

Research Article

Optimal Pricing Strategies and Customers' Equilibrium Behavior in an Unobservable M/M/1 Queueing System with Negative Customers and Repair

Doo Ho Lee

Department of Industrial and Management Engineering, Kangwon National University, 346 Joongang-ro, Samecheok-si, Gangwon-do 29513, Republic of Korea

Correspondence should be addressed to Doo Ho Lee; enjdhlee@gmail.com

Received 15 June 2017; Revised 29 September 2017; Accepted 15 October 2017; Published 8 November 2017

Academic Editor: Vladimir Turetsky

Copyright © 2017 Doo Ho Lee. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This work investigates the optimal pricing strategies of a server and the equilibrium behavior of customers in an unobservable M/M/1 queueing system with negative customers and repair. In this work, we consider two pricing schemes. The first is termed the ex-post payment scheme, where the server charges a price that is proportional to the time spent by a customer in the system. The second scheme is the ex-ante payment scheme, where the server charges a flat rate for all services. Based on the reward-cost structure, the server (or system manager) should make optimal pricing decisions in order to maximize its expected profit per time unit in each payment scheme. This study also investigates equilibrium joining/balking behavior under the server's optimal pricing strategies in the two pricing schemes. We show, given a customer's equilibrium, that the two pricing schemes are perfectly identical from an economic point of view. Finally, we illustrate the effect of several system parameters on the optimal joining probabilities, the optimal price, and the equilibrium behavior via numerical examples.

1. Introduction

Over the past several decades, queueing systems considering customers' equilibrium behavior have received much attention owing to their applications to the management of various systems. This type of economic analysis of queueing systems was pioneered by Naor [1], who was the first to study an observable single-server queueing system. Naor [1] introduced the equilibrium and social optimal strategies for the joining problem in an M/M/1 queue with a simple linear reward-cost structure. Naor's model was extended by several researchers including Yechiali [2], Johansen and Stidham [3], Stidham [4], and Mendelson and Whang [5]. Chen and Frank [6] studied Naor's model while assuming that both the customers and the server maximize their expected discounted utility using a common discount rate. Larsen [7] investigated another generalization of Naor's model assuming that the customers differ according to their service values. Erlichman and Hassin [8] analyzed a single-server Markovian queue

which allowed customers to overtake others. Edelson and Hildebrand [9] studied an unobservable case in which customers make decisions without information about the state of the system. Hassin and Haviv [10] conducted an economic analysis of a priority queue where two priority classes can be purchased, and they derived Nash equilibrium strategies (pure or mixed) of the threshold type. In contrast to classical queueing systems, customers are allowed to make decisions based on the information available upon their arrival in these systems. Therefore, the systems can be modeled as games among arriving customers and the fundamental problem is to identify the Nash equilibrium. The monograph of Hassin and Haviv [11] summarized the main approaches and several results with various observable and unobservable queueing systems, offering extensive bibliographical references.

Since the concept was introduced by Gelenbe [12], there has been a rapid increase in the amount of literature on queueing systems with negative customers due to the application to neural networks, communication systems, and

service/manufacturing systems. If a negative customer arrives at a system, it removes one ordinary customer (called a positive customer) according to a predetermined removal discipline, and the server breaks down. Because the performance measure of such a system may be influenced by server failure and repair operations, such an unreliable system is well worth analyzing in light of both queueing theory and reliability theory. Additional studies on unreliable queueing systems were conducted by Gelenbe [13], Artalejo [14], Chae et al. [15], and Lee et al. [16] and references therein.

However, few works have attempted an equilibrium analysis of an unreliable queueing system, despite the fact that server failure is a common phenomenon in everyday life. For example, equilibrium strategies with various levels of information were studied by Economou and Kanta [17]. Boudali and Economou [18, 19] and Economou and Manou [20] considered equilibrium balking strategies in an unreliable single-server queue with catastrophes (or stochastic clearing), where the server fails and all present customers in the system are forced to leave upon the catastrophe's arrival. Zhang et al. [21] analyzed optimal and equilibrium retrial rates in an unreliable queue in which the server is subject to vacations along with breakdowns and repair. Wang et al. [22] studied the equilibrium behavior of customers in a retrial queue with negative customers and derived corresponding Nash equilibrium joining strategies and a social net benefit maximization problem with respect to the levels of information available to customers upon arrival. Li et al. [23] conducted an equilibrium analysis on unreliable M/M/1 queueing systems with working breakdowns. When the system is under repair, the secondary (substitute) server provides services to customers at a slow speed instead of stopping servicing completely. Wang and Zhang [24] extended the model in [17] by assuming that the failed server experiences an exponential delay before the repair begins. Yu et al. [25] also dealt with the same model of [24] for an unobservable case.

Veltman and Hassin [26] presented an interesting paper on the pricing strategies of a server in a Markovian M/M/1 queueing system. They analyzed equilibrium pricing strategies under two types of pricing structures. In the first pricing structure, the server charges a fee that is proportional to the sojourn time (the waiting time plus the service time), referred to as an ex-post payment (EPP) scheme. In the second, however, the server charges a fixed fee for the service, meaning that the server implements an ex-ante payment (EAP) scheme. Under each pricing scheme, the equilibrium joining behavior of customers was examined and compared with socially optimal results. Later, Sun et al. [27] extended this analysis to a batch arrival queueing system, and Economou and Kanta [28], Wang and Zhang [29], and Zhang et al. [30] studied retrial cases. Recently, Ma and Liu [31] considered pricing strategies in a discrete-time Geo/Geo/1 queueing system, and Lee [32] studied the optimal pricing strategies in a discrete-time Geo/Geo/1 queueing system under the EPP and EAP schemes by assuming the sojourn time-dependent reward.

A queueing system with negative customers can be found in many real-life systems. One example is a cloud data center

where the scheduler allocates a virtual machine (VM) for a service request which is generated by a user. All arriving service requests are initially buffered in the scheduler queue, after which each is transmitted to a certain VM according to a predetermined scheduling policy built into the scheduler. In addition, the scheduler oversees service request processing and the VM status for the billing of services. Depending on the operating conditions of the cloud data center, the scheduler charges a fixed price regardless of the system usage time or charges a price that is proportional to the time. Therefore, the scheduler plays a key role in operating the cloud data center stably and maximizing revenue. Because the scheduler is the single point of failure of the entire cloud data center, a secondary scheduler must be built in preparation for a system failure, network hacking, or high availability. Once the main scheduler fails, the request in service is deleted from the scheduler and the secondary scheduler replaces the main scheduler. While the main scheduler is being repaired, the secondary scheduler stores the arriving service requests only in the scheduler queue. Therefore, in terms of maximizing the operation profits of the cloud data center, analyzing optimal and equilibrium pricing strategies is important.

Motivated by the aforementioned studies and the possible application to pricing in a cloud data center, we analyze optimal pricing strategies and the equilibrium joining/balking behavior of customers in an unobservable M/M/1 queueing system with negative customers and repair under the two pricing schemes described above. Our contributions are twofold. First, we derive not only the optimal joining strategies of customers but also the optimal prices set by the server in equilibrium under our model. To the author's knowledge, this study is the first to consider the pricing strategies together with the customer equilibrium behavior for an M/M/1 queueing system with server failure and repair. Second, under the EPP and EAP schemes, we show that the two pricing schemes do not affect customers' joining behavior and the server's maximum profits, which has not been reported in the literature on an M/M/1 queueing system with negative customers and repair.

