RAIRO-Oper. Res. 54 (2020) 569–583
https://doi.org/10.1051/ro/2019016

RAIRO Operations Research
www.rairo-ro.org

STRATEGIC BEHAVIOR IN THE CONSTANT RETRIAL QUEUE WITH A SINGLE VACATION

Yu Zhang*

Abstract. We study customers' joining strategies in an $M/M/1$ constant retrial queue with a single vacation. There is no waiting space in front of the server and a vacation is triggered when the system is empty. If an arriving customer finds the server idle, he occupies the server immediately. Otherwise, if the server is found unavailable, the customer enters a retrial pool called orbit with infinite capacity and becomes a repeated customer. According to the different information provided for customers, we consider two situations, where we investigate system characteristics and customers' joining or balk decisions based on a linear reward-cost structure. Furthermore, we establish the social welfare of the system and make comparisons between the two information levels. It is found that there exist thresholds of system parameters such that the social planner would prefer revealing more information when the system parameter is greater than or less than the corresponding threshold.

Mathematics Subject Classification. 60K25, 90B22.

Received November 26, 2018. Accepted January 23, 2019.

1. Introduction

In daily life, customers are natural to respond strategically to the service delay by deciding whether to join the service systems or not, aiming to maximize their individual utility. During the past decades, equilibrium analysis for queueing systems taking customers' joining and balking decisions into consideration has been paid significant attention, which applies to e-commerce and management in service systems. Initially, Naor [15] studied an observable $M/M/1$ queue with a simple linear reward-cost function where arriving customers are informed about the number of customers waiting in the queue, and based on that customers make their decisions on whether to join the system or to balk. In this seminal paper, both the individual equilibrium strategy and socially optimal strategy were investigated. Edelson and Hilderbrand [7] complemented Naor's study by studying the unobservable case with no information revealed to customers. From then on, there has been a growing volume of literature regarding equilibrium analysis, and the fundamental problem is to identify customers' equilibrium strategy and socially optimal strategy. The comprehensive monographs by Hassin [12], and Hassin and Haviv [13] summarized the main approaches and results on this topic. Interested readers can refer to them for more details.

Keywords. Queueing, Equilibrium strategies, Retrial, Single vacation, Social welfare.

Donlinks School of Economics and Management, University of Science and Technology Beijing, Beijing 100083, China.

* Corresponding author: yuzhang@ustb.edu.cn; yuzhang9006@163.com

In the literature, there are some works considering individual equilibrium joining and balking behavior in queueing systems with retrials, which is rather common in industrial engineering, communication systems and business management etc. In such systems, an arriving customer is served immediately if the server is available. Otherwise, if the server is found busy, the customer has to retry after some random time. For the classical $M/M/1$ retrial queue, the retrial rate of the orbit is proportional to the number of customers in the orbit. The equilibrium and socially optimal balking strategies were investigated for a classical $M/M/1$ retrial queue in Wang and Zhang [22]. To model a local area networking system, Wang and Zhang [23] studied a retrial queue with delayed vacations and the optimal pricing issues were discussed. Zhang et al. [28] gave the optimal pricing strategies in retrial queueing systems with two servers who provide complementary services. Taking the power consumption into consideration, Zhang and Wang [29] built an $M/G/1$ retrial queue with reserved idle time and setup time. The optimal reserved idle time was investigated. When the retrial rate is constant, Economou and Kanta [5] studied an $M/M/1$ constant retrial queue where customers' equilibrium strategy, social and profit maximization problems were considered under two information levels. When the server may break down at working states, Zhang et al. [26] investigated the partially observable case and fully observable case. Wang et al. [24] studied the situation where there are two types of customers, *i.e.*, primary customers and negative customers. An arriving negative customer causes the failure of server and the primary customer being served is then deleted. Recently, Wang et al. [25] proceeded the equilibrium analysis for an $M/M/1$ constant retrial queue with N-policy. With the N-policy, the server is shut down when the system is empty and turned on again when there are at least N customers in the system. Chapter 6 of Hassin [13] gives a detailed summary for equilibrium analysis of retrial queueing systems. As for some discussions of retrial models and methods, interested readers are referred to Artalejo and Gómez-Corral [2], and Falin and Templeton [8].

This paper aims to study customers' strategic behavior in an $M/M/1$ constant retrial queue with a single vacation policy. Vacation policies (summarized in Tian and Zhang [20]) apply to the situation where the server is unavailable to customers for some occasional periods of time. During such periods of time, the server may take up some other work, such as maintenance work, scanning for viruses and serving secondary customers etc. Queueing systems with server vacations are often used to model the process of many production, computer and communication systems. The vacation of server leads to the unavailability of the service, and consequently increases customers' waiting time and the complexity of the analysis for customers' equilibrium joining strategies. For this reason, queueing systems with a vacation policy have been investigated for decades. Among them, there are mainly four kinds of vacation policy; that is, single vacation policy (*e.g.* see [14]), multiple vacation policy (*e.g.* see [3] and [16] for the Markovian case, and see [6] for the non-Markovian case), N-vacation policy (*e.g.* see [9]–[11], [21]) and working vacation policy (*e.g.* see [17]–[19] and [27]). However, all these works are devoted to queueing systems without retrials. To the best of our knowledge, the only work taking customers' retrial and server's vacation into consideration is Do et al. [4] which studied customers' joining or balk decision in an $M/M/1$ constant retrial queue under the working vacation policy. In this paper, we assume the server takes a vacation once there is no customer in the system and the vacation ends after a period of time following an Exponential distribution; that is, our analysis is under a single vacation policy.

There is clearly a practical situation that motivates us to study the system described. For example, a firm has a production facility which is mainly operated in a "produce to order" mode in a competitive market environment. If the production facility is idle upon a customer's arrival, this customer can be served immediately. Otherwise, if the production facility is found unavailable, this customer will try to access it some time later. Customers decide whether or not to place their orders with this firm based on either the lead time (wait time) and production facility status (fully observable case) or long-term lead time statistics when the facility is unavailable (almost unobservable case). While processing the customer orders is the top priority, the manager also wants to reduce the idle time of the production facility which can be expensive, and take up some maintenance work after processing the orders. Such a production model can be characterized by retrial queueing systems with a single vacation policy. Although the firm can reduce the idle time and improve the utilization of the production facility with such a vacation policy, it may still want to maximize the social welfare of customers placing orders. Thus, to maximize the social welfare of customers who can decide to place an order under different information scenarios, one important issue is which information scenario is more profitable for the social planner. To gain insights into the effects of information, we consider an $M/M/1$ constant retrial queue with a single vacation policy. We derive customers' equilibrium joining strategies that offer the insights to this issue.

In summary, in addition to the modelling contribution to the queueing-game literature, our paper contains the following findings: (i) We study customers' equilibrium joining strategies under two information levels, *i.e.*, the almost unobservable case and fully observable case. In the almost unobservable case, arriving customers are only informed whether the server is idle or not, and we obtain customers' equilibrium joining strategy, which is either a pure strategy (join or balk) or a mixed strategy (join with a certain probability). In the fully observable case, both the server's state and the number of customers in the orbit are communicated to the customers upon arrival. In this situation, when the server is unavailable, there exists an optimal joining threshold policy such that a customer chooses to join if the queue length (*i.e.*, the number of customers in the orbit) is less than the threshold. In addition, when customers follow the equilibrium policy, some system characteristics and the social welfare of the system are given. (ii) Through numerical examples, we compare the two information levels in terms of social welfare. We find there exist thresholds of arrival rate, service rate, retrial rate and completion rate of the vacation such that the fully observable case benefits the social planner more than the almost unobservable case when the system parameter is beyond or below the corresponding threshold.

