

Research Article

Equilibrium Customer Strategies in the Queue with Threshold Policy and Setup Times

Peishu Chen,^{1,2} Wenhui Zhou,¹ and Yongwu Zhou¹

¹*School of Business Administration, South China University of Technology, Guangzhou, Guangdong 510000, China*

²*Department of Mathematics, Chaohu University, Chaohu, Anhui 238000, China*

Correspondence should be addressed to Peishu Chen; cps8@sina.com

Received 10 October 2014; Revised 24 May 2015; Accepted 10 June 2015

Academic Editor: John Gunnar Carlsson

Copyright © 2015 Peishu Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We consider the equilibrium behavior of customers in the $M/M/1$ queue with N policy and setup times. The server is shut down once the system is empty and begins an exponential setup time to start service again when the number of customers in the system accumulates the threshold N ($N \geq 1$). We consider the equilibrium threshold strategies for fully observable case and mixed strategies for fully unobservable case, respectively. We get various performance measures of the system and investigate some numerical examples of system size, social benefit, and expected cost function per unit time for the two different cases under equilibrium customer strategies.

1. Introduction

Due to wide applications for management in service system and modern electronic commerce, there exists an emerging trend to study the behavior of customers in queueing systems. In these models, the customers decide to join or balk in a decentralized way. Their decisions are affected by how much system information can be observed and how other customers make their decisions. This can be viewed as a game among the customers. Researches about the economic analysis of queueing models can go back at least to the pioneering work of Naor [1] who analyzed customers' optimal strategies in an observable Markovian queue with a concise reward-cost structure. Naor's model and results had already been extended by several authors; readers can refer to [2–4].

The concept of an N policy was first introduced by Yadin and Naor [5]. The server is immediately turned on whenever N ($N \geq 1$) or more customers are present in the system and turned off once there are no customers in the system. Once the server is shut down, the server may not operate until N customers are present in the model. Wang and Ke [6] studied the optimal operation of a single removable server in an $M/G/1$ queueing system with either infinite waiting room or finite waiting room by a supplementary

variable technique, where the “removable server” is merely an abbreviation for the system of turning on and turning off the server depending on the number of customers in the queue. Wang et al. [7] further consider an $M/G/1$ queueing system with N policy, where the server is typically subject to breakdowns unpredictably.

Guo and Hassin [8] first consider customers' strategic behavior and social optimization in a single Markovian queue with N policy, in which the server is activated if there are N customers in the system and turned off once there are no customers in the system. They found an interesting conclusion which is that a customer can bring positive externality to others by joining, which means that a new arrival can either keep the server operational for a longer time or stir up an inactive server to come back to work sooner, which leads to promoting other customers' benefits. Guo and Li [9] further investigated customers' strategic behavior and social optimization in $M/M/1$ queues with N policy under two partial information cases.

Queues with setup times have also been investigated by many scholars [6, 10–12]. In queueing systems with setup times, once server is reactivated, a generally random time is required for setup before it can start serving customers. As for the research on the equilibrium customer behavior

in Markovian queues with setup times was first presented by Burnetas and Economou [13]. Subsequently, Economou and Kanta [14] further investigated the Markovian queue that alternates between on and off periods in observable cases. Chen and Zhou [15] studied three different scenarios of balking strategies in the Markovian queue with setup times and server breakdowns.

Frequent setups will inevitably lead to the operating cost being too high, so it is very crucial to determine when the server should start his service in many practical queueing systems. By utilizing N policy, the server does not turn on his service until the number of customers in the system accumulates a threshold N . Selecting appropriate value N can avoid excessively frequent setups. It is more practical to analyze equilibrium strategies for the N policy system with setup times. For example, from an economic perspective, a driver may start driving a shuttle whenever a certain amount of seats is occupied. Another example is a queue-based power-saving scheme to alleviate the energy hole problem [16, 17]. Handling with sensor nodes in wireless sensor networks (WSNs) is actually crucial to understand the optimal operation of the threshold policy for power saving in a WSN. In the “ N policy” scheme, when the number of packets accumulates to N , the sensor node turns on its transmitting function of the radio server and starts transmission process for the waiting packets. The sensor node is closed as soon as all present packets are transmitted. The N policy holds effective for a power-saving mechanism since it reduces the setup power consumption required to switch between a busy mode and an idle mode of a radio server in sensor nodes. The proposed queue-based power-saving technique can be used for prolonging the lifetime of the WSN effectively and economically. To the best of our knowledge, no other scientific paper has dealt with equilibrium customer strategies in the Markovian queue with N policy and setup times. Customers make their decisions based on a nature reward-cost structure, which incorporates their desire for service as well as their unwillingness to wait. We discuss, respectively, the fully observable scenario where customers are informed not only about the server state but also about the exact number of customers in the queue and the fully unobservable scenario where an arriving customer does not observe the exact number of customers waiting for service and the state of server at all.

The structure of this paper is organized as follows. In Section 2, we describe the dynamics of the queueing system. In Section 3, we obtain the equilibrium threshold strategies for the fully observable case and stationary probabilities of the corresponding queue. In Section 4, we derive the stationary probabilities and equilibrium mixed strategies for the fully unobservable case. In Section 5, some numerical results are presented to illustrate the effect of two different information levels and N policy on the customers’ strategies. Finally, Section 6 provides some conclusions.

2. Model Description

We assume that there is a single server queue with infinite capacity where customers arrive according to a Poisson

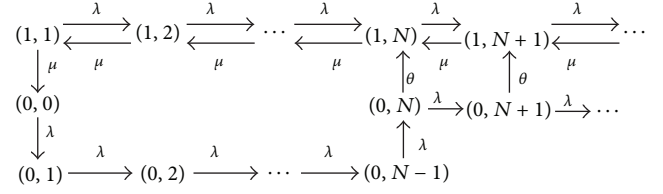


FIGURE 1: Transition rate diagram of the original model.

process with rate λ . The service times are supposed to be exponentially distributed with parameter μ . We further assume that $0 < \lambda < \mu$. The server utilizes the N policy: turn the server on as soon as N ($N \geq 1$) or more customers are present in the system and turn the server off whenever there are no customers in the system. After the server is shut down, the server cannot work until N customers are presented in the queue and a setup process begins for the server to be activated. The time needed to setup is exponentially distributed with rate θ . We further suppose that interarrival times, service times, and setup times are mutually independent.

We denote the state of the system at time t by a random vector $(I(t), N(t))$, where $I(t)$ denotes the state of the server and $N(t)$ denotes the number of customers in the system. More specifically, $(0, n)$, $0 \leq n \leq N-1$, implies that the system is on N policy with n customers in the system; $(1, n)$, $n \geq N$, means that the system is on setup times with n customers in the system; $(2, n)$, $n \geq 1$, corresponds to the server which is busy with n customers in the system. It is clear that the process $\{(I(t), N(t)) : t \geq 0\}$ is a continuous time Markov chain with state space $S = \{(0, 0), (0, 1), \dots, (0, N-1)\} \cup \{(1, N), (1, N+1), \dots\} \cup \{(2, 1), (2, 2), \dots\}$. The transition rate diagram is shown in Figure 1.