This paper is structured as follows. In Section 2, we describe the model and derive the expected sojourn time of an unobservable M/M/1 queueing system with negative customers and repair. In Sections 3 and 4, customers' equilibrium joining behavior and the server's profit maximization strategies under the two pricing schemes are driven in a partially unobservable case and in a fully unobservable case. We also deal with numerical experiments in which we investigate trends in the optimal prices and joining probabilities of EPP and EAP schemes according to various input values.

2. Preliminaries

This work considers a queueing system with the following features. Ordinary customers arrive at a single server queue, according to a Poisson process with a rate of λ ($\lambda > 0$). A server renders a service to each ordinary customer on a first-come first-out basis. The service time is exponentially distributed with a rate of μ ($\mu > 0$). Negative customers arrive only when the server is in operation according to a

Poisson process with a rate of η ($\eta > 0$). We assume the following removal discipline: each time a negative customer arrives while the server is in operation, the server breaks down and the ordinary customer being served is forced to leave the system. If the negative customer arrives at an empty system, it simply triggers a server failure because there is no customer who can leave. When the server fails, the server is turned off and a repair process begins immediately. The repair time for the failed server is exponentially distributed at a rate of γ ($\gamma > 0$). While the server is under repair, the stream of the arrival process of ordinary customers continues.

Remark 1. In many cases, a Poisson distribution can model the number of occurrences of rare events such as the arrival of requests, the service process, and service failures in information communication systems. In a cloud data center, while the number of potential users is high, each user typically submits a task at any given time with a low probability. Therefore, service requests can be adequately modeled as a Poisson process. Moreover, since a service failure is not a frequently observed event, we can assume that negative customers arrive according to a Poisson process. In the context of cloud data center applications, however, it is difficult to conclude that the service process or repair process follows a Poisson process due to the synchronized communication of packets in data centers. Nevertheless, because one of the objectives of our work is to investigate the joining behavior of customers and related pricing strategies, we assume that the service time and repair time follow the exponential distribution for analytical simplicity. As shown in the literature on cloud data center applications, Markovian queueing systems have been used to model actual system behavior and to determine the performance measures of systems [33, 34].

Whenever an ordinary customer arrives at the system, he/she decides to join or balk depending on the specific joining probability. An arriving ordinary customer joins the queue with a probability of q_0 when the server is under repair. Otherwise, he/she does with a probability of q_1 when the server is in operation.

Let $N(t)$ be the number of (ordinary) customers in the system at time t with the following definitions:

$$I(t) = \begin{cases} 1, & \text{The server is in operation at time } t, \\ 0, & \text{The server is under repair at time } t. \end{cases} \quad (1)$$

The vector process $\{(N(t), I(t)), t \geq 0\}$ then becomes a two-dimensional continuous-time Markov chain with the state space expressed as $\Omega = \{(n, 0) : n \geq 0\} \cup \{(n, 1) : n \geq 0\}$ and its nonzero transition rates are given by

$$\begin{aligned} r_{(n,i)(n+1,i)} &= \lambda q_i, \quad n \geq 0; \quad i = 0, 1; \\ r_{(n,1)(n-1,1)} &= \mu, \quad n \geq 1; \\ r_{(0,1)(0,0)} &= \eta, \\ r_{(n+1,1)(n,0)} &= \eta, \quad n \geq 0; \\ r_{(n,0)(n,1)} &= \gamma, \quad n \geq 0. \end{aligned} \quad (2)$$

The corresponding transition diagram is shown in Figure 1.

Define $p(n, i) = \lim_{t \rightarrow \infty} \Pr\{N(t) = n, I(t) = i\}$, $(n, i) \in \Omega$. The balance equations governing the system are then given by

$$(\lambda q_0 + \gamma) p(0, 0) = \eta p(0, 1) + \eta p(1, 1), \quad (3)$$

$$\begin{aligned} (\lambda q_0 + \gamma) p(n, 0) &= \eta p(n + 1, 1) \\ &+ \lambda q_0 p(n - 1, 0), \quad n \geq 1, \end{aligned} \quad (4)$$

$$(\lambda q_1 + \eta) p(0, 1) = \gamma p(0, 0) + \mu p(1, 1), \quad (5)$$

$$\begin{aligned} (\lambda q_1 + \mu + \eta) p(n, 1) &= \gamma p(n, 0) + \lambda q_1 p(n - 1, 1) \\ &+ \mu p(n + 1, 1), \quad n \geq 1, \end{aligned} \quad (6)$$

From (3)–(6), we now have the following lemma.

Lemma 2. Define $W(q_0, q_1)$ as the expected sojourn time of an ordinary customer who decides to join the queue upon his/her arrival. After some calculations, $W(q_0, q_1)$ can be obtained as

$$W(q_0, q_1) = \frac{\gamma q_1 (\lambda \eta q_1 + \gamma (\mu + \eta)) + \eta q_0 ((\mu + \eta) (\mu + \gamma + \eta) - \lambda q_1 (\mu + \gamma)) + \lambda \mu \eta q_0^2}{(\mu + \eta) (\eta q_0 + \gamma q_1) (\gamma (\mu + \eta) - \lambda (\eta q_0 + \gamma q_1))}, \quad (7)$$

where $\lambda (\eta q_0 + \gamma q_1) < \gamma (\mu + \eta)$ should hold for the stable system.

Proof. See Appendix A. □

We assume that an ordinary customer receives a reward of R ($R > 0$) units for departing the system. Moreover, he/she accumulates costs at a rate of C ($C > 0$) units per unit time during which he/she remains in the system. Ordinary customers are risk-neutral, and their decisions are irrevocable in the sense that retrials of balking customers and renegeing of entering customers are not allowed. To make the model nontrivial, we assume a condition that $R > C((\mu + \eta)^{-1} + \gamma^{-1})$,

which ensures that the reward for joining the queue exceeds the expected cost of an ordinary customer finding the system state $(0, 0)$.

Remark 3. Under our assumption, it may not sound plausible that a customer whose service due to a server failure is canceled receives the same reward of R . The author agrees; in fact, Boudali and Economou [18, 19] assumed that each customer receives either a reward of R_s units for completing the service or failure compensation of R_f units if forced to abandon the system due to a server failure. However, this paper assumes that all joining customers receive the same

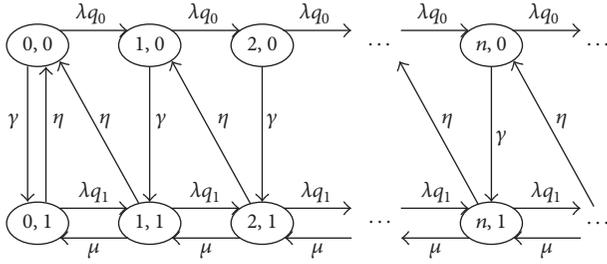


FIGURE 1: Transition diagram for the partially unobservable case.

reward for simplicity of the model and the analysis. The main purposes of this paper are to compare customer joining behavior in the EPP and EAP schemes and to determine the optimal price of each scheme under a linear reward-cost structure. We will leave the model in which customers receive different amounts of compensation depending on the completion of service as future study.

