Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 9746962, 10 pages
http://dx.doi.org/10.1155/2016/9746962

Research Article

Equilibrium and Optimal Strategies in M/M/1 Queues with Working Vacations and Vacation Interruptions

Ruiling Tian,[1] Linmin Hu,[1] and Xijun Wu[2]

[1] *College of Sciences, Yanshan University, Hebei, Qinhuangdao 066004, China*
[2] *Measurement Technology and Instrumentation Key Lab of Hebei Province, Yanshan University, Hebei, Qinhuangdao 066004, China*

Correspondence should be addressed to Ruiling Tian; tianrl@ysu.edu.cn

Received 6 October 2015; Accepted 5 January 2016

Academic Editor: Kishin Sadarangani

We consider the customers equilibrium and socially optimal joining-balking behavior in single-server Markovian queues with multiple working vacations and vacation interruptions. Arriving customers decide whether to join the system or balk, based on a linear reward-cost structure that incorporates their desire for service, as well as their unwillingness for waiting. We consider that the system states are observable, partially observable, and unobservable, respectively. For these cases, we first analyze the stationary behavior of the system and get the equilibrium strategies of the customers and compare them to socially optimal balking strategies numerically.

1. Introduction

In the past decades, there has been an emerging tendency in the literature to study queueing systems which are concerned with customers decentralized behavior and socially optimal control of arrivals. Such an economic analysis of queueing systems was pioneered by Naor [1] who studied the observable M/M/1 model with a simple linear reward-cost structure. It was assumed that an arriving customer observed the number of customers and then made his/her decision whether to join or balk. His study was complemented by Edelson and Hildebrand [2] who considered the same queueing system as that in [1] but assumed that the customers made their decisions without being informed about the state of the system. Moreover, several authors have investigated the same problem for various queueing systems incorporating diverse characteristics such as priorities, reneging, jockeying, schedules, and retrials. The fundamental results about various models can be found in the comprehensive monograph written by Hassin and Haviv [3] with extensive bibliographical references.

There are some papers in the literature that considered the strategic behavior of customers in classical vacation queueing models. Burnetas and Economou [4] studied a Markovian single-server queueing system with setup times. They derived equilibrium strategies for the customers under various levels of information and analyzed the stationary behavior of the system under these strategies. Sun and Tian [5] considered a queueing model with classical multiple vacations. They derived the equilibrium and socially optimal joining strategies in vacation and busy period of a partially observable queue, respectively. The conclusion was summarized that individual optimization led to excessive congestion without regulation. Economou and Kanta [6] considered the Markovian single-server queue that alternates between on and off periods. They derived equilibrium threshold strategies in fully observable and almost observable queues. Guo and Hassin [7] studied a vacation queue with N-policy and exhaustive service. They presented the equilibrium and socially optimal strategies for unobservable and observable queues. This work was extended by Guo and Hassin [8] to heterogenous customers and by Tian et al. [9] to M/G/1 queues. Recently, Guo and Li [10] studied strategic behavior and social optimization in partially observable Markovian vacation queues. Li et al. [11] considered equilibrium strategies in M/M/1 queues with partial breakdowns. Zhang et al. [12] studied a single-server retrial queue with two types of customers in which the server was subject to vacations along with breakdowns and repairs.

They discussed and compared the optimal and equilibrium retrial rates regarding the situations in which the customers were cooperative or noncooperative, respectively.

Recently, queueing systems with working vacations have been studied extensively where the server takes vacations once the system becomes empty (i.e., exhaustive service policy) and can still serve customers at a lower rate than regular one during the vacations. Research results of the stationary system performance on various working vacation queues can be consulted in the survey given by Tian et al. [13]. As for the work studying customers behavior in queueing systems with working vacations, Zhang et al. [14] and Sun and Li [15] obtained equilibrium balking strategies in M/M/1 queues with working vacations for four cases with respect to different levels of information. Sun et al. [16] also considered the customers equilibrium balking behavior in some single-server Markovian queues with two-stage working vacations.

However, we often encounter the situation that the server can stop the vacation once some indices of the system, such as the number of customers, achieve a certain value during a vacation. In many real life congestion situations, urgent events occur during a vacation and the server must come back to work rather than continuing to take the residual vacation. For example, if the number of customers exceeds the special value during a vacation and the server continues to take the vacation, it leads to large cost of waiting customers. Therefore, vacation interruption is more reasonable to the server vacation queues. Vacation interruption was introduced by Li and Tian [17] and Li et al. [18].

In this paper, we consider an M/M/1 queueing model with working vacations and vacation interruptions. We distinguish three cases: the observable queues, the partially observable queues, and the unobservable queues, according to the information levels regarding system states. As to the observable case, a customer can observe both the state of the server and the number of present customers before making decision. In the unobservable case, a customer cannot observe the state of the system. However, for the partially observable case, a customer only can observe the state of the server at their arrival instant and does not observe the number of customers present. The customers dilemma is whether to join the system or balk. We study the balking behavior of customers in these cases and derive Nash equilibrium strategies and socially optimal strategies.

This paper is organized as follows. Descriptions of the model are given in Section 2. In Sections 3, 4, and 5, we determine the equilibrium and socially optimal strategies in observable queues, the partially observable queues, and the unobservable queues, respectively. Conclusions are given in Section 6.

2. Model Description

Consider a classical M/M/1 queue with an arrival rate λ and a service rate μ_1. Upon the completion of service, if there is no customer in the system, the server begins a vacation and the vacation time is assumed to be exponentially distributed with the parameter θ. During a vacation period, arriving customers can be served at a mean rate of μ_0. Upon the completion of a service in the vacation, if there are also customers in the queue, the server ends the vacation and comes back to the normal working level. Otherwise, he/she continues the vacation until there are customers after a service or a vacation ends. Meanwhile, when a vacation ends, if there are no customers, another vacation is taken. Otherwise, the server also switches the rate to μ_1 and a regular busy period starts. In this service discipline, the server may come back from the vacation without completing the vacation. And the server can only go on vacations if there is no customer left in the system upon the completion of a service. Meanwhile, the vacation service rate can be only applied to the first customer that arrived during a vacation period.

We assume that the interarrival times, the service times, and the working vacation times are mutually independent. In addition, the service discipline is first in, first out (FIFO).

Denote by $N(t)$ the number of customers in the system at time t. Let $J(t) = 0$ be the state that the server is on working vacation period at time t and let $J(t) = 1$ be the state that the server is busy at time t.

Arriving customers are assumed to be identical. Our interest is in the customers' strategic response as they can decide whether to join or balk upon arrival. Assume that a customer's utility consists of a reward for receiving service minus a waiting cost. Specifically, every customer receives a reward of R units for completing service. There is a waiting cost of C units per time unit that the customer remains in the system. Customers are risk neutral and maximize their expected net benefit. Finally, we assume that there are no retrials of balking customers nor reneging of waiting customers.

