Pricing in M/M/1 queues when cost of waiting in queue differs from cost of waiting in service

Görkem Sarıyer

Department of Business Administration, Yaşar University, Turkey
gorkem.ataman@yasar.edu.tr

ABSTRACT

Service providers can adjust the entrance price to the state of the demand in real life service systems where the customers’ decision to receive the service, is based on this price, state of demand and other system parameters. We analyzed service provider’s short and long term pricing problems in unobservable M/M/1 queues having the rational customers, where, for customers, the unit cost of waiting in the queue is higher than unit cost of waiting in the service. We showed that waiting in the queue has a clear negative effect on customers’ utilities, hence the service provider’s price values. We also showed that, in the short term, monopolistic pricing is optimal for congested systems with high server utilization levels, whereas in the long term, market capturing pricing is more profitable.

1. Introduction

In real life, most service systems include customer queues. Thus, the potential customers of these systems will inevitably face waiting time before being served. Since time is a very scarce resource in the current environment [1], customers decide rationally upon receiving a service. Based on comparing the value of the received service with the total cost of receiving this service, a decision is made whether or not to join the queue; this decision generates the real state of the demand.

In order to maximize profits, the service provider takes into account the rational decision of the customers while setting the optimal system parameters, such as price or service rate. Until recently, however, pricing studies in the literature assumed that the firm was not able to adjust the price based on the state of the demand. The first model allowing the firm adjust the price to the state of the demand was analyzed by Naor [2]. In this benchmark paper, Naor, analyzed the effect of imposing an entrance fee on socially optimum, when customers will not join the system until after observing the length of the queue. A similar study by Edelson and Hildebrand, had an important difference in that customers were not able to observe the queue length before taking their decision [3]. In very similar settings, Chen and Frank analyzed the pricing problem of the monopoly, or the service provider, when the queue length is observable by customers [4], and when unobservable [5]. In the literature, other studies analyze the pricing problem of the service provider in queueing models. Knudsen allowed more than a single queue in any one firm [6]. Sarıyer compared pricing strategies of the service provider in the cases of a single server, and two servers [7]. Some other studies represented socially optimal pricing schemes for different classes of customers [8, 9, 10]. All these cited studies assumed that a queue decreases the utility of the customer from the received service. A more recent view in queueing literature suggest that customer utility will not necessarily decrease due to wait time. Debo et al. showed that expected service time, which, in the literature, was included in total waiting time, is positively correlated with the value of the received service [11]. Anand at al., Alizamir et al., and Wang et al. also considered that service value is correlated with waiting time [12, 13, 14]. Oliveras et al. proposed that customer purchase decisions are not monotonic to queue length [15]. Giebelhausen et al. concluded that longer wait time can signal greater service quality, which positively affects customer decisions [16]. Sarıyer combined these two views in the literature and represented the utility function of customers with a different structure, which allowed discrimination between the effect of waiting in the queue and waiting in the service on customer’s decision.
In this paper, we will analyze the pricing problem of the service provider in which the customer utility function is a modified version of the one represented in [17]. The model assumptions, notations, and utility function of the customer will be presented in Section 2. The pricing problem will be analyzed in both of the market capture and monopolist pricing settings, details of which will be given in Section 3. In the long term, the service provider is capable of adjusting not only the price but also the service rate. The long term analyses will be covered in Section 4. In Section 5 and 6 we respectively present the Numerical Analysis and Conclusions part.

2. Model assumptions

We analyze basic M/M/1 queues having Poisson arrivals with rate \( \lambda \), and a unique server having an exponentially distributed service rate of \( \mu \). We use monetary values to represent the system parameters, such as value of the service and cost of waiting. The value of the service is denoted by a service reward, \( R \), and a customer who receives this service incurs a linear waiting cost with \( C \) units.

In real life, service systems having queues of customers can be differentiated as observable or unobservable based on the visibility of length of the queue. Cashomat or Automatic Teller Machine (ATM), fast-food or self-service restaurants, banks are examples of service systems including observable queues, since a new arriving customer can observe the number of customers in front of him. On the other hand, some call centers, service systems taking the orders online which share services, or Automatic Teller Machine (ATM), fast food or self-service restaurants, cashomat can be differentiated as observable or unobservable systems including unobservable queues, since the customers cannot observe the length of the queue. In this paper, we assume the length of the queue is not observable upon arrival, but, based on the shared information by the service provider or the past experiences of the arriving customers, customers know the expected waiting time in the system. We also assume that the systems are in steady state. The expected waiting time combines two elements: the expected waiting time in the queue, and the expected waiting time during the service. Based on our assumption, the first element has a greater negative effect on customer utility, thus we multiply the unit cost of waiting in the queue with \( k \) and \( C \). A service provider sets a price, \( p \), for customers receiving the service. This price clearly decreases the utility of the customer. The customer decides whether or not to join the queue and receive the service. Thus, the customer decision parameter is the probability of joining, denoted by \( \alpha \). All of these are combined to derive utility function of the customer, which is denoted by \( U(\alpha) \) as:

\[
U(\alpha) = R - kCE[W_1] - CE[W_2] - p
\]

\[
= R - k\frac{\lambda \alpha}{\mu(\mu-\lambda)} - \frac{C}{\mu} - p
\]

In equation (1), \( E[W_1] \) and \( E[W_2] \) represent expected waiting time in the queue and the expected waiting time during the service respectively. The values of \( E[W_1] \) and \( E[W_2] \) are found based on the properties of \( M/M/1 \) queues.

While making their decision, customers consider this utility function. If \( R-p \) is smaller than only the service cost of a unique customer, then joining the system is not optimal, since the customer has negative utility, thus \( \alpha=0 \). If \( R-p \) is greater than total cost of waiting in the system, even if all customers join, then it remains optimal for all potential customers, since they receive positive utility, hence \( \alpha=1 \). In between, the customers join with an equilibrium probability, \( \alpha^{eq} \), where \( 0 < \alpha^{eq} < 1 \). This equilibrium probability is found by setting equation (1) to 0, and solving it for \( \alpha \). Hassin and Haviv provides a detailed review of these equilibrium analyses of rational customers [18].

3. Pricing decisions in short term

In the short term, the service provider sets the optimal entrance price to the system for the given model parameters \( \lambda \) and \( \mu \). Assuming \( \mu > \lambda \), the service rate is sufficiently high to serve all potential customers. The service provider therefore sets either the minimum price, referred to in the literature as market capturing price, in order to allow all customers to enter the system, or a higher price, denoted as monopolistic price, to push customers to join at an equilibrium rate.

3.1. Market capture pricing

The minimum price which allows all customers to join is found by giving \( \alpha \) to 1 in equation (1), setting this utility function to 0, and solving it for \( p \). Denoting this market capturing price with symbol \( p_l \), it is derived as:

\[
p_l = R - kC \frac{\lambda}{\mu(\mu-\lambda)} - \frac{C}{\mu}
\]

(2)

**Lemma 1.** \( p_l \) is increasing in \( R \) and \( \mu \), decreasing in \( \lambda \), \( k \) and \( C \).

**Proof.**

\[
\frac{dp_l}{dR} = 1 > 0, \quad \frac{dp_l}{d\mu} = \frac{kC(2\mu - \lambda)}{\mu(\mu - \lambda)^2} + \frac{C}{\mu^2} > 0
\]

\[
\frac{dp_l}{dC} = -\frac{k\lambda}{\mu(\mu - \lambda)} - \frac{1}{\mu} < 0, \quad \frac{dp_l}{dk} = -C \frac{\lambda}{\mu(\mu - \lambda)} < 0,
\]

\[
\frac{dp_l}{d\lambda} = \frac{C\mu^2}{\mu(\mu - \lambda)^2} < 0
\]

Lemma 1 can be clearly interpreted. Since the utility function of the customer increases when \( R \) increases, the service provider is able to set a higher price. On the other hand, when \( k \) and \( C \) increase, the expected cost of waiting increases, thus utility of the customer decreases, which forces the service provider to lower the price. When \( \mu \) increases, the expected waiting time in the system decreases, which generates an increase in utility function, and pushes the service provider to set a higher price; on the other hand, when \( \lambda \) increases, expected waiting time increases, causing a decrease in the entrance price.
The profit function of the service provider under this market capturing price, \( \Pi_{tp}(p_t) \), is derived as:

\[
\Pi_{tp}(p_t) = \lambda p_t = \lambda \left( R - kC + \frac{\lambda}{\mu(\mu - \lambda)} - \frac{C}{\mu} \right)
\] (3)

**Lemma 2.** \( \Pi_{tp}(p_t) \) is increasing in \( R \) and \( \mu \), decreasing in \( k \) and \( C \), and concave in \( \lambda \).

