279.

Huang, T., Y. Chen. 2015. Service systems with experience based anecdotal reasoning customers. *Production and Operations Management*, 24(5), 778-790.

Koszegi, B., Rabin, M. 2006. A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4), 1133-1165.

Kwark, Y., Chen, J., Raghunathan, S. 2014. Online product reviews: Implications for retailers and competing manufacturers. Information systems research, 25(1), 93-110.

Li, X., P. Guo, Z. Lian. 2016. Quality-speed competition in customer-intensive cervices with boundedly rational customers. *Production and Operations Management*, *Forthcoming.*

Naor, P. 1969. The regulation of queue size by levying tolls. *Econometrica: journal of the Econometric Society*, 15-24.

Shin, D., Zeevi, A. 2017. Dynamic pricing and learning with online product reviews. Technical report, Columbia University.

Sun, M. 2012. How does the variance of product ratings matter?. *Management Science*, 58(4), 696-707.

Stidham, S. 1992. Pricing and capacity decisions for a service facility: Stability and multiple local optima. *Management Science*, 38(8), 1121–1139.

Stidham, S. 2009. *Optimal Design of Queueing Systems*. CRC Press, London.

Veeraraghavan, S., L. Debo. 2009. Joining longer queues: Information externalities in queue choice. *Manufacturing & Service Operations Management*, 11(4), 543-562.

Veeraraghavan, S. K., L. Debo. 2011. Herding in queues with waiting costs: Rationality and regret. *Manufacturing & Service Operations Management*, 13(3), 329-346.

Wang, J., J. Li. 2008. A repairable M/G/1 retrial queue with Bernoulli vacation and two-phase service. *Quality Technology and Quantitative Management*, 5(2), 179-192.

Whitt, W. 1999. Improving service by informing customers about anticipated delays. *Management Science*, 45, 192-207.

Yang, L., Guo, P., Wang, Y. 2018. Service Pricing with Loss-Averse Customers. Operations Research.

Zhang, F., J. Wang, B. Liu. 2013. Equilibrium balking strategies in Markovian queues with working vacations. *Applied Mathematical Modelling*, 37(16), 8264-8282.

Online Appendix
"Cyclic Pricing When Customers Queue with Rating Information"

Appendix A: Proofs

Proof of Proposition 1

Recall that the overall effective arrival rate at price p_i is given by (5).

First, if $p_i > \eta(\boldsymbol{PT})$, all naive customers balk according to (3), i.e., $\delta_n(p_i) = 0$. Sophisticated customers maximize their expected utility according to (4), which is simplified as $\mathcal{U} = \max \left\{R - \frac{c}{\mu - \delta_s(p_i)\Lambda_s} - p_i, 0\right\}$. Then, $\delta_s(p_i) = \min\left\{\frac{1}{\Lambda_s}\left(\mu - \frac{c}{R-p_i}\right), 1\right\}$. Consequently, $\lambda(p_i) = \min\left\{\mu - \frac{c}{R-p_i}, \Lambda_s\right\}$.

Next, consider $p_i \leq \eta(\boldsymbol{PT})$. According to (3), naive customers all join, i.e., $\delta_n(p_i) = 1$. The pricing is classified into two cases:

1. If $V_n < p_i \leq \eta(\boldsymbol{PT})$, $R - \frac{c}{\mu - \Lambda_n} - p_i = V_n - p_i < 0$. Sophisticated customers know that joining leads to a negative consumption utility, and therefore, according to (4), they all balk, i.e., $\delta_s(p_i) = 0$. Consequently, $\lambda(p_i) = \Lambda_n = (1-\theta)\Lambda$.

2. If $p_i < \min\{V_n, \eta(\boldsymbol{PT})\}$, sophisticated customers decide $\delta_s(p_i) > 0$ to maximize the expected utility: $\mathcal{U} = \max \left\{R - \frac{c}{\mu - (\delta_s(p_i)\Lambda_s + \Lambda_n)} - p_i, 0\right\}$. They keep joining until $\mathcal{U} = 0$ or until all customers have joined. That is, $\delta_s(p_i) = \min\left\{\frac{1}{\Lambda_s}\left(\mu - \Lambda_n - \frac{c}{R-p_i}\right), 1\right\}$, and consequently, $\lambda(p_i) = \min\left\{\mu - \frac{c}{R-p_i}, \Lambda\right\}$

Proof of Proposition 2

When the queue is unobservable, a welfare-maximizing provider is not worse off by leaving customers an expected utility of zero (Hassin and Haviv, 2003). We can simply let $v(p_i) = p_i$ without reducing social welfare. Hence, the objective function (8) can be written as

$$\max_{\boldsymbol{PT}} \mathcal{SW} = \sum_{p_i \in \boldsymbol{PT}} v(p_i)\lambda(p_i)L_{p_i} = \sum_{p_i \in \boldsymbol{PT}} p_i \left(\mu - \frac{c}{R-p_i}\right) L_{p_i}.$$

One can show that $p_i\left(\mu - \frac{c}{R-p_i}\right) L_{p_i}$ is concave in p_i. Hence, there exists a p_k maximizing the term $p_i\left(\mu - \frac{c}{R-p_i}\right)$. Clearly, the optimal pricing strategy is to set $L_{p_k} = 1$ because $p_k\left(\mu - \frac{c}{R-p_k}\right) \geq \sum_{p_i \in \boldsymbol{PT}} p_i\left(\mu - \frac{c}{R-p_i}\right) L_{p_i}$. In short, the optimal pricing strategy is static.

After showing that the optimal welfare-maximizing pricing strategy is static, we next obtain the optimal price. Now, we can drop the subscript and simplify the objective function of the welfare-maximizing provider (8) as follows:

$$\max_{p} \mathcal{SW} = p\left(\mu - \frac{c}{R-p}\right).$$

There exists a one-to-one mapping between p and $\lambda(p)$. Since it is more straightforward to use $\lambda(p)$ as the variable, we can rewrite the above objective function as

$$\max_{\lambda(p)} \mathcal{SW} = \lambda(p)\left(R - \frac{c}{\mu - \lambda(p)}\right).$$

Based on Proposition 1, we can see that the provider inherently decides whether he wants to serve naive customer in the long run. If he does not, $\lambda(p) < \Lambda_n$; otherwise, $\lambda(p) \geq \Lambda_n$. In the following analysis, we can solve the above optimization problem under these two constraints separately. We then compare the results to obtain the optimal pricing strategy.

We first consider the case of $\lambda(p) < \Lambda_n$. It can be easily obtained that

$$\lambda(p) = \begin{cases} \Lambda_s, & \text{if } \Lambda \leq \frac{\lambda_b}{\theta} \\ \lambda_b, & \text{if } \Lambda > \frac{\lambda_b}{\theta} \end{cases}, \text{ i.e., } p = \begin{cases} V_s, & \text{if } \Lambda \leq \frac{\lambda_b}{\theta} \\ p_b, & \text{if } \Lambda > \frac{\lambda_b}{\theta} \end{cases}.$$

The corresponding joining probabilities of the customers are as follows:

$$(\delta_s(p), \delta_n(p)) = \begin{cases} (1,\ 0), & \text{if } \Lambda \leq \frac{\lambda_b}{\theta} \\ \left(\frac{\lambda_b}{\Lambda_s},\ 0\right), & \text{if } \Lambda > \frac{\lambda_b}{\theta} \end{cases}.$$

The social welfare is

$$\mathcal{SW}|_{\lambda(p)<\Lambda_n} = \begin{cases} V_s\Lambda_s, & \text{if } \Lambda \leq \frac{\lambda_b}{\theta} \\ p_b\lambda_b, & \text{if } \Lambda > \frac{\lambda_b}{\theta} \end{cases}. \tag{10}$$

Next, we consider the case of $\lambda(p) \geq \Lambda_n$. We obtain

$$\lambda(p) = \begin{cases} \Lambda, & \text{if } 0 < \Lambda \leq \lambda_b \\ \lambda_b, & \text{if } \lambda_b < \Lambda < \frac{\lambda_b}{1-\theta} \\ \Lambda_n, & \text{if } \Lambda \geq \frac{\lambda_b}{1-\theta} \end{cases}, \text{ i.e., } p = \begin{cases} R - \frac{c}{\mu-\Lambda}, & \text{if } 0 < \Lambda \leq \lambda_b \\ p_b, & \text{if } \lambda_b < \Lambda < \frac{\lambda_b}{1-\theta} \\ V_n, & \text{if } \Lambda \geq \frac{\lambda_b}{1-\theta} \end{cases}.$$

The corresponding joining probabilities of the customers are

$$(\delta_s(p), \delta_n(p)) = \begin{cases} (1,\ 1), & \text{if } 0 < \Lambda \leq \lambda_b \\ \left(\frac{\lambda_b - \Lambda_n}{\Lambda_s},\ 1\right), & \text{if } \lambda_b < \Lambda < \frac{\lambda_b}{1-\theta} \\ (0,\ 1), & \text{if } \Lambda \geq \frac{\lambda_b}{1-\theta} \end{cases}.$$

