POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT POMS

Vol. 26, No. 9, September 2017, pp. 1632–1645
ISSN 1059-1478 | EISSN 1937-5956 | 17 | 2609 | 1632

DOI 10.1111/poms.12706

Pricing in a Transportation Station with Strategic Customers

Athanasia Manou, Pelin G. Canbolat, Fikri Karaesmen
Department of Industrial Engineering, Koç University, Sariyer, Istanbul, 34450, Turkey,
amanou@ku.edu.tr, pcanbolat@ku.edu.tr, fkaraesmen@ku.edu.tr

We consider a transportation station, where customers arrive according to a Poisson process, observe the delay information and the fee imposed by the administrator and decide whether to use the facility or not. A transportation facility visits the station according to a renewal process and serves all present customers at each visit. We assume that every customer maximizes her individual expected utility and the administrator is a profit maximizer. We model this situation as a two-stage game among the customers and the administrator, where customer strategies depend on the level of delay information provided by the administrator. We consider three cases distinguished by the level of delay information: observable (the exact waiting time is announced), unobservable (no information is provided) and partially observable (the number of waiting customers is announced). In each case, we explore how the customer reward for service, the unit waiting cost, and the intervisit time distribution parameters affect the customer behavior and the fee imposed by the administrator. We then compare the three cases and show that the customers almost always prefer to know their exact waiting times whereas the administrator prefers to provide either no information or the exact waiting time depending on system parameters.

Key words: strategic customers; Nash equilibrium; pricing; stochastic transportation system; delay information
History: Received: January 2016; Accepted: February 2017 by Michael Pinedo, after 1 revision.

1. Introduction

We consider a pricing problem for a service system with strategic customers that are served in batches by a large facility such as a bus or a train. Arriving customers decide whether to join the system and wait for the next processing instance, or to balk and seek other outside options. While the main motivation for the model comes from transportation systems where vehicles visit a transportation station to pick up all waiting passengers, the model may also be appropriate for other types of batch processing.

There is a rich literature that investigates strategic customer behavior in service systems that are modeled by queues. This literature has yielded many useful insights on the design of service systems with rational customers. The analysis of these systems combines the stochastic processes modeling the queue with a game-theoretic approach modeling the customer behavior. This analysis becomes very involved and complicated as the underlying queueing model moves away from the Markovian assumptions. For instance, it is relatively recent that a queueing system with general service times and strategic customers that observe the queue length was completely analyzed in this setting (Kerner 2011). On the other hand, for better understanding the system, it is insightful to consider non-exponential models, where one may explore the effects of distributional information, in particular the variability of the processing time and its impacts on customer behavior. This is especially important because empirical work on public transportation systems reports that users are extremely sensitive to the frequency and the punctuality of the service (Stradling et al. 2007).

We assume that the administrator of the system is a profit maximizer that decides the service fee to charge each customer to maximize the station revenue. In this setting, the analysis is streamlined across three different information structures. In the observable case, arriving customers are provided the exact waiting time. In the unobservable system, customers do not have any information other than the statistical and economic parameters of the system. Finally, in the partially observable case, customers are not provided the exact waiting time, but do observe the number of customers waiting in the station. It turns out that the information structure has a significant impact on the service fee, the customer benefit and the administrator revenue. While the information infrastructure or the availability of information is not a short-term operational decision, it is insightful to understand its effects on systems with a longer-term design perspective. Comparing these three levels of information, we found out that customers never prefer the unobservable case, since in this case, the administrator sets the fee equal to the service surplus (i.e., the difference between the service reward and the expected waiting cost) and so each customer has zero expected benefit in the equilibrium. Thus, depending on the cost-reward structure and the intervisit time distribution, customers prefer either the observable or the partially observable case. On the other hand, the administrator revenue in the unobservable case is always greater than in the partially observable case (provided that the intervisit times have a strictly decreasing mean residual life distribution), since both the fee and the proportion of joining customers are higher in the former. Consequently, depending on system parameters, the administrator prefers either the observable or the unobservable case. In sum, the administrator discloses either full information or no information at all whereas customers prefer having full or partial information rather than not having any.

To gain insights into the effects of information, we consider a relatively stylized model of a transportation station which is visited by a large vehicle with generally distributed random intervisit times. In this model, customers arrive at the station randomly (according to a Poisson process) and decide whether to join and wait or to balk based on the information available and their cost-reward structure. A vehicle visits the station at random times (according to a renewal process) and picks up all the customers that are in the station, so the vehicle capacity is assumed to be infinite. This assumption reduces the system to a stochastic clearing system and differentiates it from traditional queueing systems. Such a model was first investigated by Manou et al. (2014), who established an efficient method to analyze the key performance measures with general intervisit times (at varying degrees of complexity depending on the information structure). Their results allow us to embed and explore a service-pricing problem within this model. In addition to the effects of information, we analyze the effects of intervisit time variability on the equilibrium fee, the customer utility and the administrator revenue. Two interesting and counterintuitive insights emerge from this analysis in the observable case. First, an increase in the variance of intervisit time can cause an increase in the equilibrium fee. This happens when the variance increases in a way that is favorable for customers, that is, by increasing the probability of small waiting times and not so much the probability of large waiting times. Then customers become more willing to join the system, and the administrator takes advantage of this by increasing the fee. Secondly, customers can benefit from an increase in the intervisit time variance. This happens when the administrator reduces the fee to eliminate the negative effects of increasing variance.

The rest of the study is organized as follows. Section 2 presents a literature review. Section 3 introduces the model and the notation used in the study. Sections 4–6 are devoted to the analysis of the observable case, the unobservable case and the partially observable case, respectively. Section 7 compares the results obtained for each case through analytical results and numerical examples. Section 8 concludes with some directions for future research. All proofs are included in Appendix S1.

2. Literature Review

Our work is related to three research streams: queues with strategic customers, pricing of services, and the effect of information. First, our model belongs to the class of queues with strategic customers. The study of queueing systems under a game-theoretic perspective was initiated by Naor (1969) who studied the strategic behavior of customers in M/M/1 queues assuming that an arriving customer observes the number of customers in the system and then decides whether to join or balk. Subsequently, Edelson and Hildebrand (1975) complemented this study by considering the same queueing system but assuming that customers make their decisions without being informed about the number of customers in the system. Since then, there has been a growing number of papers that studied the strategic behavior of customers in variants of M/M/1 queues. The books of Hassin and Haviv (2003), Stidham (2009) and Hassin (2016) present the main approaches and several results in the area of the economic analysis of queueing systems.

The majority of studies that consider the customer behavior in queues assume a Markovian framework. The departure from the Markovian assumption makes the analysis nontrivial while considerably enriching the insights. Relaxing the Markovian assumption makes the system more realistic and allows us to explore the effect of service variability on the behavior of customers. Economou et al. (2011) studied the behavior of customers in a vacation queue with general service and vacation times assuming that arriving customers do not observe the queue length. The extension of such an analysis to queueing systems with general service times where arriving customers observe the queue length requires calculating the expected remaining service times at arrival instants conditional on the queue length. Kerner (2008, 2011) was the first who studied M/G/1 queues with strategic customers that observe the queue length before making their decision. Kerner used the supplementary variable technique to calculate the expected conditional remaining service times. Later, Zhang et al. (2013) studied the behavior of customers in a queue with general service and setup times. They assumed that arriving customers observe the queue length and used the same technique with Kerner. Manou et al. (2014) studied the strategic behavior of customers that arrive at a transportation station that is modeled as a clearing system, observe or do not observe the number of customers in the station and decide whether to join or balk. They proposed a new probabilistic technique to calculate the performance measures of interest when the intervisit times of the transportation facility are generally distributed. Their model does not assume that the facility has infinite capacity, but their technique applies to our model which assumes a facility with infinite capacity. Incidentally, the infinite-capacity assumption differentiates our model from traditional queues, since due to this assumption, customers do not actually queue.