3. The Observable Queues

We begin with the fully observable case in which the arriving customers know not only the number of present customers $N(t)$ but also the state of the server $J(t)$ at arrival time t.

There exists a balk threshold $n(i)$ such that an arriving customer enters the system at state i $(i = 0, 1)$ if the number of the customers present upon arrival does not exceed the specified threshold. So a pure threshold strategy is specified by a pair $(n_e(0), n_e(1))$ and the balking strategy has the following form: "while arriving at time t, observe $(N(t), J(t))$; enter if $N(t) \le n_e(J(t))$ and balk otherwise."

And we have the following result.

Theorem 1. *In the observable M/M/1 queue with working vacations and vacation interruptions, there exist thresholds $(n_e(0), n_e(1))$ which are given by*

$$\begin{aligned} n_e(0) &= \left\lfloor \frac{R\mu_1}{C} - \frac{\mu_1 + \theta}{\mu_1(\mu_0 + \theta)} \right\rfloor, \\ n_e(1) &= \left\lfloor \frac{R\mu_1}{C} \right\rfloor - 1, \end{aligned} \tag{1}$$

such that a customer who observes the system at state $(N(t), J(t))$ upon his arrival enters if $N(t) \le n_e(J(t))$ and balks otherwise.

Proof. Based on the reward-cost structure which is imposed on the system, we conclude that, for an arriving customer that decides to enter, his/her expected net benefit is

$$U(n,i) = R - CT(n,i), \tag{2}$$

where $T(n,i)$ denotes his/her expected mean sojourn time given that he/she finds the system at state (n,i) upon his/her arrival. Then, we have the following equations:

$$T(0,0) = \frac{1}{\mu_0+\theta} + \frac{\theta}{\mu_0+\theta}\frac{1}{\mu_1}, \tag{3}$$

$$\begin{aligned} T(n,0) &= \frac{\mu_0}{\mu_0+\theta}\left(\frac{1}{\mu_0} + T(n-1,1)\right) \\ &\quad + \frac{\theta}{\mu_0+\theta}T(n,1), \end{aligned} \tag{4}$$

$$T(n,1) = \frac{n+1}{\mu_1}. \tag{5}$$

By iterating (4) and taking into account (5), we obtain

$$T(n,0) = \frac{n}{\mu_1} + \frac{\mu_1+\theta}{\mu_1(\mu_0+\theta)}. \tag{6}$$

We can easily check that $T(n,0)$ is strictly increasing for n. A customer strictly prefers to enter if the reward for service exceeds the expected cost for waiting (i.e., $U(n,i) > 0$) and is indifferent between entering and balking if the reward is equal to the cost (i.e., $U(n,i) = 0$).

We assume throughout the paper that

$$R > C\left(\frac{1}{\mu_0+\theta} + \frac{\theta}{\mu_0+\theta}\frac{1}{\mu_1}\right), \tag{7}$$

which ensures that the reward for service exceeds the expected cost for a customer who finds the system empty. By solving $U(n,i) \ge 0$ for n, using (5) and (6), we obtain that the customer arriving at time t decides to enter if and only if $n \le n_e(J(t))$, where $(n_e(0), n_e(1))$ are given by (1).

From Figure 1, we observe that $n_e(1)$ is larger than $n_e(0)$, and $n_e(0)$ and $n_e(1)$ are all increasing with μ_1.

For the stationary analysis, the system follows a Markov chain and the state space is

$$\Omega_{fo} = \{0,0\} \cup \{(n,i) \mid n = 1,2,\ldots,n(i)+1;\ i = 0,1\}. \tag{8}$$

And the transition diagram is illustrated in Figure 2.

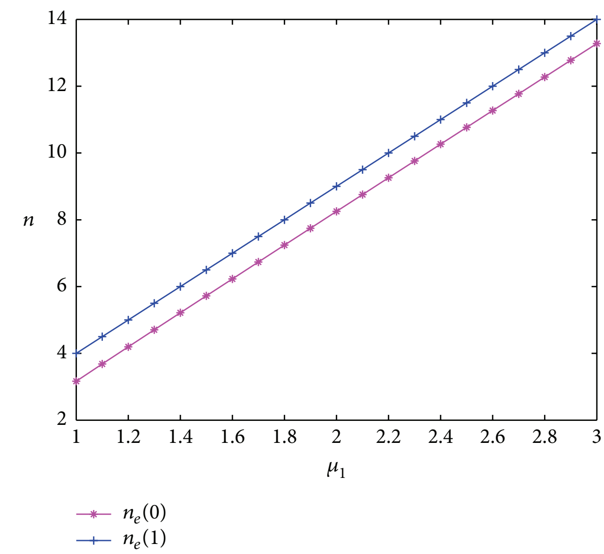

FIGURE 1: Equilibrium threshold strategies for observable case when $R = 5$, $C = 1$, $\lambda = 1$, $\mu_0 = 0.5$, and $\theta = 0.1$.

Let $\pi_{ni} = \lim_{t\to\infty} P\{N(t) = n, J(t) = i\}$ with $(n,i) \in \Omega_{fo}$; then $\{\pi_{ni} : (n,i) \in \Omega_{fo}\}$ is the stationary distribution of the process $\{(N(t), J(t)),\ t \ge 0\}$.

In steady state, we have the following system equations:

$$\lambda\pi_{00} = \mu_0\pi_{10} + \mu_1\pi_{11}, \tag{9}$$

$$(\lambda+\theta+\mu_0)\pi_{n0} = \lambda\pi_{n-1,0}, \quad n = 1,2,\ldots,n(0), \tag{10}$$

$$(\theta+\mu_0)\pi_{n(0)+1,0} = \lambda\pi_{n(0),0}, \tag{11}$$

$$(\lambda+\mu_1)\pi_{11} = \mu_1\pi_{21} + \theta\pi_{10} + \mu_0\pi_{20}, \tag{12}$$

$$\begin{aligned} (\lambda+\mu_1)\pi_{n1} &= \lambda\pi_{n-1,1} + \mu_1\pi_{n+1,1} + \theta\pi_{n0} \\ &\quad + \mu_0\pi_{n+1,0}, \\ &\qquad n = 2,3,\ldots,n(0), \end{aligned} \tag{13}$$

$$(\lambda+\mu_1)\pi_{n(0)+1,1} = \lambda\pi_{n(0)1} + \mu_1\pi_{n(0)+2} + \theta\pi_{n(0)+1,0}, \tag{14}$$

$$\begin{aligned} (\lambda+\mu_1)\pi_{n1} &= \lambda\pi_{n-1,1} + \mu_1\pi_{n+1,1}, \\ &\qquad n = n(0)+2,\ldots,n(1), \end{aligned} \tag{15}$$

$$\mu_1\pi_{n(1)+1,1} = \lambda\pi_{n(1),1}. \tag{16}$$

Define

$$\begin{aligned} \rho &= \frac{\lambda}{\mu_1}, \\ \sigma &= \frac{\lambda}{\lambda+\theta+\mu_0}. \end{aligned} \tag{17}$$

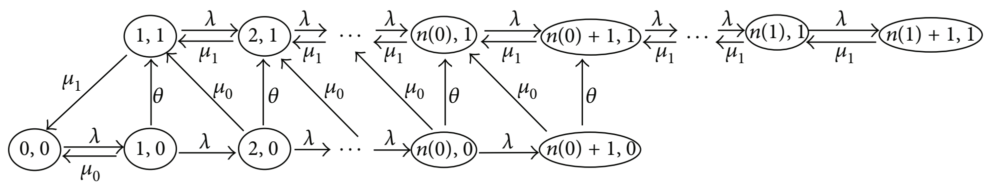

Figure 2: Transition rate diagram for the observable queues.