The corresponding social welfare is

$$\mathcal{SW}|_{\lambda(p)\geq\Lambda_n} = \begin{cases} \Lambda\left(R - \frac{c}{\mu-\Lambda}\right), & \text{if } 0 < \Lambda \leq \lambda_b \\ p_b\lambda_b, & \text{if } \lambda_b < \Lambda < \frac{\lambda_b}{1-\theta} \\ V_n\Lambda_n, & \text{if } \Lambda \geq \frac{\lambda_b}{1-\theta} \end{cases}. \tag{11}$$

Last, the provider compares (10) and (11) to obtain the optimal pricing strategy. We start with the case $\theta < \frac{1}{2}$:

$$\mathcal{SW}|_{p\geq V_n} - \mathcal{SW}|_{p<V_n} = \begin{cases} V_s\Lambda_s - \Lambda\left(R - \frac{c}{\mu-\Lambda}\right) < 0, & \text{if } 0 < \Lambda \leq \lambda_b \\ V_s\Lambda_s - p_b\lambda_b < 0, & \text{if } \lambda_b < \Lambda \leq \underline{\Lambda} = \frac{\lambda_b}{1-\theta} \\ V_s\Lambda_s - V_n\Lambda_n, & \text{if } \underline{\Lambda} < \Lambda \leq \overline{\Lambda} = \frac{\lambda_b}{\theta} \\ p_b\lambda_b - V_n\Lambda_n > 0, & \text{if } \Lambda > \overline{\Lambda} \end{cases}$$

Since both $\mathcal{SW}|_{p\geq V_n}$ and $\mathcal{SW}|_{p<V_n}$ are continuous, $\mathcal{SW}|_{p\geq V_n} - \mathcal{SW}|_{p<V_n}$ is continuous. It is worth mentioning the following two points here: at $\Lambda = \underline{\Lambda}$, $\Lambda_s V_s - \Lambda_n V_n = \Lambda_s V_s - \lambda_b p_b < 0$; while at $\Lambda = \overline{\Lambda}$, $\Lambda_s V_s - \Lambda_n V_n = \lambda_b p_b - \Lambda_n V_n > 0$. We then look into the interval $\underline{\Lambda} < \Lambda < \overline{\Lambda}$:

$$\Lambda_s V_s - \Lambda_n V_n = (2\theta - 1)\Lambda\left(R - \frac{c\mu}{(\mu - \theta\Lambda)(\mu - (1-\theta)\Lambda)}\right).$$

Define $g_{st}(\Lambda)$ as follows:

$$g_{st}(\Lambda) := R - \frac{c\mu}{(\mu - \theta\Lambda)(\mu - (1-\theta)\Lambda)}. \tag{12}$$

Function $g_{st}(\Lambda)$ is continuous in Λ. As $(2\theta - 1)\Lambda < 0$, the sign of $\Lambda_s V_s - \Lambda_n V_n$ is opposite to that of $g_{st}(\Lambda)$. Therefore, $g_{st}(\underline{\Lambda}) > 0$ and $g_{st}(\overline{\Lambda}) < 0$. Moreover,

$$\frac{dg_{st}(\Lambda)}{d\Lambda} = -\frac{c\mu(\mu - 2(1-\theta)\theta\Lambda)}{(\mu - \theta\Lambda)^2(\mu - (1-\theta)\Lambda)^2} < 0; \tag{13}$$

that is, $g_{st}(\Lambda)$ is monotonically decreasing in Λ. Therefore, $g_{st}(\Lambda)$ crosses 0 once at $\Lambda = \widehat{\Lambda}$, where $\widehat{\Lambda} \in (\overline{\Lambda}, \underline{\Lambda})$. We can easily obtain that

$$\widehat{\Lambda} = \frac{1}{2\theta(1-\theta)}\left(\mu - \sqrt{\mu\left[\mu - 4\theta(1-\theta)\left(\mu - \frac{c}{R}\right)\right]}\right).$$

Thus, when $\theta < \frac{1}{2}$, the optimal price and the corresponding social welfare are, respectively,

$$p^*_{sw} = \begin{cases} R - \frac{c}{\mu-\Lambda}, & \text{if } 0 < \Lambda \leq \lambda_b \\ p_b, & \text{if } \lambda_b < \Lambda \leq \underline{\Lambda} \\ V_n, & \text{if } \underline{\Lambda} < \Lambda \leq \widehat{\Lambda} \\ V_s, & \text{if } \widehat{\Lambda} < \Lambda \leq \overline{\Lambda} \\ p_b, & \text{if } \Lambda > \overline{\Lambda} \end{cases} \text{ and } \mathcal{SW}^* = \begin{cases} \Lambda\left(R - \frac{c}{\mu-\Lambda}\right), & \text{if } 0 < \Lambda \leq \lambda_b \\ p_b\lambda_b, & \text{if } \lambda_b < \Lambda \leq \underline{\Lambda} \\ V_n\Lambda_n, & \text{if } \underline{\Lambda} < \Lambda \leq \widehat{\Lambda} \\ V_s\Lambda_s, & \text{if } \widehat{\Lambda} < \Lambda \leq \overline{\Lambda} \\ p_b\lambda_b, & \text{if } \Lambda > \overline{\Lambda} \end{cases}. \tag{14}$$

For the case of $\theta \geq \frac{1}{2}$,

$$\mathcal{SW}|_{p\geq V_n} - \mathcal{SW}|_{p<V_n} = \begin{cases} V_s\Lambda_s - \Lambda\left(R - \frac{c}{\mu-\Lambda}\right) < 0, & \text{if } 0 < \Lambda \leq \lambda_b \\ V_s\Lambda_s - p_b\lambda_b < 0, & \text{if } \lambda_b < \Lambda \leq \underline{\Lambda} = \frac{\lambda_b}{\theta} \\ 0, & \text{if } \underline{\Lambda} < \Lambda \leq \overline{\Lambda} = \frac{\lambda_b}{1-\theta} \\ p_b\lambda_b - V_n\Lambda_n > 0, & \text{if } \Lambda > \overline{\Lambda} \end{cases}.$$

Then, it can be easily derived that the optimal price and social welfare are, respectively,

$$p^*_{sw} = \begin{cases} R - \frac{c}{\mu-\Lambda}, & \text{if } 0 < \Lambda \leq \lambda_b \\ p_b, & \text{if } \Lambda > \lambda_b \end{cases} \text{ and } \mathcal{SW}^* = \begin{cases} \Lambda\left(R - \frac{c}{\mu-\Lambda}\right), & \text{if } 0 < \Lambda \leq \lambda_b \\ p_b\lambda_b, & \text{if } \Lambda > \lambda_b \end{cases}. \quad (15)$$

Proof of Proposition 3

The Optimal Cyclic Pricing Strategy is High-Low Cyclic

A profit-maximizing provider does not leave any positive utility to the customers, and he charges a price as high as possible as long as the customers' joining decisions remain unaffected; that is, $v(p_i) \leq p_i$, where "=" holds whenever sophisticated customers join. According to Proposition 1, we consider the following two regions based on the magnitudes of V_n and $\eta(\boldsymbol{PT})$ and derive the local optima in each region.

Region 1: $\eta(\boldsymbol{PT}) \leq V_n$**.** According to Proposition 1, the effective arrival rate is as follows:

$$\lambda(p_i) = \begin{cases} \min\{\mu - \frac{c}{R-p_i}, \Lambda_s\} & \text{if } \eta(\boldsymbol{PT}) \leq V_n < p_i \text{ or } \eta(\boldsymbol{PT}) < p_i \leq V_n; \\ \min\{\mu - \frac{c}{R-p_i}, \Lambda\} & \text{if } p_i \leq \eta(\boldsymbol{PT}) < V_n. \end{cases}$$

Since sophisticated customers join with positive probability at any price p_i in this region, the equilibrium arrival rate must satisfy $v(p_i) = p_i$. Hence, the profit-maximizing provider's optimization problem is the same as that of the welfare-maximizing provider, which indicates that the (local) optimal pricing strategy is static in this region.

Region 2: $\eta(\boldsymbol{PT}) > V_n$**.** By Proposition 1, the effective arrival rate at each price p_i is

$$\lambda(p_i) = \begin{cases} \min\left\{\mu - \frac{c}{R-p_i}, \Lambda_s\right\}, & \text{if } p_i > \eta(\boldsymbol{PT}); \\ \Lambda_n = (1-\theta)\Lambda, & \text{if } V_n < p_i \leq \eta(\boldsymbol{PT}); \\ \min\left\{\mu - \frac{c}{R-p_i}, \Lambda\right\}, & \text{if } p_i \leq V_n. \end{cases}$$

Since sophisticated customers join with positive probability when $p_i > \eta(\boldsymbol{PT})$ or $p_i \leq V_n$, we have $v(p_i) = p_i$ in these two price sets. Hence, in the price sets of $p_i > \eta(\boldsymbol{PT})$ (labelled 'set A') and $p_i \leq V_n$ (labelled 'set C'), $\lambda(p_i) = \mu - \frac{c}{R-p_i}$ and thus, $p_i\left(\mu - \frac{c}{R-p_i}\right)$ is concave.