The rest of the paper is organized as follows. In Section 2, we give a detailed description of the model. In Section 3, we investigate customers' equilibrium behavior in the almost unobservable case and fully observable case and the social welfare in each case is established. In Section 4, numerical examples are illustrated to examine the effects of system parameters on customers' equilibrium strategy and we compare the two information levels to indicate which case is better from the perspective of the social planner. Finally, a brief conclusion is given in Section 5.

2. Description of the model

We consider an $M/M/1$ retrial queueing system with a single vacation policy. We assume customers arrive according to a Poisson process with rate λ and the service times are exponentially distributed with rate μ. If an arriving customer finds the server idle, he occupies the server immediately and starts being served; otherwise, if the server is unavailable upon a customer's arrival, this customer goes into a retrial pool with infinite capacity and becomes a repeated customer. The service discipline for customers in the retrial orbit is first-come-first-served (FCFS); that is, only the customer at the head of the orbit queue can repeat his request for service, which is common in literature, *e.g.* see [5], [26]. The inter-retrial times of customers follow an Exponential distribution with rate θ and the customer at the head of orbit queue continues to retry until he receives his requested service, after which, he leaves the system. Further, when the server finishes serving a customer and finds the system empty, it leaves for a vacation and the vacation time follows an Exponential distribution with rate α. After a vacation, the server stays idle to wait for customers arriving from outside or retrying successfully from the orbit.

Denote $I(t)$ as the server's state at time t. The events $I(t) = 0, 1, 2$ correspond to the server is idle, busy and on vacation, respectively. Let $N(t)$ be the number of customers in the orbit at time t, then the process $\{I(t), N(t) : t \geq 0\}$ is a two dimensional continuous-time Markov chain with state space $\{0, 1, 2\} \times \{0, 1, 2, \ldots\}$ and non-zero transition rates as follows:

$$
\begin{aligned}
q_{(0,i)(1,i)} &= \lambda, && i \geq 0\\
q_{(0,i)(1,i-1)} &= \theta, && i \geq 1\\
q_{(1,i)(1,i+1)} &= \lambda, && i \geq 0\\
q_{(1,0)(2,0)} &= \mu, && \\
q_{(1,i)(0,i)} &= \mu, && i \geq 1\\
q_{(2,i)(0,i)} &= \alpha, && i \geq 0\\
q_{(2,i)(2,i+1)} &= \lambda, && i \geq 0.
\end{aligned}
$$

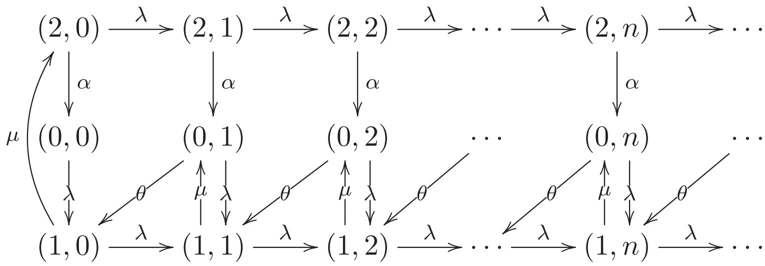

FIGURE 1. The transition rate diagram.

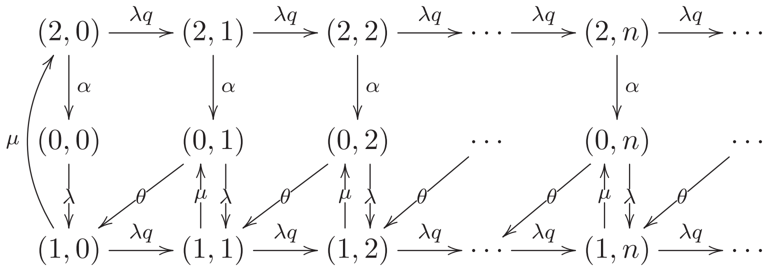

FIGURE 2. The transition rate diagram in the almost unobservable case.

The corresponding transition rate diagram is illustrated in Figure 1.

Assume each customer receives a service reward R, which reflects the added value of being served or his satisfaction towards service, and a waiting cost C is incurred when customers remain in the system. Upon arrival, each customer decides whether to join the system or not according to the different information provided for him. We assume customers are risk neutral, aiming to maximize their expected utility. In addition, customers' decisions are irrevocable; that is, neither the reneging of joining customers nor the retrial of balking customers is allowed.

In the following sections, we consider two information levels. In the first one, arriving customers can access whether the server is idle or not, which is referred to as the almost unobservable case. And in the second one, both the server's state and the number of customers in the orbit are provided for customers and we call it the fully observable case.

3. EQUILIBRIUM ANALYSIS

In this section, we will analyze customers' equilibrium joining behavior under the two information levels defined above.

3.1. Almost unobservable case

We begin with the almost unobservable situation where customers are only informed whether the server is idle or not. Since customers can begin their service immediately when the server is idle, customers finding an idle server join the system with probability 1. Otherwise, if the server is unavailable, we characterize customers' joining probability by q, where $q = 0$ or 1 represents a pure strategy and $q \in (0, 1)$ is a mixed strategy. The transition rate diagram in this situation is shown in Figure 2.

Let $P_{au}(i,j)$ be the steady state probability that the server is at state (i,j). Based on the transition rate diagram given in Figure 2, we have steady state equations given by

$$\lambda P_{au}(0,0) = \alpha P_{au}(2,0), \tag{3.1}$$
$$(\lambda+\theta)P_{au}(0,n) = \mu P_{au}(1,n) + \alpha P_{au}(1,n), \quad n = 1,2,3,\ldots, \tag{3.2}$$
$$(\lambda q+\mu)P_{au}(1,0) = \lambda P_{au}(0,0) + \theta P_{au}(0,1), \tag{3.3}$$
$$(\lambda q+\mu)P_{au}(1,n) = \lambda P_{au}(0,n) + \lambda q P_{au}(0,n-1) + \theta P_{au}(0,n+1), \quad n = 1,2,3,\ldots, \tag{3.4}$$
$$(\lambda q+\alpha)P_{au}(2,0) = \mu P_{au}(1,0), \tag{3.5}$$
$$(\lambda q+\alpha)P_{au}(2,n) = \lambda q P_{au}(2,n-1), \quad n = 1,2,3,\ldots. \tag{3.6}$$

To solve these equations, we define the partial generating functions: $Q_i(z) = \sum_{j=0}^{\infty} P_{au}(i,j)z^j, i = 0,1,2.$

Multiplying equations (3.2), (3.4), (3.6) with z^n and summing over n give that

$$(\lambda+\theta)(Q_0(z) - P_{au}(0,0)) = \mu(Q_1(z) - P_{au}(1,0)) + \alpha(Q_2(z) - P_{au}(2,0)), \tag{3.7}$$
$$(\lambda q+\mu)(Q_1(z) - P_{au}(1,0)) = \lambda(Q_0(z) - P_{au}(0,0)) + \lambda q z Q_1(z) + \frac{\theta}{z}(Q_0(z) - P_{au}(0,0)) - \alpha P_{au}(0,1), \tag{3.8}$$
$$(\lambda q+\alpha)(Q_2(z) - P_{au}(2,0)) = \lambda q z Q_2(z). \tag{3.9}$$

Substituting equation (3.3) into (3.8), we have

$$Q_1(z) = \frac{\lambda + \frac{\theta}{z}}{\mu + \lambda q - \lambda q z} Q_0(z) - \frac{\frac{\theta}{z}}{\mu + \lambda q - \lambda q z} P_{au}(0,0). \tag{3.10}$$

