Applied Mathematical Modelling 36 (2012) 6186–6202

Contents lists available at SciVerse ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

Equilibrium threshold strategies in observable queueing systems under single vacation policy

Weiqi Liu [a], Yan Ma [b], Jihong Li [a,*]

[a] College of Management, Shanxi University, Taiyuan 030006, PR China
[b] College of Mathematical Sciences, Shanxi University, Taiyuan 030006, PR China

ARTICLE INFO

Article history:
Received 18 January 2011
Received in revised form 25 January 2012
Accepted 3 February 2012
Available online 14 February 2012

Keywords:
Single vacation
Observable
Balking
Equilibrium threshold
Stationary distribution
Social benefit

ABSTRACT

This paper studies the equilibrium behavior of customers in continuous/discrete time queueing systems under single vacation policy. In the single vacation queueing system, the server can only take exactly one vacation when no customers exist in the system. This scheme is more practical under many specific circumstances. Based on the reward–cost structure, equilibrium behavior is considered in the fully observable and almost observable cases. The threshold strategies in equilibrium are obtained and the stationary system behavior is analyzed under the corresponding strategies. Finally, we illustrate the effect of the information level as well as several parameters on the equilibrium thresholds and social benefits via numerical experiments. The research results could instruct the customers to take optimal strategies and provide the managers with reference information on the pricing problem in the queueing system.

1. Introduction

For several decades, queueing systems with regard to the equilibrium behavior of customers have aroused increasing attention due to their applications for management in service system and electronic commerce. It was first introduced by Naor [1], who studied equilibrium and social optimal strategies for the joining-balking dilemma in an $M/M/1$ queue with a simple linear reward–cost structure. Afterwards, Naor's model and results were further refined and extended by several researchers such as Yechiali [2], Johansen and Stidham [3], Stidham [4] and Mendelson and Whang [5]. Chen and Frank [6] generalized Naor's model assuming that both the customers and the server maximize their expected discounted utility using a common discount rate. Larsen [7] considered another generalization of Naor's model assuming the customers differ by their service values. Erlichman and Hassin [8] discussed a single server Markovian queue allowing customers to overtake others. Various observable models can be found in the monographs of Hassin and Haviv [9].

As for the research on the equilibrium customer behavior in queues with vacations, the first was presented by Burnetas and Economou [10] who explored a single server Markovian queue with exponential setup times. Subsequently, Economou and Kanta [11] considered the Markovian queue that alternates between on and off periods in observable case. Sun et al. [12] presented the equilibrium customer behavior in an observable $M/M/1$ queue under interruptible and insusceptible setup/closedown policies. Recently, Economou et al. [13] analyzed the optimal balking strategies in single-server queues with general service and vacation times. However, there is no work concerning the equilibrium customer behavior in the single vacation queueing systems.

* Corresponding author. Tel.: +86 13485330831.
E-mail addresses: liuwq@sxu.edu.cn (W. Liu), myhicharge@126.com (Y. Ma), lijh1982@sxu.edu.cn (J. Li).

0307-904X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.apm.2012.02.003

A considerable amount of work has been done on the vacation queues in the past, due to the extensively use in communication systems, computer networks, etc. Levy and Yechiali [14,15] were first to study the $M/G/1$ and $M/M/c$ queueing models with single vacation. Later, several excellent surveys were given by Doshi [16] and Takagi [17,18]. Choudhury [19] and Sikdar and Gupta [20] dealt with an $M/G/1$ batch arrival queue and an $M/G/1$ batch service queue respectively. Madan and Al-Rawwash [21] did research on the $M^x/G/1$ queue with feedback and optional server vacations under single vacation policy. Gupta and Sikdar [22] assumed the input process is a Markovian arrival process instead of Poisson input. A comprehensive study in vacation queueing models can be found in Takagi [23] and Tian and Zhang [24].

So far, there exists no literature studying the single vacation queueing system from an economic viewpoint. It motivates us to explore the economics of single vacation queues which is more practical under many specific circumstances. Customers are in a dilemma whether to join the system or to balk. They make decisions based on a nature reward–cost structure, which incorporates their desire for service as well as their unwillingness to wait. In the observable system, customers are informed of the queue length upon arrival. As to the fully observable case, the arriving customer not only knows the number of present customers but also the state of the server. However, in the almost observable case a customer only observes the queue length before making decision.

The reminder of this paper is organized as follows. Section 2 presents the description of the model. Section 3 develops the fully and almost observable cases in continuous-time queueing system. We derive the equilibrium threshold strategies and analyze the corresponding stationary system behavior. Then in Section 4, we consider the equilibrium strategies and various performance measures in the discrete-time queueing system. The numerical discussions of thresholds and social benefits are postponed to Section 5. Finally, the concludes come in Section 6.

2. Model description

We consider a single server queueing system with infinite capacity. The server takes exactly one vacation immediately at the end of each busy period. If it finds no customers in the system upon returning from the vacation, it becomes idle until a customer arrives. When a customer arrives, it immediately starts to serve it. We assume that inter-arrival times, service times and vacation times are mutually independent. In addition, the queueing system follows First-Come-First-Served (FCFS) service discipline. Suppose Se be the mean sojourn time of a customer in equilibrium and Be be the expected net benefit.

Our interest is in the behavior of customers when they decide whether to join or to balk upon their arrival. To model the decision process we assume that every customer receives a reward of R units for completing service. This may reflect his satisfaction and the added value of being served. On the other hand, there exists a waiting cost of C units per time unit that the customer remains in the system (in queue or in service). Customers are risk neutral and maximize their expected net benefit. From now on, we assume the condition

$$R > \frac{C}{\mu} + \frac{C}{\theta}, \tag{1}$$

where μ is the service rate and θ is the vacation rate. This condition ensures that the reward for service exceeds the expected waiting cost for a customer who finds the system empty. Otherwise, after the system becomes empty for the first time no customers will ever enter. Finally, the decisions are irrevocable: retrials of balking customers and reneging of entering customers are not allowed.

3. Analysis of continuous queue

We first consider the continuous-time queueing system. Suppose that customers arrive according to a Poisson process with rate λ. Service times of the customers and vacation times are assumed to be exponentially distributed random variables. We denote the service rate and vacation rate by μ_c and θ_c respectively.

We represent the state at time t by the pair $(N(t),I(t))$, where $N(t)$ denotes the number of the customers in the system and $I(t)$ denotes the state of the server. Define the system is in a regular busy period as state 1 and in a vacation period as state 0. Thus, state $(0,1)$ indicates that the system is in the idle period; state $(n,1)$, $n = 1,2,\ldots$, indicates that the system is in the regular busy period and there are n customers; state $(n,0)$, $n = 0,1,2,\ldots$, indicates that the system is in the vacation period and there are n customers. Hereinafter, the subscript c means the continuous-time queue, fo means fully observable and ao means almost observable.

3.1. Fully observable case

We begin with the fully observable case in which the arriving customers not only know the number of present customers $N(t)$ but also the state of the server $I(t)$ at arrival time t. It is obvious that the process $\{(N(t),I(t))|t \geqslant 0\}$ is a continuous-time Markov chain with state space $\Omega_{cfo} = \{(n,i)|n \geqslant 0, i = 0,1\}$. In equilibrium, a customer who joins the system when he observes state (n,i) has the mean sojourn time $Se = \frac{n+1}{\mu_c} + \frac{1-i}{\theta_c}$, which is composed of two parts: the mean service time and the mean vacation time.

Thus his expected net benefit is

$$Be = R - \frac{C(n+1)}{\mu_c} - \frac{C(1-i)}{\theta_c}.$$

The customer strictly prefers to enter if *Be* is positive and is indifferent between entering and balking if it equals zero. We thus conclude the following.

Theorem 3.1. *In the fully observable M/M/1 queue with single vacation, there exist thresholds*

$$(n_e(0), n_e(1)) = \left(\left\lfloor \frac{R\mu_c}{C} - \frac{\mu_c}{\theta_c} \right\rfloor - 1, \ \left\lfloor \frac{R\mu_c}{C} \right\rfloor - 1 \right), \tag{2}$$

such that the strategy 'observe (N(t),I(t)), enter if $N(t) \leqslant n_e(I(t))$ and balk otherwise' is a unique equilibrium in the class of the threshold strategies.

Remark 1. $n_e(0)$ is the threshold when an arriving customer finds the system is in a vacation period and $n_e(1)$ is the threshold when it is in a regular busy period. We get $n_e(0)$ and $n_e(1)$ from the condition $Be > 0$ when $i = 0$ and 1 respectively. The symbol $\lfloor\rfloor$ indicates rounding down.

For the stationary analysis, we note that if all customers follow the threshold strategy in (2), the system follows a Markov chain with state space restricted to $\Omega_{cfo} = \{(n,i)|0 \leqslant n \leqslant n_e(i) + 1, i = 0,1\}$ and identical transition rates. The transition rate diagram is depicted in Fig. 1.

The corresponding stationary distribution $\{p(n,i)|(n,i) \in \Omega_{cfo}\}$ is obtained as the unique positive normalized solution of the following system of balance equations.

$$p(0,0)(\lambda + \theta_c) = p(1,1)\mu_c, \tag{3}$$

$$p(n,0)(\lambda + \theta_c) = p(n-1,0)\lambda, \quad n = 1,2,\ldots,n_e(0), \tag{4}$$

$$p(n_e(0)+1,0)\theta_c = p(n_e(0),0)\lambda, \tag{5}$$

$$p(0,1)\lambda = p(0,0)\theta_c, \tag{6}$$

$$p(n,1)(\lambda + \mu_c) = p(n-1,1)\lambda + p(n,0)\theta_c + p(n+1,1)\mu_c, \quad n = 1,2,\cdots,n_e(0)+1, \tag{7}$$

$$p(n,1)(\lambda + \mu_c) = p(n-1,1)\lambda + p(n+1,1)\mu_c, \quad n = n_e(0)+2,\cdots,n_e(1), \tag{8}$$

$$p(n_e(1)+1,1)\mu_c = p(n_e(1),1)\lambda. \tag{9}$$

Define $\rho = \frac{\lambda}{\mu_c}, \ \sigma = \frac{\lambda}{\lambda + \theta_c}$.