By iterating (10) and (15), taking into account (11) and (16), we obtain

$$\begin{aligned} \pi_{n0} &= \sigma^n \pi_{00}, \quad n = 0, 1, 2, \ldots, n(0), \\ \pi_{n(0)+1,0} &= \frac{\lambda \sigma^{n(0)}}{\theta + \mu_0} \pi_{00}, \\ \pi_{n1} &= \rho^{n-n(0)-1} \pi_{n(0)+1,1}, \\ &\qquad n = n(0) + 1, \ldots, n(1) + 1. \end{aligned} \tag{18}$$

From (13), it is easy to obtain that $\{\pi_{n1} \mid 1 \le n \le n(0)+1\}$ is a solution of the nonhomogeneous linear difference equation

$$\begin{aligned} \mu_1 x_{n+1} - (\lambda + \mu_1) x_n + \lambda x_{n-1} &= -\theta \pi_{n0} - \mu_0 \pi_{n+1,0} \\ &= -(\theta + \mu_0 \sigma) \sigma^n \pi_{00}. \end{aligned} \tag{19}$$

So its corresponding characteristic equation is

$$\mu_1 x^2 - (\lambda + \mu_1) x + \lambda = 0, \tag{20}$$

which has two roots at 1 and ρ. Assume that $\rho \neq 1$; then the homogeneous solution of (19) is $x_n^{\text{hom}} = A1^n + B\rho^n$. The general solution of the nonhomogeneous equation is given as $x_n^{\text{gen}} = x_n^{\text{hom}} + x_n^{\text{spec}}$, where x_n^{spec} is a specific solution. We find a specific solution $x_n^{\text{spec}} = D\sigma^n$. Substituting it into (19), we derive

$$D = -\frac{\lambda + \theta}{\lambda + \theta + \mu_0 - \mu_1} \pi_{00}. \tag{21}$$

Therefore,

$$x_n^{\text{gen}} = A + B\rho^n + D\sigma^n, \quad n = 1, 2, \ldots, n(0) + 1. \tag{22}$$

Substituting (22) into (9) and (12), it follows after some rather tedious algebra that

$$\begin{aligned} A + B\rho &= -\frac{\lambda (\lambda + \theta)}{\mu_1 (\mu_1 - \lambda - \theta - \mu_0)} \pi_{00}, \\ A + B\rho^2 &= -\frac{\lambda^2 (\lambda + \theta)}{\mu_1^2 (\mu_1 - \lambda - \theta - \mu_0)} \pi_{00}. \end{aligned} \tag{23}$$

Then, we obtain

$$\begin{aligned} A &= 0, \\ B &= -\frac{\lambda + \theta}{\mu_1 - \lambda - \theta - \mu_0} \pi_{00}. \end{aligned} \tag{24}$$

Thus, from (22), we obtain

$$\pi_{n1} = B\rho^n + D\sigma^n, \quad n = 1, 2, \ldots, n(0) + 1. \tag{25}$$

Thus, we have all the stationary probabilities in terms of π_{00}. The remaining probability, π_{00}, can be found from the normalization condition. After some algebraic simplifications, we can summarize the results in the following. □

Theorem 2. *In the observable M/M/1 queues with working vacations and vacation interruptions, the stationary distribution $\{\pi_{ni} \mid (n, i) \in \Omega_{fo}\}$ is given as follows:*

$$\begin{aligned} \pi_{n0} &= \sigma^n \pi_{00}, \quad n = 0, 1, 2, \ldots, n(0), \\ \pi_{n(0)+1,0} &= \frac{\lambda \sigma^{n(0)}}{\theta + \mu_0} \pi_{00}, \\ \pi_{n1} &= B\rho^n + D\sigma^n, \quad n = 1, 2, \ldots, n(0) + 1, \\ \pi_{n1} &= \rho^{n-n(0)-1} \left(B\rho^{n(0)+1} + D\sigma^{n(0)+1} \right), \\ &\qquad n = n(0) + 2, \ldots, n(1) + 1, \end{aligned} \tag{26}$$

where π_{00} can be solved by the normalization equation.

Because the probability of balking is equal to $\pi_{n(0)+1,0} + \pi_{n(1)+1,1}$, the social benefit per time unit when all customers follow the threshold policy $(n(0), n(1))$ equals

$$\begin{aligned} U_s(n(0), n(1)) &= \lambda R \left(1 - \pi_{n(0)+1,0} - \pi_{n(1)+1,1}\right) \\ &\quad - C \left(\sum_{n=0}^{n(0)+1} n\pi_{n0} + \sum_{n=1}^{n(1)+1} n\pi_{n1} \right). \end{aligned} \tag{27}$$

Let $(n^*(0), n^*(1))$ be the socially optimal threshold strategy. From Figure 3, we obtain $(n^*(0), n^*(1)) = (3, 4)$, while the equilibrium threshold strategy $(n_e(0), n_e(1)) = (78, 79)$. We observe that $n_e(0) > n^*(0)$ and $n_e(1) > n^*(1)$, which indicates that the individual optimization is inconsistent with the social optimization.

4. Partially Observable Queues

In this section, we turn our attention to the partially observable case, where arriving customers only can observe the state of the server at their arrival instant and do not observe the number of customers present.