There exists a unique optimal price in each of these two sets, which we denote as p_A and p_C, respectively. Moreover, in the price set of $V_n < p_i \leq \eta(\boldsymbol{PT})$ (labelled 'set B'), the arrival rate is constant as Λ_n, and hence there also exists a unique optimal price in this set, which we denote as p_B. Then, the server's profit-maximizing problem can be simplified as

$$\begin{aligned}
\max_{p_i,\ L_{p_i}, i=A,B,C} \quad & \Pi = p_A\lambda(p_A)L_{p_A} + p_B\Lambda_n L_{p_B} + p_C\lambda(p_C)L_{p_C} \\
\text{s.t.} \quad & p_A > \eta(\boldsymbol{PT}), \\
& V_n < p_B \leq \eta(\boldsymbol{PT}), \\
& p_C \leq V_n, \\
& L_{p_A} + L_{p_B} + L_{p_C} = 1,\ L_{p_i} \geq 0,
\end{aligned}$$

where

$$\eta(\boldsymbol{PT}) = \frac{p_A\lambda(p_A)L_{p_A} + V_n\Lambda_n L_{p_B} + p_C\lambda(p_C)L_{p_C}}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}}. \tag{16}$$

The Lagrangian function is as follows:

$$\begin{aligned}
\mathcal{L}(\boldsymbol{PT}, \boldsymbol{\alpha}) = {} & p_A\lambda(p_A)L_{p_A} + p_B\Lambda_n L_{p_B} + p_C\lambda(p_C)L_{p_C} \\
& + \alpha_1(p_A - \eta(\boldsymbol{PT})) - \alpha_2(V_n - p_B) - \alpha_3(p_B - \eta(\boldsymbol{PT})) - \alpha_4(p_C - V_n) \\
& - \alpha_5(L_{p_A} + L_{p_B} + L_{p_C} - 1) + \alpha_6 L_{p_A} + \alpha_7 L_{p_B} + \alpha_8 L_{p_C}.
\end{aligned}$$

We then obtain the following Kuhn-Tucker conditions:

$$\frac{\partial \mathcal{L}}{\partial p_A} = \frac{d(p_A\lambda(p_A))}{dp_A}L_{p_A} + \alpha_1\left(1 - \frac{\partial\eta(\boldsymbol{PT})}{\partial p_A}\right) + \alpha_3\frac{\partial\eta(\boldsymbol{PT})}{\partial p_A} = 0; \tag{17}$$

$$\frac{\partial \mathcal{L}}{\partial p_B} = \Lambda_n L_{p_B} - \alpha_1\frac{\partial\eta(\boldsymbol{PT})}{\partial p_B} + \alpha_2 - \alpha_3\left(1 - \frac{\partial\eta(\boldsymbol{PT})}{\partial p_B}\right) = 0; \tag{18}$$

$$\frac{\partial \mathcal{L}}{\partial p_C} = \frac{d(p_C\lambda(p_C))}{dp_C}L_{p_C} - \alpha_1\frac{\partial\eta(\boldsymbol{PT})}{\partial p_C} + \alpha_3\frac{\partial\eta(\boldsymbol{PT})}{\partial p_C} - \alpha_4 = 0; \tag{19}$$

$$\frac{\partial \mathcal{L}}{\partial L_{p_A}} = p_A\lambda(p_A) - (\alpha_1 - \alpha_3)\frac{\partial\eta(\boldsymbol{PT})}{\partial L_{p_A}} - \alpha_5 + \alpha_6 = 0; \tag{20}$$

$$\frac{\partial \mathcal{L}}{\partial L_{p_B}} = p_B\Lambda_n - (\alpha_1 - \alpha_3)\frac{\partial\eta(\boldsymbol{PT})}{\partial L_{p_B}} - \alpha_5 + \alpha_7 = 0; \tag{21}$$

$$\frac{\partial \mathcal{L}}{\partial L_{p_C}} = p_C\lambda(p_C) - (\alpha_1 - \alpha_3)\frac{\partial\eta(\boldsymbol{PT})}{\partial L_{p_C}} - \alpha_5 + \alpha_8 = 0; \tag{22}$$

$$\begin{aligned}
&\alpha_1(p_A - \eta(\boldsymbol{PT})) = 0;\ \alpha_2(V_n - p_B) = 0;\ \alpha_3(p_B - \eta(\boldsymbol{PT})) = 0;\ \alpha_4(p_C - V_n) = 0; \\
&p_A - \eta(\boldsymbol{PT}) > 0;\ p_B - V_n > 0;\ p_B - \eta(\boldsymbol{PT}) \leq 0;\ p_C - V_n \leq 0; \\
&\alpha_5(L_{p_A} + L_{p_B} + L_{p_C} - 1) = 0;\ \alpha_6 L_{p_A} = 0;\ \alpha_7 L_{p_B} = 0;\ \alpha_8 L_{p_C} = 0; \alpha_i \geq 0\ (i = 1, 2...8) \\
&L_{p_A} + L_{p_B} + L_{p_C} - 1 = 0;\ 0 \leq L_{p_A} \leq 1;\ 0 \leq L_{p_B} \leq 1;\ 0 \leq L_{p_C} \leq 1;
\end{aligned}$$

where

$$\frac{\partial \eta(\boldsymbol{PT})}{\partial p_A} = \frac{L_{p_A}}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}} \left(\frac{d(p_A\lambda(p_A))}{d\lambda(p_A)} - \eta(\boldsymbol{PT}) \right) \frac{d\lambda(p_A)}{dp_A}; \quad (23)$$

$$\frac{\partial \eta(\boldsymbol{PT})}{\partial p_B} = 0;$$

$$\frac{\partial \eta(\boldsymbol{PT})}{\partial p_C} = \frac{L_{p_C}}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}} \left(\frac{d(p_C\lambda(p_C))}{d\lambda(p_C)} - \eta(\boldsymbol{PT}) \right) \frac{d\lambda(p_C)}{dp_C}; \quad (24)$$

$$\frac{\partial \eta(\boldsymbol{PT})}{\partial L_{p_A}} = \frac{\lambda(p_A)(p_A - \eta(\boldsymbol{PT}))}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}}; \quad (25)$$

$$\frac{\partial \eta(\boldsymbol{PT})}{\partial L_{p_B}} = \frac{\Lambda_n(V_n - \eta(\boldsymbol{PT}))}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}}; \quad (26)$$

$$\frac{\partial \eta(\boldsymbol{PT})}{\partial L_{p_C}} = \frac{\lambda(p_C)(p_C - \eta(\boldsymbol{PT}))}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}}. \quad (27)$$

Note that $p_A - \eta(\boldsymbol{PT}) > 0$ and $p_B - V_n > 0$ imply that $\alpha_1 = \alpha_2 = 0$. As $\frac{\partial \eta(\boldsymbol{PT})}{\partial p_B} = 0$ and $\alpha_1 = \alpha_2 = 0$, (18) is simplified as $\frac{\partial \mathcal{L}}{\partial p_B} = \Lambda_n L_{p_B} - \alpha_3 = 0$, according to which the pricing strategy is classified into two cases: (1) $L_{p_B} = 0$ and $\alpha_3 = 0$; (2) $L_{p_B} > 0$ and $\alpha_3 > 0$.

Case 1: $L_{p_B} = 0$ and $\alpha_3 = 0$. In this case, p_B becomes irrelevant. The provider allocates the pricing cycle between p_A and p_C, and sophisticated customers join at both prices, indicating that $v(p_A) = p_A$ and $v(p_C) = p_C$. Hence, maximizing profit becomes the same as maximizing welfare; thus, the optimal pricing strategy is static.

Case 2: $L_{p_B} > 0$ and $\alpha_3 > 0$. Since $\alpha_3(p_B - \eta(\boldsymbol{PT})) = 0$, we have $p_B^* = \eta(\boldsymbol{PT})$. Then, (17) and (19) are simplified as

$$\frac{\partial \mathcal{L}}{\partial p_A} = \left(1 + \frac{\alpha_3}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}}\right) \frac{d(p_A\lambda(p_A))}{dp_A} L_{p_A} - \frac{\alpha_3 L_{p_A} \eta(\boldsymbol{PT})}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}} \frac{d\lambda(p_A)}{dp_A} = 0,$$

$$\frac{\partial \mathcal{L}}{\partial p_C} = \left(1 + \frac{\alpha_3}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}}\right) \frac{d(p_C\lambda(p_C))}{dp_C} L_{p_C} - \frac{\alpha_3 L_{p_C} \eta(\boldsymbol{PT})}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}} \frac{d\lambda(p_C)}{dp_C} - \alpha_4 = 0,$$

by inserting (23) and (24), respectively. Then, we have

$$L_{p_C}\frac{\partial \mathcal{L}}{\partial p_A} - L_{p_A}\frac{\partial \mathcal{L}}{\partial p_C} = \left[\left(R + \frac{\alpha_3(R - \eta(\boldsymbol{PT}))}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}} \right) \left(\frac{c}{(R-p_C)^2} - \frac{c}{(R-p_A)^2} \right) L_{p_C} + \alpha_4 \right] L_{p_A} = 0. \quad (28)$$

As $p_C < p_A$, $\frac{c}{(R-p_C)^2} < \frac{c}{(R-p_A)^2}$. Thus, (28) implies one of the following two cases: (2-i) $L_{p_A} = 0$; (2-ii) $L_{p_A} > 0$ and $\alpha_4 > 0$.

Case 2-i: If $L_{p_A} = 0$, the provider decides how to allocate the pricing cycle between p_B and p_C. Recall that $p_C \leq V_n$. Then from (16), we have $\eta(\boldsymbol{PT}) \leq V_n$, i.e., the aforementioned Region 1 case. Thus, according to the discussion there, the optimal pricing shall be static pricing.