Our work also relates to pricing of services. This stream is rich. For instance, Li (1988), Low (1974) and Cil et al. (2011) provided optimal pricing policies for firms that are modeled as queueing systems. In a service system with strategic customers the fee has a dual role: It helps to control the queue length and provides a revenue stream. The administrator of such a queueing system should carefully balance these consequences. Naor (1969), Yechiali (1971) and Knudsen (1972) studied several queueing systems with strategic customers that observe the queue length at their arrival instants and proved that the profit maximizing fee is greater than the socially optimal fee in these systems. Edelson and Hildebrand (1975) also studied M/M/1 queues and proved that if the arriving customers do not observe the queue length and the cost-reward structure is linear, the profit maximizing fee equals the socially optimal fee. Chen and Frank (2004) studied the pricing problem in the same model but assuming that the cost-reward function is nonlinear. In this case, the profit maximizing fee is not equal to the socially optimal fee. They also explored the effects of system parameters on the equilibrium fee. Chen and Frank (2001) and Yildirim and Hasenbein (2010) studied the state-dependent pricing problem in queues with strategic customers that observe the queue length. Zhou et al. (2013) determined the optimal uniform pricing strategy for a service system with two classes of customers.

Finally, this study is also related to the stream of studies that explore the effect of information on the performance of a system with strategic customers. Hassin (1986) assumed that the revenue maximizing server may suppress information about the queue length and proved that it is not always socially optimal to prevent this. Also, he proved that it is never socially optimal to encourage suppression when the server prefers to reveal the queue length. Guo and Zipkin (2007, 2008, 2009) studied three queueing systems with strategic customers. For each model, they analyzed the behavior of customers under several levels of information and determined which level of information is preferable for the customers, the service provider and the society depending on the system parameters.

As we mentioned above, there are plenty of studies that considered the pricing problem in Markovian systems with strategic customers. To the best of our knowledge, this is the first work that studies the pricing problem in a non-Markovian system. This is a key contribution of our study as we are able to explore the effect of intervisit time variability not only on the customer behavior but also on the equilibrium fee and the administrator revenue. This study also contributes to the research stream that explores the effect of information. Although Guo and Zipkin (2008) explored the effect of information in non-Markovian systems, this study is different in that it explores the effect of information in a system with a profit maximizing administrator.

3. The Model

We consider a transportation station with infinite waiting space, where customers arrive according to a Poisson process at rate $\lambda > 0$. A transportation facility with infinite capacity visits the station according to a renewal process. The intervisit times (i.e., the times between the successive visits) of the transportation facility are assumed to be independent and identically distributed. In the sequel, we denote the intervisit time by X and assume that it is a continuous random variable on $[0, \infty)$ with distribution function F and Laplace transform $\mathcal{L}$. Discrete random variables can also be handled with some technical changes in the proofs. We also assume that X has finite moments. At the visit epochs of the transportation facility, all customers are served instantaneously and removed from the station, so the underlying process is a stochastic clearing process.

The transportation system is managed by an administrator, who initially sets a service fee $\tau \geq 0$ for using the transportation facility. After observing this fee and the information provided to them, arriving customers decide whether to use the transportation facility or not. A customer who chooses not to use the facility earns no reward and incurs no cost whereas a customer who chooses to use it pays the service fee τ, incurs a waiting cost at rate $c > 0$ per time unit spent in the station, and earns a reward $r > 0$ upon service completion. All costs and rewards are assumed to be in utility units. Each customer makes the decision of joining the system or balking so as to maximize her individual expected utility, and the administrator wishes to maximize the expected revenue of the station per time unit. This yields a two-stage game between customers and the administrator. In the first stage of the game, the administrator selects the fee and in the second stage, arriving customers decide whether to join or to balk after observing the fee and the information provided by the administrator. The second stage is indeed a simultaneous-move game between customers.

In the two-stage game between the administrator and the customers, an administrator strategy is a nonnegative fee, and a customer strategy is a function that assigns a joining probability to each possible value of the fee and of the information provided to the customer. Since customers are assumed to be homogeneous, the analysis in the paper restricts attention to symmetric customer strategies; in other words, it assumes that customers behave the same way when they face the same fee and the same information. An *equilibrium* (more precisely, a symmetric subgame perfect Nash equilibrium) consists of a fee and a customer strategy such that:

(I) For any given fee, the customer strategy maximizes the customer's expected utility assuming all other customers adopt this same strategy. Equivalently, for any given fee, customer joining probabilities prescribed by the customer strategy form a (symmetric) Nash equilibrium of the simultaneous-move game among customers.
(II) The fee maximizes the administrator's expected revenue per time unit given the customer strategy.

The formal description of customer strategies and of the equilibrium is given in Appendix S1 to keep the exposition simple.

4. Observable Case

This section analyzes the situation where every arriving customer is informed of her exact waiting time. In this case, given the fee $\tau \geq 0$, a customer whose waiting time is $w \geq 0$ experiences a utility of $r - \tau - cw$ if she joins and a utility of zero if she balks. Consequently, a customer prefers to join if $r - cw > \tau$, to balk if $r - cw < \tau$, and is indifferent between joining and balking if $r - cw = \tau$. The administrator's problem to maximize the expected revenue then reduces to a convex optimization problem as shown in the proof of Theorem 1 in Appendix S1.

THEOREM 1. *In any equilibrium of the observable stochastic clearing system with generally distributed intervisit times, the administrator sets the fee τ^o that is the unique nonnegative solution of the equation*

$$\int_0^{\frac{r-\tau}{c}} [1 - F(x)]\mathrm{d}x = \frac{\tau}{c}\left[1 - F\left(\frac{r-\tau}{c}\right)\right], \qquad (1)$$

and a customer whose waiting time is $w \geq 0$ joins with probability

$$q(w) = \begin{cases} 1 & \text{if } 0 \leq w < \frac{r-\tau^o}{c}, \\ p & \text{if } w = \frac{r-\tau^o}{c}, \\ 0 & \text{if } w > \frac{r-\tau^o}{c}, \end{cases} \qquad (2)$$

where $p \in [0, 1]$ is arbitrary. The expected total customer utility per time unit is

$$\pi_C^o = \frac{\lambda}{E[X]} \int_0^{\frac{r-\tau^o}{c}} (r - \tau^o - cx)[1 - F(x)]\mathrm{d}x, \qquad (3)$$

the expected administrator revenue per time unit is

$$\pi_A^o = \frac{\lambda\tau^o}{E[X]} \int_0^{\frac{r-\tau^o}{c}} [1 - F(x)]\mathrm{d}x = \frac{\lambda(\tau^o)^2}{cE[X]}\left[1 - F\left(\frac{r-\tau^o}{c}\right)\right], \qquad (4)$$

and the expected social utility per time unit is

$$\pi_S^o = \frac{\lambda}{E[X]} \int_0^{\frac{r-\tau^o}{c}} (r - cx)[1 - F(x)]\mathrm{d}x. \qquad (5)$$

We note that, in the observable case, there are multiple equilibria. They differ only in the joining probability of the customer who observes waiting time equal to $(r - \tau^o)/c$. This does not affect the expected utilities Equations (3)–(5).

The following corollary provides a lower and an upper bound on the equilibrium fee when customers are provided the exact waiting time at the time of their arrival.

COROLLARY 1. *In the observable system, the equilibrium service fee τ^o satisfies $r/2 \leq \tau^o < r$.*

The lower bound on the equilibrium fee is in contrast with the unobservable system, where the fee can be arbitrarily small, as will be shown in section 5. This can be explained by considering the customers that arrive to the station just before the visit of the facility. In the observable system, these customers observe very small waiting times and are willing to join even if the fee is large, so the administrator can charge a relatively high fee and still be sure that some customers will join. In the unobservable case, however, high fees are not justified even for these customers when the expected waiting time is large, since in this case, customers do not learn their exact waiting times and decide based on their expected waiting times.