**Proof.**

\[
\frac{d\Pi_{tp}(p_t)}{dr} = \lambda > 0, \quad \frac{d\Pi_{tp}(p_t)}{d\mu} = \frac{kC\lambda^2(2u - \lambda)}{\mu(\mu - \lambda)^2} + \frac{C\lambda}{\mu} > 0
\]

\[
\frac{d\Pi_{tp}(p_t)}{dk} = -\frac{k\lambda^2}{\mu(\mu - \lambda)} - \frac{1}{\mu} < 0,
\]

\[
\frac{d\Pi_{tp}(p_t)}{dC} = -\frac{C}{\mu(\mu - \lambda)} < 0,
\]

\[
\frac{d^2\Pi_{tp}(p_t)}{d\lambda^2} = R - k\lambda \left[ \frac{2\mu(\mu - \lambda)}{\mu(\mu - \lambda)^2} - \frac{C}{\mu} \right] < 0
\]

In Lemma 2, we showed that profit value is concave in \( \lambda \). Thus, when the arrival rate increases, the profit value increases up to a certain level, after which value profit decreases, due to congestion in the system. For the other model parameters, \( R, \mu, k \) and \( C \), the profit function behaves similarly to price function, \( p_L \), as expected.

### 3.2. Monopolistic pricing

Service provider sets a higher price thus customers join with an equilibrium joining probability which is derived from equation (1) as:

\[
U(\alpha) = 0 \rightarrow R - p = \frac{C}{\mu} \left( \frac{k\lambda}{\mu - \lambda} + 1 \right)
\]

\[
\rightarrow a^eq = \frac{\mu (R - p) \mu - 1}{k (\mu - \lambda)}
\] (4)

**Lemma 3.** \( a^eq \) is increasing in \( R \) and \( \mu \), and decreasing in \( k \) and \( C \).

**Proof.**

\[
\frac{da^eq}{dR} = \frac{k\mu}{\lambda \left( k + \frac{(R - p)\mu}{\mu - 1} \right)} > 0,
\]

\[
\frac{da^eq}{d\mu} = \frac{\left( R - p \right) \mu - 1}{\lambda \left( k + \frac{(R - p)\mu}{\mu - 1} \right)} \left( k + \frac{(R - p)\mu}{\mu - 1} \right) > 0,
\]

\[
\frac{da^eq}{dk} = -\frac{\mu (R - p) \mu - 1}{\lambda (k + \frac{(R - p)\mu}{\mu - 1})} < 0
\]

\[
\frac{da^eq}{dC} = \frac{k\mu^2}{\lambda \left( k + \frac{(R - p)\mu}{\mu - 1} \right)} < 0,
\]

\[
\frac{d^2a^eq}{d\mu^2} = -\frac{k \left( (R - p) \mu - 1 \right)}{\lambda (k + \frac{(R - p)\mu}{\mu - 1})^2} < 0
\]

The interpretation of this result is very similar to Lemma 1.

We will find the expression of the monopolistic price, \( p_h \), by analyzing the profit function of the service provider in this setting. Denoting the profit function with \( \pi_{tp}(p) \), we derive it as:

\[
\pi_{tp}(p) = \lambda a^eq p = \frac{\mu (R - p) \mu - 1}{k + \frac{(R - p)\mu}{\mu - 1}} p
\] (5)

Since the derivation of \( p_h \) is messy, we will give the analysis numerically in Section 5.

### 4. Service rate and pricing decisions in long term

In this section, we assume that service provider can adjust both the price and service rate. Thus, the profit functions have two parameters, \( \mu \) and \( p \). As given in Chen and Frank [5], we assume a constant marginal cost of speeding up the service rate, \( F > 0 \).

For the market capturing price setting, the profit function given in equation (3) is rewritten as:

\[
\pi_{tp}(\mu) = \lambda p_t - F \mu = \lambda \left( R - kC + \frac{\lambda}{\mu(\mu - \lambda)} - \frac{C}{\mu} \right) - F \mu
\] (6)

**Lemma 5.** \( \pi_{tp}(\mu) \) is concave in \( \mu \).

**Proof.**

\[
\frac{d\pi_{tp}(\mu)}{d\mu} = \lambda \left( \frac{k\lambda(2\mu - \lambda)}{\mu(\mu - \lambda)^2} + \frac{C}{\mu^2} \right) - F;
\]

\[
\frac{d^2\pi_{tp}(\mu)}{d\mu^2} = -C\lambda \left( 2k\lambda \frac{3\mu^2 - 3\mu\lambda + \lambda^2}{\mu(\mu - \lambda)^3} + \frac{2}{\mu^3} \right)
\]

Since \( \mu > \lambda, 3\mu^2 - 3\mu\lambda + \lambda^2 > 0 \) for all real \( \lambda \) and \( \mu \) then the second derivative of the profit function is always negative, which shows that the profit function is concave with respect to \( \mu \).