A mixed strategy for a customer is specified by a vector (q_0, q_1), where q_i is the probability of joining when the server is in state i $(i = 0, 1)$. If all customers follow the same mixed

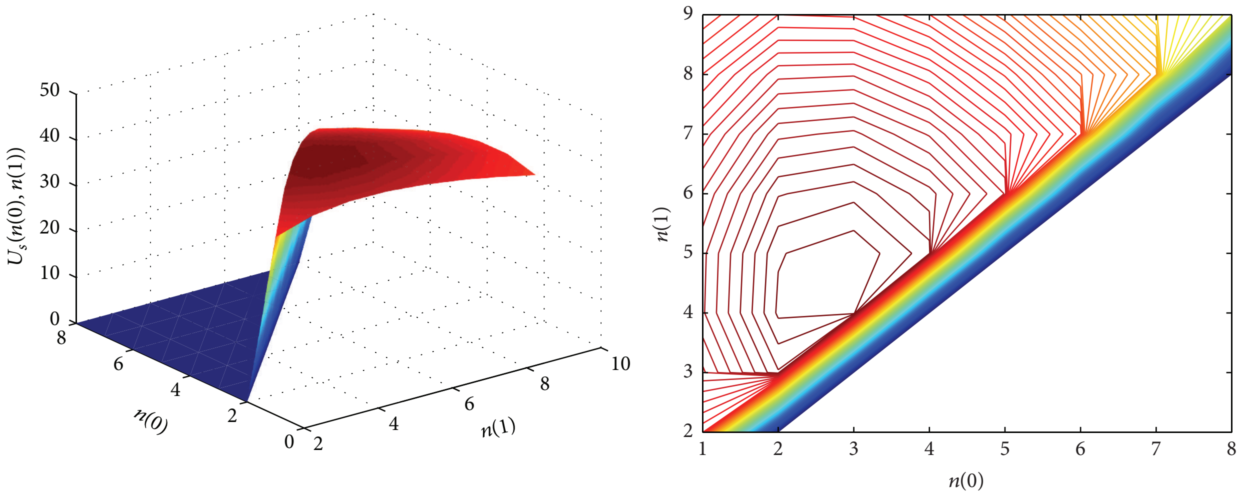

Figure 3: Social welfare for the observable case when $R = 40$, $C = 1$, $\lambda = 1.5$, $\mu_1 = 2$, $\mu_0 = 0.5$, and $\theta = 0.5$.

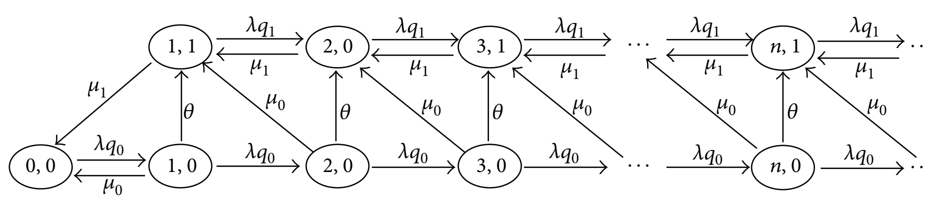
Figure 4: Transition rate diagram for the partially observable queues.

strategy (q_0, q_1), then the system follows a Markov chain where the arrival rate equals $\lambda_i = \lambda q_i$ when the server is in state i $(i = 0, 1)$. The state space is $\Omega_{po} = \{0,0\} \cup \{(n,i) \mid n \geq 1,\ i = 0,1\}$ and the transition diagram is illustrated in Figure 4.

Denote the stationary distribution as

$$\pi_{ni} = \lim_{t\to\infty} P\{N(t) = n,\ J(t) = i\}, \quad (n,i) \in \Omega_{po},$$
$$\pi_0 = \pi_{00}, \tag{28}$$
$$\boldsymbol{\pi}_n = (\pi_{n0}, \pi_{n1}), \quad n \geq 1.$$

Using the lexicographical sequence for the states, the transition rate matrix (generator) $\mathbf{Q}$ can be written as the tridiagonal block matrix:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{A}_0 & \mathbf{C}_0 & & & \\ \mathbf{B}_1 & \mathbf{A} & \mathbf{C} & & \\ & \mathbf{B} & \mathbf{A} & \mathbf{C} & \\ & & \mathbf{B} & \mathbf{A} & \mathbf{C} \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \tag{29}$$

where

$$\mathbf{A}_0 = -\lambda q_0,$$
$$\mathbf{B}_1 = \begin{pmatrix} \mu_0 \\ \mu_1 \end{pmatrix},$$
$$\mathbf{C}_1 = \begin{pmatrix} \lambda q_0 & 0 \end{pmatrix},$$
$$\mathbf{A} = \begin{pmatrix} -(\lambda q_0 + \theta + \mu_0) & \theta \\ 0 & -(\lambda q_1 + \mu_1) \end{pmatrix},$$
$$\mathbf{B} = \begin{pmatrix} 0 & \mu_0 \\ 0 & \mu_1 \end{pmatrix},$$
$$\mathbf{C} = \begin{pmatrix} \lambda q_0 & 0 \\ 0 & \lambda q_1 \end{pmatrix}. \tag{30}$$

To analyze this QBD process, it is necessary to solve for the minimal nonnegative solution of the matrix quadratic equation

$$\mathbf{R}^2\mathbf{B} + \mathbf{R}\mathbf{A} + \mathbf{C} = 0, \tag{31}$$

and this solution is called the rate matrix and denoted by $\mathbf{R}$.

Lemma 3. *If $\rho_1 = \lambda q_1/\mu_1 < 1$, the minimal nonnegative solution of (31) is given by*

$$\mathbf{R} = \begin{pmatrix} \sigma_0 & \dfrac{\lambda q_0 + \theta}{\mu_1}\sigma_0 \\ 0 & \rho_1 \end{pmatrix}, \tag{32}$$

where $\sigma_0 = \lambda q_0/(\lambda q_0 + \theta + \mu_0)$.

Proof. Because the coefficients of (31) are all upper triangular matrices, we can assume that $\mathbf{R}$ has the same structure as

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}. \tag{33}$$

Substituting $\mathbf{R}^2$ and $\mathbf{R}$ into (31) yields the following set of equations:

$$\begin{aligned} -\left(\lambda q_0+\theta+\mu_0\right) r_{11}+\lambda q_0&=0, \\ \mu_1 r_{22}^2-\left(\lambda q_1+\mu_1\right) r_{22}+\lambda q_1&=0, \\ \mu_0 r_{11}^2+\mu_1 r_{12}\left(r_{11}+r_{22}\right)+\theta r_{11}-\left(\lambda q_1+\mu_1\right) r_{12}&=0. \end{aligned} \tag{34}$$

To obtain the minimal nonnegative solution of (31), we take $r_{22}=\rho_1$ (the other roots are $r_{22}=1$) in the second equation of (34). From the first equation of (34), we obtain $r_{11}=\lambda q_0/(\lambda q_0+\theta+\mu_0)=\sigma_0$. Substituting σ_0 and ρ_1 into the last equation of (34), we get $r_{12}=((\lambda q_0+\theta)/\mu_1)\sigma_0$. This is the proof of Lemma 3. □

Theorem 4. *Assuming that $\rho_1<1$, the stationary probabilities $\{\pi_{ni}\mid(n,i)\in\Omega_{po}\}$ of the partially observable queues are as follows:*

$$\begin{aligned} \pi_{n0}&=K\sigma_0^n, \quad n\geq 0, \\ \pi_{n1}&=K\rho_1\frac{\lambda q_0+\theta}{\lambda q_0+\theta+\mu_0}\sum_{j=0}^{n-1}\rho_1^j\sigma_0^{n-1-j}, \quad n\geq 1, \end{aligned} \tag{35}$$

where

$$K=\frac{\theta+\mu_0}{\lambda q_0+\theta+\mu_0}\left(1+\frac{\lambda q_0+\theta}{\lambda q_0+\theta+\mu_0}\frac{\lambda q_0}{\mu_1-\lambda q_1}\right)^{-1}. \tag{36}$$