In this paper, we suppose that every customer gets a reward of R units for completing his service. Moreover, we suppose that there exists a waiting cost of C units per time unit that is continuously accumulated from the time that the customer joins the system till the time he leaves the system after he finished service.

3. Equilibrium Strategies and Some Performance Measures for the Fully Observable System

We now commit to the fully observable scenario, which assumes that customers observe not only the server state, but also the exact number of customers in the system. We will study the equilibrium threshold strategies and stationary probabilities for the fully observable case as well as some interesting performance measures of this system.

3.1. Equilibrium Threshold Strategies and Stationary Probabilities for the Fully Observable Case. In order to study the general threshold strategy adopted by all customers in the fully observable case, we will consider the mean overall sojourn time of the customer who is confronted with different server state upon arrival firstly. We start with the fully

observable case where customers know the exact state of the system and a customer in state (i, n) has mean sojourn time as follows.

(I) If a new customer finds the system in state $(0, n)$, where $0 \leq n \leq N - 1$, then his expected benefit is

$$R - C \left(\frac{N - (n + 1)}{\lambda} + \frac{1}{\theta} + \frac{n + 1}{\mu} \right). \quad (1)$$

Let $f(x) = R - C((N - (x + 1))/\lambda + 1/\theta + (x + 1)/\mu)$; we have $f'(x) = C(1/\lambda - 1/\mu) > 0$ (since we assume that $0 < \lambda < \mu$ in Section 2). Obviously, $f(n)$ increases by increasing n . That is to say, when the system is on N policy, the least expected benefit is the first customer in the system. For the sake of ensuring a customer finds the system in $(0, 0)$ state and still joins the system, we must have

$$R - C \left(\frac{N - 1}{\lambda} + \frac{1}{\theta} + \frac{1}{\mu} \right) \geq 0, \quad (2)$$

which means that

$$N \leq \left\lfloor 1 + \frac{\lambda R}{C} - \frac{\lambda}{\theta} - \frac{\lambda}{\mu} \right\rfloor. \quad (3)$$

(II) If a new customer finds the system in state $(1, n)$, where $n \geq N$ (this phenomenon means the system is in setup phase), then his expected benefit is

$$R - C \left(\frac{1}{\theta} + \frac{n + 1}{\mu} \right) \geq 0, \quad (4)$$

which means that if the server is currently in a setup state, the upper bound on the number of customers that an incoming customer is willing to stay behind is

$$n_e(1) = \left\lfloor \frac{R\mu}{C} - \frac{\mu}{\theta} \right\rfloor - 1. \quad (5)$$

(III) If a new customer finds the system in state $(2, n)$, then his expected benefit is

$$R - C \left(\frac{n + 1}{\mu} \right) \geq 0, \quad (6)$$

which means that the upper bound on the number of service epochs that a customer is willing to wait is

$$n_e(2) = \left\lfloor \frac{R\mu}{C} \right\rfloor - 1. \quad (7)$$

Remark 1. Because any customer in N policy state must go through setup state and at last in service state, we have $N \leq n_e(1) < n_e(2)$.

Theorem 2. In the fully observable $M/M/1$ queue with setup times and N policy, a customer who observes the server in vacation is motivated to join if $N \leq \lfloor 1 + \lambda R/C - \lambda/\theta - \lambda/\mu \rfloor$ or who observes the server in setup state is motivated to join if $n \leq n_e(1)$ or who observes the server in service is motivated to join if $n \leq n_e(2)$, or balk otherwise.

Proof. From (3), we immediately find that a customer who observes the server in vacation is motivated to join if and only if $N \leq \lfloor 1 + \lambda R/C - \lambda/\theta - \lambda/\mu \rfloor$. From (5), we find that if an arriving customer observes the server in setup state he is motivated to join on the condition that the number of customers in the system is no more than $n_e(1)$, which means that $n \leq n_e(1)$. From (7), we find that if an arriving customer observes the server in service he is motivated to join if and only if $n \leq n_e(2)$. If an arriving customer finds the number of customers in the system exceeds $n_e(2)$, which means that his expected net benefit is negative in any case, then his rational decision is to balk the system. \square

We now turn our attention to the stationary distribution probabilities. Be aware that if all customers follow the threshold strategy in Theorem 2, then the system follows a Markov chain $(I(t), N(t))$ with state space of the system

$$\begin{aligned} S_{ob} = \{ & (0, 0), (0, 1), \dots, (0, N - 1) \} \\ & \cup \{ (1, n) \mid N \leq n \leq n_e(1) + 1 \} \\ & \cup \{ (2, n) \mid 1 \leq n \leq n_e(2) + 1 \}. \end{aligned} \quad (8)$$

The transition diagram is depicted in Figure 2.

The corresponding stationary distribution $(p_{ob}(i, j) \in S_{ob})$ is given by the following theorem.

Theorem 3. In the fully observable $M/M/1$ queue with N policy, setup times, and $\rho \neq \sigma$, in which the customers follow the $(N, n_e(1), n_e(2))$ threshold strategy, the stationary probabilities $(p_{ob}(i, j) \in S_{ob})$ are as follows:

$$p_{ob}(0, n) = p_{ob}(0, 0), \quad n = 1, 2, \dots, N - 1, \quad (9)$$

$$p_{ob}(1, n) = \sigma^{n-N+1} p_{ob}(0, 0), \quad n = N, N + 1, \dots, n_e(1), \quad (10)$$

$$p_{ob}(1, n_e(1) + 1) = \frac{\sigma^{n_e(1)-N+2}}{1 - \sigma} p_{ob}(0, 0), \quad (11)$$

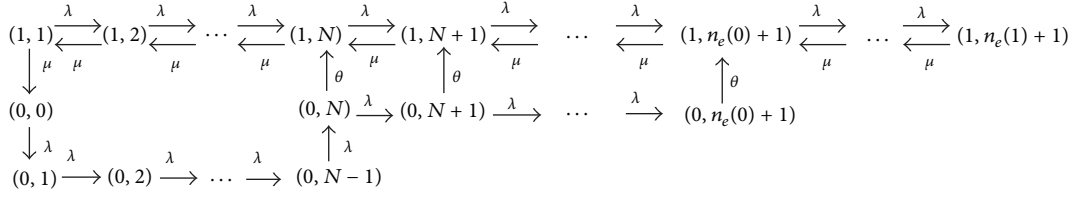
$$p_{ob}(2, n) = \frac{\rho}{1 - \rho} (1 - \rho^n) p_{ob}(0, 0), \quad n = 1, 2, \dots, N, \quad (12)$$