Thus, there is a unique optimal value for the service rate which equates the first derivative to 0, and maximizes the profit function given in equation (6).

For the monopolistic pricing, the profit function is rewritten as:
In Figure 1, we observe that, for low levels of $\rho$, i.e. higher levels of $\mu$, the market capturing pricing is more profitable, since the percentage difference is negative. This means, for low congested systems, the service provider obtains a higher profit under market capture pricing compared to monopolistic pricing. However, when the system is congested, i.e. $\lambda$ is very close to $\mu$, the service provider should set a monopolistic price to maximize profits. As seen in this Figure, congested systems are characterized by a sharp difference in profit values. When we analyze the Figure for different $k$ values, for low levels of $\rho$, we cannot observe a clear effect of $k$ values in the percentage difference of profit (i.e. the percentage differences are almost equal for $k=1$, $k=1.25$, $k=1.5$, and $k=2$ when $\rho \leq 0.85$). The reason behind this observation is, the service provider sets such a monopolistic price that pushes almost all of the customers to join the queue and receive the service when system is not congested, or for high levels of $\mu$.

5. Numerical results and discussion

Numerical results are first given for the short term. For given levels of $R$, $C$, $k$, $\lambda$, and $\mu$, we find the values of the minimum price, $p_l$, and the profit, $\Pi_{hp}(p_l)$, under this price for the market capture pricing setting. Then, for the monopolistic pricing, we will find the equilibrium joining probability, $\alpha_{eq}$, value of the monopolistic price, $p_h$, and the profit under this price $\Pi_{hp}(p_h)$. Table 1 presents the results.

In our observations presented in Table 1, we fix $\lambda=10$, and $C=5$. We take R as 20 and 30 to represent low and high values, and take $\mu$ as 10.5 and 15 to represent low and high congestion in the system. We observe that there is a clear difference in the profit values for the high utilization levels, i.e. when the service rate is very close to the arrival rate, in which profit under monopolistic pricing is much higher than under market capture pricing. On the other hand, for low utilization levels, the market capture pricing becomes increasingly more profitable compared to monopolistic pricing, since the rate of the server is sufficiently high to serve all potential arrivals within a limited wait time. As expected from the definition, the price values are always higher under monopolistic pricing setting, i.e. $p_h > p_l$. Since the expected waiting cost in the system increases in $k$, we see that the price and profit values decrease in both of the market capture and monopolistic pricing. On the other hand, in both price settings, the price and profit values increase with increases in the ratio between the reward and unit waiting cost, namely, $R/C$.

In Figure 1, we should show how the percentage change in profit value, under market capture and monopolistic pricing settings, is affected by the increase in server utilization, $\rho$. In this experiment, we fixed $R=30$, $C=5$, $\lambda=10$ and increase $\mu$ from 10.5 to 20, where $\rho$ changes between $[0.5\,0.95]$. We repeat the analysis for different $k$ values, where $k$ takes the values of 1, 1.25, 1.5, and 2 respectively for different experiments. In these experiments $k=1$ is taken in order to represent the situation given in literature, where unit cost of waiting in the queue is exactly equal to the unit cost of waiting in service. Then the values of $k$ are increased consequentially in order to show how the percentage difference in profit values under market capture and monopolistic pricing, is affected with the increased difference between unit cost of waiting in the queue and in service. Finally, as the highest value of $k$, we take it as 2, and repeat the experiment in order not to differentiate the unit cost of waiting in the queue and in service much more (to keep close to literature results). Besides, when $k=2$, as will be seen in Figure 1, the sharp difference is obtained to show the related effect. In the vertical column of the figure, we represent $(\Pi_{hp}(p_h) - \Pi_{lp}(p_l))/\Pi_{lp}(p_l)$.

In Table 1, we present the results for different levels of $\mu$ and $p_l$.