Using equations (3.1) and (3.5), it follows from equation (3.6) that

$$Q_2(z) = \frac{\lambda(\alpha+\lambda q)}{\alpha(\alpha+\lambda q - \lambda q z)} P_{au}(0,0). \tag{3.11}$$

Substituting equations (3.10) and (3.11) into (3.2), we get

$$Q_0(z) = \frac{\theta + \frac{\lambda(\alpha+\lambda q)}{\alpha+\lambda q-\lambda q z} - \frac{\theta\mu}{z(\mu+\lambda q-\lambda q z)} - \frac{\lambda(\alpha+\lambda q)}{\alpha}}{\lambda+\theta - \frac{\lambda\mu z+\theta\mu}{z(\mu+\lambda q-\lambda q z)}} P_{au}(0,0). \tag{3.12}$$

Combining equations (3.10)–(3.12) and employing the normalizing condition $Q_0(1) + Q_1(1) + Q_2(1) = 1$, we derive that

$$P_{au}(0,0) = \frac{\alpha^2[\theta\mu - \lambda q(\lambda+\theta)]}{\lambda^2 q\mu(\lambda q+\alpha) + \theta\{\alpha^2[\lambda(1-q)+\mu] + \lambda\mu\alpha + \lambda^2 q\mu\}}, \tag{3.13}$$

and thus $Q_0(z)$, $Q_1(z)$, $Q_2(z)$ are followed.

We give the stability condition and system characteristics in Lemma 3.1 and Theorem 3.2.

Lemma 3.1. *For the almost unobservable model of $M/M/1$ retrial queue with a single vacation, the system is stable if and only if $\theta\mu - \lambda q(\lambda+\theta) > 0$.*

Proof. According to Theorem 1.6 given in Anderson [1], since all states are communicating, all probabilities $P_{au}(i,j)$ $(i = 1,2,3,\ j \geq 0)$ are either all positive with sum 1, or all equal to zero, by the recurrent events theory. From the expression of $P_{au}(0,0)$ given in equation (3.13), $P_{au}(0,0)$ is positive if and only if $\theta\mu - \lambda q(\lambda+\theta) > 0$. Thus, the stability condition for the system is $\theta\mu - \lambda q(\lambda+\theta) > 0$ by the ergodicity theory. □

The ergodic condition guarantees the system not being too crowded and the existence of stationary distribution of the Markov chain. In what follows, when customers are informed whether the server is idle or not, we proceed the analysis under this ergodic condition, that is, at steady state.

Theorem 3.2. *In the almost unobservable model of $M/M/1$ retrial queue with a single vacation, at steady state,*

(i) The probabilities that the server is idle, busy or on vacation are, respectively, given by

$$P_{au}(0) = \frac{\lambda(\lambda q+\alpha)[\mu\theta-\lambda q(\lambda+\theta)]}{\lambda^2 q\mu(\lambda q+\alpha)+\theta[\lambda\mu\alpha+\lambda^2 q\mu+\alpha^2(\lambda-\lambda q+\mu)]}, \tag{3.14}$$

$$P_{au}(1) = \frac{\mu\theta\alpha^2+\lambda^3 q^2\mu+\lambda q\alpha(\lambda\mu-\alpha\theta)}{\lambda^2 q\mu(\lambda q+\theta)+\theta[\lambda\mu\alpha+\lambda^2 q\mu+\alpha^2(\lambda-\lambda q+\mu)]}, \tag{3.15}$$

$$P_{au}(2) = \frac{\lambda[\lambda^2 q(\lambda q+\alpha)+\theta(\alpha^2+\lambda q\theta+\lambda^2 q^2)]}{\lambda^2 q\mu(\lambda q+\theta)+\theta[\lambda\mu\alpha+\lambda^2 q\mu+\alpha^2(\lambda-\lambda q+\mu)]}. \tag{3.16}$$

(ii) Customers' expected waiting time in the orbit is

$$E[W_{au}(q)] = \frac{1}{\theta}+\left(1+\frac{\lambda q}{\theta}\right)\left\{\frac{1}{\alpha}+\frac{\lambda+\theta}{\theta\mu-\lambda q(\lambda+\theta)}-\frac{\lambda q+2\alpha}{\alpha^2+\alpha\mu+\lambda q\mu}\right\}. \tag{3.17}$$

Proof. Inserting $z=1$ into equations for $Q_0(z)$, $Q_1(z)$, $Q_2(z)$, we have the probabilities that the server is under different states. Regarding to customers' mean waiting time in the orbit, since customers' effective arrival rate to the retrial orbit (denoted by λ_o) is $\lambda_o = \lambda q(P_{au}(1)+P_{au}(2))$ and the expected number of customers in the orbit (denoted by $E(N_{au})$) is $E(N_{au}) = Q_0'(1)+Q_1'(1)+Q_2'(1)$, by Little's Law $E(W_{au}) = \frac{E(N_{au})}{\lambda_o}$, customers' mean waiting time in the orbit is then followed. □

Recall that customers receive a service reward R after being served and there is a waiting cost C per time unit. The expected utility of a customer if he decides to join the system is then defined as

$$U_{au}(q) = R - C\left(E[W_{au}(q)]+\frac{1}{\mu}\right). \tag{3.18}$$

Theorem 3.3. *In the almost unobservable model of $M/M/1$ retrial queue with a single vacation, customers' equilibrium joining probability, defined as q_e^{un}, is as follows:*

(i) if $R > C(E[W_{au}(1)]+\frac{1}{\mu})$, then $q_e^{au}=1$;
(ii) if $R < C(E[W_{au}(0)]+\frac{1}{\mu})$, then $q_e^{au}=0$;
(iii) if there exists q^* satisfying $R = C(E[W_{au}(q^*)]+\frac{1}{\mu}))$, then $q_e^{au}=q^*$.

Proof. If $R > C(E[W_{au}(1)]+\frac{1}{\mu})$, that means that when all the other customers choose to join, the expected utility of a tagged customer who also chooses to join is positive. So all joining is an equilibrium. Similarly, if $R < C(E[W_{au}(0)]+\frac{1}{\mu})$, that means that when all the other customers choose to balk, the expected utility of a tagged customer who chooses to join is negative. So the best response for this tagged customer is balking and thus, no customer entering is an equilibrium. If all the other customers choose to join the system with probability q^* which is a solution to equation $R = C(E[W_{au}(q^*)]+\frac{1}{\mu})$, there is no difference for a tagged customer between joining and balking. Because in this case, the expected utility of the tagged customer when choosing joining or balking is equal to zero. So $q_e^{au}=q^*$. □

Intuitively, Nash equilibrium is the one under which a tagged customer has no incentive to change his decision once other customers follow it. That means that when all the other customers adopt this strategy, the tagged customer receives no incremental benefit if he deviates from the Nash equilibrium strategy. Define the social welfare of the system as the sum of the utilities of customers joining the system. Let SW_{au} be the social welfare per time unit when all customers finding the server unavailable follow the equilibrium strategy given in Theorem 3.3. Then it can be expressed as

$$SW_{au}(q) = \lambda P_{au}(0)\left(R - \frac{C}{\mu}\right) + \lambda q_e^{au}(P_{au}(1) + P_{au}(2)) \cdot \left(R - C\left(E[W_{au}(q)] + \frac{1}{\mu}\right)\right) \cdot \tag{3.19}$$