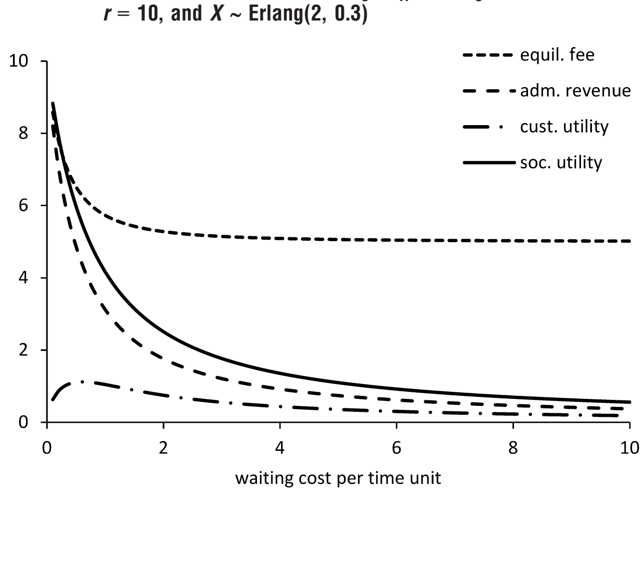
The following corollary states the effect of changing the service reward on the equilibrium.

COROLLARY 2. *In the equilibrium of the observable system, the service fee τ^o, the difference $r - \tau^o$, the expected total customer utility π_C^o, the expected administrator revenue π_A^o, and the expected social utility π_S^o are increasing in r.*

The insights provided by Corollary 2 are intuitive. The administrator takes advantage of the increasing customer satisfaction represented as an increase in the reward, and increases the fee. This increase of the fee could discourage the customers from joining, but it does not. The increase of the service reward overweighs the increase of the fee, since $r - \tau^o$ is increasing in r. Hence, as r increases, the upper bound on the waiting time for which customers join with probability 1 increases, so customers can tolerate waiting more, which yields a higher joining rate and greater expected utilities for customers and the administrator.

One might expect that the reverse of the results stated in Corollary 2 hold for waiting cost, but they do not all hold in general. In particular, the expected utility of customers can increase as a result of an increase in the waiting cost, when the latter causes a big reduction in the fee. Figure 1 illustrates the effect of the waiting cost on the equilibrium fee, the expected total customer utility, the expected administrator revenue, and the expected social utility. In this example, as the waiting cost increases, the service fee, the expected administrator revenue, and the expected social utility decrease. We observed this monotonicity in all our numerical experiments, but we were not able to prove it for the general problem. On the other hand, the expected total customer utility depicted in Figure 1 is not monotone in the waiting cost. Specifically, it is first increasing, and then decreasing. This suggests that for small waiting costs, a slight increase of the waiting cost causes a sharp decrease of the fee and this

Figure 1 **Fee τ^o, Expected Utilities π_C^o, π_A^o and π_S^o vs. c for $\lambda = 1$, $r = 10$, and $X \sim$ Erlang(2, 0.3)**



combination benefits customers. However, for larger waiting costs, the increase of the cost harms customers, as the reduction of the fee is relatively small.

We next explore the effects of changing the variance for a fixed expected intervisit time in the observable system. To do this, we consider two intervisit times X and Y with distribution F_X and F_Y respectively and assume $E[X] = E[Y]$. The effect of the variance on the fee in the observable case can be either way. The relation between the service fees τ_X^o with intervisit time X and τ_Y^o with intervisit time Y is obtained by using the concavity of the expected administrator revenue $\pi_{A,X}^o(\tau)$ in $\tau \geq 0$ (which is shown in the proof of Theorem 1) and the first-order condition. The resulting relation is that τ_X^o is less than (equal to, greater than) τ_Y^o with Y if

$$\int_0^{\frac{r-\tau_Y^o}{c}} [1 - F_X(x)]\mathrm{d}x < \frac{\tau_Y^o}{c}\left[1 - F_X\left(\frac{r - \tau_Y^o}{c}\right)\right]$$
$$(=, \ >, \text{respectively}).$$

The effect of increasing the variance on the expected administrator revenue when the intervisit times are assumed to be ordered under convex ordering is stated in the following corollary. The intervisit time X is less than Y under convex ordering if $E[g(X)] \leq E[g(Y)]$ for any convex function g. When this is the case, the expected values of X and Y coincide, and the variance of X is less than or equal to the variance of Y.

COROLLARY 3. *In the observable system, if X is less than Y under convex ordering,*

(*i*) *The expected administrator revenue with intervisit time X is greater than or equal to the expected revenue with Y.*
(*ii*) *Under the additional assumption that $\tau_X^o \leq \tau_Y^o$, the expected total customer utility with intervisit time X is greater than or equal to the expected total customer utility with Y.*

Intuitively, one would expect that customers are more willing to pay a higher fee as the variance of the intervisit time decreases when the expected intervisit time is preserved. A reduction in variance can be interpreted as an improvement in service reliability when the expected value is preserved and consequently, the service provider can reasonably charge a higher fee and possibly earn a higher revenue. Corollary 3 establishes that the latter is true in the observable case under the convex-ordering assumption; however, as discussed before the corollary, an improvement in variance does not always result in increasing fees (even under the convex-ordering assumption). As Example 1 illustrates, in some cases, decreasing the variance of the intervisit time will decrease the fee. One possible explanation for this phenomenon is that a reduction in variance can eliminate favorable deviations as well unfavorable ones. In other words, the decrease in the variance of the intervisit time may be due to decreasing the probability of small waiting times and not so much the probability of large waiting times. This, in turn, can force the administrator to lower the fee in order to attract customers. On the other hand, when the variance decreases in a favorable way by reducing the probability of large waiting times, arriving customers are given smaller waiting times more often and join the system. This encourages the administrator to increase the fee.

Unlike the expected administrator revenue, the expected total customer utility can increase or decrease as a result of a reduction in the variance of the intervisit time. Specifically, if the variance of the intervisit time and the equilibrium fee decrease, the expected total customer utility increases, as stated in Corollary 3. On the other hand, if the variance of the intervisit time decreases and the equilibrium fee increases, the expected total customer utility may increase or decrease, as illustrated in Example 1.

Example 1. (Erlang intervisit times). Let $\lambda = c = 1$ and r = 10. Suppose the intervisit time X has Erlang distribution with parameters $n \geq 1$ and $\mu > 0$. Tables 1–3 and Figure 2 exhibit the changes in the service fee, the throughput (or the effective arrival rate $\lambda_e = \pi_A^o/\tau^o$), the expected administrator revenue, the expected total customer utility, and the expected social utility as the distribution parameters vary. In each table, the expected intervisit time is set to a constant, so as n increases, the intervisit time decreases under convex ordering, e.g., the intervisit time with $n = 1$ and $\mu = 0.2$ is greater than the one with $n = 2$ and $\mu = 0.4$ under convex ordering. In Figure 2, the coefficient of variation is fixed (equivalently n is fixed), and as μ increases, the expected intervisit time decreases.

Table 1 Fee τ^o, Throughput λ_e Expected Utilities π_A^o, π_C^o and π_S^o for $\lambda = 1$, $c = 1$, $r = 10$, and $X \sim$ Erlang(n, μ) with $E[X] = 5$

n	μ	Fee	Throughput	Administrator revenue	Customer utility	Social utility
1	0.2	6.0400	0.5471	3.3043	0.9936	4.2979
2	0.4	6.0190	0.6346	3.8196	1.1083	4.9279
3	0.6	5.9850	0.6787	4.0623	1.1654	5.2277
4	0.8	5.9530	0.7070	4.2089	1.2039	5.4128
5	1.0	5.9240	0.7274	4.3090	1.2326	5.5416

Table 2 Fee τ^o, Throughput λ_e Expected Utilities π_A^o, π_C^o and π_S^o for $\lambda = 1$, $c = 1$, $r = 10$, and $X \sim$ Erlang(n, μ) with $E[X] = 4$