We distinguish the following two cases based on the information available to ordinary customers upon their arrival: (i) the partially unobservable case, where ordinary customers observe the server state $I(t)$ but not the number of customers in the system $N(t)$ and (ii) the fully unobservable case, where ordinary customers are not informed about $I(t)$ or $N(t)$.

3. Pricing Analyses in the Partially Unobservable Case

In this section, we investigate a partially unobservable case where arriving ordinary customers observe the state of the server upon their arrival, but not the number of customers in the system. There are four pure strategies available for a customer: to join the queue when the server is under repair or in operation or not to join the queue when the server is under repair or in operation. An ordinary customer's joining strategy can be described by the two probabilities of q_0 and q_1 ($0 \leq q_i \leq 1$), which are the joining probabilities.

Remark 4. In fact, the model under our study can be viewed as an M/M/1 queue under a random environment. Specifically, the external environment $I(t)$ is an irreducible continuous-time Markov chain in the finite state space $\{0, 1\}$. The infinitesimal generator of the external environment is given by

$$\mathbf{Q} = \begin{pmatrix} -\gamma & \gamma \\ \eta & -\eta \end{pmatrix}. \quad (8)$$

With (π_0, π_1) as the stationary distribution of the external environment after solving $(\pi_0, \pi_1)\mathbf{Q} = (0, 0)$, we have $\pi_0 = \eta(\gamma + \eta)^{-1}$ and $\pi_1 = \gamma(\gamma + \eta)^{-1}$. That is, π_0 (π_1) represents the stationary probability that the server is under repair (in operation). Therefore, the effective arrival rate λ_{eff} is equal to $\lambda q_0 \pi_0 + \lambda q_1 \pi_1 = \lambda(\eta q_0 + \gamma q_1)(\gamma + \eta)^{-1}$ due to the PASTA property [35].

3.1. EPP Scheme. Under the EPP scheme, the server charges the price in proportion to the customer's sojourn time. Let K_t ($K_t \geq 0$) and P_t denote the price charged by the server (or system) and the expected server profit per time unit, respectively. The expected benefit U_t then has the following relationship: $U_t = R - (K_t + C)W(q_0, q_1)$. If U_t equals zero at customer's equilibrium, we have $R = (K_t + C)W(q_0, q_1)$. At customer's equilibrium, therefore, the price can be expressed in terms of the customer's joining probabilities, denoted by $q_{0,t}$ and $q_{1,t}$:

$$K_t = \frac{R}{W(q_{0,t}, q_{1,t})} - C. \quad (9)$$

Substituting (9) into $P_t = \lambda_{\text{eff}} K_t W(q_{0,t}, q_{1,t})$, we have

$$\begin{aligned} P_t &= \frac{\lambda(q_{0,t}\eta + q_{1,t}\gamma) K_t W(q_{0,t}, q_{1,t})}{\gamma + \eta} \\ &= \frac{\lambda(q_{0,t}\eta + q_{1,t}\gamma)(R - CW(q_{0,t}, q_{1,t}))}{\gamma + \eta}. \end{aligned} \quad (10)$$

We establish the following nonlinear programming (NLP) problem to maximize P_t with respect to $q_{0,t}$ and $q_{1,t}$:

$$\begin{aligned} \max_{q_{0,t}, q_{1,t}} \quad & P_t = \frac{\lambda(q_{0,t}\eta + q_{1,t}\gamma)(R - CW(q_{0,t}, q_{1,t}))}{\gamma + \eta} \\ \text{s.t.} \quad & 0 \leq q_{i,t} \leq 1, \quad i = 0, 1, \\ & \lambda(\eta q_{0,t} + \gamma q_{1,t}) < \gamma(\mu + \eta). \end{aligned} \quad (11)$$

In (11), we want to maximize the expected server profit per time unit under the EPP scheme in the partially unobservable case. The first constraint implies that the joining probability should be bounded between 0 and 1, and the second guarantees the system to be stable. We introduce the following lemma.

Lemma 5. *The optimization problem in (11) is a convex maximization problem (CMP).*

Proof. See Appendix B. \square

From Lemma 5, we now have the following theorem.

Theorem 6. *Under the EPP scheme in the partially unobservable case, if $0 \leq q_{0,t} \leq 1$ and $0 \leq q_{1,t} \leq 1$, there exists a unique equilibrium where customers join the queue with probabilities of $q_{0,t}^*$ and $q_{1,t}^*$; the optimal price which maximizes the expected server profit per unit time is given by $K_t^* = R/W(q_{0,t}^*, q_{1,t}^*) - C$.*

Proof. Because the optimization problem in (11) is a CMP, the Lagrange multiplier method can be used to find the optimal solution. Observing (11), we find that all constraints are inequality constraints; thus, Karush-Kuhn-Tucker conditions can be used to generalize the method of Lagrange multipliers. However, the explicit form of the optimal $q_{i,t}$, $i = 0, 1$, denoted by $q_{i,t}^*$, is too long and complicated. Instead, we introduce another means by which the value of $q_{i,t}^*$ can be obtained.

According to Boyd and Vandenberghe [36], if the objective function is strictly concave in the feasible region, a unique local optimal solution obtained through the Newton method becomes a unique global optimal solution. Details about the Newton method can be found in Boyd and Vandenberghe [36]. This completes the proof. \square

One may raise the question of how we can say that $q_{0,t}^*$ and $q_{1,t}^*$ are also equilibrium joining probabilities, as they are affected by the price charged by the server. As mentioned earlier, the expected benefit U_t has the following relationship: $U_t = R - (K_t^* + C)W(q_{0,t}^*, q_{1,t}^*)$. We distinguish the following three cases.

Case 1 ($R \leq (K_t^* + C)W(0,0)$). In this case, even if no other customer joins, the expected benefit of a customer who joins is nonpositive. Therefore, the strategy of joining with probabilities $q_{0,t}^* = q_{1,t}^* = 0$ is an equilibrium strategy and no other equilibrium is possible. Moreover, in this case, not joining is the dominant strategy.

Case 2 ($R \geq (K_t^* + C)W(1,1)$). In this case, even if all potential customers join, they all enjoy a nonnegative benefit. Therefore, the strategy of joining with probability $q_{0,t}^* = q_{1,t}^* = 1$ is an equilibrium strategy and no other equilibrium is possible. Moreover, in this case, joining is the dominant strategy.

Case 3 ($(K_t^* + C)W(0,0) < R < (K_t^* + C)W(1,1)$). In this case, if $q_{0,t}^* = q_{1,t}^* = 1$, then a customer who joins suffers a negative benefit. Hence, this cannot be an equilibrium strategy. Likewise, if $q_{0,t}^* = q_{1,t}^* = 0$, a customer who joins receives a positive benefit, more than by balking. Hence, this cannot be an equilibrium strategy. Therefore, there exists unique equilibrium joining probabilities $q_{0,t}^*$ and $q_{1,t}^*$ satisfying $R = (K_t^* + C)W(q_{0,t}^*, q_{1,t}^*)$ for which customers are indifferent with regard to joining and/or balking. Assume that $R = 20$, $C = 2$, $\lambda = 1$, $\mu = 0.8$, $\gamma = 1.5$, and $\eta = 0.05$. From Theorem 6, we obtain the optimal joining probabilities $q_{0,t}^* = 0.34801$ and $q_{1,t}^* = 0.54417$ and the optimal price $K_t^* = 3.78749$. Therefore, $W(q_{0,t}^*, q_{1,t}^*) = 3.45577$ and the relationship $R = (K_t^* + C)W(q_{0,t}^*, q_{1,t}^*)$ should hold.