By iterating (4) and (8), taking into account (3), (5) and (9), we obtain

$$p(n,0) = \frac{\mu_c}{\lambda + \theta_c} \sigma^n p(1,1), \quad n = 1,2,\cdots,n_e(0), \tag{10}$$

$$p(n_e(0)+1,0) = \frac{\lambda}{\theta_c} \frac{\mu_c}{\lambda + \theta_c} \sigma^{n_e(0)} p(1,1), \tag{11}$$

$$p(n,1) = \rho^{n-n_e(0)-1} p(n_e(0)+1,1), \quad n = n_e(0)+2,\cdots,n_e(1)+1. \tag{12}$$

From (7) we observe that $\{p(n,1)|n = 1,2,\cdots,n_e(0)+1\}$ is a solution of the nonhomogeneous linear difference equation with constant coefficients

$$\mu_c x_{n+1} - (\lambda + \mu_c)x_n + \lambda x_{n-1} = -\theta_c p(n,0) = -\frac{\theta_c \mu_c}{\lambda + \theta_c} \sigma^n p(1,1), \quad n = 2,3,\cdots,n_e(0), \tag{13}$$

where the last equation is due to (10). Using the standard approach for solving such equations, we consider the corresponding characteristic equation

$$\mu_c x^2 - (\lambda + \mu_c)x + \lambda = 0,$$

which has two roots 1 and ρ. Then the general solution of the homogeneous version of (13) is $x_n^{hom} = A1^n + B\rho^n$ (we assume $\rho \neq 1$). The general solution x_n^{gen} of (13) is given as $x_n^{gen} = x_n^{hom} + x_n^{spec}$, where x_n^{spec} is a specific solution of (13). Because the

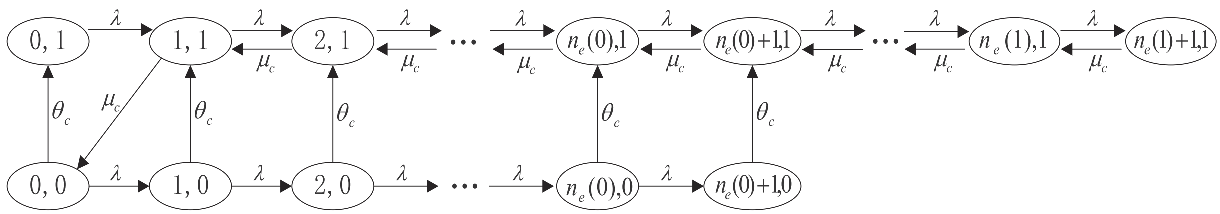

Fig. 1. Transition rate diagram for the $(n_e(0),n_e(1))$ threshold strategy in the fully observable queue under continuous case.

nonhomogeneous part of (13) is geometric with parameter σ, we can find a specific solution $C\sigma^n$ (we assume $\sigma \neq \rho$ and both are different from 1). Substituting $x_n^{spec} = C\sigma^n$ into (13) we obtain

$$C = \frac{\mu_c}{\mu_c - \lambda - \theta_c} p(1, 1). \tag{14}$$

Hence the general solution of (13) is given as

$$x_n^{gen} = A1^n + B\rho^n + C\sigma^n, \quad n = 1, 2, \cdots, n_e(0) + 1, \tag{15}$$

where C is given by (14) and A, B are to be determined.

From (15), for $n = 1$ it follows that

$$A + B\rho = \frac{\mu_c\theta_c - (\lambda + \theta_c)^2}{(\lambda + \theta_c)(\mu_c - \lambda - \theta_c)} p(1, 1). \tag{16}$$

Furthermore, substituting (15) into (7) for $n = 2$ and taking into account (3), (4) and (6), it follows after some rather tedious algebra that

$$A + B\rho^2 + \frac{\mu_c}{\mu_c - \lambda - \theta_c}\left(\frac{\lambda}{\lambda + \theta_c}\right)^2 p(1, 1) = \frac{1}{\mu_c}[(\lambda + \mu_c)p(1, 1) - \lambda p(0, 1) - \theta_c p(1, 0)],$$

which is equivalent to the form:

$$A + B\rho^2 = \frac{\lambda}{\mu_c} \frac{\mu_c\theta_c - (\lambda + \theta_c)^2}{(\lambda + \theta_c)(\mu_c - \lambda - \theta_c)} p(1, 1). \tag{17}$$

Solving (16) and (17), we obtain

$$\begin{cases} A = 0, \\ B = \dfrac{\mu_c^2\theta_c - \mu_c(\lambda + \theta_c)^2}{\lambda(\lambda + \theta_c)(\mu_c - \lambda - \theta_c)} p(1, 1). \end{cases}$$

Then, from (15),

$$p(n, 1) = \left[\frac{\mu_c\theta_c - (\lambda + \theta_c)^2}{(\lambda + \theta_c)(\mu_c - \lambda - \theta_c)}\rho^{n-1} + \frac{\mu_c}{\mu_c - \lambda - \theta_c}\sigma^n\right] p(1, 1), \quad n = 1, 2, \cdots, n_e(0) + 1. \tag{18}$$

We have thus expressed all stationary probabilities in terms of $p(1,1)$ in relations see (3), (6), (10)–(12) and (18). The remaining probability $p(1,1)$ can be found from the normalization equation

$$\sum_{n=0}^{n_e(0)+1} p(n, 0) + \sum_{n=0}^{n_e(1)+1} p(n, 1) = 1.$$

After some algebraic simplifications, we can express all stationary probabilities in terms of ρ and σ in the following theorem:

Theorem 3.2. *Consider a fully observable M/M/1 queue with single vacation and $\sigma \neq 1 \neq \rho \neq \sigma$, in which customers follow the threshold policy $(n_e(0),\ n_e(1))$ given in Theorem 3.1. The stationary probabilities $\{p_{fo}(n, i) | (n, i) \in \Omega_{cfo}\}$ are as follows:*

$$p_{fo}(1, 1) = \left[\frac{1}{\rho} + \frac{\sigma^2}{\rho(1 - \sigma)} + \frac{1 - \rho^{n_e(1)+1}}{1 - \rho} + \frac{\sigma^2(1 - \sigma^{n_e(0)+1})}{(\sigma - \rho)(1 - \sigma)} + \frac{\sigma^{n_e(0)+2}(\rho - \rho^{n_e(1)-n_e(0)+1}) - \sigma^2(1 - \rho^{n_e(1)+1})}{(\sigma - \rho)(1 - \rho)}\right]^{-1}, \tag{19}$$

$$p_{fo}(n, 0) = \frac{\sigma^{n+1}}{\rho} p_{fo}(1, 1), \quad n = 0, 1, 2, \cdots, n_e(0), \tag{20}$$

$$p_{fo}(n_e(0) + 1, 0) = \frac{1}{(1 - \sigma)\rho}\sigma^{n_e(0)+2} p_{fo}(1, 1), \tag{21}$$

$$p_{fo}(n, 1) = \frac{1}{\sigma - \rho}[(\sigma - \sigma^2 - \rho)\rho^{n-1} + \sigma^{n+1}] p_{fo}(1, 1), \quad n = 0, 1, 2, \cdots, n_e(0) + 1, \tag{22}$$

$$p_{fo}(n, 1) = \frac{1}{\sigma - \rho}\left[\sigma - \sigma^2 - \rho + \sigma\rho\left(\frac{\sigma}{\rho}\right)^{n_e(0)+1}\right]\rho^{n-1} p_{fo}(1, 1), \quad n = n_e(0) + 2, \cdots, n_e(1) + 1. \tag{23}$$

Because of the PASTA property, the probability of balking is equal to $p_{fo}(n_e(0) + 1, 0) + p_{fo}(n_e(1) + 1, 1)$. In addition, every customer receives a reward of R units for completing service and there exists a waiting cost of C units per time unit when the customer remains in the system (in queue or in service). Hence the social benefit per time unit when all customers follow the threshold policy $(n_e(0), n_e(1))$ given in Theorem 3.1 equals

$$SB_{cfo} = R\lambda(1 - p_{fo}(n_e(0)+1,0) - p_{fo}(n_e(1)+1,1)) - C\left(\sum_{n=0}^{n_e(0)+1} np_{fo}(n,0) + \sum_{n=0}^{n_e(1)+1} np_{fo}(n,1)\right).$$

3.2. Almost observable case

Next we consider the almost observable case where arriving customers only know $N(t)$ before making decisions. Hence the stationary distribution of the corresponding Markov chain is from Theorem 3.2 with $n_e(0)=n_e(1)=n_e$ and state space $\Omega_{cao} = \{n|0 \leqslant n \leqslant n_e+1\}$ as well as identical transition rates. The transition diagram is depicted in Fig. 2.