3.2. Fully observable case

In this subsection, we turn our attention to the situation where both the number of customers in the orbit and the server's state are revealed to customers. Let $T(i,j)$ denote a tagged customer's expected sojourn time in the system when he is at the jth position in the orbit and the server's state is i. Based on the reward-cost structure, the expected utility for the tagged customer is

$$U_{fo}(i,j) = R - CT(i,j). \tag{3.20}$$

By a first step argument, we have $T(i,j)$ satisfies the following equations

$$T(1,0) = \frac{1}{\mu}, \tag{3.21}$$

$$T(1,j) = \frac{1}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}T(1,j) + \frac{\mu}{\lambda+\mu}T(0,j), \quad j = 1,2,\ldots, \tag{3.22}$$

$$T(0,j) = \frac{1}{\lambda+\theta} + \frac{\lambda}{\lambda+\theta}T(1,j) + \frac{\theta}{\lambda+\mu}T(1,j-1), \quad j = 1,2,\ldots, \tag{3.23}$$

$$T(2,j) = \frac{1}{\lambda+\alpha} + \frac{\lambda}{\lambda+\alpha}T(2,j) + \frac{\alpha}{\lambda+\alpha}T(0,j), \quad j = 1,2,\ldots, \tag{3.24}$$

where $T(1,0)$ represents the expected sojourn time of a tagged customer being served. Solving equation (3.22), we get

$$T(0,j) = T(1,j) - \frac{1}{\mu} \cdot \tag{3.25}$$

Plugging equation (3.25) into (3.23) gives that

$$T(1,j) = \frac{\lambda+\mu+\theta}{\mu\theta} + T(1,j-1), \quad j = 1,2,\ldots. \tag{3.26}$$

By iterating equation (3.26), we have $T(1,j) = \frac{\lambda+\mu+\theta}{\mu\theta}j + \frac{1}{\mu}$ for $j \geq 0$ and so $T(0,j) = \frac{\lambda+\mu+\theta}{\mu\theta}j$, $j \geq 1$. Substituting the expression of $T(0,j)$ into equation (3.24), it follows that $T(2,j) = \frac{\lambda+\mu+\theta}{\mu\theta}j + \frac{1}{\alpha}$ for $j \geq 1$.

Theorem 3.4. *In the fully observable $M/M/1$ retrial queue with a single vacation, there exists a pair of threshold $(n_e(1), n_e(2))$, such that a customer who observes the state $(I(t), N(t))$ joins the orbit if $N(t) \leq n_e(I(t)) - 1$ and balks otherwise, where $(n_e(1), n_e(2)) = (\lfloor x_1 \rfloor, \lfloor x_2 \rfloor)$ and x_i $(i = 1,2)$ are the solutions to equations:*

$$U_{fo}(i,j) = R - CT(i,j), \quad i = 1,2. \tag{3.27}$$

Proof. It is obvious that $T(1,j)$ and $T(2,j)$ are increasing in j and consequently, $U_{fo}(1,j)$ and $U_{fo}(2,j)$ are decreasing. A customer enters the system if $U_{fo}(i,j) \geq 0$; otherwise, he chooses balking. Thus, there exist thresholds such that customers' expected utility of joining is negative when the queue length is greater than the corresponding threshold, and the thresholds can be obtained by $U_{fo}(i,j) = 0$, $i = 1,2$. □

Corollary 3.5. *When $R\alpha - C > 0$, $n_e(1)$ and $n_e(2)$ are nonincreasing in λ, but nondecreasing in μ and θ. Furthermore, $n_e(2)$ is nondecreasing with respect to α.*

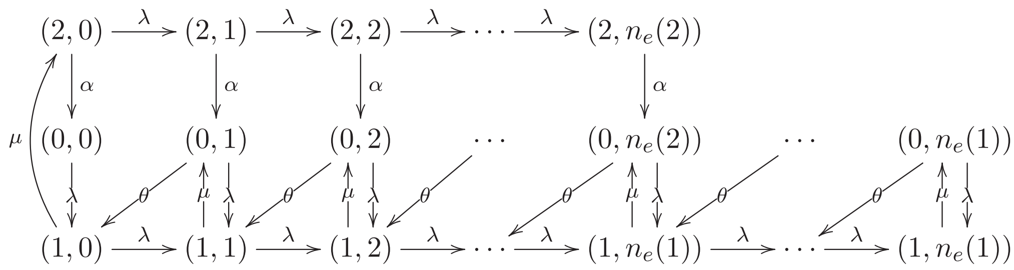

FIGURE 3. The transition rate diagram with $n_e(1) \geq n_e(2)$.

Proof. By $U_{fo}(i,j) = R - CT(i,j)$, $i = 1,2$, we have $x_1 = \frac{\mu\theta R - C}{C(\lambda+\mu+\theta)}$ and $x_2 = \frac{\mu\theta R\alpha - C\mu\theta}{C\alpha(\lambda+\mu+\theta)}$. It is obvious that x_1 and x_2 are decreasing in λ. By the first-order derivatives of x_1 and x_2 with respect to μ, θ, α, it follows that

$$\frac{\partial x_1}{\partial \mu} = \frac{R\theta(\lambda+\theta)+C}{C(\lambda+\mu+\theta)^2}, \tag{3.28}$$

$$\frac{\partial x_1}{\partial \theta} = \frac{R\mu(\lambda+\theta)+C}{\lambda+\mu+\theta}, \tag{3.29}$$

$$\frac{\partial x_2}{\partial \mu} = \frac{\theta(R\alpha - C)(\lambda+\theta)}{C\alpha(\lambda+\mu+\theta)^2}, \tag{3.30}$$

$$\frac{\partial x_2}{\partial \theta} = \frac{\mu(R\alpha - C)(\lambda+\mu)}{C\alpha(\lambda+\mu+\theta)^2}, \tag{3.31}$$

$$\frac{\partial x_2}{\partial \alpha} = \frac{C\mu\theta}{\alpha}\cdot \tag{3.32}$$

Recall that $(n_e(1), n_e(2)) = (\lfloor x_1 \rfloor, \lfloor x_2 \rfloor)$. Thus, when $R\alpha - C > 0$, we come to the conclusions in the lemma. □

This lemma gives the sensitivity analysis of customers' equilibrium threshold policy, which can be explained as follows. When customers' arrival rate λ increases, the system becomes crowded and consequently, customers have less incentive to join. In contrast, when the service rate μ, retrial rate θ and the completion rate of vacation α increase, customers' mean waiting time in the system decreases, attracting more customers to join. Recall that $T(2,j) = \frac{\lambda+\mu+\theta}{\mu\theta}j + \frac{1}{\alpha}$ for $j \geq 1$. Only when $R\alpha - C > 0$, it is possible that $n_e(2) > 0$ and thus the assumption $R\alpha - C > 0$ is natural.

When all customers follow the threshold policy given in Theorem 3.4, we have a Markov chain similar in Figure 1, but the state space is restricted to $S_{fo} = \{(0,n) \mid 0 \leq n \leq \max\{n_e(1), n_e(2)\}\} \cup \{(1,n) \mid 0 \leq n \leq \max\{n_e(1), n_e(2)\}\} \cup \{(2,n) \mid 0 \leq n \leq n_e(2)\}$, which is shown in Figures 3–4. For the stationary analysis, the stationary distribution denoted by $P_{fo}(i,j)$ $((i,j) \in S_{fo})$ can be obtained by balance equations of the system.