3.2. EAP Scheme Model. Next, We consider the EAP scheme model, where the server charges a flat rate for services. Let K_f and P_f be the fixed price charged by the server and the server's expected profit per time unit, respectively. Given that the expected benefit for a customer is expressed as $U_f = R - K_f - CW(q_0, q_1)$ and equals zero in customers' equilibrium, we obtain $R = K_f + CW(q_0, q_1)$. Let $q_{0,f}$ and $q_{1,f}$ denote the customer's joining probabilities in the EAP scheme. K_f is a function of $q_{0,f}$ and $q_{1,f}$ and has the form of

$$K_f = R - CW(q_{0,f}, q_{1,f}). \quad (12)$$

Substituting (12) into $P_f = \lambda_{\text{eff}}K_f$, we have

$$P_f = \frac{\lambda K_f (q_{0,f}\eta + q_{1,f}\gamma)}{\gamma + \eta}$$

$$= \frac{\lambda (q_{0,f}\eta + q_{1,f}\gamma) (R - CW(q_{0,f}, q_{1,f}))}{\gamma + \eta}. \quad (13)$$

We establish the following NLP problem to maximize P_f with respect to $q_{0,f}$ and $q_{1,f}$:

$$\begin{aligned} \max_{q_{0,f}, q_{1,f}} \quad & P_f \\ & = \frac{\lambda (q_{0,f}\eta + q_{1,f}\gamma) (R - CW(q_{0,f}, q_{1,f}))}{\gamma + \eta} \end{aligned} \quad (14)$$

$$\text{s.t. } 0 \leq q_{i,f} \leq 1, \quad i = 0, 1,$$

$$\lambda (\eta q_{0,f} + \gamma q_{1,f}) < \gamma (\mu + \eta).$$

In (14), we also want to maximize the expected server profit per time unit under the EAP scheme in the partially unobservable case. The first constraint implies that the joining probabilities should be bounded between 0 and 1, and the second specifies the stability condition. It should be noted that the optimization problem in (14) is perfectly identical to the CMP in (11). In other words, the optimal joining strategies for the customer and the maximum expected server profits are identical in the two different pricing schemes. We introduce the following theorem.

Theorem 7. *Under the EAP scheme in the partially unobservable case, if $0 \leq q_{0,f} \leq 1$ and $0 \leq q_{1,f} \leq 1$, there exists a unique equilibrium where customers join the queue with probability $q_{i,f}^* = q_{i,t}^*$, $i = 0, 1$, and the optimal price which maximizes the expected server profit per unit time is determined by $K_f^* = R - CW(q_{0,f}^*, q_{1,f}^*)$.*

Proof. From Lemma 8, the CMP in (14) is perfectly identical to that in (11), which means that $q_{i,f}^* = q_{i,t}^*$, $i = 0, 1$. The optimal price which maximizes the expected server profit per unit time is calculated by inserting $q_{0,f}^*$ and $q_{1,f}^*$ into (12). This completes the proof. \square

3.3. Numerical Experiments in the Partially Unobservable Case. Based on the results of the partially unobservable case, we compare the joining probability of the EPP scheme and that of the EAP scheme with various experimental parameters. We set the common experimental parameters as $(R, C) = (20, 2)$.

Figure 2(a) shows the trends of the optimal joining probabilities q_0^* and q_1^* according to the service rate μ . Note that $q_i^* = q_{i,t}^* = q_{i,f}^*$, $i = 0, 1$. We confirm that q_0^* and q_1^* are increasing functions of μ . It is clear that when the server renders faster service, more customers are likely to join the queue. We also observe that $q_0^* < q_1^*$. If an arriving customer joins the queue while the server is under repair, he/she should wait the additional remaining repair time. This has a negative impact on the customers' joining behavior. Therefore, customers prefer to join the queue when the server is in operation.

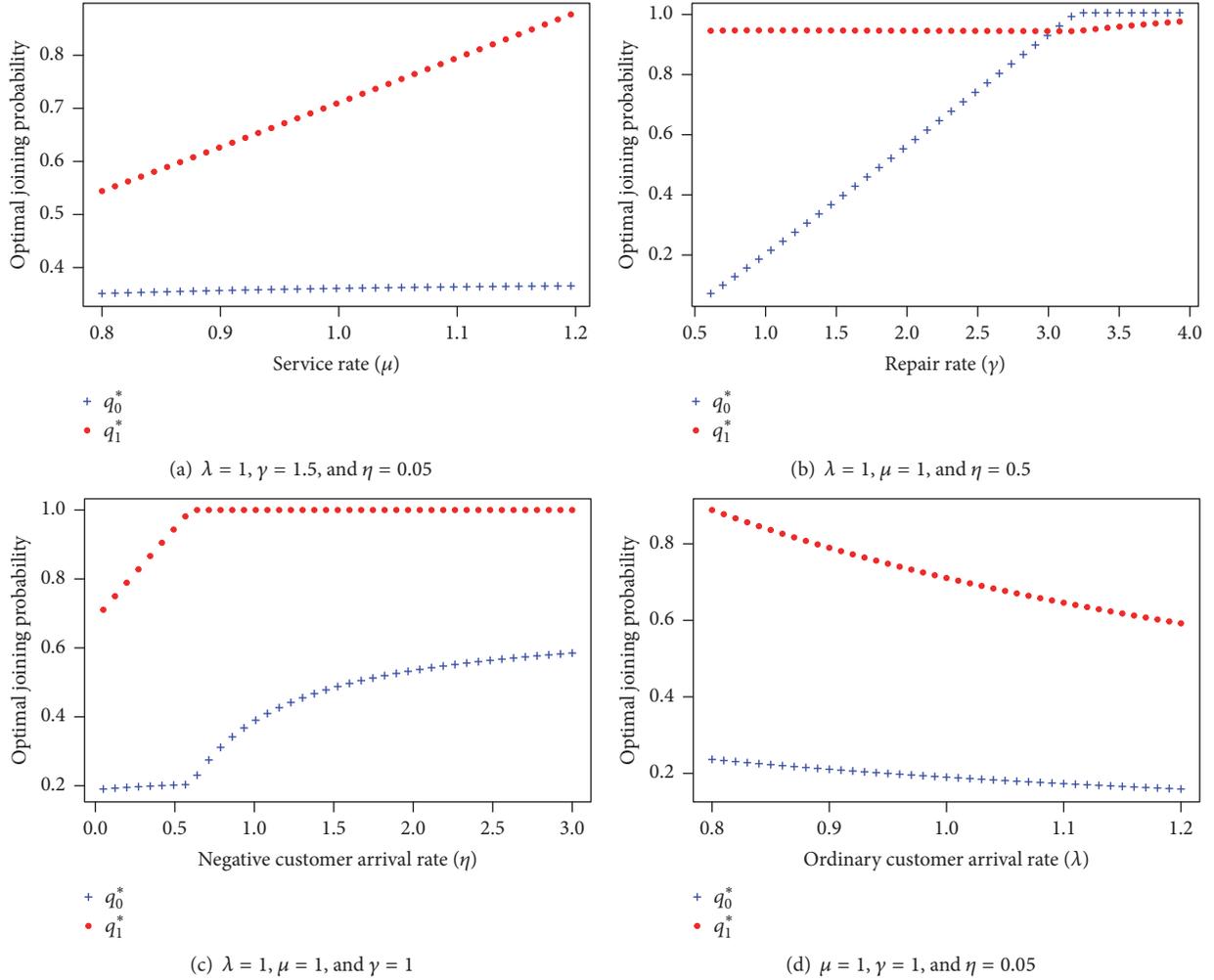


FIGURE 2: Optimal joining probabilities.