Proof. With the matrix-geometric solution method [19], we have

$$\boldsymbol{\pi}_n=\left(\pi_{n0},\pi_{n1}\right)=\left(\pi_{10},\pi_{11}\right)\mathbf{R}^{n-1}, \quad n\geq 1. \tag{37}$$

In addition, $(\pi_{00},\pi_{10},\pi_{11})$ satisfies the set of equations

$$\left(\pi_{00},\pi_{10},\pi_{11}\right)B[\mathbf{R}]=0, \tag{38}$$

where

$$\begin{aligned} B[\mathbf{R}]&=\begin{pmatrix}\mathbf{A}_0 & \mathbf{C}_0 \\ \mathbf{B}_1 & \mathbf{A}+\mathbf{R}\mathbf{B}\end{pmatrix} \\ &=\begin{pmatrix}-\lambda q_0 & \lambda q_0 & 0 \\ \mu_0 & -\left(\lambda q_0+\theta+\mu_0\right) & \lambda q_0+\theta \\ \mu_1 & 0 & -\mu_1\end{pmatrix}. \end{aligned} \tag{39}$$

Then, from (38), we derive

$$\begin{aligned} -\lambda q_0\pi_{00}+\mu_0\pi_{10}+\mu_1\pi_{11}&=0, \\ \lambda q_0\pi_{00}-\left(\lambda q_0+\theta+\mu_0\right)\pi_{10}&, \\ \left(\lambda q_0+\theta\right)\pi_{10}-\mu_1\pi_{11}&=0. \end{aligned} \tag{40}$$

Taking π_{00} as a known constant, we obtain

$$\pi_{10}=\sigma_0\pi_{00}, \tag{41}$$

$$\pi_{11}=\rho_1\frac{\lambda q_0+\theta}{\lambda q_0+\theta+\mu_0}\pi_{00}. \tag{42}$$

Note that

$$\mathbf{R}^{n-1}=\begin{pmatrix}\sigma_0^{n-1} & \dfrac{\lambda q_0+\theta}{\mu_1}\sigma_0\sum_{j=0}^{n-2}\sigma_0^j\rho_1^{n-j-2} \\ 0 & \rho_1^{n-1}\end{pmatrix}, \quad n\geq 2. \tag{43}$$

By π_{10}, π_{11}, $\mathbf{R}^{n-1}$, and (37), we get (35). Finally, π_{00} can be determined by the normalization condition. This is the proof of Theorem 4. □

Let p_i be the steady-state probability when the server is at state i $(i=0,1)$; we have

$$p_0=\sum_{n=0}^{\infty}\pi_{n0}=K\frac{\lambda q_0+\theta+\mu_0}{\theta+\mu_0}, \tag{44}$$

$$p_1=\sum_{n=1}^{\infty}\pi_{n1}=K\frac{\lambda q_0\left(\lambda q_0+\theta\right)}{\left(\mu_1-\lambda q_1\right)\left(\theta+\mu_0\right)}. \tag{45}$$

So the effective arrival rate $\overline{\lambda}$ is given by

$$\overline{\lambda}=\lambda\left(p_0q_0+p_1q_1\right). \tag{46}$$

We now consider a customer who finds the server at state i upon arrival. The conditional mean sojourn time of such a customer that decides to enter given that the others follow the same mixed strategy (q_0,q_1) is given by

$$W_i\left(q_0,q_1\right)=\frac{\sum_{n=i}^{\infty}T(n,i)\pi_{ni}}{\sum_{n=i}^{\infty}\pi_{ni}}. \tag{47}$$

By substituting (5)-(6) and (35) into (44), we obtain

$$\begin{aligned} W_0\left(q_0,q_1\right)&=\frac{\lambda q_0+\theta+\mu_1}{\mu_1\left(\theta+\mu_0\right)}, \\ W_1\left(q_0,q_1\right)&=\frac{1}{\mu_1-\lambda q_1}+\frac{\lambda q_0+\theta+\mu_0}{\mu_1\left(\mu_0+\theta\right)}. \end{aligned} \tag{48}$$

Based on the reward-cost structure, the expected net benefit of an arriving customer if he/she finds the server at state i and decides to enter is given by

$$\begin{aligned} U_0\left(q_0\right)&=R-C\frac{\lambda q_0+\theta+\mu_1}{\mu_1\left(\theta+\mu_0\right)}, \\ U_1\left(q_0,q_1\right)&=R-C\left(\frac{1}{\mu_1-\lambda q_1}+\frac{\lambda q_0+\theta+\mu_0}{\mu_1\left(\mu_0+\theta\right)}\right). \end{aligned} \tag{49}$$

We now can proceed to determine mixed equilibrium strategies of a customer in the partially observable case and have the following.

Theorem 5. *For the partially observable case, there exists a unique mixed equilibrium strategy (q_0^e,q_1^e), where the vector (q_0^e,q_1^e) is given as follows:*

(1) $C(\theta+\mu_1)/\mu_1(\theta+\mu_0)<R\leq C(\lambda+\theta+\mu_1)/\mu_1(\theta+\mu_0)$:

$$\left(q_0^e,q_1^e\right)=\begin{cases}\left(x_1,0\right), & \dfrac{\mu_1-\mu_0}{\mu_1\left(\theta+\mu_0\right)}<\dfrac{1}{\mu}, \\ \left(x_1,x_2\right), & \dfrac{1}{\mu}\leq\dfrac{\mu_1-\mu_0}{\mu_1\left(\theta+\mu_0\right)}\leq\dfrac{1}{\mu-\lambda}, \\ \left(x_1,1\right), & \dfrac{\mu_1-\mu_0}{\mu_1\left(\theta+\mu_0\right)}>\dfrac{1}{\mu-\lambda}.\end{cases} \tag{50}$$

(2) $C(\lambda+\theta+\mu_1)/\mu_1(\theta+\mu_0)<R$:

$$(q_0^e, q_1^e) = \begin{cases} (1,0), & R < \dfrac{C}{\mu_1} + \dfrac{C(\lambda+\theta+\mu_0)}{\mu_1(\theta+\mu_0)}, \\ (1,x_3), & \dfrac{C}{\mu_1} + \dfrac{C(\lambda+\theta+\mu_0)}{\mu_1(\theta+\mu_0)} \le R \le \dfrac{C}{\mu_1-\lambda} + \dfrac{C(\lambda+\theta+\mu_0)}{\mu_1(\theta+\mu_0)}, \\ (1,1), & R > \dfrac{C}{\mu_1-\lambda} + \dfrac{C(\lambda+\theta+\mu_0)}{\mu_1(\theta+\mu_0)}, \end{cases} \tag{51}$$

where

$$\begin{aligned} x_1 &= \frac{1}{\lambda}\left(\frac{R\mu_1(\theta+\mu_0)}{C} - \theta - \mu_1\right), \\ x_2 &= \frac{\mu_1}{\lambda}\left(1 - \frac{\mu_0+\theta}{\mu_1-\mu_0}\right), \\ x_3 &= \frac{1}{\lambda}\left(\mu_1 - \frac{1}{R/C - (\lambda+\theta+\mu_0)/\mu_1(\theta+\mu_0)}\right). \end{aligned} \tag{52}$$

Proof. To prove Theorem 5, we first focus on $U_0(q_0)$. Condition (1) assures that q_0^e is positive. Therefore, we have two cases.