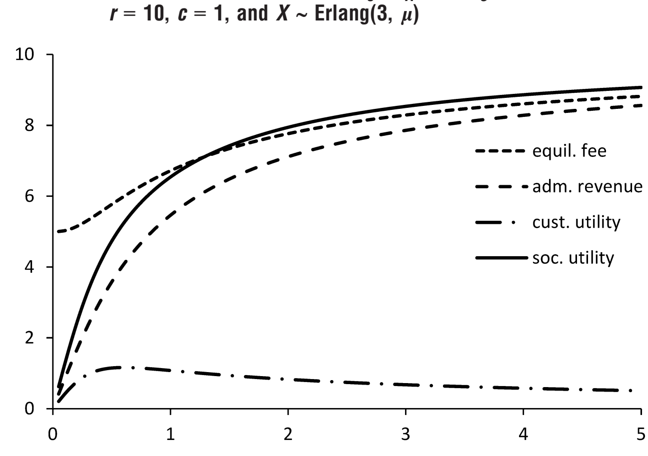
n	μ	Fee	Throughput	Administrator revenue	Customer utility	Social utility
1	0.25	6.2400	0.6094	3.8025	1.0441	4.8466
2	0.5	6.2820	0.6993	4.3932	1.1214	5.5146
4	1	6.2860	0.7716	4.8503	1.1605	6.0108
8	2	6.2650	0.8279	5.1868	1.1816	6.3684
16	4	6.2300	0.8713	5.4281	1.1965	6.6246

Table 3 Fee τ^o, Throughput λ_e Expected Utilities π_A^o, π_C^o and π_S^o for $\lambda = 1$, $c = 1$, $r = 10$, and $X \sim$ Erlang(n, μ) with $E[X] = 2$

n	μ	Fee	Throughput	Administrator revenue	Customer utility	Social utility
1	0.5	6.9934	0.7776	5.4381	1.0683	6.5064
2	1	7.2270	0.8509	6.1495	1.0026	7.1521
4	2	7.4190	0.9011	6.6851	0.8909	7.5760
6	3	7.5130	0.9221	6.9274	0.8199	7.7473
8	4	7.5710	0.9342	7.0726	0.7674	7.8400

- As stated in Corollary 1, the fee is always between $r/2 = 5$ and $r = 10$.
- As variance decreases, the fee decreases in Table 1, first increases then decreases in Table 2, and increases in Table 3.
- Consistently with Corollary 3, in each table, the expected administrator revenue increases as variance decreases (equivalently, as n increases).
- In Table 1 and in the last three rows of Table 2, the fee decreases and the expected total customer utility increases as variance decreases. This illustrates Corollary 3.
- In the first three rows of Table 2, the fee and the expected total customer utility increase as variance decreases, whereas, in Table 3, the fee increases and the expected total customer utility decreases as variance decreases.
- In all tables, the throughput and the social utility increase as variance decreases.
- In Figure 2, the coefficient of variation is fixed (equivalently n is fixed). As the expected intervisit time decreases (equivalently, as μ increases), the fee, the expected administrator revenue, and the social utility increase. The

Figure 2 Fee τ^o, Expected Utilities π_C^o, π_A^o and π_S^o vs. μ for $\lambda = 1$, $r = 10$, $c = 1$, and $X \sim$ Erlang(3, μ)



expected total customer utility appears to be first increasing then decreasing. These suggest that, in the observable case, higher service quality may harm the customers through increased fees.

5. Unobservable Case

This section considers the case where arriving customers do not receive any information about the state of the system (i.e., they do not observe the number of customers in the station and the remaining time until the next visit of the transportation facility). In this case, each arriving customer makes the decision to join or to balk based on her expected waiting time.

Since arrivals form a Poisson process, we can use PASTA property to derive the expected waiting time of an arriving customer, which is basically the expected remaining time until the next visit of the transportation facility at her arrival instant. By PASTA property, the remaining time as observed by a Poisson arrival has the same distribution as the remaining time at an arbitrary instant. The remaining time W at an arbitrary instant has density function $(1 - F(w))/E[X]$, so the expected waiting time of an arriving customer is $E[W] = E[X^2]/(2E[X])$ and the expected utility of joining is $r - \tau - cE[X^2]/(2E[X])$. To exclude the trivial case where customers have no benefit from joining even if the service is free, the rest of this section assumes $2rE[X] > cE[X^2]$. The following theorem establishes the existence and the uniqueness of equilibrium under this condition, and computes the equilibrium.

Theorem 2. *Given $2rE[X] > cE[X^2]$, in the unique equilibrium of the unobservable stochastic clearing system with generally distributed intervisit times, the administrator sets the fee*

$$\tau^u = r - c\frac{E[X^2]}{2E[X]}, \tag{6}$$

and all customers join. The expected total customer utility per time unit is

$$\pi_C^u = 0, \tag{7}$$

the expected administrator revenue per time unit is

$$\pi_A^u = \lambda\left(r - c\frac{E[X^2]}{2E[X]}\right), \tag{8}$$

and the expected social utility per time unit is

$$\pi_S^u = \lambda\left(r - c\frac{E[X^2]}{2E[X]}\right). \tag{9}$$

Theorem 4 establishes that in the unobservable case, the administrator extracts all the service surplus (viz., the difference between the utility of the service and the expected cost of waiting for it) in the form of a fee, and consequently, customers break even in the equilibrium. In this case, the equilibrium is unique and the fee can be arbitrarily small unlike in the observable case, where there are multiple equilibria and the fee is bounded below by $r/2$. The following corollary states the effects of parameter changes on the equilibrium.

Corollary 4. *Given $2rE[X] > cE[X^2]$, in the equilibrium of the unobservable system,*

(*i*) *The fee τ^u is increasing in r and $E[X]$, decreasing in c and $E[X^2]$, and independent of λ.*
(*ii*) *The expected total customer utility π_C^u is zero independently of all model parameters.*
(*iii*) *The expected administrator revenue π_A^u and the expected social utility π_S^u are increasing in λ, r and $E[X]$, and decreasing in c and $E[X^2]$.*

The statements (i) and (iii) imply that for **fixed expected intervisit time** $E[X]$, the equilibrium service fee and the expected administrator revenue are decreasing in the variance of the intervisit time, $Var[X] = E[X^2] - E^2[X]$. That is, the administrator suffers from increasing the variance of the intervisit time. This result intuitively makes sense as an increase in the variance of the intervisit time can be interpreted as a reduction in service reliability, which would make the offered service less attractive for customers; to make up for this, the administrator lowers the fee and, since the joining rate remains the same, his expected revenue decreases. Hence, in the unobservable case, the monotonicity of the fee and the expected administrator revenue in intervisit time variance holds in general unlike the observable case, where an increase in the variance can increase or decrease the fee, and the monotonicity of the expected administrator revenue is shown under the convex-ordering assumption.

Similarly, for **fixed coefficient of variation** $\sqrt{Var[X]}/E[X]$, the equilibrium service fee and the expected administrator revenue diminish as the expected intervisit time $E[X]$ increases, since

$$\tau^u = r - c\frac{E[X^2]}{2E[X]} = r - c\left(\frac{\sqrt{Var[X]}}{E[X]}\right)^2\left(\frac{E[X]}{2}\right) - c\frac{E[X]}{2}.$$

In this case, however, the reason why the service becomes less attractive for customers is the long expected service time rather than the variability.

6. Partially Observable Case

This section examines the partially observable case, where every arriving customer decides whether to join or to balk after observing the fee and the number of customers in the station. In this case, the expected utility of a joining customer depends on the expected waiting time conditional on the number of customers that are in the station at her arrival instant.

Determining the expected waiting time given the information provided to the customer was straightforward in the observable and unobservable cases, as the exact waiting time was given in the former case and PASTA property applied in the latter. The partially observable case calls for evaluating the expected waiting time conditional on the number of customers observed upon arrival. We do this by using the results of Manou et al. (2014), who developed an iterative procedure to calculate this conditional expectation for a fixed service fee and characterized the unique equilibrium joining strategy of customers under the assumption that the intervisit time has a strictly decreasing mean residual life distribution. A nonnegative random variable X is said to have a (strictly) *decreasing mean residual life* (DMRL) distribution, if the function $E[X - x|X \geq x]$ is (strictly) decreasing in $x \geq 0$. This assumption is reasonable in transportation systems, as the longer the time elapsed from the previous visit of the transportation facility is, the shorter the expected time till its next visit will be. Under this assumption, Manou et al. proved that the unique equilibrium strategy of customers is of the reverse-threshold type, that is, there exists a threshold n such that an arriving customer balks if she observes less than n customers in the station and joins with probability 1 if she observes more than n customers. This result is essential in the proof of the main theorem of this section.