Figure 2(b) shows the trends of q_0^* and q_1^* according to the repair rate γ . We observe that q_0^* and q_1^* are increasing functions of γ . As the repair speed increases, the probability that the server is in operation also increases. Therefore, more customers are likely to join the queue and receive services due to the high reliability of the server. Moreover, as γ increases, q_0^* increases more sharply than q_1^* . That is, the repair rate has a stronger positive impact on customers arriving when the server is under repair than on those arriving when the server is in operation.

Figure 2(c) shows the trends of q_0^* and q_1^* according to the negative customer arrival rate η . The interesting observation in Figure 4 is that q_0^* sharply increases after q_1^* reaches 1. This phenomenon, that is, that customers who arrive while the server is under repair tend to imitate the behavior of those who arrive while the server is in operation, is known as the “Follow-The-Crowd” behavior.

Figure 2(d) shows that both q_0^* and q_1^* are decreasing functions of the ordinary customer arrival rate λ . As more customers arrive, the congestion of the system will increase, implying that higher (external) arrival rate leads to a lower probability of actual joining.

4. Pricing Analyses for the Fully Unobservable Case

In this section, we consider the fully unobservable case, where an arriving ordinary customer cannot observe the system state at all. There are two pure strategies available for a customer: to join or not to join. An ordinary customer’s joining strategy can be described by a probability of q ($0 \leq q \leq 1$), which is the probability of joining, and the effective arrival rate (joining rate) is λq . The transition diagram for this case is illustrated in Figure 3.

In the fully observable case, all customers have the same joining probability of q ; therefore, by considering $q_0 = q_1 = q$ in (7), we have

$$W(q) = \frac{\gamma(\gamma + \eta) + \eta(\mu + \eta)}{(\gamma + \eta)(\gamma(\mu + \eta) - \lambda q(\gamma + \eta))}. \quad (15)$$

Under the EPP scheme, the server charges the price in proportion to the customer’s sojourn time. Therefore, the expected benefit has the following relationship: $U_t = R - (K_t + C)W(q)$. If U_t equals zero at customer’s equilibrium, we have

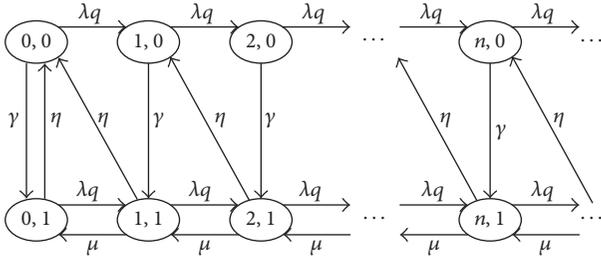


FIGURE 3: Transition diagram for the fully unobservable case.

$R = (K_t + C)W(q)$. At customer's equilibrium, the price can be expressed in terms of the customer's joining probability, denoted by q_t :

$$\begin{aligned} K_t &= \frac{R}{W(q_t)} - C \\ &= \frac{R(\gamma + \eta)(\gamma(\mu + \eta) - \lambda q_t(\gamma + \eta))}{\gamma(\gamma + \eta) + \eta(\mu + \eta)} - C. \end{aligned} \quad (16)$$

Substituting (16) into $P_t = \lambda q_t K_t W(q_t)$, we can establish the following NLP problem to maximize P_t with respect to q_t :

$$\begin{aligned} \max_{q_t} \quad & P_t = R\lambda q_t - \frac{C\lambda q_t(\gamma(\gamma + \eta) + \eta(\mu + \eta))}{(\gamma + \eta)(\gamma(\mu + \eta) - \lambda q_t(\gamma + \eta))} \\ \text{s.t.} \quad & 0 \leq q_t \leq 1, \\ & \lambda q_t(\eta + \gamma) < \gamma(\mu + \eta). \end{aligned} \quad (17)$$

In (17), we want to maximize the expected server profit per time unit under the EPP scheme in the fully unobservable case. The first constraint implies that q_t is the joining probability and the second is derived from (11) when $q_{0,t} = q_{1,t} = q_t$. We introduce the following.

Lemma 8. *The optimization problem in (17) is a convex maximization problem (CMP). In addition, the optimal joining probability is $\max\{0, \min\{x_1, 1\}\}$, where x_1 is the root of the equation $\partial P_t / \partial q_t = 0$ satisfying $\lambda x_1(\eta + \gamma) < \gamma(\mu + \eta)$.*

Proof. See Appendix C. \square

Theorem 9. *Under the EPP scheme in the fully unobservable case, if $0 \leq q_t \leq 1$, there exists a unique equilibrium where customers join the queue with probabilities of $q_t^* = \max\{0, \min\{x_1, 1\}\}$ and the optimal price which maximizes the expected server profit per unit time is given by $K_t^* = R/W(q_t^*) - C$.*

Proof. The optimal solution q_t^* of the CMP in (17) is determined by Lemma 8 and the optimal price which maximizes the expected server profit per unit time is calculated by inserting q_t^* into (16). This completes the proof. \square

Under the EAP scheme, the server charges a flat rate for services. Hence, the expected benefit has the following

relationship: $U_f = R - K_f - CW(q)$. If U_f equals zero at customer's equilibrium, we have $R = K_f + CW(q)$. At customer's equilibrium, therefore, the price can be expressed in terms of the customer's joining probability, denoted by q_f :

$$\begin{aligned} K_f &= R - CW(q_f) \\ &= R - \frac{C\gamma(\gamma + \eta) + \eta(\mu + \eta)}{(\gamma + \eta)(\gamma(\mu + \eta) - \lambda q_f(\gamma + \eta))}. \end{aligned} \quad (18)$$

Inserting (18) into $P_f = \lambda q_f K_f$, we have the following NLP problem to maximize P_f with respect to q_f :

$$\begin{aligned} \max_{q_f} \quad & P_f \\ &= R\lambda q_f - \frac{C\lambda q_f(\gamma(\gamma + \eta) + \eta(\mu + \eta))}{(\gamma + \eta)(\gamma(\mu + \eta) - \lambda q_f(\gamma + \eta))} \\ \text{s.t.} \quad & 0 \leq q_f \leq 1, \\ & \lambda q_f(\eta + \gamma) < \gamma(\mu + \eta). \end{aligned} \quad (19)$$

Clearly, the optimization problem in (19) is identical to that in (17).

Theorem 10. *Under the EAP scheme in the fully unobservable case, if $0 \leq q_f \leq 1$, there exists a unique equilibrium where customers join the queue with probability $q_f^* = q_t^*$; the optimal price which maximizes the expected server profit per unit time is given by $K_f^* = R - CW(q_f^*)$.*

Proof. Note that $q_f^* = q_t^*$ from (17) and (19). The optimal price which maximizes the expected server profit per unit time is calculated by inserting q_f^* into (18). This completes the proof. \square

Finally, we present the numerical analyses. The trends of the optimal price and the joining probability in the fully unobservable case are similar to those in the partially unobservable case. Instead of conducting the same numerical analyses, we compare the maximum server profits of the two cases. We set the common experimental parameters as $(R, C) = (20, 2)$. Let P_{pa}^* and P_{fu}^* denote the maximum server profit for the partially unobservable case and the fully unobservable case, respectively. In Figure 4, we find that P_{pa}^* is greater than P_{fu}^* with respect to various values of η and λ , indicating that informing arriving customers of the state of the server is more beneficial in terms of the server's profit.