Theorem 3.3. *Consider an almost observable M/M/1 queue with single vacation and $\sigma \neq 1 \neq \rho \neq \sigma$, in which customers follow threshold policy n_e. The stationary probabilities $\{p_{ao}(n)|n \in \Omega_{cao}\}$ are as follows:*

$$p_{ao}(n) = K\left[\frac{\sigma^{n+1}}{\rho} + \frac{(\sigma-\sigma^2-\rho)\rho^{n-1}+\sigma^{n+1}}{\sigma-\rho}\right], \quad n = 0,1,2,\cdots,n_e,$$
$$p_{ao}(n_e+1) = K\left[\frac{\sigma^{n_e+2}}{(1-\sigma)\rho} + \frac{(\sigma-\sigma^2-\rho)\rho^{n_e}+\sigma^{n_e+2}}{\sigma-\rho}\right],$$

where

$$p_{ao}(n) = p_{fo}(n,0) + p_{fo}(n,1), \quad n = 0,1,2,\cdots,n_e+1,$$

and

$$K = \left[\frac{1}{\rho} + \frac{\sigma^2}{\rho(1-\sigma)} + \frac{\sigma^2(1-\sigma^{n_e+1})}{(\sigma-\rho)(1-\sigma)} + \frac{(\sigma-\sigma^2-\rho)(1-\rho^{n_e+1})}{(\sigma-\rho)(1-\rho)}\right]^{-1}.$$

The mean sojourn time of a customer who finds n customers in the system is

$$Se = \frac{n+1}{\mu_c} + \frac{\pi^-_{I|N}(0|n)}{\theta_c},$$

where $\pi^-_{I|N}(0|n)$ is the probability that an arriving customer finds the server at state 0, given that there are n customers. Therefore, if such a customer decides to enter, the expected net benefit is

$$Be = R - \frac{C(n+1)}{\mu_c} - \frac{C\pi^-_{I|N}(0|n)}{\theta_c}. \tag{24}$$

We get $\pi^-_{I|N}(0|n) = \frac{\lambda p_{fo}(n,0)}{\lambda p_{fo}(n,0)+\lambda p_{fo}(n,1)}$, $n = 0,1,2,\cdots,n_e+1$. Using the various forms of $p_{fo}(n,i)$ from (19)–(23) we obtain

$$\pi^-_{I|N}(0|n) = \left[1 + \frac{\rho}{\sigma-\rho} + \left(\frac{1-\sigma}{\sigma} - \frac{\rho}{\sigma-\rho}\right)\left(\frac{\rho}{\sigma}\right)^n\right]^{-1}, \quad n = 0,1,2,\cdots,n_e, \tag{25}$$
$$\pi^-_{I|N}(0|n_e+1) = \left[1 + (1-\sigma)\left(\frac{\rho}{\sigma-\rho} + \left(\frac{1-\sigma}{\sigma} - \frac{\rho}{\sigma-\rho}\right)\left(\frac{\rho}{\sigma}\right)^{n_e+1}\right)\right]^{-1}. \tag{26}$$

In light of (24)–(26), we introduce the function

$$f(n,x) = R - \frac{C(n+1)}{\mu_c} - \frac{C}{\theta_c}\left[1 + x\left(\frac{\rho}{\sigma-\rho} + \left(\frac{1-\sigma}{\sigma} - \frac{\rho}{\sigma-\rho}\right)\left(\frac{\rho}{\sigma}\right)^n\right)\right]^{-1}, \quad x \in [1-\sigma,1],\ n = 0,1,2,\cdots, \tag{27}$$

which will allow us to prove the existence of equilibrium threshold strategies and derive the corresponding thresholds. Let

$$f_U(n) = f(n,1) = R - \frac{C(n+1)}{\mu_c} - \frac{C}{\theta_c}\left[1 + \frac{\rho}{\sigma-\rho} + \left(\frac{1-\sigma}{\sigma} - \frac{\rho}{\sigma-\rho}\right)\left(\frac{\rho}{\sigma}\right)^n\right]^{-1}, \quad n = 0,1,2,\cdots, \tag{28}$$
$$f_L(n) = f(n,1-\sigma) = R - \frac{C(n+1)}{\mu_c} - \frac{C}{\theta_c}\left[1 + (1-\sigma)\left(\frac{\rho}{\sigma-\rho} + \left(\frac{1-\sigma}{\sigma} - \frac{\rho}{\sigma-\rho}\right)\left(\frac{\rho}{\sigma}\right)^n\right)\right]^{-1}, \quad n = 0,1,2,\cdots, \tag{29}$$

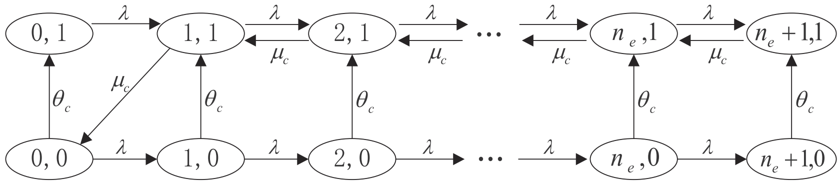

Fig. 2. Transition rate diagram for the n_e threshold strategy in the almost observable queue under continuous case.

It is easy to see that

$$f_U(0) = f(0,1) = R - \frac{C}{\mu_c} - \frac{C}{\theta_c}\sigma > 0,$$

and

$$f_L(0) = f(0,1-\sigma) = R - \frac{C}{\mu_c} - \frac{C}{\theta_c}\left[1 + \frac{(1-\sigma)^2}{\sigma}\right]^{-1} > 0.$$

In addition,

$$\lim_{n\to\infty} f_U(n) = \lim_{n\to\infty} f_L(n) = -\infty.$$

Hence there exists n_U such that

$$f_U(0), f_U(1), f_U(2), \cdots, f_U(n_U) > 0 \text{ and } f_U(n_U+1) \leqslant 0. \tag{30}$$

The function $f(n,x)$ is clearly increasing with respect to x for every fixed n, thus we get the relation $f_L(n) \leqslant f_U(n)$, $n = 0,1,2,\cdots$. In particular, $f_L(n_U+1) \leqslant 0$ while $f_L(0) > 0$. Hence, there exists $n_L \leqslant n_U$ such that

$$f_L(n_L) > 0 \text{ and } f_L(n_L+1), \cdots, f_L(n_U), f_L(n_U+1) \leqslant 0. \tag{31}$$

We can now establish the existence of the equilibrium threshold policies in the almost observable case and give the following theorem.

Theorem 3.4. *In the almost observable M/M/1 queue with single vacation, all pure threshold strategies 'observe $N(t)$, enter if $N(t) \leqslant n_e$ and balk otherwise' for n_e=n_L,n_L + 1, $\cdots$,n_U are equilibrium strategies.*

Proof. Consider a tagged customer at his arrival instant and assume all other customers follow the same threshold strategy 'observe $N(t)$, enter if $N(t) \leqslant n_e$ and balk otherwise' for some fixed $n_e \in \{n_L, n_L+1, \cdots, n_U\}$. Then $\pi^-_{I|N}(0|n)$ is given by (25), (26).

If the tagged customer finds $n \leqslant n_e$ customers and decides to enter, his expected net benefit is equal to

$$R - \frac{C(n+1)}{\mu_c} - \frac{C}{\theta_c}\left[1 + \frac{\rho}{\sigma-\rho} + \left(\frac{1-\sigma}{\sigma} - \frac{\rho}{\sigma-\rho}\right)\left(\frac{\rho}{\sigma}\right)^n\right]^{-1} = f_U(n) > 0,$$

because of (24), (25), (27), (28) and (30). So in this case the customer prefers to enter.

If the tagged customer finds $n = n_e + 1$ customers and decides to enter, his expected net benefit is

$$R - \frac{C(n_e+2)}{\mu_c} - \frac{C}{\theta_c}\left[1 + (1-\sigma)\left(\frac{\rho}{\sigma-\rho} + \left(\frac{1-\sigma}{\sigma} - \frac{\rho}{\sigma-\rho}\right)\left(\frac{\rho}{\sigma}\right)^{n_e+1}\right)\right]^{-1} = f_L(n_e+1) \leqslant 0,$$

because of (24), (26), (27), (29) and (31). So in this case the customer prefers to balk. □

Because the probability of balking is equal to $p_{ao}(n_e+1)$, the social benefit per time unit when all customers follow the threshold policy n_e given in Theorem 3.4 equals

$$SB_{cao} = R\lambda(1 - p_{ao}(n_e+1)) - C\left(\sum_{n=0}^{n_e+1} np_{a_o}(n)\right).$$

4. Analysis of discrete queue

Now we discuss the equilibrium customer behavior in discrete-time queue. Assume that customer arrivals occur at the end of slot $t = n^-$, $n = 0,1,\cdots$. Inter-arrival times are independent and identical distributed sequences following a geometric distribution with rate p. The beginning and ending of service occur at slot division point $t = n$, $n = 0,1,\cdots$. The service times S and vacation times V are geometrically distributed with rates μ_d and θ_d respectively. Thus,

$$P(T=k) = p\bar{p}^{k-1}, \quad k \geqslant 1,\ 0<p<1,\ \bar{p} = 1-p,$$
$$P(S=k) = \mu_d\overline{\mu_d}^{k-1}, \quad k \geqslant 1,\ 0<\mu_d<1,\ \overline{\mu_d} = 1-\mu_d,$$
$$P(V=k) = \theta_d\overline{\theta_d}^{k-1}, \quad k \geqslant 1,\ 0<\theta_d<1,\ \overline{\theta_d} = 1-\theta_d.$$

Let L_n be the number of customers in the system at time n^+. According to the assumptions above, a customer who finishes service and leaves at $t = n^+$ does not reckon on L_n while arrives at $t = n^-$ should reckon on L_n. We assume

$$J_n = \begin{cases} 0, & \text{the system is in a vacation period at time } n^+, \\ 1, & \text{the system is in a service period at time } n^+. \end{cases}$$

Parallel to the analysis of continuous queue, we show that there also exist equilibrium threshold strategies for customers. However, the stationary behaviors of the system are more difficult to acquire than those in the continuous queue. Hereinafter, the subscript d means the discrete-time queue.

4.1. Fully observable case

The arriving customers know both the queue length L_n and the state of the server J_n at arrival instant. Evidently, $\{(L_n,J_n)|n \geqslant 0\}$ is a discrete-time Markov chain with state space $\Omega_{dfo} = \{(k,j)|k \geqslant 0, j = 0,1\}$. The mean sojourn time and expected net benefit when a customer observes state (k,j) are

$$Se = \frac{k+1}{\mu_d} + \frac{1-j}{\theta_d},$$
$$Be = R - \frac{C(k+1)}{\mu_d} - \frac{C(1-j)}{\theta_d}.$$

The customer strictly prefers to enter if Be is positive and is indifferent between entering and balking if it equals zero, which is identical to the continuous case. We thus conclude the following theorem.