### Table 1: Price and profit values – short term

<table>
<thead>
<tr>
<th>R/C</th>
<th>k</th>
<th>$\lambda/\mu$ ($\rho$)</th>
<th>$p_l$</th>
<th>$\Pi_{hp}(p_l)$</th>
<th>$\alpha_{eq}$</th>
<th>$p_h$</th>
<th>$\Pi_{hp}(p_h)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.25</td>
<td>0.95</td>
<td>7.62</td>
<td>76.20</td>
<td>0.87</td>
<td>16.70</td>
<td>144.82</td>
</tr>
<tr>
<td>4</td>
<td>1.25</td>
<td>0.67</td>
<td>18.83</td>
<td>188.33</td>
<td>0.97</td>
<td>18.90</td>
<td>183.68</td>
</tr>
<tr>
<td>4</td>
<td>2.00</td>
<td>0.95</td>
<td>0.48</td>
<td>4.76</td>
<td>0.82</td>
<td>16.05</td>
<td>132.26</td>
</tr>
<tr>
<td>4</td>
<td>2.00</td>
<td>0.67</td>
<td>18.33</td>
<td>183.33</td>
<td>0.98</td>
<td>18.40</td>
<td>180.83</td>
</tr>
<tr>
<td>6</td>
<td>1.25</td>
<td>0.95</td>
<td>17.62</td>
<td>176.19</td>
<td>0.91</td>
<td>25.90</td>
<td>233.58</td>
</tr>
<tr>
<td>6</td>
<td>1.25</td>
<td>0.67</td>
<td>28.83</td>
<td>288.33</td>
<td>0.97</td>
<td>28.90</td>
<td>280.86</td>
</tr>
<tr>
<td>6</td>
<td>2.00</td>
<td>0.95</td>
<td>10.48</td>
<td>104.76</td>
<td>0.87</td>
<td>25.00</td>
<td>216.85</td>
</tr>
<tr>
<td>6</td>
<td>2.00</td>
<td>0.67</td>
<td>28.33</td>
<td>283.33</td>
<td>0.98</td>
<td>28.40</td>
<td>279.10</td>
</tr>
</tbody>
</table>

$$\pi_{hp}(\mu, p) = \lambda \alpha_{eq} p - F \mu = \mu \frac{(R-p\mu)}{k+\alpha_{eq} \mu - 1} p - F \mu$$  \hspace{1cm} (7)
Thus, the profit under monopolistic pricing gets very close to the profit under market capture pricing. However, for congested systems, since the effect of waiting time is sharply increased, under market capturing pricing (i.e. all customers join), the sharp percentage difference in profit values can be observed for different $k$ values. In this case, or for lower levels of $\mu$, the percentage difference in profit values under market capture and monopolistic pricing settings, increases for increasing $k$ values. Thus, especially for congested systems, which is generally the case of real life situation, it can be understood that unit cost of waiting in the queue has a direct effect on a service provider’s pricing decision, and hence on profit.

We now turn our attention to long term decisions. For this, marginal cost of increasing the service rate was added to the calculations. We give optimal service rate, $\mu^*$, market capturing price given this service rate, $p_l^*$, and profit given these parameters, $\Pi_{lp}(\mu^*, p_l^*)$ in Table 2. The values of $R$, $C$, and $\lambda$ are taken as given in the previous numerical experiments.

### Table 2: Service rate, price and profit values of market capture pricing – long term

<table>
<thead>
<tr>
<th>$R/C$</th>
<th>$k$</th>
<th>$F$</th>
<th>$\mu^*$</th>
<th>$p_l^*$</th>
<th>$\Pi_{lp}(\mu^<em>, p_l^</em>)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.25</td>
<td>5</td>
<td>13.60</td>
<td>18.36</td>
<td>115.56</td>
</tr>
<tr>
<td>4</td>
<td>1.25</td>
<td>10</td>
<td>12.40</td>
<td>17.50</td>
<td>50.97</td>
</tr>
<tr>
<td>4</td>
<td>2.00</td>
<td>5</td>
<td>14.40</td>
<td>18.07</td>
<td>108.75</td>
</tr>
<tr>
<td>4</td>
<td>2.00</td>
<td>10</td>
<td>13.20</td>
<td>17.25</td>
<td>40.54</td>
</tr>
<tr>
<td>6</td>
<td>1.25</td>
<td>5</td>
<td>13.60</td>
<td>28.36</td>
<td>215.56</td>
</tr>
<tr>
<td>6</td>
<td>1.25</td>
<td>10</td>
<td>12.40</td>
<td>27.50</td>
<td>150.97</td>
</tr>
<tr>
<td>6</td>
<td>2.00</td>
<td>5</td>
<td>14.40</td>
<td>28.07</td>
<td>208.75</td>
</tr>
<tr>
<td>6</td>
<td>2.00</td>
<td>10</td>
<td>13.20</td>
<td>27.25</td>
<td>140.54</td>
</tr>
</tbody>
</table>