Figure 4(a) tells us that P_{pa}^* increases when $\eta \leq \mu$ and decreases when $\eta > \mu$. If the failure rate is greater than the service rate, more customers will leave the system due to server failures, resulting in frequent repairs. Hence, the sojourn time becomes longer and more customers are reluctant to join the queue. This has a negative impact on the server's profit. Moreover, P_{fu}^* is not influenced by the value of η ; thus, we conclude that η is not a considering factor in the fully unobservable case.

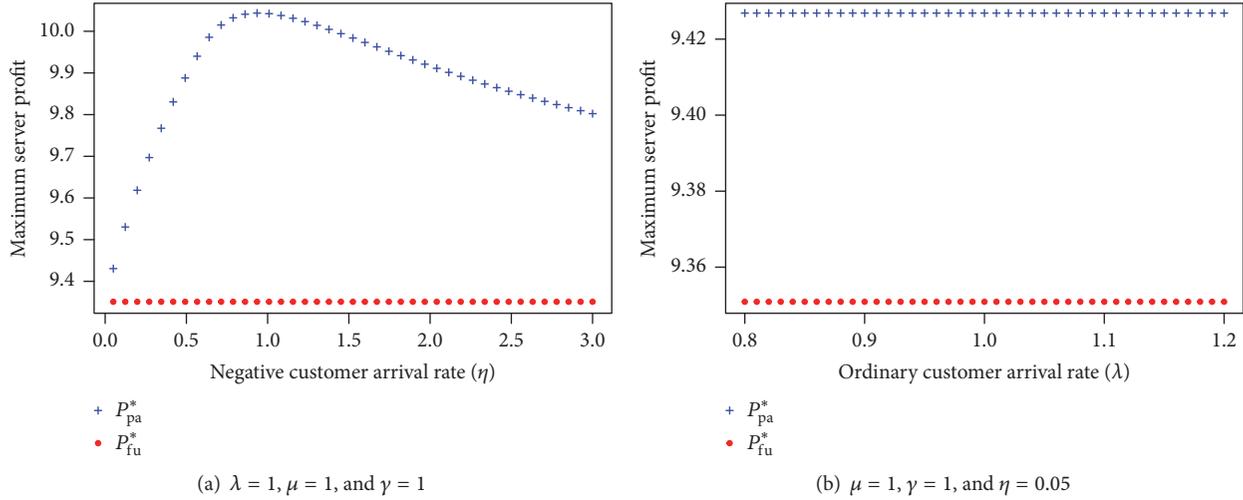


FIGURE 4: Maximum server profit for the partially unobservable and the fully unobservable cases.

Figure 4(b) shows that λ has no effect on the server’s profit in both the partially and fully unobservable cases. According to our parameter settings, $P_{pa}^* = 9.42641$ and $P_{fu}^* = 9.35089$ regardless of the value of λ . This implies that the server always maintains the same maximum profit by adjusting the joining probabilities and the prices for a service, even if the value of λ varies. Thus, λ is not a considering factor in both the partially and fully unobservable cases.

5. Conclusions

In this work, we studied customers’ joining/balking behavior at equilibrium and the server’s optimal pricing strategies in partially and fully unobservable M/M/1 queueing systems with negative customers and repair. We introduced two different pricing models: the EPP scheme and the EAP scheme. In the EPP scheme, the server imposes a price that is proportional to ordinary customer’s sojourn time, while in the EAP scheme, the server charges a fixed price for the service. We set the convex maximization problem and determined the optimal joining probabilities and prices that maximize the server’s expected profits for each pricing scheme. We also determined that the two different pricing schemes present identical outcomes with regard to customers’ equilibrium joining/balking behavior and the server’s maximum profits. Moreover, through various numerical experiments, we found that the maximum profit of the server in the partially unobservable case is greater than that in the fully unobservable case. In conclusion, providing information about the state of the server to arriving customers is advantageous in terms of maximizing the server’s profit.

This work could provide queue managers with useful guidelines for making pricing decisions. It can also aid customers with regard to optimal strategies. Future studies can include the extension of this work to other queueing systems with, for instance, multiple servers, a vacation policy, and/or a sojourn time-dependent reward-cost structure. The

M/G/1-type or GI/M/1-type pricing analyses can also be considered as future research topics.

Appendix

A. The Proof of Lemma 2

We define the partial probability generating functions: $G_i(z) = \sum_{n=0}^{\infty} p(n, i)z^n$, $i = 0, 1$. Note that the normalizing condition is given by $G_0(1)+G_1(1) = 1$. Multiplying (3)–(6) by appropriate powers of z and summing over $n \geq 0$, we obtain

$$G_0(z) = \frac{\eta z^{-1}G_1(z) + \eta p(0, 1)(1 - z^{-1})}{\lambda q_0(1 - z) + \gamma}, \tag{A.1}$$

$$G_1(z) = \frac{\mu p(0, 1)(1 - z^{-1}) + \gamma G_0(z)}{\lambda q_1(1 - z) + \mu(1 - z^{-1}) + \eta}.$$

Solving (A.1) with respect to $G_i(z)$, $i = 0, 1$, we have

$$G_0(z) = \frac{\eta p(0, 1)(\lambda q_1(1 - z) + \mu + \eta)}{\gamma(\eta + \mu) + \lambda q_0(\mu(1 - z) - \eta z) - \lambda q_1 z(\lambda q_0(1 - z) + \gamma)}, \tag{A.2}$$

$$G_1(z) = \frac{p(0, 1)(\lambda \mu q_0(1 - z) + \gamma(\mu + \eta))}{\gamma(\eta + \mu) + \lambda q_0(\mu(1 - z) - \eta z) - \lambda q_1 z(\lambda q_0(1 - z) + \gamma)}.$$

By normalizing, the only unknown value $p(0, 1)$ in (A.2) is determined by

$$p(0, 1) = \frac{\gamma(\mu + \eta) - \lambda(\eta q_0 + \gamma q_1)}{(\gamma + \eta)(\mu + \eta)}. \tag{A.3}$$

Since all states are communicating, from the theory of recurrent events it can be deduced that the probabilities $p(n, i)$ ($n \geq 0, i = 0, 1$) are either all positive or, alternatively,

all equal to zero. If the Markov chain $\{(N(t), I(t)), t \geq 0\}$ is positive recurrent, then all the probabilities $p(n, i)$ are positive. Therefore, $p(0, 1) > 0$ and we have $\lambda(\eta q_0 + \gamma q_1) < \gamma(\mu + \eta)$ from (A.3). Conversely, if $\lambda(\eta q_0 + \gamma q_1) < \gamma(\mu + \eta)$, then $p(0, 1) > 0$. We conclude that all the probabilities $p(n, i)$ are positive from the ergodicity theory for continuous-time Markov chains. Thus, the vector process $\{(N(t), I(t)), t \geq 0\}$ is stable if and only if $\lambda(\eta q_0 + \gamma q_1) < \gamma(\mu + \eta)$.