Case 2-ii: If $L_{p_A} > 0$ and $\alpha_4 > 0$, by $\alpha_4(p_C - V_n) = 0$, we have $p_C = V_n$. Based on (25) and (27), taking the difference between (20) and (22) yields

$$\begin{aligned}
&p_A\lambda(p_A) - p_C\lambda(p_C) + \alpha_3\left(\frac{\partial\eta(\boldsymbol{PT})}{\partial L_{p_A}} - \frac{\partial\eta(\boldsymbol{PT})}{\partial L_{p_C}}\right) + \alpha_6 - \alpha_8 \\
&= \left(1 + \frac{\alpha_3}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}}\right)(p_A\lambda(p_A) - p_C\lambda(p_C)) \\
&\qquad - \frac{\alpha_3(\lambda(p_A) - \lambda(p_C))\eta(\boldsymbol{PT})}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}} + \alpha_6 - \alpha_8 \\
&:= \mathcal{A} + \alpha_6 - \alpha_8 = 0
\end{aligned}$$

Next, we consider the sign of $\mathcal{A}$, based on which we further have three sub-cases:

1. $\mathcal{A} < 0$. Then, $\alpha_6 > 0$. By $\alpha_6 L_{p_A} = 0$, it indicates $L_{p_A} = 0$. This is the same as Case 2-i stated above. That is, the optimal pricing shall be static pricing.

2. $\mathcal{A} > 0$. Then, $\alpha_8 > 0$. By $\alpha_8 L_{p_C} = 0$, it indicates $L_{p_C} = 0$. Thus, the two prices of the cyclic pricing strategy satisfy $p_A > \eta(\boldsymbol{PT})$ and $p_B = \eta(\boldsymbol{PT})$.

3. $\mathcal{A} = 0$. Then, $\alpha_6 = \alpha_8$. We further consider the difference between (21) and (22). Based on (26) and (27), and considering $p_C = V_n$, we then have

$$\begin{aligned}
&p_B\Lambda_n - p_C\lambda(p_C) + \alpha_3\left(\frac{\partial\eta(\boldsymbol{PT})}{\partial L_{p_B}} - \frac{\partial\eta(\boldsymbol{PT})}{\partial L_{p_C}}\right) + \alpha_7 - \alpha_8 \\
&= (p_B - V_n)\Lambda_n\left(1 + \frac{\alpha_3}{\lambda(p_A)L_{p_A} + \Lambda_n L_{p_B} + \lambda(p_C)L_{p_C}}\right) + \alpha_7 - \alpha_8 = 0.
\end{aligned}$$

   As $p_B > V_n$, the above equation indicates $\alpha_8 > 0$. Since here $\alpha_6 = \alpha_8$, this implies $\alpha_6 > 0$. By $\alpha_6 L_{p_A} = 0$ and $\alpha_8 L_{p_C} = 0$, we then have $L_{p_A} = 0$ and $L_{p_C} = 0$. Therefore, $L_{p_B} = 1$; that is, the optimal pricing strategy becomes static pricing.

In summary, there are at most two prices in the optimal pricing strategy, namely, the optimal pricing strategy is either static or high-low cyclic. Provided that it is high-low cyclic, hereafter we can simplify the notations by denoting the higher price as p_h and the lower price as p_l. The two prices satisfy $p_h > \eta(\boldsymbol{PT})$ and $p_l = \eta(\boldsymbol{PT})$. Otherwise, the optimal pricing strategy always reduces to static pricing.

The Optimal Profit-Maximizing Pricing Strategy

If the profit-maximizing provider adopts a static pricing strategy, i.e., $N = 1$, since $p = v(p)$, his optimization problem shall be equivalent to that of the welfare-maximizing provider. Thus, the optimal profit-maximizing static price shall be the same as that under welfare maximization, which is given in Proposition 2. In the following analysis, we focus on obtaining the optimal cyclic pricing strategy.

We drop the subscript of L_{p_i} $(i = h,\ l)$ and use L and $1 - L$ to denote the proportion of time remaining at p_h and p_l, respectively. Then,

$$\boldsymbol{PT} = \{(p_h, L), (p_l, 1 - L)\}.$$

Since we have shown above that under the optimal cyclic pricing strategy, $p_l = \eta(\boldsymbol{PT})$, the provider's decision under a cyclic pricing strategy becomes deciding p_h and L. Therefore, his optimization problem under the cyclic pricing strategy can be simplified as follows:

$$\begin{aligned}
&\max_{p_h, L}\ \Pi_{cy} = p_h \left(\mu - \frac{c}{R - p_h}\right) L + \Lambda_n \eta(\boldsymbol{PT})(1 - L), \\
&\text{s.t.} \qquad p_h > \eta(\boldsymbol{PT}), && (29)\\
&\qquad\qquad \mu - \frac{c}{R - p_h} \leq \Lambda_s, && (30)\\
&\qquad\qquad 0 < L < 1,
\end{aligned}$$

where

$$\eta(\boldsymbol{PT}) = \frac{v(p_h)\lambda(p_h)L + V_n\Lambda_n(1 - L)}{\lambda(p_h)L + \Lambda_n(1 - L)} = \frac{p_h\left(\mu - \frac{c}{R - p_h}\right)L + V_n\Lambda_n(1 - L)}{\left(\mu - \frac{c}{R - p_h}\right)L + \Lambda_n(1 - L)}. \tag{31}$$

Since there exists a one-to-one mapping between $\lambda(p_h) = \mu - \frac{c}{R - p_h}$ and p_h, deciding $\lambda(p_h)$ is equivalent to deciding p_h. Hereafter, we use $\lambda(p_h)$ as a direct decision variable since it is more straightforward. Moreover, (29) and (30) can be simplified as:

$$\begin{cases} \lambda(p_h) \leq \Lambda_s & \text{if } \theta < \frac{1}{2} \\ \lambda(p_h) < \Lambda_n & \text{if } \theta \geq \frac{1}{2} \end{cases}.$$

Scenario $\theta < \frac{1}{2}$. The Lagrangian function can be written as

$$\mathcal{L}(p_h, L, \alpha_1, \alpha_2, \alpha_3) = p_h\lambda(p_h)L + \Lambda_n\eta(\boldsymbol{PT})(1 - L) - \alpha_1(\lambda(p_h) - \Lambda_s) + \alpha_2 L - \alpha_3(L - 1).$$

The Kuhn-Tucker conditions are as follows:

$$\frac{\partial \mathcal{L}}{\partial \lambda(p_h)} = \frac{d(\lambda(p_h)p_h)}{d\lambda(p_h)}L + \frac{\partial \eta(\boldsymbol{PT})}{\partial \lambda(p_h)}\Lambda_n(1-L) - \alpha_1 = \frac{\partial \Pi_{cy}}{\partial \lambda(p_h)} - \alpha_1 = 0; \tag{32}$$

$$\frac{\partial \mathcal{L}}{\partial L} = \lambda(p_h)p_h - \Lambda_n\eta(\boldsymbol{PT}) + \frac{\partial \eta(\boldsymbol{PT})}{\partial L}\Lambda_n(1-L) + \alpha_2 - \alpha_3 = 0; \tag{33}$$

$$\lambda(p_h) \le \Lambda_s; \alpha_1(\lambda(p_h) - \Lambda_s) = 0; \tag{34}$$

$$0 < L < 1;\ \ \alpha_2 L = 0;\ \ \alpha_3(L-1) = 0; \alpha_1, \alpha_2, \alpha_3 \ge 0,$$

where

$$\frac{\partial \eta(\boldsymbol{PT})}{\partial \lambda(p_h)} = \frac{L}{\Lambda_n(1-L) + \lambda(p_h)L}\left(\frac{d(\lambda(p_h)p_h)}{d\lambda(p_h)} - \eta(\boldsymbol{PT})\right); \tag{35}$$

$$\frac{\partial \eta(\boldsymbol{PT})}{\partial L} = \frac{\lambda(p_h)(p_h - \eta(\boldsymbol{PT})) + \Lambda_n(\eta(\boldsymbol{PT}) - V_n)}{\Lambda_n(1-L) + \lambda(p_h)L} = \frac{\lambda(p_h)\Lambda_n(p_h - V_n)}{(\Lambda_n(1-L) + \lambda(p_h)L)^2}. \tag{36}$$

A solution $(\lambda(p_h^*), L^*)$ that maximizes the provider's profit shall satisfy all above conditions. Note that in order to make a cyclic pricing strategy viable, we shall have $0 < L < 1$; that is, it requires $\alpha_2 = \alpha_3 = 0$. Otherwise, cyclic pricing degenerates into static pricing.