When $\mu \geq \alpha$ (*i.e.*, $n_e(1) \geq n_e(2)$), we have the following balance equations

$$\lambda P_{fo}(0,0) = \alpha P_{fo}(2,0), \tag{3.33}$$

$$(\lambda+\theta)P_{fo}(0,i) = \mu P_{fo}(1,i) + \alpha P_{fo}(2,i), \quad i = 1,2,\ldots,n_e(2), \tag{3.34}$$

$$(\lambda+\theta)P_{fo}(0,i) = \mu P_{fo}(1,i), \quad i = n_e(2)+1,\ldots,n_e(1), \tag{3.35}$$

$$(\lambda+\mu)P_{fo}(1,0) = \lambda P_{fo}(0,0) + \theta P_{fo}(0,1), \tag{3.36}$$

$$(\lambda+\mu)P_{fo}(1,i) = \lambda P_{fo}(1,i-1) + \lambda P_{fo}(0,i) + \theta P_{fo}(0,i+1), \quad i = 1,2,\ldots,n_e(1)-1, \tag{3.37}$$

$$\mu P_{fo}(1,n_e(1)) = \lambda P_{fo}(0,n_e(1)) + \lambda P_{fo}(1,n_e(1)-1), \tag{3.38}$$

$$(\lambda+\alpha)P_{fo}(2,0) = \mu P_{fo}(1,0), \tag{3.39}$$

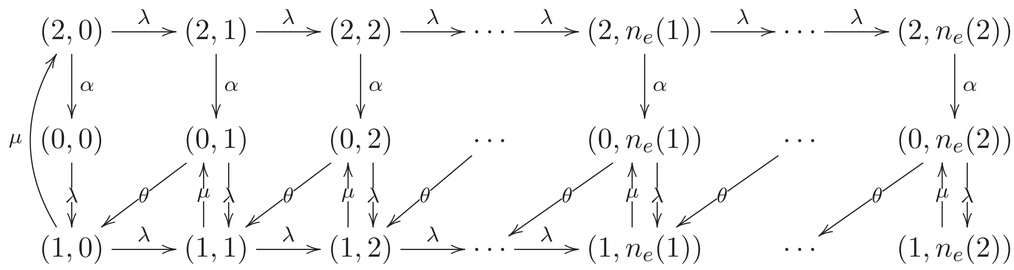

FIGURE 4. The transition rate diagram with $n_e(1) < n_e(2)$.

$$(\lambda+\alpha)P_{fo}(2,i) = \lambda P_{fo}(2,i-1), \quad i = 1,2,\ldots,n_e(2)-1, \tag{3.40}$$
$$\alpha P_{fo}(2,n_e(2)) = \lambda P_{fo}(2,n_e(2)-1). \tag{3.41}$$

By iterating equation (3.40), using (3.33) and (3.41), we have

$$P_{fo}(2,i) = \frac{\lambda}{\alpha}\left(\frac{\lambda}{\lambda+\alpha}\right)^i P_{fo}(0,0), \quad i = 0,1,\ldots,n_e(2)-1, \tag{3.42}$$
$$P_{fo}(2,n_e(2)) = \left(\frac{\lambda}{\alpha}\right)^2\left(\frac{\lambda}{\lambda+\alpha}\right)^{n_e(2)-1} P_{fo}(0,0). \tag{3.43}$$

Substituting equations (3.42)–(3.43) into (3.34) gives that

$$P_{fo}(0,i) = \frac{\mu}{\lambda+\theta}P_{fo}(1,i) + \frac{\lambda}{\lambda+\theta}\left(\frac{\lambda}{\lambda+\alpha}\right)^i P_{fo}(0,0), \quad i = 1,2,\ldots,n_e(2)-1, \tag{3.44}$$

$$P_{fo}(0,n_e(2)) = \frac{\mu}{\lambda+\theta}P_{fo}(1,n_e(2)) + \frac{\alpha}{\lambda+\theta}\left(\frac{\lambda}{\alpha}\right)^2\left(\frac{\lambda}{\lambda+\alpha}\right)^{n_e(2)-1} P_{fo}(0,0), \tag{3.45}$$

and it follows from (3.35) that

$$P_{fo}(0,i) = \frac{\mu}{\lambda+\theta}P_{fo}(1,i), \quad i = n_e(2)+1,\ldots,n_e(1). \tag{3.46}$$

Plugging equations (3.44)–(3.46) into (3.37), we have

$$\frac{\mu\theta}{\lambda+\theta}P_{fo}(1,i+1) + \left(\frac{\lambda\mu}{\lambda+\theta} - \lambda - \mu\right)P_{fo}(1,i) + \lambda P_{fo}(1,i-1)$$
$$= -\left(\frac{\lambda}{\lambda+\alpha}\right)^i\left(\frac{\lambda^2}{\lambda+\theta} + \frac{\lambda^2\theta}{(\lambda+\theta)(\lambda+\alpha)}\right)P_{fo}(0,0), \quad i = 1,2,\ldots,n_e(2)-2, \tag{3.47}$$
$$\frac{\mu\theta}{\lambda+\theta}P_{fo}(1,n_e(2)) + \left(\frac{\lambda\mu}{\lambda+\theta} - \lambda - \mu\right)P_{fo}(1,n_e(2)-1) + \lambda P_{fo}(1,n_e(2)-2)$$
$$= -\left(\frac{\lambda}{\lambda+\alpha}\right)^{n_e(2)-1}\left[\frac{\lambda^2}{\lambda+\theta} + \frac{\alpha\theta}{\lambda+\theta}(\frac{\lambda}{\alpha})^2\right]P_{fo}(0,0), \tag{3.48}$$
$$\frac{\mu\theta}{\lambda+\theta}P_{fo}(1,n_e(2)+1) + \left(\frac{\lambda\mu}{\lambda+\theta} - \lambda - \mu\right)P_{fo}(1,n_e(2)) + \lambda P_{fo}(1,n_e(2)-1)$$

$$= -\frac{\lambda\alpha}{\lambda+\theta}\left(\frac{\lambda}{\alpha}\right)^2\left(\frac{\lambda}{\lambda+\alpha}\right)^{n_e(2)-1} P_{fo}(0,0), \tag{3.49}$$

$$\frac{\mu\theta}{\lambda+\theta}P_{fo}(1,i+1)+\left(\frac{\lambda\mu}{\lambda+\theta}-\lambda-\mu\right)P_{fo}(1,i)+\lambda P_{fo}(1,i-1)=0,$$
$$i=n_e(2)+1,\ldots,n_e(1)-1. \tag{3.50}$$

For equation (3.47), $P_{fo}(1,i)$ $(i=0,1,\ldots,n_e(2)-1)$ are the solutions to the nonhomogeneous linear difference equation with constant coefficients

$$\frac{\mu\theta}{\lambda+\theta}x_{i+1}+\left(\frac{\lambda\mu}{\lambda+\theta}-\lambda-\mu\right)x_i+\lambda x_{i-1}$$
$$=-\left(\frac{\lambda}{\lambda+\alpha}\right)^i\left(\frac{\lambda^2}{\lambda+\theta}+\frac{\lambda^2\theta}{(\lambda+\theta)(\lambda+\alpha)}\right)P_{fo}(0,0), \quad i=1,2,\ldots,n_e(2)-2. \tag{3.51}$$

