Study of Client Reject Policies under Lead-Time and Price Dependent Demand

Abduh Sayid Albana, Yannick Frein, Ramzi Hammami

The Interest in Client Reject Policy under the Lead-Time and Price Dependent Demand System

Abduh Sayid Albana[a,*], Yannick Frein[a,c], Ramzi Hammami[b]
[a]*Université Grenoble Alpes, G-SCOP, F-38000 Grenoble, France*
[b]*ESC Rennes School of Business, 35065 Rennes, France*
[c]*CNRS, G-SCOP, F-38000 Grenoble, France*

Abstract - *Delivery lead-time has become a factor of competitiveness for companies and an important criterion of purchase for the customers today. This is particularly true in the context B-to-B where the customers are always in search of shorter delivery lead-times to reduce the uncertainties on the supply and have more flexibility during the signing of the orders. Some research indicates that companies risk to lose markets if they are not capable of respecting the delivery lead-times imposed by the customers. The demand of the customers is thus very sensitive to the delivery lead-times. And many research has been done in the demand that depend in the lead time since 1998. However, all of those researchers use queuing theory M/M/1 approach. This M/M/1 approach isn't suitable for the real case. In the real case customer have a tendency to leave if the firms quoted long lead-time. In another word, the costumer have a tendency to balk if the queue is too long. Therefore, the use of the M/M/1/K is proposed to accommodate the balking behavior of those customer.*

Keywords: *Lead-time quotation, Pricing, M/M/1/K, M/M/1.*

1. Introduction

The lead-time, which represents the elapsed time between the placement of the order by the customer and the receipt of this order, has also became a factor of competitiveness for companies and an important criterion of purchase for the customers today. A lead-time promise that is too optimistic, in order to lure the customer into placing a firm order, will increase congestion in the facility and if the overly optimistic promise results in late delivery, this may affect the firm's reputation for on-time delivery and deter future customers (Slotnick, 2014). In addition, companies risk to lose markets if they are not capable of respecting the delivery lead-times imposed by the customers (Kapuscinski and Tayur, 2007). It is generally agreed that timely customer service is also a major determinant in gaining competitive advantage in today's markets. Geary and Zonnenberg (2000) reported that top performers among 110 organizations conducted initiatives not only to reduce costs and maintain reliability, but also to improve delivery speed and flexibility. Lawton (2007) reported that in its competition to win customers from Dell, Hewlett Packard responded to customer complaints about late deliveries by improving its delivery performance by 30%, as well as informing customers about availability of products on order. Baker et al. (2001) also found that less than 10% of end consumers and less than 30% of corporate customers base their purchasing decisions on price only; for a substantial majority of purchasers both price and the delivery lead time are crucial factors that determine their purchase decisions. Thus, in order to improve consumer's satisfaction and their competitiveness, firms should focus on fast, reliable delivery as well as sale price.

In the past 2 decades, a considerable number of researchers in economics and different functional areas of business have studied the lead time- and price-dependent demand. These researchers have developed a variety of mathematical functions to characterize the demand (Huang et al., 2013). Huang et al. (2013) suggests that there may exist further research opportunities for using lead time–dependent demand as they had found that a large number of publications belong to the categories of price- and quality-dependent demand models, but fewer publications belong to lead time–dependent categories.

In lead time- and price-dependent domain, starting from Palaka et al. (1998), they present a model to study the lead-time setting, capacity utilization, and pricing decisions of a profit maximizing firm that incurs congestion related costs and faces demand which is dependent upon both price and lead-time. One of their objective was to develop a model that captures the full profit impact of stochasticity in demand arrivals as well as the production process when customers are delay sensitive and the firm incurs congestion costs and lateness penalties. They considered a firm that serves customers in a make-to-order fashion. In their research, they restricted the attention to a short time horizon, and hence capacity was assumed to be constant while price, quoted lead-time, and demand (or, equivalently, capacity utilization) were treated as decision variables. Customers are served on a first come-first served basis. They assumed that the arrival pattern of customers can be described by a Poisson process. Further, the processing times of the customer orders were assumed to be exponentially distributed. These assumptions allow them to use an M/M/l model of the firm's operations.

Following Palaka et al. (1998), many research were done based on M/M/l approach in modeling the system of the firms in order to capture the interrelationships among pricing, lead-time quotation, and demand. So and Song (1998) develop an analytical framework for a firm to understand the strong interrelationships among pricing, delivery time guarantees, demand, and the overall profitability of offering the services. Different from Palaka et al. (1998) who studied the impact of quoted lead time and pricing on demands where a linear demand function was used, So and Song (1998) use the log linear function to model the demand as a function of both price and delivery time where demand also depends on the delivery reliability level. Pekgün et al. (2008) who studied centralization and decentralization of pricing decisions and lead-time decisions of a Make-To-Order (MTO) firm, used the M/M/l of Palaka et al. (1998) to model their centralized policy. There are many other research that use the M/M/1 to model their MTO system in order to model the lead-time- and price-dependent demand. (see Ho and Zheng, 2004; Liu et al., 2007; Ray and Jewkes, 2004; Zhao et al., 2012).