For the analysis of customer behavior in the partially observable case, we exclude the trivial cases where an arriving customer has no benefit from joining even when all customers balk and the service fee is zero. Manou et al. (2014) proved that the expected waiting time of a joining customer given that all other customers balk is $E[X^2]/(2E[X])$, so the nontriviality condition becomes $2rE[X] > cE[X^2]$, which is the nontriviality condition in the unobservable case. The following theorem characterizes the equilibria in the partially observable case under this condition.

THEOREM 3. *Given* $2rE[X] > cE[X^2]$, *in any equilibrium of the partially observable stochastic clearing system with intervisit times that follow a strictly DMRL distribution, the administrator sets the fee* τ^p *that maximizes*

$$\lambda\tau\left(\frac{\frac{r-\tau}{c}\lambda+1}{\frac{r-\tau}{c}\lambda+\frac{1}{q(\tau)}}\right) \tag{10}$$

over $\max\left\{0,\, r - c\left[\frac{E[X]}{1-\mathcal{L}(\lambda)} - \frac{1}{\lambda}\right]\right\} \leq \tau < r - c\frac{E[X^2]}{2E[X]}$ *with* $q(\tau) \in (0, 1]$ *being the unique solution of*

$$r - \tau - c\left[\frac{E[X]}{1-\mathcal{L}(\lambda q)} - \frac{1}{\lambda q}\right] = 0 \tag{11}$$

with respect to $q \in [0, 1]$ *for given* τ. *Arriving customers join with probability*

$$q_n = \begin{cases} q(\tau^p) & \text{if } n = 0, \\ 1 & \text{if } n = 1, 2, \ldots. \end{cases} \tag{12}$$

The expected total customer utility per time unit is

$$\pi_C^p = \lambda\left(r - \tau^p - c\frac{E[X^2]}{2E[X]}\right), \tag{13}$$

the expected administrator revenue per time unit is

$$\pi_A^p = \lambda\tau^p\left(\frac{\frac{r-\tau^p}{c}\lambda+1}{\frac{r-\tau^p}{c}\lambda+\frac{1}{q_0}}\right), \tag{14}$$

and the expected social utility per time unit is

$$\pi_S^p = \lambda\left[r - \tau^p\left(\frac{\frac{1}{q_0}-1}{\frac{r-\tau^p}{c}\lambda+\frac{1}{q_0}}\right) - c\frac{E[X^2]}{2E[X]}\right]. \tag{15}$$

The rest of this section presents several numerical examples that shed light on the effect of changing system parameters on the equilibrium fee and the expected utilities. In these examples, we used MATLAB to solve (2) *and to maximize* (1).

EXAMPLE 2. (EFFECT OF SERVICE REWARD). Let the intervisit time follow Erlang(2, 0.5) distribution, the arrival rate be $\lambda = 2$ and the waiting cost be $c = 1$. The nontriviality condition $2rE[X] > cE[X^2]$ is satisfied if and only if $r > 3$. Figure 3 presents the fee τ^p, the expected total customer utility π_C^p, the expected administrator revenue π_A^p and the expected social utility π_S^p for $r \in (3, 10]$. We have the following observations.

- τ^p, π_C^p, π_A^p and π_S^p are nondecreasing in r.
- When the service reward r is small:
  - The fee τ^p is strictly between $r - c\left[\frac{E[X]}{1-\mathcal{L}(\lambda)} - \frac{1}{\lambda}\right]$ and $r - c\frac{E[X^2]}{2E[X]}$.
  - An arriving customer who finds the station empty joins with probability $0 < q_0 < 1$.
- When the service reward r is large:
  - All customers join with probability 1.
  - The fee is $\tau^p = r - c\left[\frac{E[X]}{1-\mathcal{L}(\lambda)} - \frac{1}{\lambda}\right] = r - \frac{11}{3}$, which is linearly increasing in r.
  - The expected total customer utility $\pi_C^p = \lambda\left(r - \tau^p - c\frac{E[X^2]}{2E[X]}\right) = \frac{4}{3}$ does not depend on r.

Figure 3 **Fee τ^p and Expected Utilities π_C^p, π_A^p and π_S^p vs. r for $\lambda = 2$, $c = 1$, and $X \sim$ Erlang(2, 0.5)**

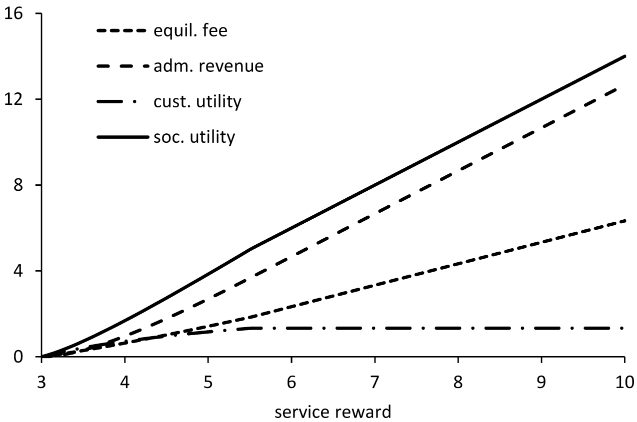
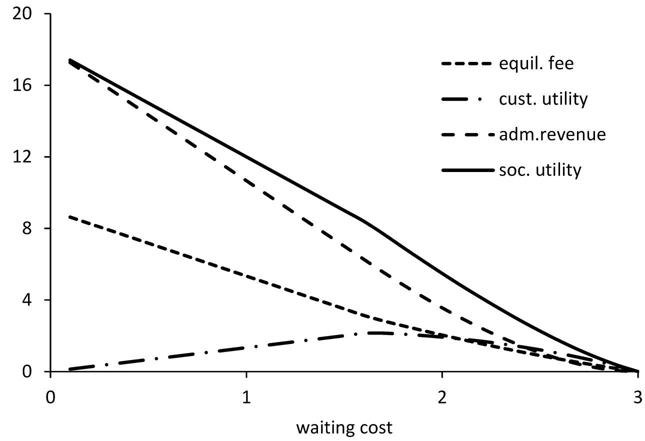


Figure 4 **Fee τ^p and Expected Utilities π_C^p, π_A^p and π_S^p vs. c for $\lambda = 2$, $r = 9$, and $X \sim$ Erlang(2, 0.5)**



- The expected administrator revenue $\pi_A^p = \lambda\tau^p = 2(r - \frac{11}{3})$ is linearly increasing in r.
- The expected social utility $\pi_S^p = \pi_C^p + \pi_A^p = 2(r - 3)$ is linearly increasing in r.
- If the service reward increases by Δ, the administrator, knowing that all customers will join, increases the fee by the same amount Δ. Thus, the administrator receives all the extra utility that results from the increase of the reward. Customers do not benefit from the increase of the service reward, as their expected utility remains the same.

Example 3. (Effect of waiting cost). Let the intervisit time follow Erlang(2, 0.5) distribution, the arrival rate be $\lambda = 2$ and the service reward be $r = 9$. The nontriviality condition is satisfied if and only if $c < 3$. Figure 4 presents the fee τ^p, the expected total customer utility π_C^p, the expected administrator revenue π_A^p and the expected social utility π_S^p for $c \in (0, 3)$. We have the following observations.