To solve this equation, we consider the corresponding characteristic equation $\frac{\mu\theta}{\lambda+\theta}x^2+(\frac{\lambda\mu}{\lambda+\theta}-\lambda-\mu)x+\lambda=0$, which has two roots, *i.e.*, 1 and $\frac{\lambda(\lambda+\theta)}{\mu\theta}$. For the homogeneous version of equation (3.51), the general solution to it is $x_i^{\text{hom}}=A_1+B_1(\frac{\lambda(\lambda+\theta)}{\mu\theta})^i$. If we define x_i^{spc} as a special solution to equation (3.51), x_i^{spc} has the form of $C_1(\frac{\lambda}{\lambda+\alpha})^i$ based on the nonhomogeneous part. Substituting $x_i^{\text{spc}}=C_1(\frac{\lambda}{\lambda+\alpha})^i$ into equation (3.51), we obtain

$$C_1=-\frac{[\frac{\lambda^2}{\lambda+\theta}+\frac{\lambda^2\theta}{(\lambda+\alpha)(\lambda+\theta)}]\frac{\lambda}{\lambda+\alpha}}{\frac{\mu\theta}{\lambda+\theta}\cdot(\frac{\lambda}{\lambda+\alpha})^2+(\frac{\lambda\mu}{\lambda+\theta}-\lambda-\mu)\frac{\lambda}{\lambda+\alpha}+\lambda}P_{fo}(0,0). \tag{3.52}$$

So the general solution to equation (3.51) is

$$x_i^{\text{gen}}=A_1+B_1\left(\frac{\lambda(\lambda+\theta)}{\mu\theta}\right)^i+C_1\left(\frac{\lambda}{\lambda+\alpha}\right)^i, \quad i=0,1,\ldots,n_e(2)-1, \tag{3.53}$$

where C_1 is given in equation (3.52). That means

$$P_{fo}(1,i)=A_1+B_1\left(\frac{\lambda(\lambda+\theta)}{\mu\theta}\right)^i+C_1\left(\frac{\lambda}{\lambda+\alpha}\right)^i, \quad i=0,1,\ldots,n_e(2)-1. \tag{3.54}$$

Similarly, we have the solution to equation (3.50), that is,

$$P_{fo}(1,i)=A_2+B_2\left(\frac{\lambda(\lambda+\theta)}{\mu\theta}\right)^i, \quad i=n_e(2),\ldots,n_e(1). \tag{3.55}$$

Thus, by now, we have expressed all stationary probabilities in terms of A_1, B_1, A_2, B_2, $P_{fo}(0,0)$, in relations see equations (3.42)–(3.46) and (3.54)–(3.55). The remaining variables A_1, B_1, A_2, B_2 and $P_{fo}(0,0)$ can be found from (3.36), (3.38), (3.48)–(3.49) and the normalization equation

$$\sum_{i=0}^{n_e(1)}(P_{fo}(0,i)+P_{fo}(1,i))+\sum_{i=0}^{n_e(2)}P_{fo}(2,i)=1. \tag{3.56}$$

Similarly, if $\mu < \alpha$, we have the transition rate diagram illustrated in Figure 4 and the balance equations are followed

$$\lambda P_{fo}(0,0) = \alpha P_{fo}(2,0), \tag{3.57}$$
$$(\lambda+\theta) P_{fo}(0,i) = \mu P_{fo}(1,i) + \alpha P_{fo}(2,i), \quad i=1,2,\ldots,n_e(2), \tag{3.58}$$
$$(\lambda+\mu) P_{fo}(1,0) = \lambda P_{fo}(0,0) + \theta P_{fo}(0,1), \tag{3.59}$$
$$(\lambda+\mu) P_{fo}(1,i) = \lambda P_{fo}(1,i-1) + \lambda P_{fo}(0,i) + \theta P_{fo}(0,i+1), \quad i=1,2,\ldots,n_e(1), \tag{3.60}$$
$$\mu P_{fo}(1,i) = \lambda P_{fo}(0,i) + \theta P_{fo}(0,i+1), \quad i=n_e(1)+1,\ldots,n_e(2)-1, \tag{3.61}$$
$$\mu P_{fo}(1,n_e(2)) = \lambda P_{fo}(0,n_e(2)), \tag{3.62}$$
$$(\lambda+\alpha) P_{fo}(2,0) = \mu P_{fo}(1,0), \tag{3.63}$$
$$(\lambda+\alpha) P_{fo}(2,i) = \lambda P_{fo}(2,i-1), \quad i=1,2,\ldots,n_e(2)-1, \tag{3.64}$$
$$\alpha P_{fo}(2,n_e(2)) = \lambda P_{fo}(2,n_e(2)-1). \tag{3.65}$$

Adopting the same method as in the case with $\mu \geq \alpha$, we derive that

$$P_{fo}(0,i) = \frac{\mu}{\lambda+\theta} P_{fo}(1,i) + \frac{\lambda}{\lambda+\theta}\left(\frac{\lambda}{\lambda+\alpha}\right)^i P_{fo}(0,0), \quad i=1,2,\ldots,n_e(2), \tag{3.66}$$

$$P_{fo}(1,i) = A_3 + B_3\left(\frac{\lambda(\lambda+\theta)}{\mu\theta}\right)^i + C_1\left(\frac{\lambda}{\lambda+\alpha}\right)^i, \quad i=0,1,\ldots,n_e(1)+1, \tag{3.67}$$

$$P_{fo}(1,i) = P_{fo}(1,i-1) - \left(\frac{\lambda}{\lambda+\alpha}\right)^{i-1}\left[\frac{\lambda^2}{\mu\theta} + \frac{\lambda^2}{\mu(\lambda+\alpha)}\right] P_{fo}(0,0), \quad i=n_e(1)+2,\ldots,n_e(2), \tag{3.68}$$

$$P_{fo}(2,i) = \frac{\lambda}{\alpha}\left(\frac{\lambda}{\lambda+\alpha}\right)^i P_{fo}(0,0), \quad i=0,1,\ldots,n_e(2)-1, \tag{3.69}$$

$$P_{fo}(2,n_e(2)) = \left(\frac{\lambda}{\alpha}\right)^2\left(\frac{\lambda}{\lambda+\alpha}\right)^{n_e(2)-1} P_{fo}(0,0). \tag{3.70}$$

Combining equations (3.59), (3.62) and the normalization equation $\sum_{i=0}^{n_e(2)} (P_{fo}(0,i) + P_{fo}(1,i) + P_{fo}(2,i)) = 1$, the variables A_3, B_3 and $P_{fo}(0,0)$ can then be determined.

In line with the definition for the social welfare in the almost unobservable case, we still define it as the total benefit of customers joining the system and denote the social welfare per time unit by SW_{fo}. When all customers follow the equilibrium threshold policy $(n_e(1), n_e(2))$, we have

$$\begin{aligned} SW_{fo} = {} & \lambda R(1 - P_{fo}(1,n_e(1)) - P_{fo}(2,n_e(2))) \\ & -C\Big(\sum_{i=1}^{\max\{n_e(1),n_e(2)\}} iP_{fo}(0,i) + \sum_{i=1}^{\max\{n_e(1),n_e(2)\}} iP_{fo}(1,i) + \sum_{i=1}^{n_e(2)} iP_{fo}(2,i)\Big). \end{aligned} \tag{3.71}$$

4. Numerical examples

In this section, we first investigate the effects of some system parameters on customers' equilibrium behavior in the almost unobservable and fully observable cases. Then when all customers adopt the equilibrium strategies, we compare the two information levels in terms of the social welfare, to indicate whether the social planner should reveal some information to customers to gain more profit of the system.