We start with the requirements of $\alpha_2 = \alpha_3 = 0$, under which (33) is simplified as:

$$\frac{\partial \mathcal{L}}{\partial L} = \frac{\partial \Pi_{cy}}{\partial L} = \lambda(p_h)p_h - \Lambda_n\eta(\boldsymbol{PT}) + \frac{\partial \eta(\boldsymbol{PT})}{\partial L}\Lambda_n(1-L) = 0. \tag{37}$$

Moreover,

$$\frac{\partial^2 \eta(\boldsymbol{PT})}{\partial L^2} = 2\frac{\Lambda_n - \lambda(p_h)}{\Lambda_n(1-L) + \lambda(p_h)L}\frac{\partial \eta(\boldsymbol{PT})}{\partial L};$$

$$\frac{\partial^2 \mathcal{L}}{\partial L^2} = \frac{\partial^2 \Pi_{cy}}{\partial L^2} = -2\Lambda_n\frac{\partial \eta(\boldsymbol{PT})}{\partial L} + \frac{\partial^2 \eta(\boldsymbol{PT})}{\partial L^2}\Lambda_n(1-L) = -\frac{2\Lambda_n\lambda(p_h)}{\Lambda_n(1-L) + \lambda(p_h)L}\frac{\partial \eta(\boldsymbol{PT})}{\partial L} < 0, \tag{38}$$

since $\frac{\partial \eta(\boldsymbol{PT})}{\partial L} > 0$ (see (36)) whenever $\theta < \frac{1}{2}$, as $\lambda(p_h) \le \Lambda_s < \Lambda_n$ implies $p_h > V_n$. This implies that $\frac{\partial \Pi_{cy}}{\partial L}$ $(\frac{\partial \mathcal{L}}{\partial L})$ decreases in L. Moreover, by (31), $\lim_{L\to 1} \eta(\boldsymbol{PT}) = p_h$ and $\lim_{L\to 0} \eta(\boldsymbol{PT}) = V_n$, and by (36),

$$\lim_{L\to 1} \frac{\partial \eta(\boldsymbol{PT})}{\partial L} = \frac{\Lambda_n}{\lambda(p_h)}(p_h - V_n),\ \ \lim_{L\to 0} \frac{\partial \eta(\boldsymbol{PT})}{\partial L} = \frac{\lambda(p_h)}{\Lambda_n}(p_h - V_n).$$

If we define the following function

$$\phi(L, \Lambda) := \frac{\partial \Pi_{cy}}{\partial L} = \lambda(p_h)p_h - \Lambda_n\eta(\boldsymbol{PT}) + \frac{\partial \eta(\boldsymbol{PT})}{\partial L}\Lambda_n(1-L),$$

then $\phi(L, \Lambda)$ decreases in L, and

$$\phi(0, \Lambda) = \lambda(p_h)(2p_h - V_n) - V_n\Lambda_n; \tag{39}$$

$$\phi(1, \Lambda) = p_h(\lambda(p_h) - \Lambda_n).$$

Thus, to ensure that equation (37) has a solution satisfying $0 < L < 1$ (the cyclic pricing requirement), $\phi(0,\Lambda) > 0$ and $\phi(1,\Lambda) < 0$ are required. Note that $\phi(1,\Lambda) < 0$ always holds. We only need to determine the market condition under which $\phi(0,\Lambda) > 0$ holds.

Next, consider the requirement on α_1. By (32), we have $\alpha_1 = \frac{\partial \Pi_{cy}}{\partial \lambda(p_h)}$, and hence,

$$
\begin{aligned}
\frac{\partial \alpha_1}{\partial \lambda(p_h)} &= \frac{\partial^2 \Pi_{cy}}{\partial \lambda(p_h)^2} \\
&= \frac{L}{\Lambda_n(1-L)+\lambda(p_h)L}\Bigg[(2\Lambda_n(1-L)+\lambda(p_h)L)\frac{d^2(p_h\lambda(p_h))}{d\lambda(p_h)^2} \\
&\qquad - \frac{2\Lambda_n(1-L)L}{\Lambda_n(1-L)+\lambda(p_h)L}\left(\frac{d(p_h\lambda(p_h))}{d\lambda(p_h)} - \eta(\boldsymbol{PT})\right)\Bigg] \\
&< \frac{2\Lambda_n(1-L)L}{(\Lambda_n(1-L)+\lambda(p_h)L)^2}\Bigg[(\Lambda_n(1-L)+\lambda(p_h)L)\frac{d^2(p_h\lambda(p_h))}{d\lambda(p_h)^2} \\
&\qquad - L\left(\frac{d(p_h\lambda(p_h))}{d\lambda(p_h)} - \eta(\boldsymbol{PT})\right)\Bigg] \\
&< \frac{2\Lambda_n(1-L)L^2}{(\Lambda_n(1-L)+\lambda(p_h)L)^2}\left(\lambda(p_h)\frac{d^2(p_h\lambda(p_h))}{d\lambda(p_h)^2} - \frac{d(p_h\lambda(p_h))}{d\lambda(p_h)} + \eta(\boldsymbol{PT})\right) \\
&< \frac{2\Lambda_n(1-L)L^2}{(\Lambda_n(1-L)+\lambda(p_h)L)^2}\left(\lambda(p_h)\frac{d^2(p_h\lambda(p_h))}{d\lambda(p_h)^2} - \frac{d(p_h\lambda(p_h))}{d\lambda(p_h)} + p_h\right) \\
&= -\frac{2\Lambda_n(1-L)L^2}{(\Lambda_n(1-L)+\lambda(p_h)L)^2}\frac{c\lambda(p_h)(\mu+\lambda(p_h))}{(\mu-\lambda(p_h))^3} < 0.
\end{aligned}
\tag{40}
$$

The first and second "<" follow $\frac{d^2(p_h\lambda(p_h))}{d\lambda(p_h)^2} = -\frac{2c\mu}{(\mu-\lambda(p_h))^3} < 0$, and the third follows $\eta(\boldsymbol{PT}) < p_h$. Thus, α_1 decreases in $\lambda(p_h)$.

Whenever the cyclic pricing strategy is feasible, by (32) and (35), we have

$$
\frac{d(p_h\lambda(p_h))}{d\lambda(p_h)}|_{(p_h^*,\ L^*)} = \left(\frac{\Lambda_n(1-L)\eta(\boldsymbol{PT})}{\lambda(p_h)L+2\Lambda_n(1-L)} + \frac{\Lambda_n(1-L)+\lambda(p_h)L}{\lambda(p_h)L+2\Lambda_n(1-L)}\frac{\alpha_1}{L}\right)|_{(p_h^*,\ L^*)}. \tag{41}
$$

As $\alpha_1 \geq 0$ and $0 < L < 1$, $\frac{d(p_h\lambda(p_h))}{d\lambda(p_h)}|_{(p_h^*,\ L^*)} > 0$. Recall that $\frac{d(p\lambda(p))}{d\lambda(p)}|_{\lambda(p)=\lambda(p_b)} = 0$, where $p_b = R - \sqrt{\frac{cR}{\mu}}$. Therefore, $\lambda(p_h^*) < \lambda(p_b)$. This implies that $p_h^* > p_b$.

We consider the following three cases according to whether $\lambda(p_h) \leq \Lambda_s$ is binding or not.

Case 1: $\alpha_1 > 0$. It indicates $\lambda(p_h) = \Lambda_s$, or equivalently, $p_h^* = V_s$. Then, through some simple algebra, (39)can be simplified as:

$$
\phi(0,\Lambda) = (2\theta-1)\Lambda\left(R - \frac{c(\mu+\theta\Lambda)}{(\mu-\theta\Lambda)(\mu-(1-\theta)\Lambda)}\right).
$$

Define

$$
g_{cy}(\Lambda) := R - \frac{c(\mu+\theta\Lambda)}{(\mu-\theta\Lambda)(\mu-(1-\theta)\Lambda)}. \tag{42}
$$

It is easy to show $g_{cy}(\Lambda)$ is continuous and decreasing in Λ whenever $V_n \geq 0$. Let $g_{cy}(\dot{\Lambda}) = 0$. We obtain

$$\dot{\Lambda} = \frac{1}{2\theta(1-\theta)} \left(\mu + \frac{c\theta}{R} - \sqrt{\left(\mu + \frac{c\theta}{R}\right)^2 - 4\theta(1-\theta)\mu\left(\mu - \frac{c}{R}\right)} \right)$$

on its domain. Since $g_{cy}(\Lambda)$ is decreasing and $g_{cy}(\dot{\Lambda}) = 0$, $g_{cy}(\Lambda) < 0$ when $\Lambda > \dot{\Lambda}$. Also, considering that $(2\theta - 1)\Lambda < 0$ $(\theta < \frac{1}{2})$, the sign of $\phi(0, \Lambda)$ is opposite to that of $g_{cy}(\Lambda)$. Therefore, $\phi(0, \Lambda) > 0$ when $\Lambda > \dot{\Lambda}$. In other words, only if $\Lambda > \dot{\Lambda}$ will there be a feasible cyclic pricing strategy.