Let π_0 and π_1 denote the probability that the server is under repair and in operation, respectively. By the definition of $G_i(z)$, we have that $\pi_i = \sum_{n=0}^{\infty} p(n, i) = G_i(1)$, $i = 0, 1$.

From Remark 3, $\pi_0 = G_0(1) = \eta(\gamma + \eta)^{-1}$ and $\pi_1 = G_1(1) = \gamma(\gamma + \eta)^{-1}$. Therefore, we have

$$G(z) = G_0(z) + G_1(z) = \frac{\gamma(\mu + \eta) - \lambda(\eta q_0 + \gamma q_1)}{(\gamma + \eta)(\mu + \eta)} \cdot \frac{\lambda(1-z)(\eta q_1 + \mu q_0) + (\gamma + \eta)(\mu + \eta)}{\gamma(\eta + \mu) + \lambda q_0(\mu(1-z) - \eta z) - \lambda q_1 z(\lambda q_0(1-z) + \gamma)}. \quad (\text{A.4})$$

Finally, using Little's formula, the expected sojourn time of an ordinary customer who decides to join the queue upon his/her arrival is

$$W(q_0, q_1) = \frac{G'(1)}{\lambda_{\text{eff}}} = \frac{\gamma q_1(\lambda \eta q_1 + \gamma(\mu + \eta)) + \eta q_0((\mu + \eta)(\mu + \gamma + \eta) - \lambda q_1(\mu + \gamma)) + \lambda \mu \eta q_0^2}{(\mu + \eta)(\eta q_0 + \gamma q_1)(\gamma(\mu + \eta) - \lambda(\eta q_0 + \gamma q_1))}. \quad (\text{A.5})$$

B. The Proof of Lemma 5

According to the definition of the convex maximization in Boyd and Vandenberghe [36], if the objective function can be proved concave in the feasible region and the set of constraints can be proved convex, the maximization problem is a CMP.

The constraint functions in (11) are all real-valued linear functions; therefore, the set of constraints is convex.

Let $\mathbf{H}(q_{0,t}, q_{1,t})$ be the Hessian matrix of the objective function in (11), and it has the form of

$$\mathbf{H}(q_{0,t}, q_{1,t}) = \begin{pmatrix} \frac{\partial^2 P_t}{\partial q_{0,t}^2} & \frac{\partial^2 P_t}{\partial q_{0,t} \partial q_{1,t}} \\ \frac{\partial^2 P_t}{\partial q_{1,t} \partial q_{0,t}} & \frac{\partial^2 P_t}{\partial q_{1,t}^2} \end{pmatrix}, \quad (\text{B.1})$$

where

$$\begin{aligned} \frac{\partial^2 P_t}{\partial q_{0,t}^2} &= -\frac{2C\lambda^2\gamma\eta((\mu + \eta - \lambda q_{1,t})^2 + \lambda\eta q_{1,t})}{(\gamma(\mu + \eta) - \lambda(\eta q_{0,t} + \gamma q_{1,t}))^3}, \\ \frac{\partial^2 P_t}{\partial q_{0,t} \partial q_{1,t}} &= \frac{\partial^2 P_t}{\partial q_{1,t} \partial q_{0,t}} \\ &= -\frac{C\lambda^2\gamma\eta(\gamma(\mu + \eta + \lambda q_{1,t}) + \lambda q_{0,t}(\eta + 2\mu - 2\lambda q_{1,t}))}{(\gamma(\mu + \eta) - \lambda(\eta q_{0,t} + \gamma q_{1,t}))^3}, \\ \frac{\partial^2 P_t}{\partial q_{1,t}^2} &= -\frac{2C\lambda^2\gamma(\gamma^2(\mu + \eta) + \lambda\eta q_{0,t}(\lambda q_{0,t} - \gamma))}{(\gamma(\mu + \eta) - \lambda(\eta q_{0,t} + \gamma q_{1,t}))^3}. \end{aligned} \quad (\text{B.2})$$

Define D_k as the leading principal minor determinant of order k for $\mathbf{H}(q_{0,t}, q_{1,t})$. Then, we have

$$D_1 = \frac{\partial^2 P_t}{\partial q_{0,t}^2} = -\frac{2C\lambda^2\gamma\eta((\mu + \eta - \lambda q_{1,t})^2 + \lambda\eta q_{1,t})}{(\gamma(\mu + \eta) - \lambda(\eta q_{0,t} + \gamma q_{1,t}))^3} < 0,$$

$$\begin{aligned} D_2 &= \left(\frac{\partial^2 P_t}{\partial q_{0,t}^2} \right) \left(\frac{\partial^2 P_t}{\partial q_{1,t}^2} \right) - \left(\frac{\partial^2 P_t}{\partial q_{0,t} \partial q_{1,t}} \right)^2 \\ &= \frac{C^2\lambda^4\gamma^2\eta(4\mu + 3\eta)}{(\gamma(\mu + \eta) - \lambda(\eta q_{0,t} + \gamma q_{1,t}))^4} > 0. \end{aligned} \quad (\text{B.3})$$

Therefore, $\mathbf{H}(q_{0,t}, q_{1,t})$ in (B.1) is negative definite and the objective function in (11) should be strictly concave in the feasible region, which leads to (11) being a CMP.

C. The Proof of Lemma 8

The constraint functions in (17) are all real-valued linear functions; therefore, the set of constraints is convex. The second derivative of the objective function in (17) can be expressed as

$$\frac{\partial^2 P_t}{\partial q_t^2} = -\frac{2C\lambda^2(\mu + \eta)(\gamma(\gamma + \eta) + \eta(\mu + \eta))}{(\gamma(\mu + \eta) - \lambda q_t(\eta + \gamma))^3}. \quad (\text{C.1})$$

Since $\partial^2 P_t / \partial q_t^2 < 0$ in the feasible region, the objective function P_t should be strictly concave and (17) is a CMP.

Let x_1 and x_2 be the roots of the equation $\partial P_t / \partial q_t = 0$, and x_1 and x_2 are written as

$$\begin{aligned} x_1 &= \frac{\gamma(\mu + \eta)}{\lambda(\eta + \gamma)} \\ &\quad - \frac{\sqrt{RC\gamma\lambda^2(\mu + \eta)(\eta + \gamma)^3(\gamma(\gamma + \eta) + \eta(\mu + \eta))}}{R\lambda^2(\eta + \gamma)^3}, \\ x_2 &= \frac{\gamma(\mu + \eta)}{\lambda(\eta + \gamma)} \\ &\quad + \frac{\sqrt{RC\gamma\lambda^2(\mu + \eta)(\eta + \gamma)^3(\gamma(\gamma + \eta) + \eta(\mu + \eta))}}{R\lambda^2(\eta + \gamma)^3}. \end{aligned} \quad (\text{C.2})$$

We can verify that $\lambda x_1(\eta + \gamma) < \gamma(\mu + \eta)$ and $\lambda x_2(\eta + \gamma) > \gamma(\mu + \eta)$. Let q_t^* be the optimal joining probability. Then, we

have the following conclusions: if $0 \leq x_1 \leq 1$, then $q_t^* = x_1$; if $x_1 > 1$, then $q_t^* = 1$; implying that $q_t^* = \max\{0, \min\{x_1, 1\}\}$.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