Case 1 ($U_0(0)>0$ and $U_0(1)\le 0$). That is, $C(\theta+\mu_1)/\mu_1(\theta+\mu_0) < R \le C(\lambda+\theta+\mu_1)/\mu_1(\theta+\mu_0)$). In this case, if all customers who find the system empty enter the system with probability $q_0^e = 1$, then the tagged customer suffers a negative expected benefit if he/she decides to enter. Hence, $q_0^e = 1$ does not lead to an equilibrium. Similarly, if all customers use $q_0^e = 0$, then the tagged customer receives a positive benefit from entering; thus, $q_0^e = 0$ also cannot be part of an equilibrium mixed strategy. Therefore, there exists unique q_0^e satisfying

$$R - C\frac{\lambda q_0^e + \theta + \mu_1}{\mu_1(\theta+\mu_0)} = 0, \tag{53}$$

for which customers are indifferent between entering and balking. This is given by

$$q_0^e = \frac{1}{\lambda}\left(\frac{R\mu_1(\theta+\mu_0)}{C} - \theta - \mu_1\right) = x_1. \tag{54}$$

In this situation, the expected benefit U_1 is given by

$$U_1(q_0^e, q_1) = \frac{C(\mu_1-\mu_0)}{\mu_1(\theta+\mu_0)} - \frac{C}{\mu - \lambda q_1}. \tag{55}$$

Using the standard methodology of equilibrium analysis in unobservable queueing models, we derive from (55) the following equilibria:

(1) If $U_1(q_0^e, 0) < 0$, then the equilibrium strategy is $q_1^e = 0$.

(2) If $U_1(q_0^e, 0) \ge 0$ and $U_1(q_0^e, 1) \le 0$, there exists unique q_1^e, satisfying

$$U_1(q_0^e, q_1^e) = 0; \tag{56}$$

then the equilibrium strategy is $q_1^e = x_2$.

(3) If $U_1(q_0^e, 1) > 0$, then the equilibrium strategy is $q_1^e = 1$.

Case 2 ($U_0(1)\ge 0$). That is, $C(\lambda+\theta+\mu_1)/\mu_1(\theta+\mu_0) < R$). In this case, for every strategy of the other customers, the tagged customer has a positive expected net benefit if he/she decides to enter. Hence, $q_0^e = 1$.

In this situation, the expected benefit U_1 is given by

$$U_1(1, q_1) = R - \frac{C}{\mu - \lambda q_1} - \frac{C(\lambda+\theta+\mu_0)}{\mu_1(\theta+\mu_0)}. \tag{57}$$

To find q_1^e in equilibrium, we also consider the following subcases:

(1) If $U_1(1,0) < 0$, then the equilibrium strategy is $q_1^e = 0$.

(2) If $U_1(1,0) \ge 0$ and $U_1(1,1) \le 0$, there exists unique q_1^e, satisfying

$$U_1(1, q_1^e) = 0; \tag{58}$$

then the equilibrium strategy is $q_1^e = x_3$.

(3) If $U_1(1,1) > 0$, then the equilibrium strategy is $q_1^e = 1$.

By rearranging Cases 1 and 2, we can obtain the results of Theorem 5. This completes the proof. □

From Theorem 5, we can compare q_0^e with q_1^e. Figure 5 shows that q_0^e is not always less than q_1^e. Namely, sometimes the information that the server takes a working vacation does not make the customers less willing to enter the system, because in the vacation the server provides a lower service to customers. But when the vacation time and waiting time become longer, customers do not want to enter the system in vacation.

Then, from Theorem 4, the mean queue length is given by

$$\begin{aligned} E[L] &= \sum_{n=0}^{\infty} n\pi_{n0} + \sum_{n=0}^{\infty} n\pi_{n1} \\ &= K\frac{\sigma_0}{(1-\sigma_0)^2}\left(1 + \frac{\lambda q_0 + \theta}{\mu_1}\frac{1-\rho_1\sigma_0}{(1-\rho_1)^2}\right). \end{aligned} \tag{59}$$

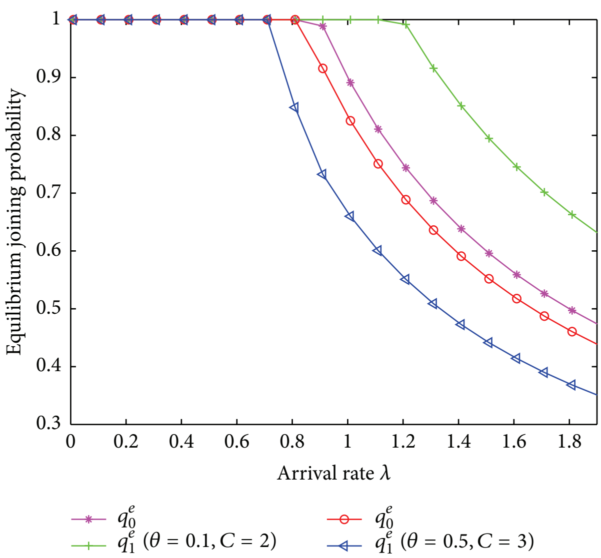

FIGURE 5: Equilibrium mixed strategies for the partially observable case when $R = 5$, $\mu_1 = 2$, and $\mu_0 = 0.5$.

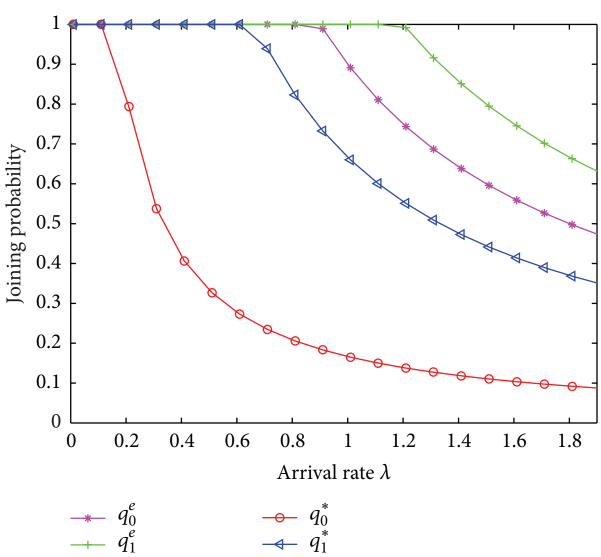

FIGURE 6: Equilibrium and socially optimal mixed strategies for the partially observable case when $R = 5, C = 2, \mu_1 = 2, \mu_0 = 0.5$, and $\theta = 0.1$.