Denote the optimal pricing decisions provided $\alpha_1 > 0$ and $\Lambda > \dot{\Lambda}$ as $(p_h^*, L^*) = (V_s, L_b)$. Then L_b solves $\frac{\partial \Pi_{cy}}{\partial L}|_{p_h^*=V_s} = 0$ where $\Lambda > \dot{\Lambda}$; specifically,

$$V_s\Lambda_s + \frac{\Lambda_n(1-L_b)(V_s\Lambda_s - V_n\Lambda_n)}{\Lambda_n(1-L_b) + \Lambda_s L_b} - \frac{\Lambda_s\Lambda_n(V_s\Lambda_s L_b + V_n\Lambda_n(1-L_b))}{(\Lambda_n(1-L_b) + \Lambda_s L_b)^2} = 0. \tag{43}$$

Case 2: $\alpha_1 = 0$ and $\lambda(p_h) = \Lambda_s = \theta\Lambda$. This captures the "binding but irrelevant" case, under which the optimal solution is $(p_h^*,\ L^*) = (V_s,\ L_b)$, where L_b is determined by (43). Plugging $(p_h^*,\ L^*) = (V_s,\ L_b)$ into $\alpha_1 = \frac{\partial \Pi_{cy}}{\partial \lambda(p_h)}$ and let $\alpha_1|_{\Lambda=\ddot{\Lambda}} = 0$, we get

$$\left(R - \frac{c\mu}{(\mu - \Lambda_s)^2} - \frac{\Lambda_n(1-L_b)}{2\Lambda_n(1-L_b) + \Lambda_s L_b} \frac{V_s\Lambda_s L_b + V_n\Lambda_n(1-L_b)}{\Lambda_n(1-L_b) + \Lambda_s L_b} \right)|_{\Lambda=\ddot{\Lambda}} = 0. \tag{44}$$

Similar to the aforementioned analysis, we can show that the cyclic pricing strategy is feasible only when $\Lambda > \dot{\Lambda}$. Recall that α_1 decreases in $\lambda(p_h)$. As $\lambda(p_h) = \theta\Lambda$, α_1 decreases in Λ. In the above case 1, $\alpha_1|_{\lambda(p_h)=\theta\dot{\Lambda}} > 0$ and here $\alpha_1|_{\lambda(p_h)=\theta\ddot{\Lambda}} = 0$. These imply that $\dot{\Lambda} < \ddot{\Lambda}$. The optimal solution is indeed feasible. Moreover, since $\lambda(p_h^*) = \theta\ddot{\Lambda} < \lambda(p_b) = \lambda_b$, $\ddot{\Lambda} < \frac{\lambda_b}{\theta}$. As $\overline{\Lambda} = \frac{\lambda_b}{\theta}$ if $\theta < \frac{1}{2}$, we can see $\ddot{\Lambda} < \overline{\Lambda}$.

Case 3: $\alpha_1 = 0$ and $\lambda(p_h) < \Lambda_s$. This captures the unbinding case, where $\lambda(p_h) = \mu - \frac{c}{R-p_h} < \Lambda_s$. As $\alpha_1|_{\lambda(p_h)=\theta\ddot{\Lambda}} = 0$ and $\alpha_1|_{\lambda(p_h)=\mu-\frac{c}{R-p_h}<\theta\Lambda} = 0$, we get that $\Lambda > \ddot{\Lambda}$ is required. The optimal solution, if existing, shall be obtained in the interior, solving (37) and the following equation simultaneously:

$$\frac{\partial \mathcal{L}}{\partial \lambda(p_h)} = \frac{d(\lambda(p_h)p_h)}{d\lambda(p_h)} L + \frac{\partial \eta(\boldsymbol{PT})}{\partial \lambda(p_h)} \Lambda_n(1-L) = 0. \tag{45}$$

Denote the interior solution as (p_h^0, L^0).

Last, we check the Kuhn-Tucker conditions on L. We derive from (39) that

$$\frac{\partial \phi(0, \Lambda)}{\partial \lambda(p_h)} = 2\frac{d(\lambda(p_h)p_h)}{d\lambda(p_h)} - V_n.$$

Under the unbinding case, by (41) and $\alpha_1 = 0$, the effective arrival rate at the high price always satisfies

$$\frac{d(p_h\lambda(p_h))}{d\lambda(p_h)} = \frac{\Lambda_n(1-L)\eta(\boldsymbol{PT})}{\lambda(p_h)L + 2\Lambda_n(1-L)} > \frac{V_n}{2}.$$

Therefore, $\frac{\partial\phi(0,\Lambda)}{\partial\lambda(p_h)} > 0$. Besides, from the above analysis, we have that $g_{cy}(\Lambda)$ decreases in Λ and $g_{cy}(\dot{\Lambda}) = 0$. As $\ddot{\Lambda} > \dot{\Lambda}$, $g_{cy}(\ddot{\Lambda}) < g_{cy}(\dot{\Lambda}) = 0$. Consequently, it can be shown that $\phi(0,\ddot{\Lambda}) > \phi(0,\dot{\Lambda}) = 0$. Since $\phi(0,\Lambda)$ is increasing in $\lambda(p_h)$ and $\lambda(p_h)$ weakly increases in Λ, we have $\phi(0,\Lambda) > 0$ when $\Lambda > \ddot{\Lambda}$. This ensures that indeed there exists an optimal $L^0 \in (0,1)$. Thus, the interior solution exists when $\Lambda > \ddot{\Lambda}$.

Scenario $\theta \geq \frac{1}{2}$. The Kuhn-Tucker conditions remain the same as those of Scenario $\theta < \frac{1}{2}$ except that the constraints in (34) change to

$$\lambda(p_h) - \Lambda_n < 0;\ \alpha_1(\lambda(p_h) - \Lambda_n) = 0.$$

Hence, it requires $\alpha_1 = 0$. Below we consider the following two cases: *binding but irrelevant* and *unbinding*.

Under binding but irrelevant case, $\alpha_1 = 0$ and $\lambda(p_h) = \Lambda_n$. (33) then can be simplified as:

$$\frac{\partial\mathcal{L}}{\partial L} = \alpha_2 - \alpha_3 = 0,$$

which requires $\alpha_2 = \alpha_3$. As $0 < L < 1$, $\alpha_2 L = 0$ and $\alpha_3(1-L) = 0$ are required, we then have $\alpha_2 = \alpha_3 = 0$. The Kuhn-Tucker conditions on L are all satisfied. Next, consider the Kuhn-Tucker conditions on p_h (32). By (32) and $\lambda(p_h) = \Lambda_n$,

$$\alpha_1 = L\left((2-L)\left(R - \frac{c\mu}{(\mu-\Lambda_n)^2}\right) - V_n(1-L)\right) := L\psi(\Lambda, L).$$

As $0 < L < 1$, solving $\alpha_1 = 0$ is equivalent to solving $\psi(\Lambda = \breve{\Lambda}, L) = 0$, from which we get

$$\breve{\Lambda} = \frac{1}{1-\theta}\left(\mu + \frac{c(1-L)}{R} - \sqrt{\left(\frac{c(1-L)}{2R}\right)^2 + \frac{c\mu(2-L)}{R}}\right).$$

We can also show that

$$\frac{\partial\psi(\Lambda,L)}{\partial L} = \frac{c\Lambda_n}{(\mu-\Lambda_n)^2} > 0;\quad \frac{\partial\psi(\Lambda,L)}{\partial\Lambda} = -\frac{c\,(1-\theta)}{(\mu-\Lambda_n)^2}\left((2-L)\frac{2\mu}{\mu-\Lambda_n} - (1-L)\right) < 0.$$

Then, by the implicit function theorem, we have:

$$\frac{d\breve{\Lambda}}{dL} = -\left.\frac{\frac{\partial\psi(\Lambda,L)}{\partial L}}{\frac{\partial\psi(\Lambda,L)}{\partial\Lambda}}\right|_{\Lambda=\breve{\Lambda}} > 0.$$

That is, $\breve{\Lambda}$ increases in L. Furthermore, when $L \to 1$, $\breve{\Lambda} \to \frac{1}{1-\theta}\left(\mu - \sqrt{\frac{c\mu}{R}}\right) = \overline{\Lambda}$, as $\theta \geq \frac{1}{2}$. Therefore, $\breve{\Lambda} < \overline{\Lambda}$.

Next, consider the unbinding case, where $\alpha_1 = 0$ and $\lambda(p_h) < \Lambda_n$. Similar to the analysis of Case 3 under Scenario $\theta < \frac{1}{2}$, here we can show that to ensure that $\lambda(p_h) < \Lambda_n$ holds, $\Lambda > \breve{\Lambda}$ is required. Besides, since $\lambda(p_h) = \Lambda_n$ when $\Lambda = \breve{\Lambda}$ as shown above, it can be easily obtained that $\phi(0, \ \breve{\Lambda}) = 0$ (see (39)). When $\alpha_1 = 0$, recall that as shown in Case 3 of Scenario $\theta < \frac{1}{2}$, $\frac{\partial \phi(0,\Lambda)}{\partial \lambda(p_h)} > 0$ and $\lambda(p_h)$ weakly increases in Λ. Thus, $\phi(0, \Lambda) > 0$ when $\Lambda > \breve{\Lambda}$. This ensures that indeed there exists an optimal $L^0 \in (0, 1)$. All the optimization conditions are satisfied. Thus, the interior solution does exist when $\Lambda > \breve{\Lambda}$.

To summarize, we show that the cyclic pricing strategy is feasible, and

1. if $\theta < \frac{1}{2}$ and $\dot{\Lambda} < \Lambda \leq \ddot{\Lambda}$ ($\ddot{\Lambda} < \overline{\Lambda}$), it is determined by a corner solution $(p_h^*, L^*) = (V_s, L_b)$, where L_b is determined by (43) and $p_h^* > p_b$.

2. if $\theta < \frac{1}{2}$ and $\Lambda > \ddot{\Lambda}$, or if $\theta \geq \frac{1}{2}$ and $\Lambda > \breve{\Lambda}$, it is determined by an interior solution $(p_h^*, L^*) = (p_h^0, L^0)$, where p_h^0 and L^0 are determined by (37) and (45) simultaneously.