$$\begin{aligned} p_{ob}(2, n) &= \frac{\rho \{ (1 - \rho) \sigma^{n+1-N} - \rho^n [\sigma - \rho + \rho^{1-N} (1 - \sigma)] \} p_{ob}(0, 0)}{(1 - \rho)(\sigma - \rho)} \\ & \quad n = N + 1, N + 2, \dots, n_e(1) + 1, \end{aligned} \quad (13)$$

$$\begin{aligned} p_{ob}(2, n) &= p_{ob}(2, n_e(1) + 1) \rho^{n-n_e(1)-1}, \\ & \quad n = n_e(1) + 2, n_e(1) + 3, \dots, n_e(2) + 1, \end{aligned} \quad (14)$$

$$\begin{aligned} p_{ob} = & \left\{ N + \sum_{n=N}^{n_e(1)} \sigma^{n-N+1} + \frac{\sigma^{n_e(1)-N+2}}{1 - \sigma} + \sum_{n=1}^N \frac{\rho(1 - \rho^n)}{1 - \rho} \right. \\ & + \sum_{n=N+1}^{n_e(1)+1} \frac{\rho \{ (1 - \rho) \sigma^{n+1-N} - \rho^{n+1-N} [1 - \sigma + \rho^{N-1} (\sigma - \rho)] \}}{(1 - \rho)(\sigma - \rho)} \\ & \left. + p_{ob}(1, n_e(1) + 1) \sum_{n=n_e(1)+2}^{n_e(2)+1} \rho^{n-n_e(1)-1} \right\}^{-1}, \end{aligned} \quad (15)$$

where $\rho = \lambda/\mu$ and $\sigma = \lambda/(\lambda + \theta)$.

FIGURE 2: Transition rate diagram for the $(N, n_e(1), n_e(2))$ threshold strategy.

Proof. The stationary distribution of the system is gained from the following balance equations:

$$\lambda p(0, 0) = \mu p(2, 1), \quad (16)$$

$$\lambda p(0, n) = \lambda p(0, n-1), \quad n = 1, 2, \dots, N-1, \quad (17)$$

$$(\lambda + \theta) p(1, N) = \lambda p(0, N-1), \quad (18)$$

$$(\lambda + \theta) p(1, n) = \lambda p(1, n-1), \quad (19)$$

$$n = N, N+1, \dots, n_e(1),$$

$$\lambda p(1, n_e(1)) = \theta p(1, n_e(1) + 1), \quad (20)$$

$$(\lambda + \mu) p(2, 1) = \mu p(2, 2), \quad (21)$$

$$(\lambda + \mu) p(2, n) = \lambda p(2, n-1) + \mu p(2, n+1), \quad (22)$$

$$n = 2, 3, \dots, N-1,$$

$$(\lambda + \mu) p(2, n) = \theta p(1, n) + \lambda p(2, n-1) + \mu p(2, n+1), \quad (23)$$

$$n = N, N+1, \dots, n_e(1) + 1,$$

$$(\lambda + \mu) p(2, n) = \lambda p(2, n-1) + \mu p(2, n+1), \quad (24)$$

$$n = n_e(1) + 2, n_e(1) + 3, \dots, n_e(2) + 1,$$

$$\lambda p(2, n_e(2)) = \mu p(2, n_e(2) + 1), \quad (25)$$

$$\sum_{n=0}^{N-1} p(0, n) + \sum_{n=N}^{n_e(1)+1} p(1, n) + \sum_{n=1}^{n_e(2)+1} p(2, n) = 1. \quad (26)$$

From (17), we easily get (9) in Theorem 3. From (18) and (19), we get (10) in Theorem 3. We further put $p_{\text{ob}}(1, n_e(1))$ in (20) and then we get exact expression of $p_{\text{ob}}(1, n_e(1) + 1)$ in (11). From (16) and (21), we get $p_{\text{ob}}(2, 1)$ and $p_{\text{ob}}(2, 2)$ and put them into (22); we immediately get expression (12) in Theorem 3. We further put $p_{\text{ob}}(1, n_e(1))$ in (20); then we get exact expression of $p_{\text{ob}}(1, n_e(1) + 1)$ in (11). From (16) and (21), we get $p_{\text{ob}}(2, 1)$ and $p_{\text{ob}}(2, 2)$ and then put them into (22); we immediately get expression (12) in Theorem 3. Next, we put

$p_{\text{ob}}(2, N-1)$ and $p_{\text{ob}}(2, N)$ into expression (23); then we have (13) in Theorem 3. From (24) and (25), we have

$$\lambda p(2, n-1) = \mu p(2, n), \quad (27)$$

$$n = n_e(1) + 2, n_e(1) + 3, \dots, n_e(2) + 1.$$

By putting $p_{\text{ob}}(2, n_e(1) + 1)$ into above recursive equation, we have (14). By using the normalizing equation (26), we get the expression of $p_{\text{ob}}(0, 0)$ in Theorem 3 by putting (9)–(14) into (26). \square

Remark 4. Theorem 3 holds for the stationary distribution corresponding to any threshold policy (N, n_1, n_2) , $N \leq n_1 \leq n_2$, and not merely to the individually optimal policy specified by $n_1 = n_e(1)$ and $n_2 = n_e(2)$.

3.2. The System Performance Measures for the Fully Observable Case. We now study an expected cost function per unit time for the $(N, n_e(1), n_e(2))$ threshold strategy system. Our objectives are to analyze performance measures of the fully observable system and determine the optimum value of the important control threshold N by numerical computations.

(1) The Actual Arrival Rate. Because an arrival balks whenever he finds the system at state $(1, n_e(1) + 1)$ or $(2, n_e(2) + 1)$, then the system's effective arrival rate is given by

$$\lambda(n_e(1), n_e(2)) = \lambda [1 - p_{\text{ob}}(1, n_e(1) + 1) - p_{\text{ob}}(2, n_e(2) + 1)]. \quad (28)$$

(2) The Expected Number of Customers. Let L_{ob} denote the expected number of customers in the $(N, n_e(1), n_e(2))$ threshold strategy system. Then we obtain

$$L_{\text{ob}} = \frac{N(N-1)p_{\text{ob}}(0, 0)}{2} + \sum_{n=N}^{n_e(1)+1} n p_{\text{ob}}(1, n) + \sum_{n=1}^{n_e(2)+1} n p_{\text{ob}}(2, n). \quad (29)$$

(3) *The Probability of the Server in Different State.* We further give the probability that the server in idle, setup, and busy steady state is as follows:

$$\begin{aligned} P(i=0) &= Np_{\text{ob}}(0,0); \\ P(i=1) &= \sum_{n=N}^{n_e(1)+1} p_{\text{ob}}(1,n); \\ P(i=2) &= \sum_{n=1}^{n_e(2)+1} p_{\text{ob}}(2,n). \end{aligned} \quad (30)$$

(4) *The Expected Busy Cycle.* Let the expected length of the idle period be denoted by $E(i=0)$. Applying the memoryless property of the Poisson process, the length of the idle period is the sum of N exponential random variables, each having a mean of $1/\lambda$. Hence, the expected length of the idle period is presented by $E(i=0) = N/\lambda$.