In Figures 5–6, we observe customers' equilibrium joining probability in the almost unobservable case (*i.e.*, q_e^{au}) and threshold policy in the fully observable case (*i.e.*, $n_e(1)$ and $n_e(2)$) are nonincreasing with respect

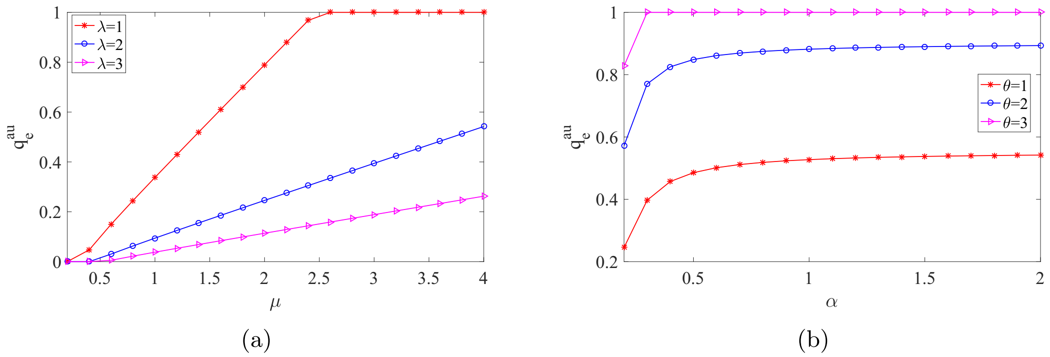

FIGURE 5. Equilibrium joining probability in the almost unobservable case for $\theta = 1$, $\alpha = 2$ (*panel a*); $\lambda = 1$, $\mu = 4$ (*panel b*); assuming $R = 10$, $C = 1$.

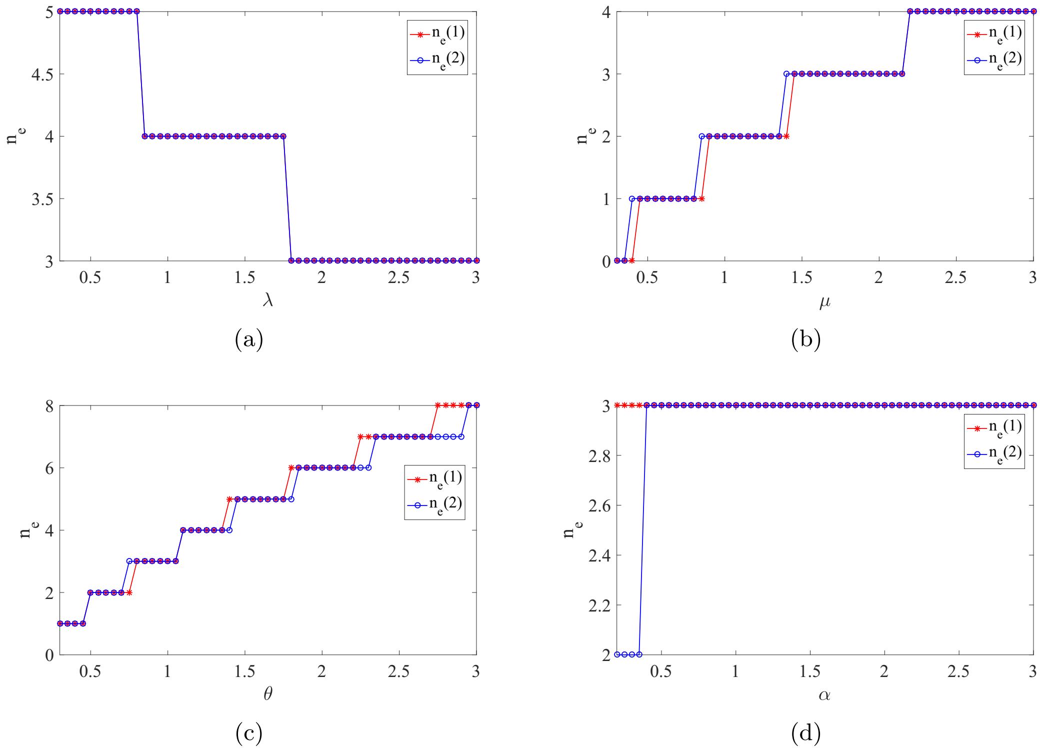

FIGURE 6. Threshold policy in the fully observable case for $\mu = 2$, $\theta = 1$, $\alpha = 2$ (*panel a*); $\lambda = 2$, $\theta = 1$, $\alpha = 2$ (*panel b*); $\lambda = 2$, $\mu = 2$, $\alpha = 2$ (*panel c*); $\lambda = 2$, $\mu = 2$, $\theta = 1$ (*panel d*); assuming $R = 10$, $C = 1$.

to λ, but nondecreasing in μ, θ and α. That is because the system becomes more crowded as λ increases, so that customers enter reluctantly. When μ and θ increase, the server can serve more customers per time unit and customers in the orbit can try to access the server more frequently. In this situation, it is intuitive that an arriving customer prefers to enter. If the completion rate of vacation, *i.e.*, α, increases, the server converts to the normal working state more quickly so as to decrease customers' delay. So more customers choose to join.

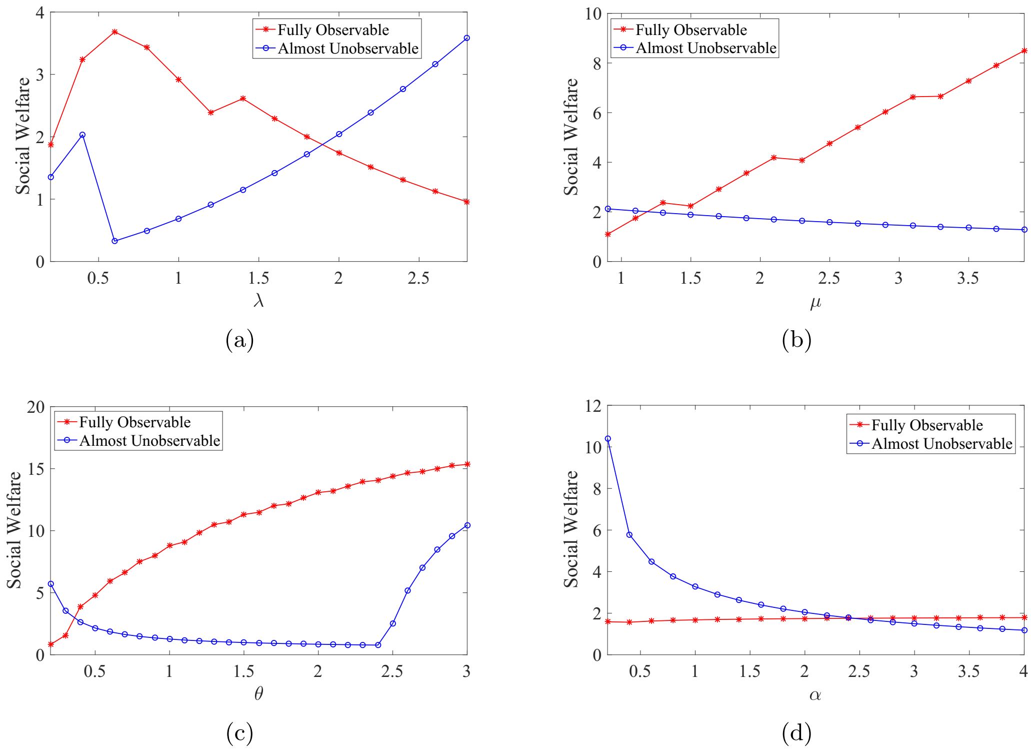

FIGURE 7. Comparison between the equilibrium social welfare in the almost unobservable and fully observable cases for $\mu = 1.1$, $\theta = 1$, $\alpha = 2$ (*panel a*); $\lambda = 2$, $\theta = 1$, $\alpha = 2$ (*panel b*); $\lambda = 2$, $\mu = 4$, $\alpha = 2$ (*panel c*); $\lambda = 2$, $\mu = 1.1$, $\theta = 1$ (*panel d*); assuming $R = 10$, $C = 1$.