And the social benefit when all customers follow a mixed policy (q_0, q_1) can now be easily computed as follows:

$$\begin{aligned} U_s(q_0, q_1) &= \bar{\lambda} R - CE[L] \\ &= \lambda (p_0 q_0 + p_1 q_1) R \\ &\quad - \frac{CK\sigma_0}{(1-\sigma_0)^2}\left(1 + \frac{\lambda q_0 + \theta}{\mu_1}\frac{1-\rho_1\sigma_0}{(1-\rho_1)^2}\right). \end{aligned} \tag{60}$$

The goal of a social planner is to maximize overall social welfare. Let (q_0^*, q_1^*) be the socially optimal mixed strategy. Figure 6 compares customers equilibrium mixed strategy (q_0^e, q_1^e) and their socially optimal mixed strategy (q_0^*, q_1^*). We also observe that $q_0^e > q_0^*$ and $q_1^e > q_1^*$. This ordering is typical when customers make individual decisions maximizing their own profit. Then, they ignore those negative externalities that they impose on later arrivals, and they tend to overuse the system. It is clear that these externalities should be taken into account when we aim to maximize the total revenue.

5. The Unobservable Queues

In this section, we consider the fully unobservable case, where arriving customers cannot observe the system state.

There are two pure strategies available for a customer, that is, to join the queue or not to join the queue. A pure or mixed strategy can be described by a fraction q $(0 \le q \le 1)$, which is the probability of joining, and the effective arrival rate, or joining rate, is λq. The transition diagram is illustrated in Figure 7.

To identify the equilibrium strategies of the customers, we should first investigate the stationary distribution of the system when all customers follow a given strategy q, which can be obtained by Theorem 4 by taking $q_0 = q_1 = q$. So we have

$$\begin{aligned} \pi_{n0} &= K_1 \sigma_q^n, \quad n \ge 0, \\ \pi_{n1} &= K_1 \rho_q \frac{\lambda q + \theta}{\lambda q + \theta + \mu_0} \sum_{j=0}^{n-1} \rho_q^j \sigma_q^{n-1-j}, \quad n \ge 1, \end{aligned} \tag{61}$$

where $\rho_q = \lambda q/\mu_1$, $\sigma_q = \lambda q/(\lambda q + \theta + \mu_0)$, and

$$K_1 = \frac{\theta + \mu_0}{\lambda q + \theta + \mu_0}\left(1 + \frac{\lambda q + \theta}{\lambda q + \theta + \mu_0}\frac{\lambda q}{\mu_1 - \lambda q}\right)^{-1}. \tag{62}$$

Then, the mean number of the customers in the system is

$$\begin{aligned} E[L] &= \sum_{n=0}^{\infty} n\pi_{n0} + \sum_{n=0}^{\infty} n\pi_{n1} \\ &= \frac{K_1\sigma_q}{(1-\sigma_q)^2}\left(1 + \frac{\lambda q + \theta}{\mu_1}\frac{1-\rho_q\sigma_q}{(1-\rho_q)^2}\right). \end{aligned} \tag{63}$$

Hence, the mean sojourn time of a customer who decides to enter upon his/her arrival can be obtained by using Little's law:

$$\begin{aligned} E[W] &= \frac{K_1\sigma_q}{\lambda q(1-\sigma_q)^2}\left(1 + \frac{\lambda q + \theta}{\mu_1}\frac{1-\rho_q\sigma_q}{(1-\rho_q)^2}\right) \\ &= \frac{1}{\theta + \mu_0}\left(1 \vphantom{\frac{1}{1}}\right. \\ &\quad \left. + \frac{\mu_1(\theta+\mu_0)}{\mu_1 - \lambda q}\frac{1}{\mu_0((\mu_1 - \lambda q)/(\lambda q + \theta)) + \mu_1}\right). \end{aligned} \tag{64}$$

And if $\lambda < \mu_1$, $E[W]$ is strictly increasing for $q \in [0, 1]$.

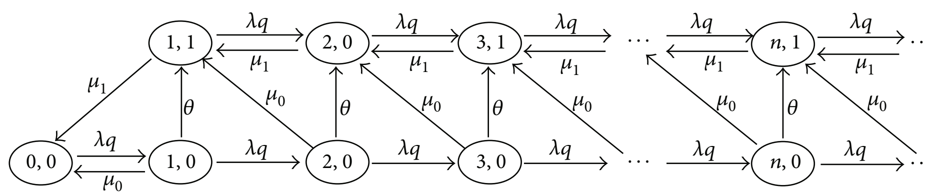

Figure 7: Transition rate diagram for the unobservable queues.

We consider a tagged customer. If he/she decides to enter the system, his/her expected net benefit is

$$U(q) = R - CE[W] = R - \frac{C}{\theta+\mu_0}\left(1 + \frac{\mu_1(\theta+\mu_0)}{\mu_1-\lambda q}\frac{1}{\mu_0((\mu_1-\lambda q)/(\lambda q+\theta))+\mu_1}\right). \quad (65)$$

We note that $U(q)$ is strictly decreasing in q, and it has a unique root q_e^*. So there exists a unique mixed equilibrium strategy q_e and $q_e = \min\{q_e^*, 1\}$.

We now turn our attention to social optimization. The social benefit per time unit can now be easily computed as

$$U_s(q) = \lambda q(R - CE[W]) = \lambda q\left(R - \frac{C}{\theta+\mu_0}\left(1 + \frac{\mu_1(\theta+\mu_0)}{\mu_1-\lambda q}\frac{1}{\mu_0((\mu_1-\lambda q)/(\lambda q+\theta))+\mu_1}\right)\right). \quad (66)$$

As for the equilibrium social welfare per time unit $U_s(q_e)$, Figure 8 shows that $U_s(q_e)$ first increases and then decreases with λ, and the social benefit achieves a maximum for intermediate values of this parameter. The reason for this behavior is that when the arrival rate is small, the system is rarely crowded; therefore, as more customers arrive, they are served and the social benefit improves. However, with any small increase of the arrival rate λ, the expected waiting time increases, which has a detrimental effect on the social benefit.

Let x^* be the root of the equation $S'(q) = 0$, and let q^* be the optimal joining probability. Then, we have the following conclusions: if $0 < x^* < 1$ and $S''(q) \le 0$, then $q^* = x^*$; if $0 < x^* < 1$ and $S''(q) > 0$ or $x^* \ge 1$, then $q^* = 1$.

From Figure 9, we observe that $q_e > q^*$. As a result, individual optimization leads to queues that are longer than socially desired. Therefore, it is clear that the social planner wants a toll to discourage arrivals.