The busy cycle for the fully observable case is denoted by $E(C_N)$, which is the length of time from the beginning of completing all the customers in system to the beginning of the next completing of all the customers in system, which included vacation period, setup period, and busy period. From the relationship

$$\frac{E(i=0)}{E(C_N)} = P(i=0) = Np_{\text{ob}}(0,0), \quad (31)$$

we can get the expected length of the busy cycle $E(C_N) = 1/[\lambda p_{\text{ob}}(0,0)]$.

(5) *The Social Benefit per Time.* The social benefit per time unit as soon as all customers follow the equilibrium policy $(N, n_e(0), n_e(1))$ equals

$$\begin{aligned} S_{\text{ob}} &= \lambda R [1 - p_{\text{ob}}(1, n_e(1) + 1) - p_{\text{ob}}(2, n_e(2) + 1)] \\ &\quad - C \left[\sum_{n=0}^{N-1} n p_{\text{ob}}(0, n) + \sum_{n=0}^{n_e(1)+1} n p_{\text{ob}}(1, n) \right. \\ &\quad \left. + \sum_{n=1}^{n_e(2)+1} n p_{\text{ob}}(2, n) \right]. \end{aligned} \quad (32)$$

(6) *The Expected Cost Function.* We next define the following:

C_h = holding cost per unit time for every customer present in the system;

C_s = cost incurred per busy cycle;

C_0 = cost per unit time for keeping the server off;

C_1 = setup cost per unit time for turning the server on;

C_2 = cost per unit time for keeping the server on and in operation;

C_3 = a fixed cost for every lost customer.

By using the definitions of all costs listed above, the expected cost function per unit time per customer is given by

$$\begin{aligned} F_{\text{ob}}(N) &= C_h L_N + \frac{C_s}{E(C_N)} + C_0 P(i=0) + C_1 P(i=1) \\ &\quad + C_2 P(i=2) \\ &\quad + \lambda C_3 [p_{\text{ob}}(1, n_e(1) + 1) + p_{\text{ob}}(2, n_e(2) + 1)]. \end{aligned} \quad (33)$$

4. Equilibrium Strategies and Some Performance Measures for the Fully Unobservable System

We turn our attention to the fully unobservable case, where the customers do not observe the state and system size at all when they join the system. We will study the equilibrium arrival rate and stationary probabilities for the fully unobservable case as well as some interesting performance measures of this system.

4.1. Equilibrium Mixed Strategies and Some Stationary Probabilities for the Fully Unobservable Case. We will prove that there exists a mixed equilibrium strategy. An arriving customer joins the fully unobservable system with a certain probability q ; then the effective arrival rate is $\lambda_0 = \lambda q$. If all customers adopt the same mixed strategy, then the system follows a Markov chain which is similar to that described in Figure 1. The state space S_{un} for the fully unobservable case is identical to the original state space S . In addition to substituting λ_0 for λ , the state transfer diagram is exactly the same as Figure 1.

Let $(p_{\text{un}}(i, n) : (i, n) \in S_{\text{un}})$ be the stationary distribution of the corresponding system. Let $p_{\text{un}}(0, 0)$ be the probability that the system is empty.

Theorem 5. *In the fully unobservable M/M/1 queue with N policy, setup times, and $\sigma_0 \neq \rho_0$, in which the customers enter the system with probability q , the stationary probabilities of the system are given by*

$$p_{\text{un}}(0, n) = p_{\text{un}}(0, 0), \quad n = 1, 2, \dots, N-1, \quad (34)$$

$$p_{\text{un}}(1, n) = \sigma_0^{n-N+1} p_{\text{un}}(0, 0), \quad n = N, N+1, \dots, \quad (35)$$

$$p_{\text{un}}(2, n) = \frac{\rho_0}{1 - \rho_0} (1 - \rho_0^n) p_{\text{un}}(0, 0), \quad n = 1, 2, \dots, N, \quad (36)$$

$$p_{un}(2, n) = \frac{\rho_0 \left\{ (1 - \rho_0) \sigma_0^{n+1-N} - \rho_0^{n+1-N} [1 - \sigma_0 + \rho_0^{N-1} (\sigma_0 - \rho_0)] \right\} p_{un}(0, 0)}{(1 - \rho_0)(\sigma_0 - \rho_0)} \quad n = N + 1, N + 2, \dots, \quad (37)$$

$$p_{un}(0, 0) = \frac{(1 - \rho_0)(1 - \sigma_0)}{N - (N - 1)\sigma_0}, \quad (38)$$

where $\lambda_0 = \lambda q$, $\rho_0 = \lambda_0/\mu$, and $\sigma_0 = \lambda_0/(\lambda_0 + \theta)$.

Proof. The balance equations are presented below:

$$\lambda_0 p(0, 0) = \mu p(2, 1), \quad (39)$$

$$\lambda_0 p(0, n) = \lambda_0 p(0, n - 1), \quad n = 1, 2, \dots, N - 1, \quad (40)$$

$$(\lambda_0 + \theta) p(1, N) = \lambda_0 p(0, N - 1), \quad (41)$$

$$(\lambda_0 + \theta) p(1, n) = \lambda_0 p(1, n - 1), \quad n = N + 1, N + 2, \dots, \quad (42)$$

$$(\lambda_0 + \mu) p(2, 1) = \mu p(2, 2), \quad (43)$$

$$(\lambda_0 + \mu) p(2, n) = \lambda_0 p(2, n - 1) + \mu p(2, n + 1), \quad n = 2, 3, \dots, N - 1, \quad (44)$$

$$(\lambda_0 + \mu) p(2, n) = \theta p(1, n) + \lambda_0 p(2, n - 1) + \mu p(2, n + 1), \quad n = N, N + 1, \dots, \quad (45)$$

$$\sum_{n=0}^{N-1} p(0, n) + \sum_{n=N}^{\infty} p(1, n) + \sum_{n=1}^{\infty} p(2, n) = 1. \quad (46)$$

From (40), we easily get (34) in Theorem 5. From (41) and (42), we can get (35) in Theorem 5. From (39) and (43), we get $p_{un}(2, 1)$ and $p_{un}(2, 2)$. By substituting them into (44), we immediately get expression (36) in Theorem 5. Next, we put $p_{un}(2, N - 1)$ and $p_{un}(2, N)$ into expression (45); we can get (37) in Theorem 5. By putting (34)–(37) into (46), we obtain the expression of $p_{un}(0, 0)$ at last. \square