Theorem 4.1. *In the fully observable Geo/Geo/1 queue with single vacation, there exist thresholds*

$$(L_e(0), L_e(1)) = \left(\left\lfloor \frac{R\mu_d}{C} - \frac{\mu_d}{\theta_d} \right\rfloor - 1, \ \left\lfloor \frac{R\mu_d}{C} \right\rfloor - 1\right), \tag{32}$$

such that the strategy 'observe (L_n,J_n), enter if $L_n \leqslant L_e(J_n)$ and balk otherwise' is a unique equilibrium in the class of the threshold strategies.

For the stationary analysis in discrete time, note that if all customers follow the threshold strategy in (32), the system follows a Markov chain with state space restricted to $\Omega_{dfo} = \{(k,0)|0 \leqslant k \leqslant L_e(0)+1\} \cup \{(k,1)|0 \leqslant k \leqslant L_e(1)+1\}$. We show the transition diagram in Fig. 3. The one-step transition probabilities of (L_n,J_n) are as follows:

Case 1: if $X_n = (k,0)$, $0 \leqslant k \leqslant L_e(0)$,

$$X_{n+1} = \begin{cases} (k,0), & \text{with probability } \overline{\theta_d}\bar{p} \\ (k,1), & \text{with probability } \theta_d\bar{p} \\ (k+1,0), & \text{with probability } \overline{\theta_d}p \\ (k+1,1), & \text{with probability } \theta_d p \end{cases}$$

Case 2: if $X_n = (L_e(0)+1, 0)$,

$$X_{n+1} = \begin{cases} (L_e(0)+1,0), & \text{with probability } \overline{\theta_d} \\ (L_e(0)+1,1), & \text{with probability } \theta_d \end{cases}$$

Case 3: if $X_n = (0,1)$,

$$X_{n+1} = \begin{cases} (0,1), & \text{with probability } \bar{p} \\ (1,1), & \text{with probability } p \end{cases}$$

Case 4: if $X_n = (1,1)$,

$$X_{n+1} = \begin{cases} (0,0), & \text{with probability } \bar{p}\mu_d \\ (1,1), & \text{with probability } 1 - p\overline{\mu_d} - \bar{p}\mu_d \\ (2,1), & \text{with probability } p\overline{\mu_d} \end{cases}$$

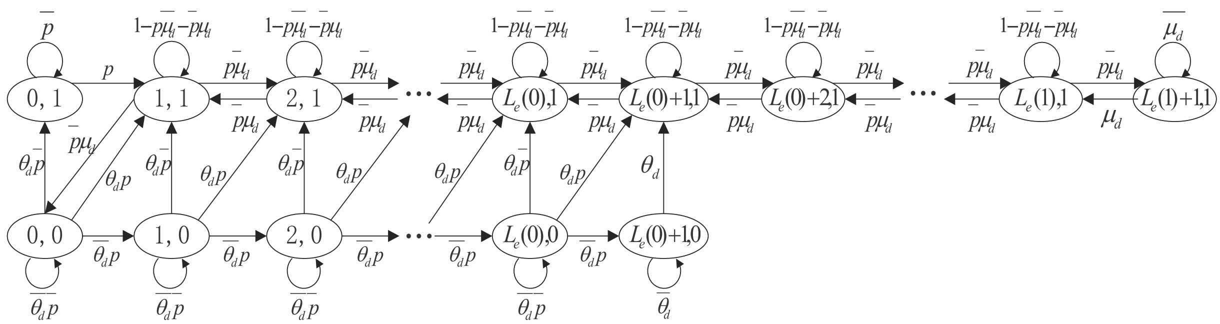

Fig. 3. Transition rate diagram for the $(L_e(0),L_e(1))$ threshold strategy in the fully observable queue under discrete case.

Case 5: if $X_n = (k,1)$, $2 \leqslant k \leqslant L_e(1)$

$$X_{n+1} = \begin{cases} (k-1,\,1), & \text{with probability } \bar{p}\mu_d \\ (k,\,1), & \text{with probability } 1-p\overline{\mu_d}-\bar{p}\mu_d \\ (k+1,\,1), & \text{with probability } p\overline{\mu_d} \end{cases}$$

Case 6: if $X_n = (L_e(1)+1, 1)$,

$$X_{n+1} = \begin{cases} (L_e(1),\,1), & \text{with probability } \mu_d \\ (L_e(1)+1,\,1), & \text{with probability } \overline{\mu_d}. \end{cases}$$

Based on the one-step transition situation analysis, using the lexicographical sequence for the states, the one-step transition probability matrix of (L_n,J_n) can be written as

$$\widetilde{\boldsymbol{P}} = \begin{bmatrix} \boldsymbol{B}_0 & \boldsymbol{A}_0 & & & & & & & & \\ \boldsymbol{B}_1 & \boldsymbol{A}_1 & \boldsymbol{C}_1 & & & & & & & \\ & \boldsymbol{B}_2 & \boldsymbol{A}_1 & \boldsymbol{C}_1 & & & & & & \\ & & \ddots & \ddots & \ddots & & & & & \\ & & & \boldsymbol{B}_2 & \boldsymbol{A}_1 & \boldsymbol{C}_1 & & & & \\ & & & & \boldsymbol{B}_2 & \boldsymbol{A}_2 & \boldsymbol{C}_2 & & & \\ & & & & & \boldsymbol{B}_3 & \boldsymbol{A}_3 & \boldsymbol{C}_3 & & \\ & & & & & & \ddots & \ddots & \ddots & \\ & & & & & & & \boldsymbol{B}_3 & \boldsymbol{A}_3 & \boldsymbol{C}_3 \\ & & & & & & & & \boldsymbol{B}_4 & \boldsymbol{A}_4 \end{bmatrix}, \tag{33}$$

where

$$\boldsymbol{B}_0 = \begin{bmatrix} \overline{\theta_d}\bar{p} & \theta_d\bar{p} \\ 0 & \bar{p} \end{bmatrix}, \qquad \boldsymbol{A}_0 = \begin{bmatrix} \overline{\theta_d}p & \theta_d p \\ 0 & p \end{bmatrix},$$

$$\boldsymbol{B}_1 = \begin{bmatrix} 0 & 0 \\ \bar{p}\mu_d & 0 \end{bmatrix}, \quad \boldsymbol{A}_1 = \begin{bmatrix} \overline{\theta_d}\bar{p} & \theta_d\bar{p} \\ 0 & 1-p\overline{\mu_d}-\bar{p}\mu_d \end{bmatrix}, \quad \boldsymbol{C}_1 = \begin{bmatrix} \overline{\theta_d}p & \theta_d p \\ 0 & p\overline{\mu_d} \end{bmatrix},$$

$$\boldsymbol{B}_2 = \begin{bmatrix} 0 & 0 \\ 0 & \bar{p}\mu_d \end{bmatrix}, \quad \boldsymbol{A}_2 = \begin{bmatrix} \overline{\theta_d} & \theta_d \\ 0 & 1-p\overline{\mu_d}-\bar{p}\mu_d \end{bmatrix}, \quad \boldsymbol{C}_2 = \begin{bmatrix} 0 \\ p\overline{\mu_d} \end{bmatrix}$$

and

$$\boldsymbol{B}_3 = \bar{p}\mu_d, \quad \boldsymbol{A}_3 = 1-p\overline{\mu_d}-\bar{p}\mu_d, \quad \boldsymbol{C}_3 = p\overline{\mu_d}, \quad \boldsymbol{B}_4 = \mu_d, \quad \boldsymbol{A}_4 = \overline{\mu_d}.$$

Let (L,J) be the stationary limit of (L_n,J_n) and its distribution is denoted as

$$\pi_{kj} = P\{L=k, J=j\}, \quad (k,j) \in \Omega_{dfo}$$

$$\pi_k = \begin{cases} (\pi_{k0},\,\pi_{k1}), & 0 \leqslant k \leqslant L_e(0)+1, \\ \pi_{k1}, & L_e(0)+2 \leqslant k \leqslant L_e(1)+1 \end{cases}$$

$$\boldsymbol{\pi} = (\pi_0,\,\pi_1,\,\cdots,\,\pi_{L_e(1)+1}).$$

We solve for the stationary distribution π_{kj} by noting that the vector $\boldsymbol{\pi}$ satisfies the equation $\boldsymbol{\pi}\widetilde{\boldsymbol{P}} = \boldsymbol{\pi}$ and have the following system of steady-state equations:

$$\pi_{00} = \overline{\theta_d}\bar{p}\pi_{00} + \bar{p}\mu_d\pi_{11}, \tag{34}$$
$$\pi_{10} = \overline{\theta_d}p\pi_{00} + \overline{\theta_d}\bar{p}\pi_{10}, \tag{35}$$
$$\pi_{k0} = \overline{\theta_d}p\pi_{k-1,0} + \overline{\theta_d}\bar{p}\pi_{k0}, \quad k = 2, 3, \cdots, L_e(0), \tag{36}$$
$$\pi_{L_e(0)+1,0} = \overline{\theta_d}p\pi_{L_e(0),0} + \overline{\theta_d}\pi_{L_e(0)+1,0}, \tag{37}$$
$$\pi_{01} = \theta_d\bar{p}\pi_{00} + \bar{p}\pi_{01}, \tag{38}$$
$$\pi_{11} = \theta_d p\pi_{00} + p\pi_{01} + \theta_d\bar{p}\pi_{10} + (1 - p\overline{\mu_d} - \bar{p}\mu_d)\pi_{11} + \bar{p}\mu_d\pi_{21}, \tag{39}$$
$$\pi_{k1} = \theta_d p\pi_{k-1,0} + p\overline{\mu_d}\pi_{k-1,1} + \theta_d\bar{p}\pi_{k0} + (1 - p\overline{\mu_d} - \bar{p}\mu_d)\pi_{k1} + \bar{p}\mu_d\pi_{k+1,1}, \quad k = 2, 3, \cdots, L_e(0), \tag{40}$$
$$\pi_{L_e(0)+1,1} = \theta_d p\pi_{L_e(0),0} + p\overline{\mu_d}\pi_{L_e(0),1} + \theta_d\pi_{L_e(0)+1,0} + (1 - p\overline{\mu_d} - \bar{p}\mu_d)\pi_{L_e(0)+1,1} + \bar{p}\mu_d\pi_{L_e(0)+2,1}, \tag{41}$$
$$\pi_{k1} = p\overline{\mu_d}\pi_{k-1,1} + (1 - p\overline{\mu_d} - \bar{p}\mu_d)\pi_{k1} + \bar{p}\mu_d\pi_{k+1,1}, \quad k = L_e(0) + 2, \cdots, L_e(1) - 1, \tag{42}$$
$$\pi_{L_e(1),1} = p\overline{\mu_d}\pi_{L_e(1)-1,1} + (1 - p\overline{\mu_d} - \bar{p}\mu_d)\pi_{L_e(1),1} + \mu_d\pi_{L_e(1)+1,1}, \tag{43}$$
$$\pi_{L_e(1)+1,1} = p\overline{\mu_d}\pi_{L_e(1),1} + \overline{\mu_d}\pi_{L_e(1)+1,1}. \tag{44}$$