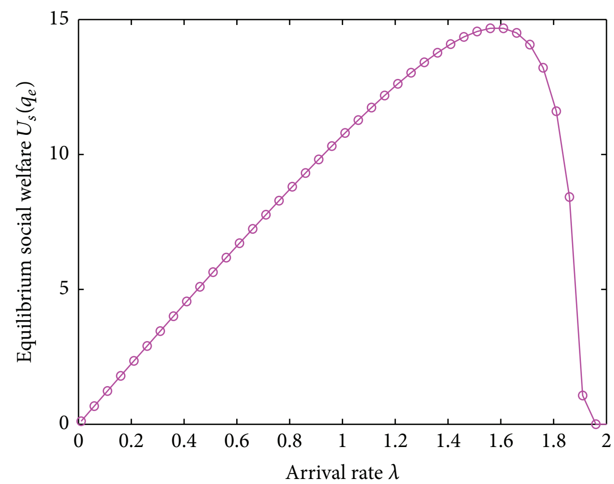

Figure 8: Equilibrium social welfare for the unobservable case when $R = 10$, $C = 1$, $\mu_1 = 2$, $\mu_0 = 1$, and $\theta = 1$.

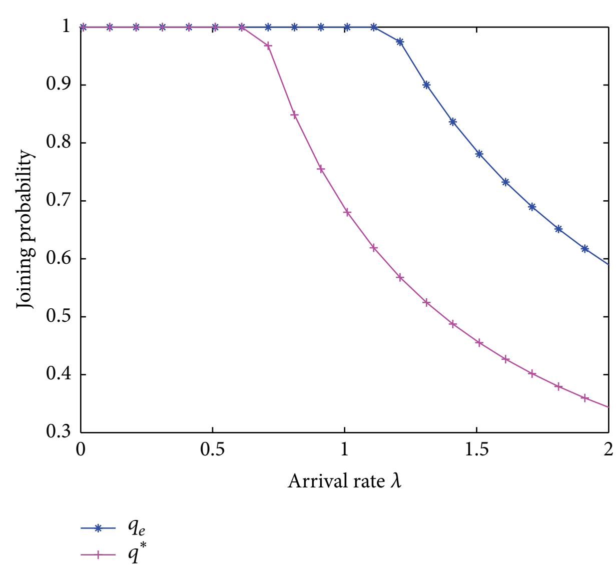

Figure 9: Equilibrium and socially optimal strategies for the unobservable case when $R = 3$, $C = 2$, $\mu_1 = 2$, $\mu_0 = 0.5$, and $\theta = 0.01$.

6. Conclusion

In this paper, we analyzed the customer strategic behavior in the M/M/1 queueing system with working vacations and vacation interruptions where arriving customers have option to decide whether to join the system or balk. Three different cases with respect to the levels of information provided to arriving customers have been investigated extensively and the equilibrium strategies for each case were derived. And we found that the equilibrium joining probability of the busy period may be smaller than that in vacation time. We also compared the equilibrium strategy with the socially optimal strategy numerically, and we observed that customers make individual decisions maximizing their own profit and tend to overuse the system. For the social planner, a toll will be adopted to discourage customers from joining.

Furthermore, an interesting extension would be to incorporate the present model in a profit maximizing framework, where the owner or manager of the system imposes an entrance fee. The problem is to find the optimal fee that maximizes the owners total profit subject to the constraint that customers independently decide whether to enter or not and take into account the fee, in addition to the service reward and waiting cost. Another direction for future work is to concern the non-Markovian queues with various vacation policies.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is partially supported by the Natural Foundation of Hebei Province (nos. A2014203096 and C2014203212), China, and the Young Faculty Autonomous Research Project of Yanshan University (nos. 13LGB032, 13LGA017, and 13LGB013), China.

References

[1] P. Naor, "The regulation of queue size by levying tolls," *Econometrica*, vol. 37, no. 1, pp. 15–24, 1969.

[2] N. M. Edelson and D. K. Hildebrand, "Congestion tolls for Poisson queuing processes," *Econometrica*, vol. 43, no. 1, pp. 81–92, 1975.

[3] R. Hassin and M. Haviv, *Equilibrium Behavior in Queueing Systems: To Queue or Not to Queue*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2003.

[4] A. Burnetas and A. Economou, "Equilibrium customer strategies in a single server Markovian queue with setup times," *Queueing Systems*, vol. 56, no. 3-4, pp. 213–228, 2007.

[5] W. Sun and N. Tian, "Contrast of the equilibrium and socially optimal strategies in a queue with vacations," *Journal of Computational Information Systems*, vol. 4, no. 5, pp. 2167–2172, 2008.

[6] 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.

[7] P. Guo and R. Hassin, "Strategic behavior and social optimization in Markovian vacation queues," *Operations Research*, vol. 59, no. 4, pp. 986–997, 2011.

[8] P. Guo and R. Hassin, "Strategic behavior and social optimization in Markovian vacation queues: the case of heterogeneous customers," *European Journal of Operational Research*, vol. 222, no. 2, pp. 278–286, 2012.

[9] R. Tian, D. Yue, and W. Yue, "Optimal balking strategies in an M/G/1 queueing system with a removable server under N-policy," *Journal of Industrial and Management Optimization*, vol. 11, no. 3, pp. 715–731, 2015.

[10] P. Guo and Q. Li, "Strategic behavior and social optimization in partially-observable Markovian vacation queues," *Operations Research Letters*, vol. 41, no. 3, pp. 277–284, 2013.

[11] 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.

[12] 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.

[13] N. Tian, J. Li, and Z. G. Zhang, "Matrix analysis method and working vacation queuesa survey," *International Journal of Information and Management Sciences*, vol. 20, pp. 603–633, 2009.

[14] F. Zhang, J. Wang, and B. Liu, "Equilibrium balking strategies in Markovian queues with working vacations," *Applied Mathematical Modelling*, vol. 37, no. 16-17, pp. 8264–8282, 2013.

[15] W. Sun and S. Li, "Equilibrium and optimal behavior of customers in Markovian queues with multiple working vacations," *TOP*, vol. 22, no. 2, pp. 694–715, 2014.

[16] W. Sun, S. Li, and Q.-L. Li, "Equilibrium balking strategies of customers in Markovian queues with two-stage working vacations," *Applied Mathematics and Computation*, vol. 248, pp. 195–214, 2014.

[17] J. Li and N. Tian, "The M/M/1 queue with working vacations and vacation interruptions," *Journal of Systems Science and Systems Engineering*, vol. 16, no. 1, pp. 121–127, 2007.

[18] J.-H. Li, N.-S. Tian, and Z.-Y. Ma, "Performance analysis of GI/M/1 queue with working vacations and vacation interruption," *Applied Mathematical Modelling*, vol. 32, no. 12, pp. 2715–2730, 2008.

[19] M. Neuts, *Matrix-Geometric Solution in Stochastic Models*, Johns Hopkins University Press, Baltimore, Md, USA, 1981.