Theorem 6. In the fully unobservable $M/M/1$ queue with N policy and setup times, where the customers enter the system with probability q , when a customer enters the system in state i ($i = 0, 1, 2$), his expected sojourn time W_i is given below:

$$\begin{aligned} W_0 &= \frac{1}{\theta} + \frac{1}{2} \left(\frac{N+1}{\mu} + \frac{N-1}{\lambda_0} \right), \\ W_1 &= \frac{N+1}{\mu} + \frac{1}{\theta} \left(1 + \frac{\lambda_0}{\mu} \right), \\ W_2 &= \frac{N}{2\mu} + \frac{1}{\mu - \lambda_0} + \frac{1}{\mu(1 - \sigma_0)} \\ &\quad - \frac{N}{2\mu[N - (N - 1)\sigma_0]}, \end{aligned} \quad (47)$$

and the average sojourn time of an arriving customer is

$$W(\lambda_0) = \frac{1}{\mu - \lambda_0} + \frac{1}{\theta} + \frac{(N - 1)N\theta}{2\lambda(\theta N + \lambda_0)}. \quad (48)$$

Proof. Let $p_{un}(i)$ be the probability that a customer arrives during state i , where $i = 0, 1, 2$. From Theorem 5, we can easily get the steady state probabilities of the server:

$$\begin{aligned} P_{un}(i = 0) &= NP_{un}(0, 0), \\ P_{un}(i = 1) &= \sum_{n=N}^{\infty} p_{un}(0, n) = \frac{\sigma_0 p_{un}(0, 0)}{1 - \sigma_0}, \\ P_{un}(i = 2) &= \sum_{n=1}^N p_{un}(1, n) + \sum_{n=N+1}^{\infty} p_{un}(1, n) = \rho_0. \end{aligned} \quad (49)$$

Let $p(n | i)$ be the stationary probability that the queue length is n conditional on observing the servers status which is i , where $i = 0, 1, 2$ and $n = 0, 1, \dots$. The conditional probabilities can be written by

$$\begin{aligned} p(n | 0) &= \frac{p_{un}(0, n)}{p_{un}(i = 0)} = \frac{1}{N}, \quad n = 1, 2, \dots, N - 1, \\ p(n + N | 1) &= \frac{p_{un}(1, n + N)}{p_{un}(i = 1)} = (1 - \sigma_0) \sigma_0^n, \quad n = 0, 1, \dots, \\ p(n | 2) &= \frac{p_{un}(2, n)}{p_{un}(i = 2)} = \frac{(1 - \sigma_0)(1 - \rho_0^n)}{N - (N - 1)\sigma_0}, \quad n = 1, 2, \dots, N, \quad (50) \\ p(n + N | 2) &= \frac{p_{un}(2, n + N)}{p_{un}(i = 2)} \\ &= \frac{(1 - \sigma_0) \left\{ (1 - \rho_0) \sigma_0^{n+1} - \rho_0^{n+1} [1 - \sigma_0 + \rho_0^{N-1} (\sigma_0 - \rho_0)] \right\}}{(\sigma_0 - \rho_0) [N - (N - 1)\sigma_0]}, \quad n = 1, 2, \dots \end{aligned}$$

Note that if the queue length is n and the server is busy, the expected sojourn time of the new arriving customer is denoted as $w(n | 2)$, which is the expected service time of the $n + 1$ customer in the system, that is, $w(n | 2) = (n + 1)/\mu$, which means the expected sojourn time upon seeing a busy server is

$$w(n | 2) = \frac{n + 1}{\mu}, \quad n = 1, 2, \dots \quad (51)$$

The expected sojourn time upon seeing a setup server is

$$w(n | 1) = \frac{1}{\theta} + \frac{n + N + 1}{\mu}, \quad n = 0, 1, \dots \quad (52)$$

The expected sojourn time upon seeing the server on N policy is

$$w(n|0) = \frac{N - (n+1)}{\lambda_0} + \frac{1}{\theta} + \frac{n+1}{\mu}, \quad (53)$$

$$n = 0, 1, \dots, N-1.$$

Then the expected sojourn time W_i ($i = 0, 1, 2$) is given as follows:

$$\begin{aligned} W_0 &= \sum_{n=0}^{N-1} w(n|0) p(n|0) \\ &= \frac{1}{\theta} + \frac{1}{2} \left(\frac{N+1}{\mu} + \frac{N-1}{\lambda_0} \right), \\ W_1 &= \sum_{n=0}^{\infty} w(n+N|1) p(n+N|1) \\ &= \frac{N+1}{\mu} + \frac{1}{\theta} \left(1 + \frac{\lambda_0}{\mu} \right), \\ W_2 &= \sum_{n=1}^{\infty} w(n|2) p(n|2) \\ &= \sum_{n=1}^N \frac{n+1}{\mu} \frac{(1-\sigma_0)(1-\rho_0^n)}{N - (N-1)\sigma_0} \\ &\quad + \sum_{n=1}^{\infty} \frac{n+N+1}{\mu} p(n+N|2) \\ &= \frac{N}{2\mu} + \frac{1}{\mu - \lambda_0} + \frac{1}{\mu(1-\sigma_0)} \\ &\quad - \frac{N}{2\mu[N - (N-1)\sigma_0]}. \end{aligned} \quad (54)$$

Then the average sojourn time of an arriving customer is

$$\begin{aligned} W(\lambda_0) &= W_0 P(i=0) + W_1 P(i=1) + W_2 P(i=2) \\ &= \frac{1}{\mu - \lambda_0} + \frac{1}{\theta} + \frac{(N-1)N\theta}{2\lambda_0(\theta N + \lambda_0)}. \end{aligned} \quad (55)$$

□

Next, according to the value of N , we divide N into two cases as follows.

Case 1. When $N = 1$, the system becomes the fully unobservable $M/M/1$ queue with setup times and $\lambda < \mu$. The average sojourn time reduces to

$$W(\lambda_0) = \frac{1}{\theta} + \frac{1}{\mu - \lambda_0}. \quad (56)$$

We consider a tagged customer at his arrival instant, if he decides to enter his expected net benefit which is

$$R - C \left(\frac{1}{\theta} + \frac{1}{\mu - \lambda_0} \right), \quad \text{where } \lambda_0 = \lambda q. \quad (57)$$

When $R \in (C/\mu + C/\theta, C/(\mu - \lambda) + C/\theta)$, we find that $R - C(1/\theta + 1/(\mu - \lambda q)) = 0$ has a unique root in $(0, 1)$. When $R > C/(\mu - \lambda) + C/\theta$, we have $R - C(1/\theta + 1/(\mu - \lambda q)) > 0$ for any $q \in [0, 1]$. The best response for an arriving customer is to join the system. Then we have the following proposition.