As reported by Huang et al. (2013) in their literature review study. From 1998 to 2013, all of the researchers used the M/M/1 to model the system. One of the characteristics of the M/M/1 is accepting all client. No clients are rejected even when lead-time is very long. This all-client acceptance policy evokes a question whether it is the optimal firm’s policy or not. Hence, we propose an alternative, the customer rejection policy. We model this rejection policy as M/M//1/K system. We will study whether this new policy is more profitable or not by comparing with the existing M/M/1.

2. Literature Review

In this section, we will explain about the queuing systems in general including M/M/1 & M/M/1/K system, then related work that use the M/M/1 to model their MTO system in order to capture the interrelationships among pricing, lead-time quotation, and demand.

2.1 The Queuing Systems

A queueing system can be described as customers arriving for service, waiting for service if it is not immediate, and if having waited for service, leaving the system after being served (Gross et al., 2008). In most cases, six basic characteristics of queueing processes provide an adequate description of a queueing system: (1) arrival pattern of customers, (2) service pattern of servers, (3) queue discipline, (4) system capacity, (5) number of service channels, and (6) number of service stages.

In usual queueing situations, the process of arrivals is stochastic, and it is thus necessary to know the probability distribution describing the times between successive customer arrivals (inter-arrival times). Similar to the arrival pattern, a probability distribution is needed to describe the sequence of customer service times.

Queue discipline refers to the manner in which customers are selected for service when a queue has formed. The most common discipline that can be observed in everyday life is first come, first served (FCFS). However, this is certainly not the only possible queue discipline. Some others are last come, first served (LCFS), selection for service in random order (RSS); and a variety of *priority* schemes.

In some queueing processes there is a physical limitation to the system capacity such as the amount of waiting room, so that when the line reaches a certain length, no further customers are allowed to enter until space becomes available as the result of a service completion. These are referred to as finite queueing situations; that is, there is a finite limit to the maximum system size. A queue with limited waiting room can be viewed as one with forced balking where a customer is forced to balk if it arrives when the queue size is at its limit. This is a simple case, since it is known exactly under what circumstances arriving customers must balk.

Number of service channels can be multi-server or a single line. And for the number of service stages, a queueing system may have only a single stage of service or several stages. An example of a multistage queueing system would be a physical examination procedure, where each patient must proceed through several stages, such as medical history.

Generally, there are three types of performance measurement of systems: (1) the waiting time that a typical customer might be forced to endure; (2) length of queue that indicate the manner in which customers may accumulate; and (3) the idle time of the servers.

There are several type of simple queueing models, such as M/M/1 and M/M/1/K. The M/M/1 is a queueing system with one server and an infinite queue where the inter-arrival and the service times have exponential probability densities. The average time between arriving customers is $1/\lambda$ and the average service time is $1/\mu$ (Thomopoulos, 2012). This could be cars arriving to a drive-through lane at a fast-food restaurant during the morning hours. While, the M/M/1/K is a system with one server and a finite queue where the maximum number of units allowed in the system is K, and where the inter-arrival and the service times have exponential probability densities. Further, the average time between arriving customers is $1/\lambda$ average service time is $1/\mu$ (Thomopoulos, 2012). A finite queue example is a three line telephone service with one operator receiving information from a caller, while two lines are open for customers waiting to talk to the operator. When all three lines are full, potential new calls are lost.

2.2 The Use of M/M/1 under Lead-Time- and Price-Dependent Demand

There are several research that use the M/M/1 to model the system in order to capture the relationship between demand, price and lead time. Some of them are : Palaka et al. (1998), So and Song (1998), and Pekgün et al. (2008). In their cases, they consider a firm that serves customers in a make-to-order fashion. Price, quoted lead-time, and demand are treated as decision variables. The customers are delay sensitive and hence the expected demand is a function of quoted lead-time as well as price. Based on Huang et al. (2013), this can be formulated as several type of model. Customers are served on a first come-first served basis (FCFS). The arrival pattern of customers were described by a Poisson process. Further, the processing times of the customer orders were assumed to be exponentially distributed. These assumptions allow to use an M/M/l model of the firm's operations. Employing an M/M/I queueing model to study the queueing effects of capacity and demand volume decisions is a traditional abstraction and makes the problem tractable without a significant loss in accuracy, especially in deriving qualitative patterns and managerial insights (Palaka et al., 1998).