References

- [1] P. Naor, "The regulation of queue size by levying tolls," *Econometrica*, vol. 37, no. 1, pp. 15–24, 1969.
- [2] U. Yechiali, "On optimal balking rules and toll charges in the GI/M/1 queue," *Operations Research*, vol. 19, no. 2, pp. 349–370, 1971.
- [3] S. G. Johansen and S. Stidham Jr., "Control of arrivals to a stochastic input-output system," *Advances in Applied Probability*, vol. 12, no. 4, pp. 972–999, 1980.
- [4] S. Stidham Jr., "Optimal control of admission to a queueing system," *Institute of Electrical and Electronics Engineers Transactions on Automatic Control*, vol. 30, no. 8, pp. 705–713, 1985.
- [5] H. Mendelson and S. Whang, "Optimal incentive-compatible priority pricing for the M/M/1 queue," *Operations Research*, vol. 38, no. 5, pp. 870–883, 1990.
- [6] H. Chen and M. Z. Frank, "State dependent pricing with a queue," *Institute of Industrial Engineers (IIE). IIE Transactions*, vol. 33, no. 10, pp. 847–860, 2001.
- [7] C. Larsen, "Investigating sensitivity and the impact of information on pricing decisions in an M/M/1/∞ queueing model," *International Journal of Production Economics*, vol. 56–57, pp. 365–377, 1998.
- [8] J. Erlichman and R. Hassin, "Equilibrium solutions in the observable M/M/1 queue with overtaking," in *Proceedings of the 4th International ICST Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS '09)*, article no. 64, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Pisa, Italy, October, 2009.
- [9] N. M. Edelson and D. K. Hildebrand, "Congestion tolls for Poisson queueing processes," *Econometrica*, vol. 43, no. 1, pp. 81–92, 1975.
- [10] R. Hassin and M. Haviv, "Equilibrium threshold strategies: The case of queues with priorities," *Operations Research*, vol. 45, no. 6, pp. 966–973, 1997.
- [11] R. Hassin and M. Haviv, *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*, International Series in Operations Research & Management Science, Kluwer Academic Publishers, Dordrecht, the Netherlands, 2003.
- [12] E. Gelenbe, "Random neural networks with negative and positive signals and product form solution," *Neural Computation*, vol. 1, no. 4, pp. 502–510, 1989.
- [13] E. Gelenbe, "G-networks: a unifying model for neural and queueing networks," *Annals of Operations Research*, vol. 48, no. 5, pp. 433–461, 1994.
- [14] J. R. Artalejo, "G-networks: A versatile approach for work removal in queueing networks," *European Journal of Operational Research*, vol. 126, no. 2, pp. 233–249, 2000.
- [15] K. C. Chae, H. M. Park, and W. S. Yang, "A GI/Geo/1 queue with negative and positive customers," *Applied Mathematical Modelling*, vol. 34, no. 6, pp. 1662–1671, 2010.
- [16] D. H. Lee, W. S. Yang, and H. M. Park, "Geo/G/1 queues with disasters and general repair times," *Applied Mathematical Modelling*, vol. 35, no. 4, pp. 1561–1570, 2011.
- [17] A. Economou and S. Kanta, "Equilibrium balking strategies in the observable single-server queue with breakdowns and repairs," *Operations Research Letters*, vol. 36, no. 6, pp. 696–699, 2008.
- [18] O. Boudali and A. Economou, "Optimal and equilibrium balking strategies in the single server Markovian queue with catastrophes," *European Journal of Operational Research*, vol. 218, no. 3, pp. 708–715, 2012.
- [19] O. Boudali and A. Economou, "The effect of catastrophes on the strategic customer behavior in queueing systems," *Naval Research Logistics (NRL)*, vol. 60, no. 7, pp. 571–587, 2013.
- [20] A. Economou and A. Manou, "Equilibrium balking strategies for a clearing queueing system in alternating environment," *Annals of Operations Research*, vol. 208, no. 1, pp. 489–514, 2013.
- [21] F. Zhang, J. Wang, and B. Liu, "On the optimal and equilibrium retrial rates in an unreliable retrial queue with vacations," *Journal of Industrial and Management Optimization*, vol. 8, no. 4, pp. 861–875, 2012.
- [22] F. Wang, J. Wang, and F. Zhang, "Strategic behavior in the single-server constant retrial queue with individual removal," *Quality Technology and Quantitative Management*, vol. 12, no. 3, pp. 325–342, 2015.
- [23] L. Li, J. Wang, and F. Zhang, "Equilibrium customer strategies in Markovian queues with partial breakdowns," *Computers & Industrial Engineering*, vol. 66, no. 4, pp. 751–757, 2013.
- [24] J. Wang and F. Zhang, "Equilibrium analysis of the observable queues with balking and delayed repairs," *Applied Mathematics and Computation*, vol. 218, no. 6, pp. 2716–2729, 2011.
- [25] S. Yu, Z. Liu, and J. Wu, "Equilibrium strategies of the unobservable M/M/1 queue with balking and delayed repairs," *Applied Mathematics and Computation*, vol. 290, pp. 56–65, 2016.
- [26] A. Veltman and R. Hassin, "Equilibrium in queueing systems with complementary products," *Queueing Systems*, vol. 50, no. 2–3, pp. 325–342, 2005.
- [27] W. Sun, S.-y. Li, N.-s. Tian, and H.-k. Zhang, "Equilibrium analysis in batch-arrival queues with complementary services," *Applied Mathematical Modelling: Simulation and Computation for Engineering and Environmental Systems*, vol. 33, no. 1, pp. 224–241, 2009.
- [28] A. Economou and S. Kanta, "Equilibrium customer strategies and social-profit maximization in the single-server constant retrial queue," *Naval Research Logistics (NRL)*, vol. 58, no. 2, pp. 107–122, 2011.
- [29] J. Wang and F. Zhang, "Monopoly pricing in a retrial queue with delayed vacations for local area network applications," *IMA Journal of Management Mathematics*, vol. 27, no. 2, pp. 315–334, 2016.
- [30] Y. Zhang, J. Wang, and F. Wang, "Equilibrium pricing strategies in retrial queueing systems with complementary services," *Applied Mathematical Modelling: Simulation and Computation for Engineering and Environmental Systems*, vol. 40, no. 11–12, pp. 5775–5792, 2016.
- [31] Y. Ma and Z. Liu, "Pricing analysis in GEO/GEO/1 queueing system," *Mathematical Problems in Engineering*, vol. 2015, Article ID 181653, 5 pages, 2015.
- [32] D. H. Lee, "A note on the optimal pricing strategy in the discrete-time Geo/Geo/1 queueing system with sojourn time-dependent reward," *Operations Research Perspectives*, vol. 4, pp. 113–117, 2017.

- [33] B. Yang, F. Tan, and Y. S. Dai, "Performance evaluation of cloud service considering fault recovery," *The Journal of Supercomputing*, vol. 65, no. 1, pp. 426–444, 2013.
- [34] J. Cao, K. Hwang, K. Li, and A. Y. Zomaya, "Optimal multiserver configuration for profit maximization in cloud computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 6, pp. 1087–1096, 2013.
- [35] R. W. Wolff, "Poisson arrivals see time averages," *Operations Research. The Journal of the Operations Research Society of America*, vol. 30, no. 2, pp. 223–231, 1982.
- [36] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