Proposition 7. *In the fully unobservable $M/M/1$ queue with setup times and $\lambda < \mu$, there exists a unique mixed equilibrium strategy “enter with probability q_e ,” where the vector q_e is given by*

$$q_e = \begin{cases} \frac{1}{\lambda} \left(\mu - \frac{C}{R - C/\theta} \right), & \text{if } R \in \left(\frac{C}{\mu} + \frac{C}{\theta}, \frac{C}{\mu - \lambda} + \frac{C}{\theta} \right); \\ 1, & \text{if } R \in \left(\frac{C}{\mu - \lambda} + \frac{C}{\theta}, +\infty \right). \end{cases} \quad (58)$$

This conclusion is consistent with Theorem 5 in [13].

Case 2. When $N \geq 2$, we have

$$\begin{aligned} W''(\lambda_0) &= (N-1) \left[\frac{1}{\lambda_0^3} - \frac{1}{(\theta N + \lambda_0)^3} + \frac{2}{(\mu - \lambda_0)^3} \right] > 0, \quad (59) \\ \lambda_0 &\in (0, \mu), \end{aligned}$$

which means that $W(\lambda_0)$ is strictly convex in λ_0 and $\lim_{\lambda_0 \rightarrow 0^+} W(\lambda_0) = \lim_{\lambda_0 \rightarrow \mu^-} W(\lambda_0) = +\infty$, and then there is a unique minimum point $\tilde{\lambda} \in (0, \mu)$, so that $W(\tilde{\lambda})$ is the minimum sojourn time and $\tilde{\lambda}$ satisfies the following equation:

$$W'(\lambda_0) = \frac{N-1}{2(\theta N + \lambda_0)^2} - \frac{N-1}{2\lambda_0^2} + \frac{1}{(\mu - \lambda_0)^2} = 0. \quad (60)$$

It is clear that when $R < CW(\tilde{\lambda})$, there is no positive equilibrium arrival rate; when $R = CW(\tilde{\lambda})$, there exists one positive equilibrium arrival rate $\tilde{\lambda} \in (0, \mu)$ iff $\tilde{\lambda} < \lambda$; when $R > CW(\tilde{\lambda})$, that is, $R \in (CW(\tilde{\lambda}), +\infty)$, the equation

$$R = C \left[\frac{1}{\mu - \lambda_0} + \frac{1}{\theta} + \frac{(N-1)N\theta}{2\lambda_0(\theta N + \lambda_0)} \right] \quad (61)$$

has two real roots $\lambda_1 \in (0, \tilde{\lambda})$ and $\lambda_2 \in (\tilde{\lambda}, \mu)$. We analyze the strictly convex $W(\lambda_0)$ carefully and then find that the equilibrium with λ_1 is unstable, for, with any small increase of the arrival rate, the expected waiting time decreases and more customers will arrive in the system; this will further increase λ_1 . The equilibrium with λ_2 and λ is stable; that is, if there is a small perturbation to them, the system will converge back to them. We denote the stable equilibrium arrival rate as λ^* ; then we have the following proposition.

Proposition 8. (a) *If $R < CW(\tilde{\lambda})$, there is no positive equilibrium arrival rate.* (b) *If $R = CW(\tilde{\lambda})$, there is one positive*

TABLE 1: Numerical examples for $\mu = 0.42$ and $\lambda_2 < \lambda$.

$u = 0.42, R = 30, \theta = 0.2, C = 1, \lambda = 0.4$										
N	1	2	3	4	5	6	7	8	9	10
λ_1	$q_e = 0.95$	0.0211146	0.0418168	0.0626556	0.0837833	0.105319	0.127408	0.150264	0.174238	0.2
λ_2		0.378885	0.377214	0.375112	0.372582	0.369567	0.365949	0.361532	0.355973	0.34861
λ^*	0.38	0.378885	0.377214	0.375112	0.372582	0.369567	0.365949	0.361532	0.355973	0.34861

TABLE 2: Numerical examples for $\mu = 0.52$ and $\lambda_1 < \lambda < \lambda_2$.

$u = 0.52, R = 30, \theta = 0.2, C = 1, \lambda = 0.4$										
N	1	2	3	4	5	6	7	8	9	10
λ_1	$q_e = 1$	0.0206732	0.0408608	0.0610673	0.0814019	0.101923	0.122683	0.14374	0.165172	0.187081
λ_2		0.479226	0.478046	0.476574	0.474841	0.472846	0.470565	0.467952	0.464943	0.461439
λ^*	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4

TABLE 3: Numerical examples for $\mu = 1$ and $\lambda_1 < \lambda < \lambda_2$.

$u = 1, R = 30, \theta = 0.2, C = 1, \lambda = 0.4$										
N	1	2	3	4	5	6	7	8	9	10
λ_1	$q_e = 1$	0.0391792	0.0583989	0.0583989	0.0776041	0.096822	0.116065	0.135342	0.154658	0.174019
λ_2		0.95934	0.958828	0.958828	0.958219	0.957533	0.95678	0.955963	0.955084	0.954143
λ^*	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4

equilibrium arrival rate $\tilde{\lambda} \in (0, \mu)$ iff $\tilde{\lambda} < \lambda$. (c) If $R > CW(\tilde{\lambda})$, there exists positive equilibrium arrival rate

$$\lambda_e = \begin{cases} \lambda_1 \text{ or } \lambda_2, & \text{if } \lambda_2 < \lambda; \\ \lambda_1 \text{ or } \lambda, & \text{if } \lambda_1 < \lambda < \lambda_2; \\ \text{nonexistent}, & \text{if } \lambda < \lambda_1, \end{cases} \quad (62)$$

and the stable positive equilibrium arrival rate

$$\lambda^* = \begin{cases} \lambda_2, & \text{if } \lambda_2 < \lambda; \\ \lambda, & \text{if } \lambda_1 < \lambda < \lambda_2; \\ \text{nonexistent}, & \text{if } \lambda < \lambda_1, \end{cases} \quad (63)$$

where $\tilde{\lambda}$ is the positive solution of λ_0 in (60) and $\lambda_1 \in (0, \tilde{\lambda})$ and $\lambda_2 \in (\tilde{\lambda}, \mu)$ are the positive solutions of λ_0 in (61).

According to Proposition 7 and case (c) of Proposition 8, we give concrete numerical forms as shown in Tables 1, 2, and 3.

4.2. The System Performance Measures for the Fully Unobservable Case. Our objective is to give some interesting performance measures of the fully unobservable system and

determine the optimum value of the control threshold N by numerical computations.

(1) *The Expected Number of Customers.* Let L_{un} denote the expected number of customers in the (N, q) equilibrium strategy. Then we obtain

$$L_{\text{un}} = \frac{N(N-1)p_{\text{un}}(0,0)}{2} + \sum_{n=N}^{+\infty} np_{\text{un}}(1,n) + \sum_{n=1}^{+\infty} np_{\text{un}}(2,n). \quad (64)$$

(2) *The Probability of the Server in Idle, Setup, and Busy Steady State as Follows.* Consider

$$\begin{aligned} P(i=0) &= Np_{\text{un}}(0,0); \\ P(i=1) &= \sum_{n=N}^{+\infty} p_{\text{un}}(1,n); \\ P(i=2) &= \sum_{n=1}^{+\infty} p_{\text{un}}(2,n). \end{aligned} \quad (65)$$

(3) *The Expected Busy Cycle.* Similar to the fully observable case, the expected length of the idle period is given by $E(i=0) = N/\lambda_0$ and the expected length of the busy cycle $E(C_N) = 1/[\lambda_0 p_{\text{un}}(0,0)]$.