Define

$$\alpha = \frac{p\overline{\mu_d}}{\bar{p}\mu_d}, \quad \beta = \frac{\overline{\theta_d}p}{1 - \overline{\theta_d}\bar{p}}.$$

By iterating (36) and (42), taking into account (35), (37), (43) and (44), we obtain

$$\pi_{k0} = \beta^k\pi_{00}, \quad k = 1, 2, \cdots, L_e(0) \tag{45}$$
$$\pi_{L_e(0)+1,0} = \frac{1 - \overline{\theta_d}\bar{p}}{\theta_d}\beta^{L_e(0)+1}\pi_{00}, \tag{46}$$
$$\pi_{k1} = \alpha^{k-L_e(0)-1}\pi_{L_e(0)+1,1}, \quad k = L_e(0) + 2, \cdots, L_e(1), \tag{47}$$
$$\pi_{L_e(1)+1,1} = \bar{p}\alpha^{L_e(1)-L_e(0)}\pi_{L_e(0)+1,1}. \tag{48}$$

From (40) we observe that $\{\pi_{k1} | k = 1, 2, \cdots, L_e(0) + 1\}$ is a solution of the nonhomogeneous linear difference equation with constant coefficients

$$\bar{p}\mu_d x_{k+1} - (p\overline{\mu_d} + \bar{p}\mu_d)x_k + p\overline{\mu_d}x_{k-1} = -\theta_d p\pi_{k-1,0} - \theta_d\bar{p}\pi_{k0} = -\frac{\theta_d}{\overline{\theta_d}}\beta^k\pi_{00}, \quad k = 2, 3, \cdots, L_e(0), \tag{49}$$

where the last equation is due to (45). We consider the corresponding characteristic equation

$$\bar{p}\mu_d x^2 - (p\overline{\mu_d} + \bar{p}\mu_d)x + p\overline{\mu_d} = 0,$$

which has two roots at 1 and α. Then the general solution of the homogeneous version of (49) is $x_k^{hom} = D1^k + E\alpha^k$ (we assume $\alpha \neq 1$). The general solution x_k^{gen} of (49) is given as $x_k^{gen} = x_k^{hom} + x_k^{spec}$, where x_k^{spec} is a specific solution of (49). Because the nonhomogeneous part of (49) is geometric with parameter β, we can find a specific solution $F\beta^k$ (we assume $\beta \neq \alpha$ and both are different from 1). Substituting $x_k^{gen} = F\beta^k$ into (49) we obtain

$$F = \frac{1 - \overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p} - \overline{\mu_d}}\pi_{00}. \tag{50}$$

Hence the general solution of (49) is given as

$$x_k^{gen} = D1^k + E\alpha^k + F\beta^k, \quad k = 1, 2, \cdots, L_e(0) + 1, \tag{51}$$

where F is given by (50) and D, E are to be determined.

From (51) for $k = 1$, we obtain

$$D + E\alpha + F\beta = \frac{1 - \overline{\theta_d}\bar{p}}{\bar{p}\mu_d}\pi_{00}. \tag{52}$$

Furthermore, substituting (51) into (39) for $k = 2$, it follows after some rather tedious algebra that

$$D + E\alpha^2 + F\beta^2 = \frac{(p\overline{\mu_d} + \bar{p}\mu_d)(1 - \overline{\theta_d}\bar{p})^2 - \bar{p}\mu_d\theta_d(1 - \overline{\theta_d}\bar{p}^2)}{(\bar{p}\mu_d)^2(1 - \overline{\theta_d}\bar{p})}\pi_{00}. \tag{53}$$

Solving (52) and (53), we obtain

$$\begin{cases} D = 0, \\ E = G\pi_{00}, \\ G = \dfrac{\overline{\mu_d}\theta_d\overline{\theta_d}^2\bar{p}^2(\overline{\mu_d} - \bar{p} - p\bar{p}\overline{\theta_d} + 2p) + \overline{\mu_d}\theta_d\overline{\theta_d}\bar{p}(p^2\mu_d + 2\mu_d - 3p) + \mu_d\theta_d^2\overline{\theta_d}\bar{p}^2(1 - p\mu_d - \overline{\theta_d}\bar{p}^2) + \overline{\mu_d}\theta_d(p - \mu_d)}{p\overline{\mu_d}\theta_d(\bar{p}\overline{\theta_d} - \overline{\mu_d})(\bar{p}\overline{\theta_d} - 1)(p - \mu_d)}. \end{cases}$$

Then, from (51),

$$\pi_{k1} = \left(G\alpha^k + \frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\beta^k\right)\pi_{00}, \quad k = 1,2,\cdots,L_e(0)+1. \tag{54}$$

We have thus expressed all stationary probabilities in terms of π_{00} in relations see (38), 45,46,47,48 and (54). The remaining probability π_{00} can be found from the normalization equation

$$\sum_{k=0}^{L_e(0)+1}\pi_{k0} + \sum_{k=0}^{L_e(1)+1}\pi_{k1} = 1.$$

After some algebraic simplification, we can express all stationary probabilities in the following theorem.

Theorem 4.2. *Consider a fully observable Geo/Geo/1 queue with single vacation and $\beta \neq 1 \neq \alpha \neq \beta$, in which customers follow the threshold policy $(L_e(0), L_e(1))$ given in Theorem 4.1. The stationary probabilities $\{\pi_{kj}|(k,j)\in\Omega_{dfo}\}$ are as follows:*

$$\pi_{00} = \left\{\frac{1-\overline{\theta_d}\bar{p}}{\theta_d} + \frac{\theta_d\bar{p}}{p} + G\left(\frac{p\overline{\mu_d}}{\mu_d - p} - \frac{p\bar{p}}{\mu_d - p}\alpha^{L_e(1)+1}\right) + \frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\left[\frac{\overline{\theta_d}p}{\theta_d} - \left(\frac{\overline{\theta_d}p}{\theta_d} - \frac{p\overline{\mu_d}}{\mu_d - p} + \frac{p\bar{p}}{\mu_d - p}\alpha^{L_e(1)-L_e(0)}\right)\beta^{L_e(0)+1}\right]\right\}^{-1}, \tag{55}$$

$$\pi_{k0} = \beta^k\pi_{00}, \quad k = 1,2,\cdots,L_e(0), \tag{56}$$

$$\pi_{L_e(0)+1,0} = \frac{\beta^{L_e(0)+1}}{1-\beta}\pi_{00}, \tag{57}$$

$$\pi_{01} = \frac{\theta_d\bar{p}}{p}\pi_{00}, \tag{58}$$

$$\pi_{k1} = \left(G\alpha^k + \frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\beta^k\right)\pi_{00}, \quad k = 1,2,\cdots,L_e(0)+1, \tag{59}$$

$$\pi_{k1} = \left(G\alpha^k + \frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\alpha^{k-L_e(0)-1}\beta^{L_e(0)+1}\right)\pi_{00}, \quad k = L_e(0)+2,\cdots,L_e(1), \tag{60}$$

$$\pi_{L_e(1)+1,1} = \bar{p}\left(G\alpha^{L_e(1)+1} + \frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\alpha^{L_e(1)-L_e(0)}\beta^{L_e(0)+1}\right)\pi_{00}, \tag{61}$$

where

$$\alpha = \frac{p\overline{\mu_d}}{\bar{p}\mu_d}, \quad \beta = \frac{\overline{\theta_d}p}{1-\overline{\theta_d}\bar{p}},$$

$$G = \frac{\overline{\mu_d}\theta_d\overline{\theta_d}^2\bar{p}^2(\overline{\mu_d}-\bar{p}-p\bar{p}\overline{\theta_d}+2p) + \overline{\mu_d}\theta_d\overline{\theta_d}\bar{p}(p^2\mu_d+2\mu_d-3p) + \mu_d\theta_d^2\overline{\theta_d}\bar{p}^2(1-p\mu_d-\overline{\theta_d}\bar{p}^2) + \overline{\mu_d}\theta_d(p-\mu_d)}{p\overline{\mu_d}\theta_d(\bar{p}\overline{\theta_d}-\overline{\mu_d})(\bar{p}\overline{\theta_d}-1)(p-\mu_d)}.$$

Because of the PASTA property, the probability of balking is equal to $\pi_{L_e(0)+1,0} + \pi_{L_e(1)+1,1}$. In addition, every customer receives a reward of R units for completing service and there exists a waiting cost of C units per time unit when the customer remains in the system (in queue or in service). Hence the social benefit per time unit when all customers follow the threshold policy $(L_e(0), L_e(1))$ given in Theorem 4.1 equals

$$SB_{dfo} = Rp(1-\pi_{L_e(0)+1,0}-\pi_{L_e(1)+1,1}) - C\left(\sum_{k=0}^{L_e(0)+1}k\pi_{k0} + \sum_{k=0}^{L_e(1)+1}k\pi_{k1}\right).$$

4.2. Almost observable case

Finally we focus on the almost observable case, where arriving customers only observe the queue length at their arriving instant. Similar to the discussion in the case of continuous-time queue, the stationary distribution of the corresponding Markov chain is from Theorem 4.2 with $L_e(0) = L_e(1) = L_e$ and state space $\Omega_{dao}=\{k|0 \leqslant k \leqslant L_e+1\}$. The transition diagram is depicted in Fig. 4.