With regard to the monotonicity of the social welfare, recall that the social welfare is the sum of entering customers' utilities. Assume customers follow the equilibrium strategy. When customers' potential arrival rate λ increases, the number of customers entering the system per time unit grows (positive effect), but it also pushes the system towards higher states (with more customers in the system). Since customers at the higher states need to wait longer, their expected utility at such states is lower (negative effect). So the monotonicity of the social welfare per time unit with respect to λ is determined by which effect dominates. Similar explanations work for the situation with respect to μ, θ and α. When μ and θ increase, the probability that the system is at higher states decreases (negative effect), while customers' waiting time at these states declines and thus their expected utility grows (positive effect). In addition, under this circumstance, customers have more incentive to enter due to the decreasing waiting time, which benefits the social planner. As α increases, more customers choose to join (positive effect). The probability that the server is on vacation decreases (negative effect), whereas the system is more possible at the states with an idle server (positive effect). Furthermore, customers have higher expected utilities at the states with a server on vacation (positive effect), and their utilities at the states with an idle server stay constant as α increases. Therefore, as λ, μ, θ and α increase, which effect works significantly plays a decisive role for the monotonicity of the social welfare.

When comparing the social welfare of the system under different information levels, from Figure 7, we observe there exist thresholds such that the social welfare is higher in the almost unobservable case when μ, θ, α (or λ) are less (higher) than the corresponding threshold, while the fully observable situation benefits the social planner more when μ, θ, α (or λ) are higher (lower) than the corresponding threshold.

5. Conclusions

In this paper, we analyzed customer strategic behavior in constant retrial queues with a single vacation policy where arriving customers decide whether to join the system or not according to the different information provided for them and a linear reward-cost function. Specifically, we considered two information levels. In the almost unobservable case, only the information whether the server is idle or not is provided for customers, and we obtained customers' equilibrium joining strategy which can be either a pure strategy or a mixed strategy. In the fully observable case, arriving customers are informed of the server's state and the number of customers in the orbit, and the optimal threshold policy was derived under different server's states. When arriving customers adopt equilibrium strategies, we gave the social welfare which is the sum of the utilities of entering customers. Through numerical examples, the two information levels were compared in terms of the social welfare. We found there exist thresholds such that the social welfare is higher in the almost unobservable case when the service rate μ, retrial rate θ, completion rate of the vacation α (or the arrival rate λ) are less (higher) than the corresponding threshold, while the fully observable situation benefits the social planner more when μ, θ, α (or λ) are higher (lower) than the corresponding threshold.

References

[1] W.J. Anderson, Continuous-time Markov chains: an applications-oriented approach. Springer Science & Business Media (2012).

[2] J.R. Artalejo and A. Gómez-Corral, Retrial queueing systems. Springer-Verlag, Heidelberg (2008).

[3] A. Burnetas and A. Economou, Equilibrium customer strategies in a single server Markovian queue with setup times. *Queueing Syst.* **56** (2007) 213–228.

[4] N.H. Do, T. Van Do and A. Melikov, Equilibrium customer behavior in the $M/M/1$ retrial queue with working vacations and a constant retrial rate. To appear in: *Oper. Res.* (2018) 1–20.

[5] A. Economou and S. Kanta, Equilibrium customer strategies and social-profit maximization in the single-server constant retrial queue. *Naval. Res. Logist.* **58** (2011) 107–122.

[6] A. Economou, A. Gómez-Corral and S. Kanta, Optimal balking strategies in single-server queues with general service and vacation times. *Perform. Eval.* **68** (2011) 967–982.

[7] N.M. Edelson and D.K. Hilderbrand, Congestion tolls for poisson queuing processes. *Econometrica* **43** (1975) 81–92.

[8] G.I. Falin and J.G.C. Templeton, Retrial queues. Chapman and Hall, London (1997).

[9] P. Guo and R. Hassin, Strategic behavior and social optimization in Markovian vacation queues. *Oper. Res.* **59** (2011) 986–997.

[10] P. Guo and R. Hassin, Strategic behavior and social optimization in Markovian vacation queues: the case of heterogeneous customers. *Eur. J. Oper. Res.* **222** (2012) 278–286.

[11] P. Guo and Q. Li, Strategic behavior and social optimization in partially-observable Markovian vacation queues. *Oper. Res. Lett.* **41** (2013) 277–284.

[12] R. Hassin, Rational queueing. CRC Press, Boca Raton (2016).

[13] R. Hassin and M. Haviv, To queue or not to queue: equilibrium behavior in queueing systems. Kluwer Academic Publishers, Boston (2003).

[14] W. Liu, Y. Ma and J. Li, Equilibrium threshold strategies in observable queueing systems under single vacation policy. *Appl. Math. Model.* **36** (2012) 6186–6202.

[15] P. Naor, The regulation of queue size by levying tolls. *Econometrica* **37** (1969) 15–24.

[16] W. Sun and N. Tian, Contrast of the equilibrium and socially optimal strategies in a queue with vacations. *J. Comput. Inf. Syst.* **4** (2008) 2167–2172.

[17] W. Sun and S. Li, Equilibrium and optimal behavior of customers in Markovian queues with multiple working vacations. *TOP* **22** (2014) 694–715.

[18] W. Sun, S. Li and Q. Li, Equilibrium balking strategies of customers in Markovian queues with two-stage working vacations. *Appl. Math. Comput.* **248** (2014) 195–214.

[19] W. Sun, S. Li and N. Tian, Equilibrium and optimal balking strategies of customers in unobservable queues with double adaptive working vacations. *Qual. Tech. Quant. Manag.* **14** (2016) 94–113.

[20] N. Tian and Z.G. Zhang, Vacation queueing models: theory and applications. Springer Science & Business Media (2006).

[21] R. Tian, D. Yue and W. Yue, Optimal balking strategies in an $M/G/1$ queueing system with a removable server under N-policy. *J. Ind. Manag. Optim.* **11** (2015) 715–731.

[22] J. Wang and F. Zhang, Strategic joining in $M/M/1$ retrial queues. *Eur. J. Oper. Res.* **240** (2017) 76–87.

[23] J. Wang and F. Zhang, Monopoly pricing in a retrial queue with delayed vacations for local area network applications. *IMA J. Manag. Math.* **27** (2016) 315–334.

[24] F. Wang, J. Wang and F. Zhang, Strategic behavior in the single-server constant retrial queue with individual removal. *Qual. Tech. Quant. Manag.* **12** (2015) 325–342.

[25] J. Wang, X. Zhang and P. Huang, Strategic behavior and social optimization in a constant retrial queue with the N-policy. *Eur. J. Oper. Res.* **256** (2017) 841–849.

[26] Z. Zhang, J. Wang and F. Zhang, Equilibrium customer strategies in the single-server constant retrial queue with breakdowns and repairs. *Math. Probl. Eng.* **2014** (2014) 379572.

[27] F. Zhang, J. Wang and B. Liu, Equilibrium balking strategies in Markovian queues with working vacations. *Appl. Math. Model.* **37** (2013) 8264–8282.

[28] Y. Zhang, J. Wang and F. Wang, Equilibrium pricing strategies in retrial queueing systems with complementary services. *Appl. Math. Model.* **40** (2016) 5775–5792.

[29] Y. Zhang and J. Wang, Equilibrium pricing in an $M/G/1$ retrial queue with reserved idle time and setup time. *Appl. Math. Model.* **49** (2017) 514–530.