- τ^p, π_A^p and π_S^p are nonincreasing in c.
- When the waiting cost is small:
  - All customers join with probability 1.
  - The fee is $\tau^p = r - c\left[\frac{E[X]}{1-\mathcal{L}(\lambda)} - \frac{1}{\lambda}\right] = 9 - \frac{11}{3}c$, which is linearly decreasing in c.
  - The expected total customer utility $\pi_C^p = \lambda\left(r - \tau^p - c\frac{E[X^2]}{2E[X]}\right) = \frac{4}{3}c$ is linearly increasing in c.
  - The expected administrator revenue $\pi_A^p = \lambda\tau^p = 2(9 - \frac{11}{3}c)$ is linearly decreasing in c.
  - The expected social utility $\pi_S^p = \pi_C^p + \pi_A^p = 2(9 - 3c)$ is linearly decreasing in c.
  - If the waiting cost increases by Δ, the administrator reduces the fee by more than Δ to attract all the customers. This fee reduction harms the administrator, but benefits the customers.
- When the waiting cost is large:
  - The fee τ^p is strictly between $r - c\left[\frac{E[X]}{1-\mathcal{L}(\lambda)} - \frac{1}{\lambda}\right]$ and $r - c\frac{E[X^2]}{2E[X]}$. It decreases as c increases, but not as sharply as in the case when the waiting cost is small.
  - The expected utilities π_C^p, π_A^p and π_S^p are all decreasing in c.
  - Although the administrator reduces the fee as c increases, customers do not benefit from this reduction. This happens because the negative effect (*increase of the waiting cost*) *overweighs the positive effect* (*fee reduction*).

Example 4. (Effect of variance). Let the intervisit time follow Erlang(n, μ) distribution with $n \geq 1$ and $\mu > 0$, the arrival rate be $\lambda = 2$, the service reward be $r = 5$, and the waiting cost be $c = 1$. The mean is fixed at $E[X] = 4$. Table 4 presents the fee τ^p, the throughput (or the effective arrival rate $\lambda_e = \pi_A^p/\tau^p$), the expected total customer utility π_C^p, the expected administrator revenue π_A^p and the expected social utility π_S^p for n = 1, 2, 4, 8, 16. As n increases, the variance of the intervisit time decreases, which can be interpreted as an improvement in service reliability.

- The administrator takes advantage of a variance reduction by increasing the fee.
- The increase of the fee, in turn, makes the customers less willing to join. Indeed, the throughput λ_e decreases as variance decreases.
- Although the throughput decreases, the expected total customer utility and the administrator revenue increase.

Example 5. (Effect of expected intervisit time). Let the intervisit time follow Erlang(2, μ) distribution

Table 4 Fee τ^p, Throughput λ_e Expected Utilities π_A^p, π_C^p and π_S^o for $\lambda = 2$, $c = 1$, $r = 5$, and $X \sim$ Erlang(n, μ) with $E[X] = 4$

n	μ	Fee	Throughput	Administrator revenue	Customer utility	Social utility
1	0.25	0.9990	2.0000	1.9980	0.0020	2.0000
2	0.5	1.4200	1.8958	2.6921	1.1600	3.8521
4	1	1.7340	1.7649	3.0604	1.5320	4.5924
8	2	1.8990	1.7068	3.2413	1.7020	4.9433
16	4	1.9850	1.6767	3.3283	1.7800	5.1083

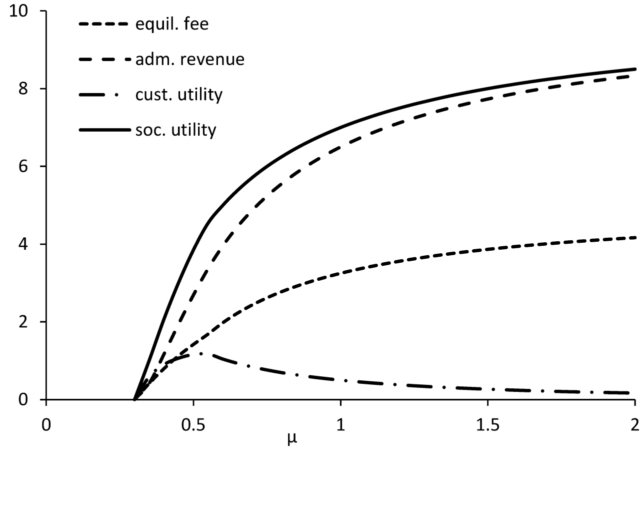
with $\mu > 0$, the arrival rate be $\lambda = 2$, the service reward be $r = 5$, and the waiting cost be $c = 1$. In this case, the coefficient of variation $\sqrt{Var[X]}/E[X]$ remains fixed at $1/\sqrt{2}$. The nontriviality condition is satisfied if $\mu > 0.3$. Figure 5 presents the fee τ^p, the expected total customer utility π_C^p, the expected administrator revenue π_A^p and the expected social utility π_S^p for $\mu > 0.3$.

- As μ increases, that is, the frequency of the visits of the transportation facility increases, the fee and the administrator revenue increase.
- The customers benefit from a higher service frequency when μ is small. However, when μ is large, an increase of the visit frequency harms the customers as the administrator imposes a very high fee.

7. Comparison

In this section, we aim to determine which level of information is preferable for the customers, which is preferable for the administrator and which is socially preferable. In general, the answer depends on the specific values of the system parameters. Subsection 7.1 compares the three cases assuming that the intervisit times are exponentially distributed. Subsection 7.2 compares the three cases under the assumption that the intervisit times have a strictly DMRL distribution.

Figure 5 Fee τ^p and Expected Utilities π_C^p, π_A^p and π_S^p vs. μ for $\lambda = 2$, $c = 1$, $r = 5$, and $X \sim$ Erlang(2, μ)



7.1. Exponential Intervisit Times

When the intervisit times are exponential, the partially observable case reduces to the unobservable case, so this section compares the observable and unobservable cases. Let μ denote the rate of the intervisit times. For the observable case, by Theorem 1, the fee τ^o is the unique nonnegative solution to the equation

$$e^{\mu\left(\frac{r-\tau^o}{c}\right)} = 1 + \frac{\mu\tau^o}{c},$$

and the expected utilities are

$$\pi_A^o = \frac{\lambda\mu(\tau^o)^2}{c} e^{-\mu\left(\frac{r-\tau^o}{c}\right)} = \frac{\lambda\mu(\tau^o)^2}{c+\mu\tau^o},$$

$$\pi_S^o = \lambda\mu \int_0^{\frac{r-\tau^o}{c}} (r - cx)e^{-\mu x}dx = \lambda r - \frac{2\lambda c\tau^o}{c+\mu\tau^o},$$

$$\pi_C^o = \pi_S^o - \pi_A^o = \lambda(r-\tau^o) - \frac{\lambda c\tau^o}{c+\mu\tau^o}.$$

By Theorem 4, for the unobservable case under the condition $r > c/\mu$, the fee is $\tau^u = r - (c/\mu)$, and the expected utilities are $\pi_C^u = 0$, and $\pi_A^u = \pi_S^u = \lambda(r - (c/\mu))$.

The value of $(r - (c/\mu))/r$, which is increasing in r and μ, and decreasing in c, turns out to be determinative in the comparison of these expressions. We refer to this ratio as the normalized service surplus and denote it by δ. We note that $0 < \delta < 1$ when the nontriviality assumption holds. Proposition 1 compares the fee and the expected utilities in the observable and unobservable cases. These findings are summarized in Table 5.