Then, it can be easily shown that the interior solution (p_h^0, L^0) is obtained whenever $\Lambda \geq \overline{\Lambda}$, regardless of the magnitude of θ as $\ddot{\Lambda} < \overline{\Lambda}$ and $\breve{\Lambda} < \overline{\Lambda}$. Note that the above solution offers a necessary condition for the optimal cyclic pricing strategy. Considering $\frac{\partial^2 \Pi_{cy}}{\partial \lambda(p_h(\Lambda))^2} < 0$ and $\frac{\partial^2 \Pi_{cy}}{\partial L^2} < 0$ (see (38) and (40), respectively), the solution we obtain from the above Lagrangian method is unique. Also, note that the static pricing strategy is actually a boundary case of the cyclic pricing strategy with $L = 0$ or 1. Following the technique used by Chen and Wan (2003) and Yang et al. (2018), if the cyclic pricing strategy we obtain from the above Kuhn-Tucker conditions outperforms the boundary cases (i.e., the static pricing strategy), it indicates that it is indeed optimal. Below, we identify market conditions under which the cyclic pricing strategy we obtain from the above Lagrangian method outperforms the static pricing strategy and hence is indeed optimal.

Define $\Pi_{cy}^*(\Lambda)$ as the profit under the cyclic pricing strategy, i.e.,

$$\Pi_{cy}^*(\Lambda) = \Pi_{cy}(p_h^*(\Lambda), L^*(\Lambda)).$$

Function $\Pi_{cy}^*(\Lambda)$ is continuous in Λ. Taking the first- and second-order derivatives of $\Pi_{cy}^*(\Lambda)$

with respect to Λ, we obtain

$$\begin{aligned}\frac{d\Pi^*_{cy}(\Lambda)}{d\Lambda} &= \frac{\partial \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial L^*(\Lambda)}\frac{dL^*(\Lambda)}{d\Lambda} + \frac{\partial \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \lambda(p^*_h(\Lambda))}\frac{d\lambda(p^*_h(\Lambda))}{d\Lambda}\\ &\quad + \frac{\partial \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \Lambda_n}\frac{d\Lambda_n}{d\Lambda},\\ \frac{d^2\Pi^*_{cy}(\Lambda)}{d\Lambda^2} &= \frac{\partial^2 \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial L^*(\Lambda)^2}\left(\frac{dL^*(\Lambda)}{d\Lambda}\right)^2 + \frac{\partial \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial L^*(\Lambda)}\frac{d^2L^*(\Lambda)}{d\Lambda^2}\\ &\quad + \frac{\partial^2 \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \lambda(p^*_h(\Lambda))^2}\left(\frac{d\lambda(p^*_h(\Lambda))}{d\Lambda}\right)^2 + \frac{\partial \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \lambda(p^*_h(\Lambda))}\frac{d^2\lambda(p^*_h(\Lambda))}{d\Lambda^2}\\ &\quad + \frac{\partial^2 \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \Lambda_n^2}\left(\frac{d\Lambda_n}{d\Lambda}\right)^2 + \frac{\partial \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \Lambda_n}\frac{d^2\Lambda_n}{d\Lambda^2}. \qquad (46)\end{aligned}$$

Obviously, $\Lambda_n = (1-\theta)\Lambda$, and hence $\frac{d^2\Lambda_n}{d\Lambda^2} = 0$. Moreover,

$$\begin{aligned}\frac{\partial \eta(\boldsymbol{PT})}{\partial \Lambda_n} &= \frac{1-L}{\Lambda_n(1-L)+\lambda(p_h)L}\left(R - \frac{c\mu}{(\mu-\Lambda_n)^2} - \eta(\boldsymbol{PT})\right) < 0, \qquad (47)\\ \frac{\partial^2 \eta(\boldsymbol{PT})}{\partial \Lambda_n^2} &= \frac{1-L}{\Lambda_n(1-L)+\lambda(p_h)L}\left(-\frac{2c\mu}{(\mu-\Lambda_n)^3} - 2\frac{\partial \eta(\boldsymbol{PT})}{\partial \Lambda_n}\right),\\ \frac{\partial \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \Lambda_n} &= \left.\left(\eta(\boldsymbol{PT})(1-L) + \Lambda_n(1-L)\frac{\partial \eta(\boldsymbol{PT})}{\partial \Lambda_n}\right)\right|_{(p^*_h, L^*)},\\ \frac{\partial^2 \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \Lambda_n^2} &= \left.\left(2(1-L)\frac{\partial \eta(\boldsymbol{PT})}{\partial \Lambda_n} + \Lambda_n(1-L)\frac{\partial^2 \eta(\boldsymbol{PT})}{\partial \Lambda_n^2}\right)\right|_{(p^*_h, L^*)} < 0. \qquad (48)\end{aligned}$$

The "<" in (47) and (48) follows from the fact that $R - \frac{c\mu}{(\mu-\Lambda_n)^2} < R - \frac{c}{\mu-\Lambda_n} = V_n$ and the requirement of $V_n < \eta(\boldsymbol{PT})$ for the cyclic pricing strategy to be feasible. In addition, we have

$$\frac{\partial \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial L^*(\Lambda)} = 0 \text{ and } \frac{\partial^2 \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial L^*(\Lambda)^2} < 0;$$

see (37) and (38). When a corner solution is obtained, $\lambda(p^*_h(\Lambda)) = \theta\Lambda$, indicating $\frac{d^2\lambda(p^*_h(\Lambda))}{d\Lambda^2} = 0$, whereas when an interior solution is obtained,

$$\frac{\partial \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \lambda(p^*_h(\Lambda))} = 0.$$

Moreover, from (40), we have

$$\frac{\partial^2 \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \lambda(p^*_h(\Lambda))^2} < 0.$$

Therefore, we simplify (46) as:

$$\begin{aligned}\frac{d^2\Pi^*_{cy}(\Lambda)}{d\Lambda^2} &= \frac{\partial^2 \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial L^*(\Lambda)^2}\left(\frac{dL^*(\Lambda)}{d\Lambda}\right)^2 + \frac{\partial^2 \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \lambda(p^*_h(\Lambda))^2}\left(\frac{d\lambda(p^*_h(\Lambda))}{d\Lambda}\right)^2\\ &\quad + \frac{\partial^2 \Pi_{cy}(p^*_h(\Lambda), L^*(\Lambda))}{\partial \Lambda_n^2}\left(\frac{d\Lambda_n}{d\Lambda}\right)^2 < 0. \qquad (49)\end{aligned}$$

That is, $\Pi^*_{cy}(\Lambda)$ is concave in Λ.

We have obtained that the interior solution (p_h^0, L^0) is obtained whenever $\Lambda \geq \overline{\Lambda}$, regardless of the magnitude of θ. Moreover, recall that we have obtained in Proposition 2 that when $\Lambda \geq \overline{\Lambda}$, $\Pi^*_{st}(\Lambda) = p_b\lambda_b$ regardless of the magnitude of θ. Therefore, when $\Lambda \geq \overline{\Lambda}$, we have

$$\Pi^*_{cy}(\Lambda) = \Pi_{cy}(p_h^0(\Lambda), L^0(\Lambda)) > \max_{p_h} \Pi_{cy}(p_h(\Lambda), 1) = p_b\lambda_b = \Pi^*_{st}(\Lambda). \tag{50}$$

That is, the cyclic pricing strategy is always preferred when $\Lambda \geq \overline{\Lambda}$.

When $\Lambda < \overline{\Lambda}$, we consider the following two cases according to the magnitude of θ.

Case 1: $\theta < \frac{1}{2}$ and $\dot{\Lambda} < \Lambda < \overline{\Lambda}$

Comparing $g_{cy}(\Lambda)$ defined in (42) with $g_{st}(\Lambda)$ defined in (12), we can easily obtain that $g_{cy}(\Lambda) < g_{st}(\Lambda)$. Thus, $g_{cy}(\dot{\Lambda}) = 0$ indicates $g_{st}(\dot{\Lambda}) > 0$. Recall that $g_{st}(\Lambda)$ decreases in Λ (see (13)) and $g_{st}(\widehat{\Lambda}) = 0$. Thus, $\dot{\Lambda} < \widehat{\Lambda}$. Moreover, it can be easily shown that $\dot{\Lambda} > \lambda_b$. That is, $\lambda_b < \dot{\Lambda} \leq \widehat{\Lambda}$.