Theorem 4.3. *Consider an almost observable Geo/Geo/1 queue with the single vacation and $\beta \neq 1 \neq \alpha \neq \beta$, in which customers follow the threshold policy L_e. The stationary probabilities $\{\pi'_k|k\in\Omega_{dao}\}$ are as follows:*

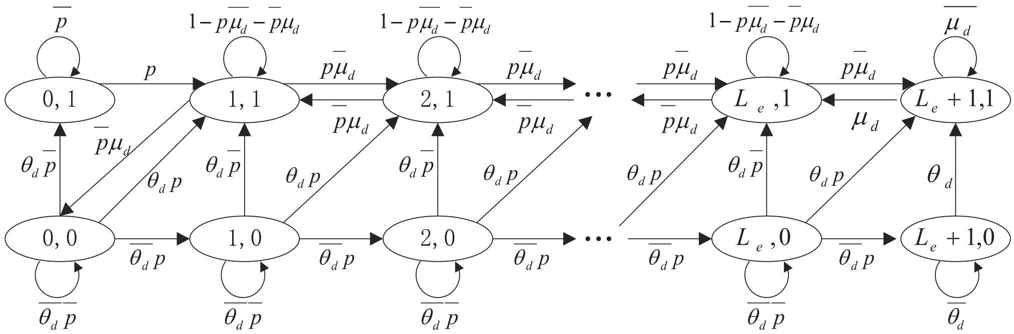

Fig. 4. Transition rate diagram for the L_e threshold strategy in the almost observable queue under discrete case.

$$\pi'_0 = K'\left(1+\frac{\theta_d\bar{p}}{p}\right),$$
$$\pi'_k = K'\left(G\alpha^k+\frac{\mu_d}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\beta^k\right),\quad k=1,2,\cdots,L_e,$$
$$\pi'_{L_e+1} = K'\left[\bar{p}G\alpha^{L_e+1}+\left(\bar{p}\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}+\frac{1-\overline{\theta_d}\bar{p}}{\theta_d}\right)\beta^{L_e+1}\right],$$

where $\pi'_k=\pi_{k0}+\pi_{k1},\ k=0,1,\cdots,L_e+1$, G is the same in Theorem 4.2 and

$$K'=\left[1+\frac{\theta_d\bar{p}}{\theta_d}+\frac{\mu\overline{\theta_d}p}{\theta_d(\overline{\theta_d}\bar{p}-\mu)}+G\left(\frac{p\overline{\mu_d}}{\mu_d-p}-\frac{p\bar{p}}{\mu_d-p}\alpha^{L_e+1}\right)+\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\left(-\frac{p}{\theta_d}\beta^{L_e+1}\right)\right]^{-1}.\quad\square$$

The following analysis is similar with the almost observable case in continuous queue. Because the expected net benefit of a customer who finds k customers in the system, if he decides to enter, is

$$Be=R-\frac{C(k+1)}{\mu_d}-\frac{C\pi_{\bar{J}|L}(0|k)}{\theta_d},\tag{62}$$

where $\pi_{\bar{J}|L}(0|k)$ is the probability that an arriving customer finds the server in a vacation time, given that there are k customers. Using the various forms of π_{kj} from (55)–(61), we get

$$\begin{cases}\pi_{\bar{J}|L}(0|0)=\frac{p\pi_{00}}{p\pi_{00}+p\pi_{01}}=\left(1+\frac{\theta_d\bar{p}}{p}\right)^{-1},\\ \pi_{\bar{J}|L}(0|k)=\frac{p\pi_{k0}}{p\pi_{k0}+p\pi_{k1}}=\left[1+\left(\frac{\theta_d\bar{p}}{p\overline{\mu_d}}-\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\right)\left(\frac{\alpha}{\beta}\right)^k+\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\right]^{-1},\quad k=1,2,\cdots,L_e,\\ \pi_{\bar{J}|L}(0|L_e+1)=\left\{1+(1-\beta)\bar{p}\left[\left(\frac{\theta_d\bar{p}}{p\overline{\mu_d}}-\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\right)\left(\frac{\alpha}{\beta}\right)^{L_e+1}+\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\right]\right\}^{-1}.\end{cases}\tag{63}$$

In light of (62), (63), we introduce the function

$$g(k,y)=R-\frac{C(k+1)}{\mu_d}-\frac{C}{\theta_d}\left\{1+y\left[\left(\frac{\theta_d\bar{p}}{p\overline{\mu_d}}-\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\right)\left(\frac{\alpha}{\beta}\right)^k+\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\right]\right\}^{-1},$$
$$y\in[(1-\beta)\bar{p},1],\quad k=0,1,2,\cdots,\tag{64}$$

which will allow us to prove the existence of equilibrium threshold strategies and derive the corresponding thresholds. Let

$$g_U(k)=g(k,1)=R-\frac{C(k+1)}{\mu_d}-\frac{C}{\theta_d}\left[1+\left(\frac{\theta_d\bar{p}}{p\overline{\mu_d}}-\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\right)\left(\frac{\alpha}{\beta}\right)^k+\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\right]^{-1},\quad k=0,1,2,\cdots,\tag{65}$$

$$g_L(k)=g(k,(1-\beta)\bar{p})=R-\frac{C(k+1)}{\mu_d}-\frac{C}{\theta_d}\left\{1+(1-\beta)\bar{p}\left[\left(\frac{\theta_d\bar{p}}{p\overline{\mu_d}}-\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\right)\left(\frac{\alpha}{\beta}\right)^k+\frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p}-\overline{\mu_d}}\right]\right\}^{-1},\quad k=0,1,2\cdots\tag{66}$$

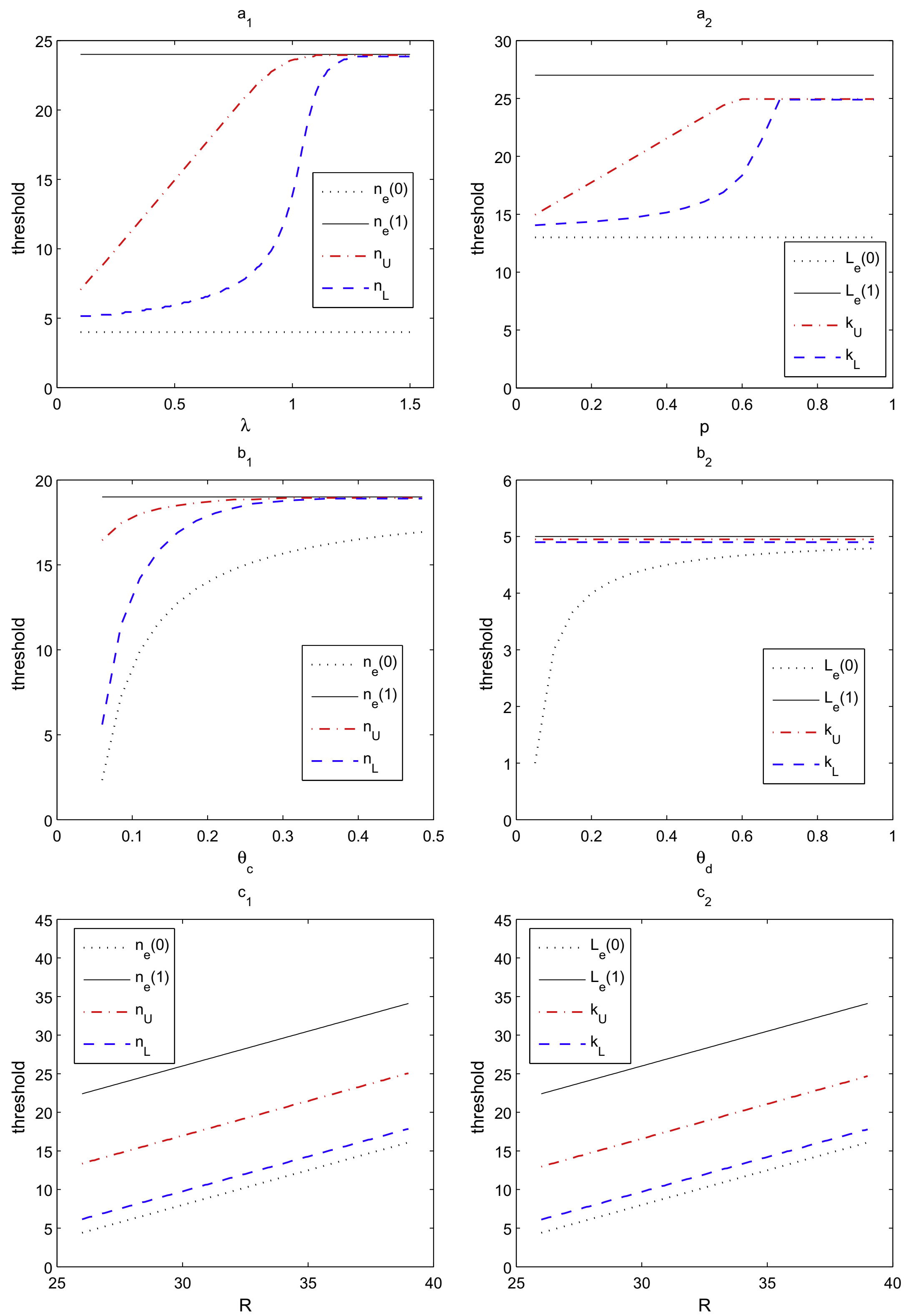

Fig. 5. Equilibrium thresholds for the fully observable and almost observable systems. The three figures on the left are in continuous case. Sensitivity with respect to: a_1. λ, for $\mu = 1$, $\theta_c = 0.05$, $C = 1$, $R = 25$; b_1. θ_c, for $\lambda = 0.8$, $\mu = 1$, $C = 1$, $R = 20$; c_1. R, for $\lambda = 0.4$, $\mu = 0.9$, $\theta_c = 0.05$, $C = 1$. The three figures on the right are in discrete case. Sensitivity with respect to: a_2. p, for $\mu = 0.7$, $\theta_d = 0.05$, $C = 1$, $R = 40$; b_2. θ_d, for $\mu = 0.2$, $p = 0.6$, $C = 1$, $R = 30$; c_2. R, for $p = 0.4$, $\mu = 0.9$, $\theta_d = 0.05$, $C = 1$.