Proposition 1. *Given exponentially distributed intervisit times with rate μ and the normalized service surplus* $0 < \delta = \frac{r-\frac{c}{\mu}}{r} < 1$,

(*i*) *The fee in the unobservable case is less than (equal to, greater than) the fee in the observable case if* $\delta < 1 - e^{-1} \approx 0.6321$ *(=, >, respectively).*
(*ii*) *The expected total customer utility is always greater in the observable case than in the unobservable case.*
(*iii*) *The expected administrator revenue in the unobservable case is less than (equal to, greater than) the expected administrator revenue in the observable case if* $\delta < K \approx 0.2483$ *(=, >, respectively).*
(*iv*) *The expected social utility in the unobservable case is less than (equal to, greater than) the expected social utility in the observable case if* $\delta < \frac{\ln 2}{1+\ln 2} \approx 0.4094$ *(=, >, respectively).*

Table 5 System in Which the Corresponding Value is Greater Depending on $\delta = \frac{r - \frac{c}{\mu}}{r}$

	$\delta \in (0, 0.2483)$	$\delta \in (0.2483, 0.4094)$	$\delta \in (0.4094, 0.6321)$	$\delta \in (0.6321, 1)$
Fee	o	o	o	u
Customer utility	o	o	o	o
Administrator revenue	o	u	u	u
Social utility	o	o	u	u

According to Table 5, customers always prefer having information whereas what is best for the administrator and what is socially best depend on the value of the normalized service surplus. If the normalized service surplus is small enough, the administrator prefers disclosing information, which turns out to be also socially optimal. Hassin (1986) showed these same results for the comparison of the unobservable and partially observable (which he refers to as observable) M/M/1 queues with strategic customers and a profit maximizing administrator. For large values of the normalized service surplus, Hassin proved that the administrator's preference and what is socially best depend also on the arrival rate λ. Our results do not depend on λ, since we consider a clearing system. Table 5 shows that when the normalized service surplus is sufficiently large, the administrator prefers hiding information, which is also socially optimal. We also note that our results compare the observable and unobservable cases whereas Hassin compares the unobservable and partially observable cases, which are equivalent for the clearing model with exponential intervisit times. Guo and Zipkin (2007) compared the three cases for the M/M/1 queue with heterogeneous customers without pricing, showed that the preferences of customers and of the administrator depend on the shape of the customer-type distribution.

7.2. Strictly DMRL Intervisit Times

This subsection explores the effect of information when the intervisit times have a strictly DMRL distribution. Proposition 2 compares the fee and the expected utilities in the unobservable and partially observable cases. We compare the observable case with the other two cases numerically and present our observations in Table 6.

Proposition 2. *Given intervisit times with a strictly DMRL distribution and* $2rE[X] > cE[X^2]$,

(*i*) *The fee in the unobservable case is greater than the fee in the partially observable case.*
(*ii*) *The expected total customer utility is greater in the partially observable case than in the unobservable case.*
(*iii*) *The expected administrator revenue is more in the unobservable case than in the partially observable case.*
(*iv*) *The expected social utility in the unobservable case is at least as much as the expected social utility in the partially observable case.*

Proposition 2 establishes that when intervisit times have a strictly DMRL distribution, customers always prefer to have partial information over not having any information, and the administrator always prefers hiding information over disclosing it partially. Hiding information is also socially better than disclosing it partially. Hence, the assumption of strictly DMRL intervisit times in the clearing system eliminates the possibility that customers' preference between no information and partial information coincides with the administrator's preference. This differs the results of Hassin (1986), and Guo and Zipkin (2007) for the M/M/1 queue. Both papers showed that when the system parameters satisfy certain conditions, both the customers and the administrator prefer the partially observable case over the unobservable one.

Example 6. (Erlang intervisit times). Let the intervisit time follow Erlang(2, μ) distribution with $\mu \in (0, 3)$, the arrival rate be $\lambda = 1$, the service reward be $r = 5$, and the waiting cost be $c = 1$. Figures 6–8 present the expected total customer utility, the expected administrator revenue and the expected social utility per time unit, respectively, for the three levels of information.

- For large values of expected waiting time (equivalently, for small μ):
  - Customers prefer the observable case, that is, they want to know their exact waiting time. This can be explained by the large utilities experienced by customers that arrive just before the visit of the facility.

Table 6 System in Which the Corresponding Value is Greater Depending on $\delta = \left(r - \frac{cE[X^2]}{E[X]}\right)/r$

	Small δ	Large δ
Fee	o	u
Customer utility	o	o
Administrator revenue	o	u
Social utility	o	u

Figure 6 Expected Total Customer Utility π_C^u, π_C^o and π_C^p vs. μ for $\lambda = 1$, $c = 1$, $r = 5$, and $X \sim$ Erlang(2, μ)

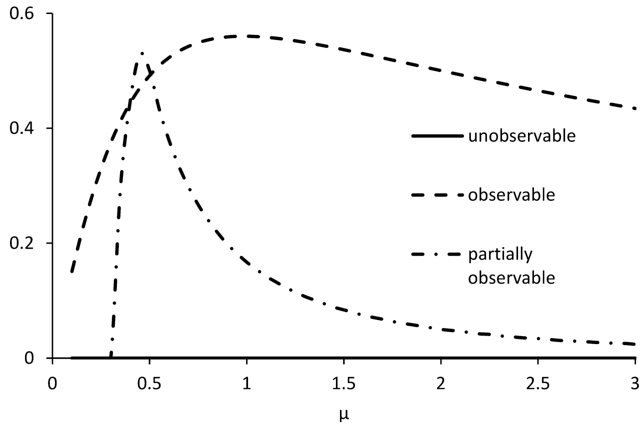
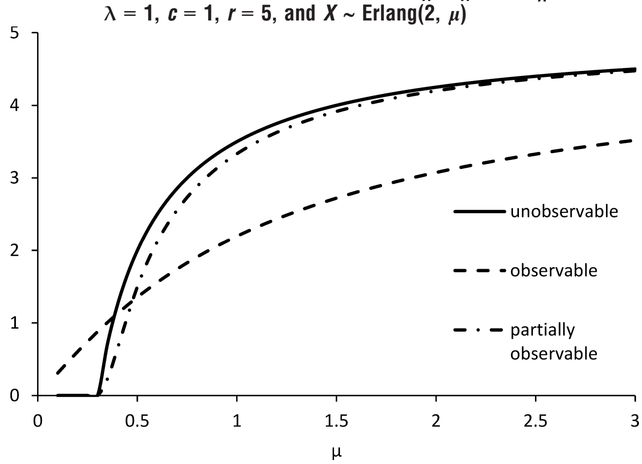


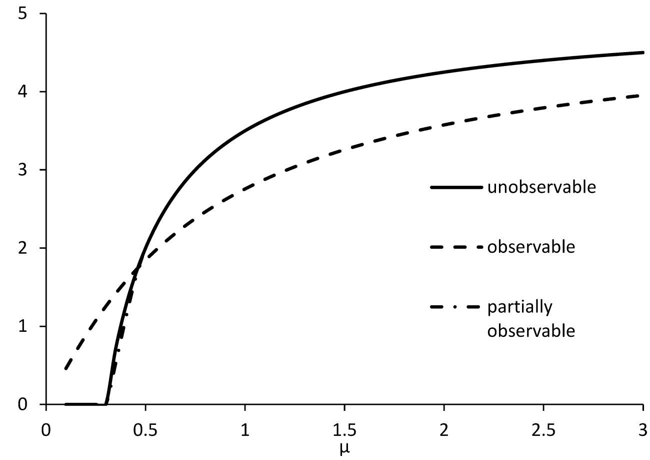
Figure 7 Expected Administrator Revenue π_A^u, π_A^o and π_A^p vs. μ for $\lambda = 1$, $c = 1$, $r = 5$, and $X \sim$ Erlang(2, μ)



- The administrator prefers to reveal the full information because in the observable case, he receives a high fee from the few customers that join.
- The socially optimal level of information is provided in the observable case.

- For small values of expected waiting time (equivalently, for large μ):
  - Customers prefer to be fully informed because the imposed fee is lower in the observable case.
  - The administrator prefers to hide the information, since in the unobservable case, all customers join and the fee is high enough.
  - The socially optimal level of information is provided in the unobservable case.
- For intermediate values of expected waiting time (equivalently, for intermediate μ):
  - Customers prefer the partially observable case, where the fee almost equals the fee in the observable case, but the fraction of joining customers is greater.

Figure 8 Expected Social Utility π_S^u, π_S^o and π_S^p vs. μ for $\lambda = 1$, $c = 1$, $r = 5$, and $X \sim$ Erlang(2, μ)



  - The administrator prefers to hide the information.
  - The socially optimal level of information is provided in the unobservable or in the observable case.

In our numerical experiments, we noticed that for small and large values of r and c, customers prefer to know their exact waiting time, whereas for intermediate values of r and c, they prefer to know only the number of customers in the system. Also, in these experiments, the observable case turns out to be optimal for the administrator and the society when the service reward is small and the waiting cost is large. On the other hand, the unobservable case turns out to be optimal for the administrator and the society when the service reward is large and the waiting cost is small. Table 6 summarizes some of our numerical findings for small and large values of the normalized service surplus $\delta = \left(r - \frac{cE[X^2]}{E[X]}\right)/r$. Observations summarized in Table 6 for strictly DMRL intervisit times are same as the results for small and large values of the normalized service surplus when intervisit times are exponential as presented in Table 5.