### Table 3: Service rate, price and profit values of monopolistic pricing – long term

<table>
<thead>
<tr>
<th>$R/C$</th>
<th>$K$</th>
<th>$F$</th>
<th>$\alpha^{eq}$</th>
<th>$\mu^*$</th>
<th>$p_h^*$</th>
<th>$\Pi_{hp}(\mu^<em>, p_h^</em>)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.25</td>
<td>5.00</td>
<td>0.99</td>
<td>13.25</td>
<td>18.20</td>
<td>114.85</td>
</tr>
<tr>
<td>4</td>
<td>1.25</td>
<td>10.00</td>
<td>0.99</td>
<td>12.25</td>
<td>17.35</td>
<td>50.63</td>
</tr>
<tr>
<td>4</td>
<td>2.00</td>
<td>5.00</td>
<td>0.99</td>
<td>14.25</td>
<td>18.00</td>
<td>108.68</td>
</tr>
<tr>
<td>4</td>
<td>2.00</td>
<td>10.00</td>
<td>0.99</td>
<td>13.00</td>
<td>17.10</td>
<td>40.24</td>
</tr>
<tr>
<td>6</td>
<td>1.25</td>
<td>5.00</td>
<td>0.99</td>
<td>13.25</td>
<td>28.20</td>
<td>214.36</td>
</tr>
<tr>
<td>6</td>
<td>1.25</td>
<td>10.00</td>
<td>0.99</td>
<td>12.25</td>
<td>27.35</td>
<td>150.42</td>
</tr>
<tr>
<td>6</td>
<td>2.00</td>
<td>5.00</td>
<td>0.99</td>
<td>14.25</td>
<td>28.00</td>
<td>208.65</td>
</tr>
<tr>
<td>6</td>
<td>2.00</td>
<td>10.00</td>
<td>0.99</td>
<td>13.00</td>
<td>27.10</td>
<td>139.79</td>
</tr>
</tbody>
</table>
Similarly, in Table 3, we give the equilibrium joining probability, $\alpha^{eq}$, optimal service rate, $\mu^*$, monopolistic price given this service rate, $p_h^*$, and profit given these parameters, $\Pi_{hp}^*(\mu^*, p_h^*)$.

We observe that the optimal strategy for the service provider in the long term, when there is the option to set the service rate, is to set the market capturing price, which has higher profit values compared to monopolistic pricing. So in the long term, the service provider's optimal decision is to set a service rate that can serve all potential customers (until it is profitable, i.e. when $F_0$ does not exceed the revenue received from the customer). This result is very similar to the findings of Chen and Frank [5]. We also observe that in monopolistic pricing, the optimal action pushes the service provider to set price and service rate values which allow almost all to join the system, i.e. $\alpha^{eq}$ values are almost 1. In both settings, the optimal price and service rate values increase in $R/C$, and decrease in $k$ and $F$ values.

6. Conclusion

In this paper, we analyzed the pricing decision of the service provider. The system is designed as $M/M/1$ queues with rational customers who are unable to observe the length of the queue prior to making decisions. These customers wait in the queue to be served. In contrast to previous studies, we assume that waiting in the queue has a greater negative effect on the utility of the customers, compared to waiting during the service. We have shown an increase in the unit cost of waiting in the queue has a negative effect on the utility of the customer. Since customers are rational, the profit maximizer should be decided by considering customer utility. Thus, service provider's pricing decisions are similarly negatively affected by the customers' unit cost of waiting in the queue.

We have analyzed the service provider's pricing problem in both the short and long terms. In the short term, when the service provider optimally sets the entrance price for the given service rate, monopolistic pricing was observed as the most efficient setting for high server utilization periods, i.e. when the system is congested. In the long term, however, when the service provider optimally sets the service rate and entrance price, market capture pricing is found to be the most profitable setting. Hence, in the long term, the service provider's optimal action is to set a service rate that allows all customers to join the queue and receive the service at the minimum price.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments.

References


Görkem Sarıyer has been an Assistant Professor in the Faculty of Economics and Administrative Sciences in Yasar University. She has a BA in Software Engineering (Izmir University of Economics/Turkey), M.Sc. in Applied Statistics (Izmir University of Economics/ Turkey), and PhD degree in Industrial Engineering and Operations Management (Koç University/ Turkey). Her research interest includes decision making models of strategic customers, health care services and management, modeling in queues, service operations and management.