(4) *The Social Benefit per Time.* We assume the stable equilibrium arrival rate $\lambda_0 = \lambda^*$ and all customers follow

the equilibrium strategy (N, q) ; then the social benefit per time unit when all customers follow the equilibrium equals

$$\begin{aligned}
 S_{un} &= \lambda_0 [P(i=0)B(0) + P(i=1)B(1) + P(i=2) \\
 &\quad \cdot B(2)] = \lambda_0 N p_{un}(0,0) \left\{ R - C \left[\frac{1}{\theta} \right. \right. \\
 &\quad \left. \left. + \frac{1}{2} \left(\frac{N+1}{\mu} + \frac{N-1}{\lambda_0} \right) \right] \right\} + \lambda_0 \frac{\sigma_0 p_{un}(0,0)}{(1-\sigma_0)} \left\{ R \right. \\
 &\quad \left. - C \left[\frac{N+1}{\mu} + \frac{1}{\theta} \left(1 + \frac{\lambda_0}{\mu} \right) \right] \right\} + \lambda_0 \rho_0 \left\{ R \right. \\
 &\quad \left. - C \left(\frac{N}{2\mu} + \frac{1}{\mu - \lambda_0} + \frac{1}{\mu(1-\sigma_0)} \right. \right. \\
 &\quad \left. \left. - \frac{N}{2\mu[N - (N-1)\sigma_0]} \right) \right\}. \quad (66)
 \end{aligned}$$

(5) *The Expected Cost Function.* By using the definition of each cost listed before, the expected cost function per unit time per customer is given by

$$\begin{aligned}
 F_{un}(N) &= C_h L_N + \frac{C_s}{E(C_N)} + C_0 P(i=0) \\
 &\quad + C_1 P(i=1) + C_2 P(i=2) + (\lambda - \lambda_0) C_3. \quad (67)
 \end{aligned}$$

5. Numerical Examples

In this section, we obtain some numerical experiments to show the different effects of the fully observable and unobservable information systems. We let $R = 30$, $C = 1$, $\lambda = 0.4$, and $\theta = 0.2$ in all the figures. For the fully observable case, from inequality (3), we have $N \leq \lfloor 1 + \lambda R / C - \lambda / \theta - \lambda / \mu \rfloor = 10$ for any $\mu > 0.4$. We vary the values of positive integer N from 1 to 10 in Figures 3, 4, and 5.

In Figure 3, we select $\mu = 0.42, 0.52, 1$ and vary the values of positive integer N from 1 to 10. We find that, for the fully observable case, the mean customers L_{ob} are an increasing function of N and a decreasing function of service rate μ . For the fully unobservable case, when $\mu = 0.42$, from Table 1, we find that the stable arrival rate λ^* is reducing by increasing N , which means more customers are balking the system, so the system size is reducing by increasing N ; when $\mu = 0.52$ and 1, we find that the stable arrival rate λ^* is a constant value and does not change by varying the values of N from 1 to 10; the system size which is an increasing function of N on account of more customers is blocked. When μ and N are fixed, we find, as for $\mu = 0.42$ and 0.52, the system size of fully observable L_{ob} is smaller than corresponding fully unobservable L_{un} . When $\mu = 1$, L_{ob} and L_{un} are almost equal. It is clear that when customers know the system size and server state, they may be more rational to decide whether to enter the system, which brings more effective regulation to the system size. When the service rate is very larger than arrival rate and the service reward is generous, which means that customers get fast service, good reward and very few customers balk the system, then such good advantages inevitably lead to the

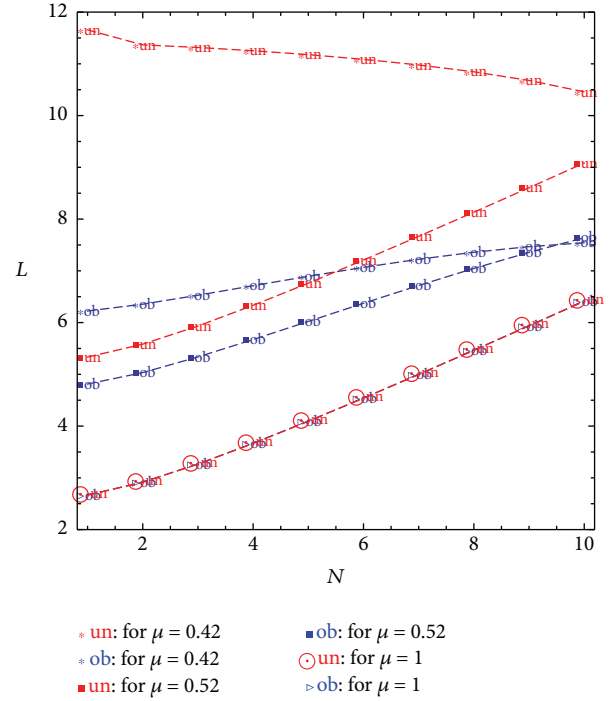


FIGURE 3: System size L versus N for different μ .

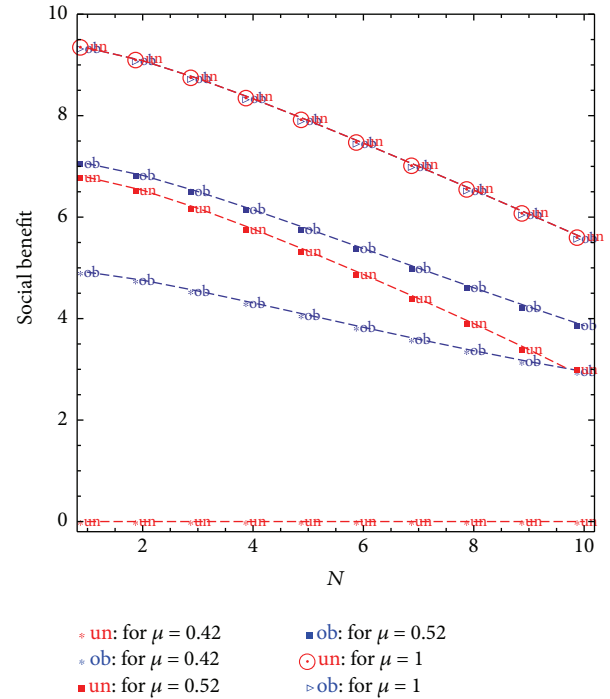


FIGURE 4: Social benefit S versus N for different μ .

disclosures of information (the server state and system size) having a little effect on the system size.