8. Conclusion

In this study, we considered the strategic behavior of the customers and the administrator in a transportation station. We modeled the situation as a two-stage game for the three different levels of delay information. In each one of the observable, unobservable, and partially observable cases, we studied the equilibrium behavior, assuming generally distributed intervisit times of the transportation facility and we explored the effects of the changes in system parameters on the administrator and the customers. Our analytical results in the observable and unobservable cases showed that an increase in the service reward can benefit some or all of the involved parties, but it does not harm anyone. However, a reduction in the waiting cost, an increase in the service frequency, or a reduction in the intervisit time variance benefits the administrator and the society, but can harm the customers as illustrated by our numerical examples. The comparison of the three cases with different levels of information provided to customers upon their arrival suggests that in most cases, the customers prefer to know their exact waiting times. For the administrator and the society, either the observable or the unobservable case is optimal depending on the system parameters. In particular, when the unobservable case is socially optimal, the administrator also prefers to hide the information.

We next discuss some extensions of our work and some research directions for future.

Heterogeneous Customers. In this study, we assumed that customers have the same reward and cost parameters. In practice, customers can have different service rewards and/or waiting costs. In a forthcoming paper, we extend our results to the setting with heterogeneous customers assuming that the fee is fixed.

State-Dependent Fees. Our analysis of the observable and partially observable cases assumed that the administrator charges the same fee independently of the state of the system. This assumption is quite realistic for public transportation systems, but it is also interesting to explore the effect of state-dependent pricing on the equilibrium. If the administrator can charge a fee that depends on the waiting time in the observable case, he will charge $r - cw$ for the waiting time $w < r/c$ and receive all the service surplus. Consequently, the expected total customer utility becomes zero as in the unobservable case. The expected administrator revenue $\pi_A^{o,sd}$ and the expected social utility $\pi_S^{o,sd}$ are equal in the observable case with state-dependent fees and

$$\pi_A^{o,sd} = \frac{\lambda}{E[X]} \int_0^{\frac{r}{c}} (r - cx)[1 - F(x)]\mathrm{d}x$$
$$\geq \frac{\lambda}{E[X]} \int_0^{\frac{r-\tau^o}{c}} (r - cx)[1 - F(x)]\mathrm{d}x = \pi_S^o \geq \pi_A^o.$$

Hence, allowing the administrator to charge fees that depend on the waiting time improves the expected administrator revenue and the expected social utility; however, it harms customers by allocating all the social utility to the administrator. Also, a similar argument yields $\pi_A^{o,sd} \geq \pi_A^u$, so the administrator generates a greater revenue in the observable case with state-dependent fees as compared to the unobservable case. Finding out the effects of state-dependent pricing for the partially observable case is more complicated, so we leave it for future research.

Administrator Objectives and Constraints. Our analysis assumed that the administrator is a profit maximizer. In practice, transportation administrators may not be pure profit maximizers and their actions can be subject to governmental regulations such as price caps, subsidies, and quality constraints. Studying the effect of such regulations and other objectives pursued by the administrator is an interesting research question.

Capacity Constraints. An assumption that greatly simplified the analysis in this study was that the service facility has infinite capacity. This assumption can be justified when the transportation facility has a very large capacity. But relaxing it would make the model much more realistic in general; however, the analysis of the relaxed model would be very complicated.

Information Structures. Including different information structures such as those providing real-time information to customers is another possible research direction. As the information structure gets more complicated, obtaining analytical results is likely to get more challenging.

Competition among Service Providers. We restricted attention to the modeling and analysis of a single transportation station. In practice, there are usually several transportation alternatives, so strategic customers may also choose between the alternatives. This adds another layer of competition to the model. How this competition affects the fees and the expected utilities is another interesting research question.

Acknowledgments

The authors are grateful to the editor and the referees for their constructive comments and suggestions, which have helped us to improve both the quality and the exposition of the study. Athanasia Manou was supported by AXA Research Fund. The second author was supported by Marie Curie Career Integration Grant from the European Union's Seventh Framework Programme (FP7-PEOPLE-2013-CIG, Proposal No. 618853, Acronym RISK).

References

Chen, H., M. Frank. 2001. State dependent pricing with a queue. *IIE Trans.* **33**(10): 847–860.

Chen, H., M. Frank. 2004. Monopoly pricing when customers queue. *IIE Trans.* **36**(6): 1–13.

Cil, E. B., F. Karaesmen, L. Ormeci. 2011. Dynamic pricing and scheduling in a multi-class single-server queueing system. *Queueing Syst.* **67**(4): 305–331.

Economou, A., A. Gomez-Corral, S. Kanta. 2011. Optimal balking strategies in single-server queues with general service and vacation times. *Perform. Eval.* **68**(10): 967–982.

Edelson, N. M., K. Hildebrand. 1975. Congestion tolls for Poisson queueing processes. *Econometrica* **43**(1): 81–92.

Guo, P., P. Zipkin, 2007. Analysis and comparison of queues with different levels of delay information. *Management Sci.* **53**(6): 962–970.

Guo, P., P. Zipkin. 2008. The effects of information on a queue with balking and phase-type service times. *Nav. Res. Logistics* **55**(5): 406–411.

Guo, P., P. Zipkin. 2009. The effects of the availability of waiting-time information on a balking queue. *Eur. J. Oper. Res.* **198**(1): 199–209.

Hassin, R. 1986. Consumer information in markets with random products quality: The case of queues and balking. *Econometrica* **54**(5): 1185–1195.

Hassin, R. 2016. *Rational Queueing*. CRC Press, Taylor and Francis Group, Boca Raton, FL.

Hassin, R., M. Haviv. 2003. *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*. Kluwer Academic Publishers, Boston.

Kerner, Y. 2008. The conditional distribution of the residual service time in the Mn/G/1 queue. *Stoch. Models* **24**(3): 364–375.

Kerner, Y. 2011. Equilibrium joining probabilities for an M/G/1 queue. *Games Econ. Behav.* **71**(2): 521–526.

Knudsen, N. C. 1972. Individual and social optimization in a multi-server queue with a general cost-benefit structure. *Econometrica* **40**(3): 515–528.

Li, L. 1988. A stochastic theory of the firm. *Math. Oper. Res.* **13**(3): 447–466.

Low, D. W. 1974. Optimal dynamic pricing policies for an M/M/s queue. *Oper. Res.* **22**(3): 545–561.

Manou, A., A. Economou, F. Karaesmen. 2014. Strategic customers in a transportation station: When is it optimal to wait? *Oper. Res.* **62**(4): 910–925.

Naor, P. 1969. The regulation of queue size by levying tolls. *Econometrica* **37**(1): 15–24.

Stidham Jr., S. 2009. *Optimal Design of Queueing Systems*. CRC Press, Taylor and Francis Group, Boca Raton, FL.

Stradling, S., M. Carreno, T. Rye, A. Noble. 2007. Passenger perceptions and ideal urban bus journey experience. *Transport Policy* **14**(4): 283–292.

Yechiali, U. 1971. On optimal balking rules and toll charges in the GI/M/1 queue. *Oper. Res.* **19**(2): 349–370.

Yildirim, U., J. J. Hasenbein. 2010. Admission control and pricing in a queue with batch arrivals. *Oper. Res. Lett.* **38**(5): 427–431.

Zhang, F., J. Wang, B. Liu. 2013. Equilibrium joining probabilities in observable queues with general service and setup times. *J. Ind. Manag. Optimiz.* **9**(4): 901–917.

Zhou, W., X. Chao, X. Gong. 2013. Optimal uniform pricing strategy of a service firm when facing two classes of customers. *Prod. Oper. Manag.* **23**(4): 676–688.

Supporting Information

Additional supporting information may be found online in the supporting information tab for this article:

Appendix S1: On Line Supplement.