From (14), we can show that in the interval $(\lambda_b,\ \overline{\Lambda})$, the profit under the optimal static pricing strategy $\Pi^*_{st}(\Lambda)$, equal to the corresponding social welfare, first (weakly) decreases and then increases in Λ. Recall that $\Pi^*_{cy}(\Lambda)$ is concave in Λ; see (49). Hence, $\Pi^*_{st}(\Lambda)$ and $\Pi^*_{cy}(\Lambda)$ cross each other at most twice when $\dot{\Lambda} < \Lambda \leq \overline{\Lambda}$. Recall that $\phi(0, \dot{\Lambda}) = 0$ when $\theta < \frac{1}{2}$, which implies $L^* = 0$ at $\Lambda = \dot{\Lambda}$. Thus, based on (14) and by $\lambda_b < \dot{\Lambda} \leq \widehat{\Lambda}$, we have:

$$\lim_{\Lambda \to \dot{\Lambda}^-} \Pi^*_{cy}(\Lambda) = \lim_{\Lambda \to \dot{\Lambda}^-} \Pi_{cy}(p_h^*(\Lambda), L^*(\Lambda)) = \Pi_{cy}(V_s,\ 0) = V_n\Lambda_n \leq \lim_{\Lambda \to \dot{\Lambda}^-} \Pi^*_{st}(\Lambda) = \Pi^*_{st}(\dot{\Lambda}).$$

We also show in (50) that when $\Lambda \geq \overline{\Lambda}$, $\Pi^*_{cy}(\Lambda) > \Pi_{st}(\Lambda)$. Thus, considering the continuity of $\Pi_{cy}(\Lambda)$, we obtain

$$\lim_{\Lambda \to \overline{\Lambda}^+} \Pi^*_{cy}(\Lambda) = \Pi^*_{cy}(\overline{\Lambda}) > \lim_{\Lambda \to \overline{\Lambda}^+} \Pi^*_{st}(\Lambda) = \Pi^*_{st}(\overline{\Lambda}).$$

Hence, we can see that $\Pi^*_{cy}(\Lambda)$ crosses $\Pi^*_{st}(\Lambda)$ once from below. Therefore, when $\theta < \frac{1}{2}$, there exists a $\dot{\Lambda} < \widetilde{\Lambda} < \overline{\Lambda}$ such that $\Pi_{cy}(\Lambda) \geq \Pi_{st}(\Lambda)$ when $\widetilde{\Lambda} < \Lambda < \overline{\Lambda}$ and $\Pi_{cy}(\Lambda) < \Pi_{st}(\Lambda)$ when $\dot{\Lambda} < \Lambda < \widetilde{\Lambda}$.

Case 2: $\theta \geq \frac{1}{2}$ and $\breve{\Lambda} < \Lambda < \overline{\Lambda}$

When $\theta \geq \frac{1}{2}$, the optimal static profit is given by (15), which is weakly increasing and concave in Λ. Since $\Pi^*_{cy}(\Lambda)$ is also concave in Λ; see (49), $\Pi^*_{st}(\Lambda)$ and $\Pi^*_{cy}(\Lambda)$ cross each other at most twice. Recall from the analysis of Scenario $\theta \geq \frac{1}{2}$ above, $\lambda(p_h^*(\Lambda)) = \Lambda_n$ when $\Lambda = \breve{\Lambda}$. Thus,

$$\lim_{\Lambda \to \breve{\Lambda}^-} \Pi_{cy}^*(\breve{\Lambda}) = \max_L \Pi_{cy}(p_h^*(\breve{\Lambda}), L) = \Lambda_n V_n \leq \lim_{\Lambda \to \breve{\Lambda}^-} \Pi_{st}^*(\Lambda) = \Pi_{st}^*(\breve{\Lambda}).$$

At the same time, by the continuity of $\Pi_{cy}(\Lambda)$ and using (50), we have:

$$\lim_{\Lambda \to \overline{\Lambda}^+} \Pi_{cy}^*(\Lambda) = \Pi_{cy}^*(\overline{\Lambda}) > \lim_{\Lambda \to \overline{\Lambda}^+} \Pi_{st}^*(\Lambda) = \Pi_{st}^*(\overline{\Lambda}).$$

Therefore, when $\theta \geq \frac{1}{2}$, there exists $\breve{\Lambda} < \widetilde{\Lambda} < \overline{\Lambda}$ such that $\Pi_{cy}(\Lambda) \geq \Pi_{st}(\Lambda)$ when $\widetilde{\Lambda} \leq \Lambda < \overline{\Lambda}$ and $\Pi_{cy}(\Lambda) < \Pi_{st}(\Lambda)$ when $\breve{\Lambda} < \Lambda < \widetilde{\Lambda}$.

Proposition 3 is thus completely proved.

Proof of Corollary 1

When $\theta \geq 0.5$, we have $\widetilde{\Lambda} < \overline{\Lambda}$ (see the proof of Proposition 3). Clearly, $\lambda(p_h^*) < \lambda(p_l^*)$. And the effective arrival rate at the low price phase satisfies $\lambda(p_l^*) = \Lambda_n = (1-\theta)\Lambda < (1-\theta)\overline{\Lambda} = \lambda_b$.

Proof of Proposition 4

When observing a price p_i, a naive customer joins if and only if

$$\eta'(\boldsymbol{PT}) + \bar{p} - p_i \equiv \eta(\boldsymbol{PT}) - p_i \geq 0, \tag{51}$$

where $\bar{p} - p_i > 0$ (< 0, respectively) is the additional gain (loss, respectively) the customer obtains due to the price reduction (increase, respectively) with respect to the average one. Obviously, (51) is equivalent to (3). Sophisticated customers, on the other hand, makes their joining-or-balking decisions to maximize their expected utilities based on the service-related information by taking naive customers' joining decisions into consideration, which still can be written as (4). In other words, the decision rules of both types of customers stay the same as those stated in §3.1. Thus, the joining decisions of both types of customers with ratings on both the price and the net utility remain the same as those stated in Proposition 1 when customers rate over the consumption utility.

Appendix B: Discussion on the Average Rating and How it is Formed

In this paper, we directly assume that incoming customers are informed about the average rating of the system, which is static in the long run. Here, we relax this assumption and illustrate that the average rating can still be achieved through a convergence process when customers adopt exponential smoothing to aggregate the recent review data with the historical data. We have the following assumptions about customers:

1. A customer observes all the rating information up to her arrival time and adopts an exponential smoothing method to compute the "average rating".

2. All customers, regardless of their types, understand the randomness embedded in the service process; thus, naive customers will not consider joining a service until a considerable number of ratings have been accumulated so that the fluctuation in the service process is averaged out.

Customers adopt an exponential smoothing method in calculating their expected rating. Consider a customer arriving at t. She divides all the historical ratings up to t into two parts. The first part consists of ratings within the time period $[t-T, t]$, where T is the same as the aforementioned pricing cycle length, and the second part includes ratings in the time period $[0, t-T]$. In the example of Yelp.com, the pricing cycle length is one month.

In this continuous-time model, at any time t, the average rating for the current time period $[t-T, t]$ is always the average rating of an intact pricing cycle; see Figure 7 for the illustration.

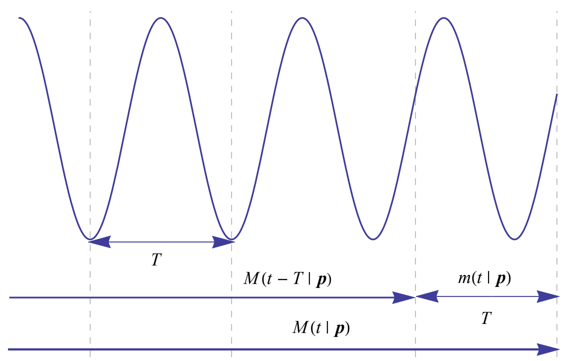

Figure 7: An Illustration of Average Rating Computation

Denote the average rating during the time period $[t-T, t]$ by $m(t|\boldsymbol{PT})$ and the average rating up to time t by $M(t|\boldsymbol{PT})$. Then, we have

$$M(t|\boldsymbol{PT}) = \alpha m(t|\boldsymbol{PT}) + (1-\alpha)M(t-T|\boldsymbol{PT}), \tag{52}$$

where α $(0 < \alpha < 1)$ is the smoothing factor and represents the weight the customer puts on the recent rating information.

Now, we can show that the long-run average rating $M(t|\boldsymbol{PT})$ converges to the average rating in a pricing cycle $\eta(\boldsymbol{PT})$. For a customer arriving at any time t, the first part of the average rating she observes, $m(t|\boldsymbol{PT})$, equals $\eta(\boldsymbol{PT})$, the expected average rating in each pricing cycle. According to (52),

$$\begin{aligned} M(t|\boldsymbol{PT}) &= \alpha\eta(\boldsymbol{PT}) + (1-\alpha)M(t-T|\boldsymbol{PT}) \\ &= \alpha\eta(\boldsymbol{PT}) + (1-\alpha)(\alpha\eta(\boldsymbol{PT}) + (1-\alpha)M(t-2T|\boldsymbol{PT})) \\ &= \ldots \\ &= \frac{\alpha\eta(\boldsymbol{PT})(1-(1-\alpha)^n)}{1-(1-\alpha)} + (1-\alpha)^n M(t-nT|\boldsymbol{PT}) \\ &= \eta(\boldsymbol{PT}) + (1-\alpha)^n[M(t-nT|\boldsymbol{PT}) - \eta(\boldsymbol{PT})]. \end{aligned} \tag{53}$$

As $t \to \infty$, $n \to \infty$, and thus, $(1-\alpha)^n \to 0$. Then, $\lim_{t\to\infty} M(t|\boldsymbol{PT}) = \eta(\boldsymbol{PT})$. Hence, the long-run average rating $M(t|\boldsymbol{PT})$ at any given time t converges to $\eta(\boldsymbol{PT})$, where $\eta(\boldsymbol{PT})$ is determined by the pricing strategy $\boldsymbol{PT}$ and given by (1).