In Figure 4, we find that, along with the increase of N , the social benefits S_{ob} and S_{un} are on the decline. S_{ob} is bigger than S_{un} when $\mu = 0.42, 0.52$ and select the same N ; S_{ob} is almost equal to S_{un} when $\mu = 1$ and select the same N .

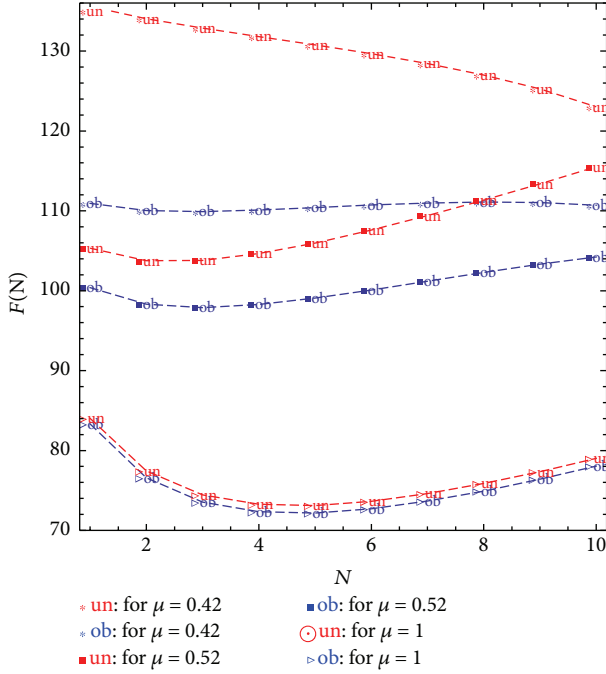


FIGURE 5: Cost function $F(N)$ versus N for different μ .

For unobservable case, when $\mu = 0.42$, from Table 1, we know that the equilibrium arrival strategy taken by all the customers is the cost of sojourn time which equals the service reward, which leads to $S_{un} = 0$.

In Figure 5, we still select $\mu = 0.42, 0.52, 1$ and vary the values of positive integer N from 1 to 10. We further fix the following cost parameters: $C_h = 5$, $C_s = 200$, $C_0 = 10$, $C_1 = 50$, $C_2 = 80$, and $C_3 = 30$. We find that when $\mu = 0.42, 0.52$ and selecting the same N , the value of $F_{ob}(N)$ is always bigger than $F_{un}(N)$; when $\mu = 1$ and selecting the same N , $F_{ob}(N)$ is almost equal to $F_{un}(N)$. If we select appropriate N , we can reduce the expected cost for all the fully observable and unobservable cases.

6. Conclusions

In this paper, we studied the equilibrium behavior of customers in $M/M/1$ queue with N policy and server setup times. We provide fully observable and fully unobservable scenarios with respect to system size and system state provided to arriving customers and derived the equilibrium balking strategies for each case. We mainly study how the N policy affects system size, social benefit, and system cost under the condition of equilibrium strategies taken by fully observable and fully unobservable arriving customers.

Conflict of Interests

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

Acknowledgments

This research is partially supported by CPSF (2015M572327), NSFC (71131003, 71371075, and 71271089), and NSFA (KJ2013B162, KJ2013A194, and KJ2014A174). It is also supported by Chaohu University Scientific Research Fund. The authors would like to sincerely thank the anonymous referees and editors for their valuable comments and suggestions which are very helpful for them to improve the presentation of the paper.

References

- [1] P. Naor, "The regulation of queue size by levying tolls," *Econometrica*, vol. 37, no. 1, pp. 15–24, 1969.
- [2] A. Economou and S. Kanta, "Equilibrium customer strategies and social-profit maximization in the single-server constant retrial queue," *Naval Research Logistics*, vol. 58, no. 2, pp. 107–122, 2011.
- [3] R. Hassin and M. Haviv, *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*, Kluwer Academic, Boston, Mass, USA, 2003.
- [4] H. Chen and M. Frank, "Monopoly pricing when customers queue," *IIE Transactions*, vol. 36, no. 6, pp. 569–581, 2004.
- [5] M. Yadin and P. Naor, "Queueing systems with a removable service station," *Journal of the Operational Research Society*, vol. 14, no. 4, pp. 393–405, 1963.
- [6] K.-H. Wang and J.-C. Ke, "A recursive method to the optimal control of an $M/G/1$ queueing system with finite capacity and infinite capacity," *Applied Mathematical Modelling*, vol. 24, no. 12, pp. 899–914, 2000.
- [7] K.-H. Wang, T.-Y. Wang, and W. L. Pearn, "Optimal control of the N policy $M/G/1$ queueing system with server breakdowns and general startup times," *Applied Mathematical Modelling*, vol. 31, no. 10, pp. 2199–2212, 2007.
- [8] P. Guo and R. Hassin, "Strategic behavior and social optimization in Markovian vacation queues," *Operations Research*, vol. 59, no. 4, pp. 986–997, 2011.
- [9] P. Guo and Q. Li, "Strategic behavior and social optimization in partially-observable Markovian vacation queues," *Operations Research Letters*, vol. 41, no. 3, pp. 277–284, 2013.
- [10] G. Choudhury, "An $M^X/G/1$ queueing system with a setup period and a vacation period," *Queueing Systems*, vol. 36, pp. 23–38, 2000.
- [11] W. Bischof, "Analysis of $M/G/1$ -queues with setup times and vacations under six different service disciplines," *Queueing Systems*, vol. 39, no. 4, pp. 265–301, 2001.
- [12] A. Allahverdi, C. T. Ng, T. C. E. Cheng, and M. Y. Kovalyov, "A survey of scheduling problems with setup times or costs," *European Journal of Operational Research*, vol. 187, no. 3, pp. 985–1032, 2008.
- [13] A. Burnetas and A. Economou, "Equilibrium customer strategies in a single server Markovian queue with setup times," *Queueing Systems*, vol. 56, no. 3–4, pp. 213–228, 2007.
- [14] A. Economou and S. Kanta, "Equilibrium balking strategies in the observable single-server queue with breakdowns and repairs," *Operations Research Letters*, vol. 36, no. 6, pp. 696–699, 2008.
- [15] P. Chen and Y. Zhou, "Equilibrium balking strategies in the single server queue with setup times and breakdowns," *Operational Research*, vol. 15, no. 2, pp. 213–231, 2015.

- [16] F.-C. Jiang, D.-C. Huang, C.-T. Yang, and F.-Y. Leu, "Lifetime elongation for wireless sensor network using queue-based approaches," *The Journal of Supercomputing*, vol. 59, no. 3, pp. 1312–1335, 2012.
- [17] D. H. Lee and W. S. Yang, "The N -policy of a discrete time $Geo/G/1$ queue with disasters and its application to wireless sensor networks," *Applied Mathematical Modelling*, vol. 37, no. 23, pp. 9722–9731, 2013.