It is easy to see that

$$g_U(0) = R - \frac{C}{\mu_d} - \frac{C}{\theta_d}\left[1 + \frac{\theta_d \bar{p}}{p\overline{\mu_d}}\right]^{-1} > 0$$

and

$$g_L(0) = R - \frac{C}{\mu_d} - \frac{C}{\theta_d}\left[1 + (1-\beta)\bar{p}\frac{\theta_d \bar{p}}{p\overline{\mu_d}}\right]^{-1} > 0.$$

In addition,

$$\lim_{k\to\infty} g_U(k) = \lim_{k\to\infty} g_L(k) = -\infty$$

Hence there exists k_U such that

$$g_U(0),\ g_U(1),\ g_U(2), \cdots, g_U(k_U) > 0 \text{ and } g_U(k_U+1) \leqslant 0. \tag{67}$$

The function $g(k,y)$ is clearly increasing with respect to y for every fixed k, thus we get the relation $g_L(k) \leqslant g_U(k)$, $k = 0, 1, 2, \cdots$. In particular, $g_L(k_U + 1) \leqslant 0$ while $g_L(0) > 0$. Hence, there exists $k_L \leqslant k_U$ such that

$$g_L(k_L) > 0 \text{ and } g_L(k_L+1), \cdots, g_L(k_U),\ g_L(k_U+1) \leqslant 0. \tag{68}$$

We can now establish the existence of the equilibrium threshold policies in the almost observable case and give the following theorem.

Theorem 4.4. *In the almost observable Geo/Geo/1 queue with single vacation, all pure threshold strategies 'observe L_n, enter if $L_n \leqslant L_e$ and balk otherwise' for L_e=$k_L, k_L + 1, \cdots, k_U$ are equilibrium strategies.*

Proof. Consider a tagged customer at his arrival instant and assume all other customers follow the same threshold strategy 'observe L_n, enter if $L_n \leqslant L_e$ and balk otherwise ' for some fixed $L_e \in \{k_L, k_L + 1, \cdots, k_U\}$. Then $\pi^-_{J|L}(0|k)$ is given by (63).

If the tagged customer finds $k \leqslant L_e$ customers and decides to enter, his expected net benefit is equal to

$$R - \frac{C(k+1)}{\mu_d} - \frac{C}{\theta_d}\left[1 + \left(\frac{\theta_d \bar{p}}{p\overline{\mu_d}} - \frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p} - \overline{\mu_d}}\right)\left(\frac{\alpha}{\beta}\right)^k + \frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p} - \overline{\mu_d}}\right]^{-1} = g_U(k) > 0$$

because of (62),(63),(64),(65) and (67). So in this case the customer prefers to enter.

If the tagged customer finds $k = L_e + 1$ customers and decides to enter, his expected net benefit is

$$R - \frac{C(L_e+2)}{\mu_d} - \frac{C}{\theta_d}\left\{1 + (1-\beta)\bar{p}\left[\left(\frac{\theta_d \bar{p}}{p\overline{\mu_d}} - \frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p} - \overline{\mu_d}}\right)\left(\frac{\alpha}{\beta}\right)^{L_e+1} + \frac{1-\overline{\theta_d}\bar{p}}{\overline{\theta_d}\bar{p} - \overline{\mu_d}}\right]\right\}^{-1} = g_L(L_e+1) \leqslant 0$$

because of (62),(63),(64), (66) and (68). Therefore in this case the customer prefers to balk. □

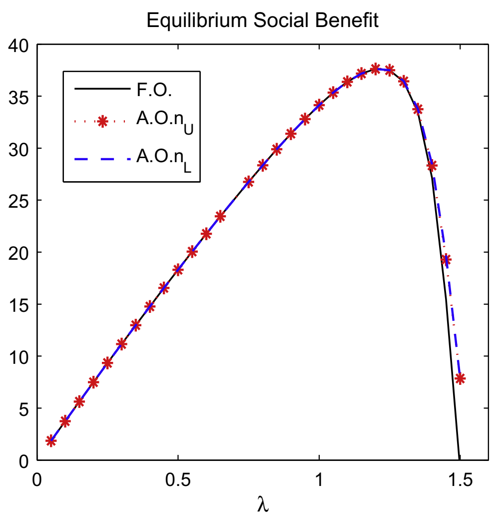

Fig. 6. Equilibrium social benefit in continuous case. Sensitivity with respect to λ, for $\mu = 1.5$, $\theta_c = 0.8$, $C = 1$, $R = 40$.

Because the probability of balking is equal to π'_{L_e+1}, the social benefit per time unit when all customers follow the threshold policy L_e given in Theorem 4.4 equals

$$SB_{dao} = Rp\left(1 - \pi'_{L_e+1}\right) - C\left(\sum_{k=0}^{L_e+1} k\pi'_k\right).$$

5. Numerical experiments

In this section, according to the results obtained, we present a set of numerical experiments. Here we concern about the values of the equilibrium thresholds and the social benefits per unit time in both the continuous-time queue and discrete-time queue. We illustrate the effect of the information level as well as several parameters on the equilibrium thresholds and the social benefits.

In Fig. 5 we observe the effect of the information level on the equilibrium thresholds and explore the sensitivity of the equilibrium thresholds with respect to λ, p, θ and R. It shows that in all of six diagrams the equilibrium thresholds $\{n_L, \cdots, n_U\}$ and $\{k_L, \cdots, k_U\}$ for the almost observable case always locate respectively in $(n_e(0), n_e(1))$ and $(L_e(0), L_e(1))$ for the fully observable case. In other words, no matter in continuous case or discrete case, the thresholds in the almost observable model have

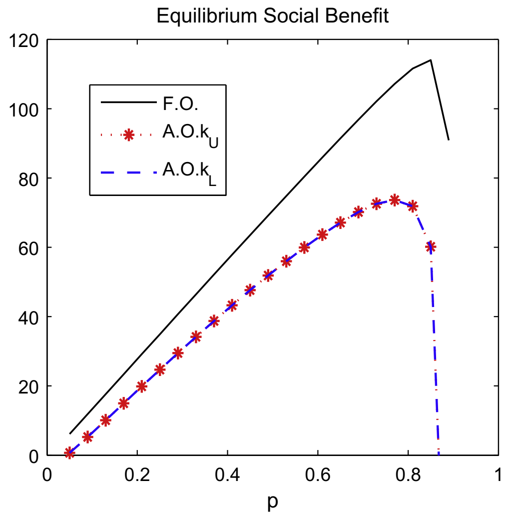

Fig. 7. Equilibrium social benefit in discrete case. Sensitivity with respect to p, for $\mu = 0.9$, $\theta_d = 0.5$, $C = 0.2$, $R = 150$.

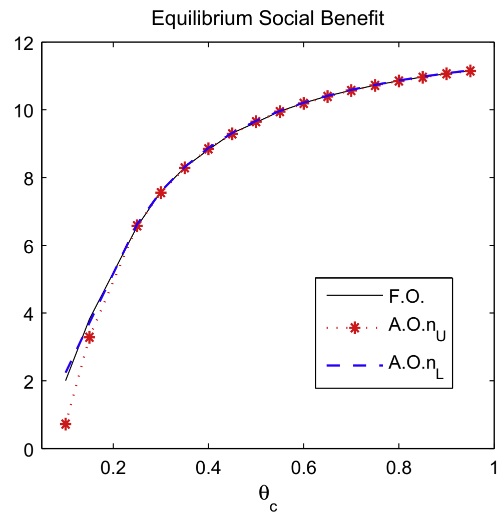

Fig. 8. Equilibrium social benefit in continuous case. Sensitivity with respect to θ_c, for $\lambda = 0.8$, $\mu = 1$, $C = 1$, $R = 25$.

intermediate values between the two separate thresholds when the arriving customers are informed of the state of the server. This property is also held for the models with multiple vacations or interruptible setup/closedown times (see references [10] and [12]).

Concerning the sensitivity of the equilibrium thresholds, we can make the following observations. Comparing the figures in continuous case with discrete case in Fig. 5, we find that the curves have the same shape and trend. Under varying arrival rate the fully observable thresholds remain fixed since the arrival rate is irrelevant to the customer's decision when he has full state information. On the other hand, the almost observable thresholds increase with the arrival rate, which means that if an arriving customer is told the information of the present queue length, then he is more likely to enter when the arrival rate is higher. This phenomenon that customers in equilibrium tend to imitate the behaviors of other customers in the almost observable model is of the 'Follow-The-Crowd' (FTC) type. The reason is that when the arrival rate is high, it is probably that the server is active, therefore the expected delay from server vacation is reduced. However, the vacation rate varies, all thresholds increase, expect that the equilibrium threshold of the server being working remains constant. Evidently, this is intuitive, because when the server vacation gets shorter, customers generally have a greater incentive to enter both in the fully and almost observable models. Finally, along with the rising of the service reward R, the thresholds increase in a linear fashion. In addition, the values of R are identical when the parameters are as the same in continuous case as in discrete case. It is easy to explain, because the continuous-time model can be looked as the limit of the discrete-time model.

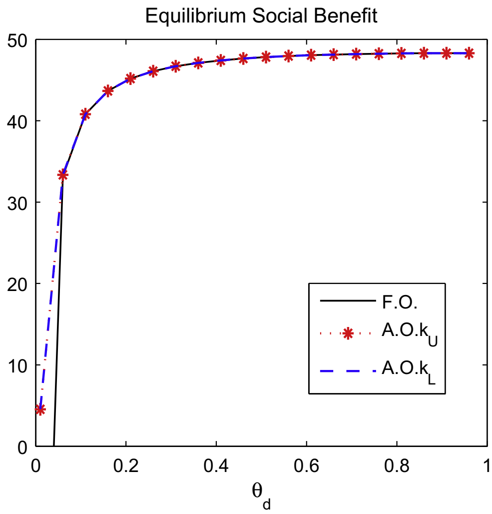

Fig. 9. Equilibrium social benefit in discrete case. Sensitivity with respect to θ_d, for $p = 0.5$, $\mu = 0.9$, $C = 0.1$, $R = 100$.

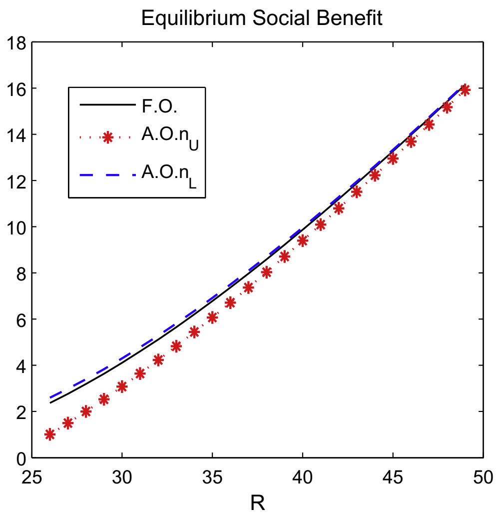

Fig. 10. Equilibrium social benefit in continuous case. Sensitivity with respect to R, for $\lambda = 0.8$, $\mu = 1$; $\theta_c = 0.1$, $C = 1$.

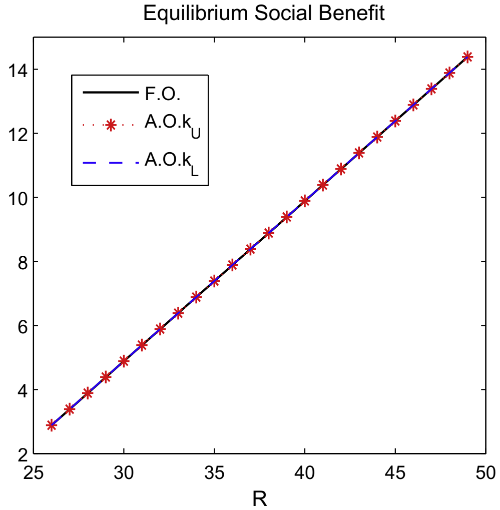

Fig. 11. Equilibrium social benefit in discrete case. Sensitivity with respect to R, for $p = 0.5$, $\mu = 0.9$; $\theta_d = 0.1$, $C = 0.1$.

Furthermore, we consider the social benefits for the fully and almost observable cases under the corresponding equilibrium strategies in both the continuous case and discrete case. The results are presented in Figs. 6–11. In the almost observable case, we have seen that there are multiple equilibrium strategies corresponding to thresholds $\{n_L, \cdots, n_U\}$ in continuous model or thresholds $\{k_L, \cdots, k_U\}$ in discrete model. For this reason we only present the social benefits under the two extreme thresholds. In figures we observe that the curves in continuous case have roughly the same trend as in discrete case. The difference in equilibrium social benefit is almost invisible between the fully and almost observable case, especially in Fig. 11 where the three pieces of curves nearly overlap each other. Moreover, Fig. 7 shows that in discrete model the equilibrium social benefit in the fully observable case is bigger than that in the almost observable extreme cases. However, in Fig. 10 the equilibrium social benefit in the fully observable case has an intermediate value between the social benefits under the two extreme thresholds. Even in Figs. 6 and 9, some certain values of the social benefit in the fully observable case are smaller than the other two values in the almost observable case. Thus, it may be argued that the additional information on the server state is not always helpful for increasing equilibrium social benefit.

As to the sensitivity of the equilibrium social benefit, we find that it increases with respect to the vacation rate and service reward R, both of which are intuitive. Regarding the arrival rate, regardless of whether in continuous case or discrete case, the social benefit achieves a maximum for intermediate values. This can be explained that when the arrival rate is small the system is rarely crowded, therefore customers can be served soon and the social benefit improves. However, with the increasing of the arrival rate, a smaller percentage tends to enter because of the longer delays, which bring negative effect to the social benefit.

6. Conclusion

In this paper we studied the equilibrium customer behavior in observable $M/M/1$ and $Geo/Geo/1$ queueing systems under single vacation policy. To the best of the authors' knowledge, this is the first time analyzing the queueing system with single vacation from an economic perspective. Customers have the right to decide whether to join or to balk according to the accurate situation, which is more sensible than the classical viewpoint that the decisions are made by the servers and the customers are forced to follow them. In observable models, customers are informed of the queue length at arriving instant. We classified two subcases: fully observable and almost observable, depending on the additional information, or lack thereof, the state of the server. The equilibrium strategies and resulting stationary system behavior were explored from continuous and discrete perspective and we comparatively analyzed the results of both. Moreover, we discussed the effect of the information level as well as several parameters on the equilibrium thresholds and social benefit.

The study could offer the customers with optimal strategies to reduce the loss of queueing. Besides, the results could provide reference information on the pricing issues in queueing systems. Further extensions would be researched about the equilibrium customer strategies in unobservable models. Furthermore, we can also explore equilibrium behavior under working vacation policy.

Acknowledgements

The authors thank editor and the referees for their valuable comments and acknowledge that this work was supported by the Humanities and Social Science Fund in Ministry of Education in China (No. 10YJC630114) and Post-Doctoral Research Project of China (No. 92169).

References

[1] P. Naor, The regulation of queue size by levying tolls, Econometrica 37 (1969) 15–24.
[2] U. Yechiali, On optimal balking rules and toll charges in the GI/M/1 queue, Oper. Res. 19 (1971) 349–370.
[3] S.G. Johansen, S. Stidham Jr., Control of arrivals to a stochastic input-output system, Adv. Appl. Probab. 12 (1980) 972–999.
[4] S. Stidham Jr., Optimal control of admissions to a queueing system, IEEE Trans. Automatic Control 30 (1985) 705–713.
[5] H. Mendelson, S. Whang, Optimal incentive-compatible priority pricing for the M/M/1 queue, Oper. Res. 38 (1990) 870–883.
[6] H. Chen, M. Frank, State dependent pricing with a queue, IIE Tran. 33 (2001) 847–860.
[7] C. Larsen, Investigating sensitivity and the impact of information on pricing decisions in an M/M/1/ queueing model, Int. J. Prod. Econ. 56–57 (1998) 365–377.
[8] J. Erlichman, R. Hassin, Equilibrium solutions in the observable M/M/1 queue with overtaking, Naval Res. Logistics 58 (2011) 107–122.
[9] R. Hassin, M. Haviv, Equilibrium Behavior in Queueing Systems: to Queue or not to Queue, Kluwer Academic Publishers, Dordrecht, 2003.
[10] A. Burnetas, A. Economou, Equilibrium customer strategies in a single server Markovian queue with setup times, Queueing Syst. 56 (2007) 213–228.
[11] A. Economou, S. Kanta, Equilibrium balking strategies in the observable single-server queue with breakdowns and repairs, Oper. Res. Lett. 36 (2008) 696–699.
[12] W. Sun, P. Guo, N. Tian, Equilibrium threshold strategies in observable queueing systems with setup/closedown times, CEJOR 18 (2010) 241–268.
[13] A. Economou, A. Gómez-Corral, S. Kanta, Optimal balking strategies in single-server queues with general service and vacation times, Perform. Eval. 68 (2011) 967–982.
[14] Y. Levy, U. Yechiali, Utilization of idle time in an M/G/1 queueing system, Manag. Sci. 22 (1975) 202–211.
[15] Y. Levy, U. Yechiali, An M/M/c queue with servers vacations, INFOR 14 (1976) 153–163.
[16] B.T. Doshi, Queueing systems with vacations-a survey, Queueing Syst. 1 (1986) 29–66.
[17] H. Takagi, Queueing Analysis, Vacation Systems, vol. 1, Elsevier Science, Amsterdam, 1991.
[18] H. Tagagi, Queueing Analysis, Discrete-Time Systems, vol. 3, North-Holland Elsevier, Amsterdam, 1993.
[19] G. Choudhury, A batch arrival queue with a vacation time under single vacation policy, Comput. Oper. Res. 29 (2002) 1941–1955.
[20] K. Sikdar, U.C. Gupta, Analytic and numerical aspects of batch service queues with single vacation, Comput. Oper. Res. 32 (2005) 943–966.
[21] K.C. Madan, M. Al-Rawwash, On the Mx/G/1 queue with feedback and optional server vacations based on a single vacation policy, Appl. Math. Comput. 160 (2005) 909–919.
[22] U.C. Gupta, K. Sikdar, Computing queue length distributions in MAP/G/1/N queue under single and multiple vacation, Appl. Math. Comput. 174 (2006) 1498–1525.
[23] H. Takagi, Queueing Analysis, Vacation and Priority Systems, vol. 1, North-Holland Elsevier, Amsterdam, 1991.
[24] N. Tian, Z.G. Zhang, Vacation Queueing Models, Theory and Applications, Springer, New York, 2006.