Since the system is assumed as an M/M/l queueing system with mean service rate, p, and mean arrival rate (or, demand), λ, the expected number of customers in the system, L_s, is given by $L_s = \lambda/(\mu - \lambda)$ and the actual lead-time (time in the system) is exponentially distributed with mean $1/(\mu - \lambda)$ (Kleinrock, 1975). The probability that the firm is not able to meet the quoted lead-time, l, is given by $e^{-(\mu-\lambda)l}$ and the expected lateness of a late job is $1/(\mu - \lambda)$ - the same as the expected lead-time due to the memoryless property of the exponential distribution.

Those model of the firm is a good representation of any manufacturing firm that fills custom orders on a first come-first served basis. An example of this type of firm would be mail order companies that sell personal computers. The model is also suitable for representing service firms that incur congestion related costs. Examples of both manufacturing and service firms that offer compensation for late orders can be found in the literature. See for example, So and Song (1998).

As stated by Gross et al. (2008); Kleinrock (1975); Thomopoulos (2012), one of the characteristics of M/M/1 is the infinite system capacity. Thus, all clients are accepted. This policy can cause a long lead-time. Several authors handle this issue by adding a compensation for late delivery, such as Palaka et al. (1998); So and Song, (1998). This became a question whether this is a better policy than rejecting some clients or not. Hence, we propose clients rejection policy as an alternative to the existing modeling the M/M/1. The details of this new approach are explained in the section 3 and we will compare it with the existing M/M/1 in section 5 and section 7.

3. Proposed Model (The M/M/1/K)

As in Palaka et al. (1998), We consider a make-to-order firm where the capacity is assumed to be constants while price, quoted lead-time and demand are the decision variables. Customers are served in first-come, first-served basis (FCFS). The arrival proses are assumed to be Poisson process as well as the processing times of the customer are assumed as exponentially distributed. Yet, the system has capacity where the maximum number of customer allowed in the system is K. This assumption allow us to model the systems as M/M/1/K.

The customers are delay sensitive and the expected demand is a function of quoted lead time and price. As in Liu et al. (2007); Palaka et al. (1998); and Pekgün et al., (2008), the demand is assumed to be linear function of price and quoted lead-time.

$$\Lambda(p,l) = a - b_1 p - b_2 l, \tag{1}$$

where:

p = price of the good/service set by the firm,

l = quoted lead-time,

$\Lambda(p,l)$ = expected demand for the good/service at price p and quoted lead-time l,

a = demand corresponding to zero price and zero quoted lead-time,

b_1 = price sensitivity of demand,

b_2 = lead-time sensitivity of demand,

Since the demand is downward sloping in both price and quoted lead-time, b_1 and b_2 are restricted to be non-negative (Palaka et al., 1998).

According to Palaka et al. (1998), this linear demand function is tractable and has several desirable properties. The first desirable properties is the price elasticity of demand is increasing in both price and quoted lead-time. Customers would be more sensitive to long lead-times when they are paying more for the goods or service. Similarly, customers would be more sensitive to high prices when they also have longer waiting times. It is also noteworthy that the more complex Cobb-Douglas and exponential demand functions do not exhibit these desirable properties. Second desirable property of the linear demand function is separability of price and quoted lead-time. This separability is desirable because an underlying premise of time-based competition is that customers perceive time and money as substitutes, which would imply that the demand function should be separable in price and quoted lead-time.

As in the case of Palaka et al. (1998), to prevent the firms from quoting unrealistically short lead-times, we assume that the firm maintains a certain minimum service level. The service level is defined as the probability of meeting the quoted lead-time ($P\ (W \le l) \ge s$).

Since we assume an M/M/l/K queueing system with mean service rate, μ, and mean arrival rate (or, demand), λ, throughput rate (effective demand), $\overline{\lambda}$, the expected number of customers in the system, L_s (see eq. (2)), and the actual lead-time (time in the system) is exponentially distributed with mean $L_s / \overline{\lambda}$ (see eq. (3)). The probability that the firm is not able to meet the quoted lead-time l ($P\ (W > l)$) is given in eq. (4). The expected lateness of a late job is given in eq. (3), as stated by Palaka et al. (1998), the same as the expected lead-time due to the memoryless property of the exponential distribution The expected lateness of a late job.

$$L_s = \frac{\rho}{1-\rho} - \frac{(K+1)\rho^{K+1}}{1-\rho^{K+1}} \text{ with } \rho = \frac{\lambda}{\mu} \tag{2}$$

$$W = \frac{L_s}{\overline{\lambda}} \text{ with } \overline{\lambda} = \lambda(1-P_K) \text{ and } P_K = \frac{1-\rho}{1-\rho^{K+1}}\rho^K \tag{3}$$

$$P(W > l) = \sum_{k=0}^{K-1}\left(\sum_{i=0}^{k}\frac{(\mu \cdot l)^i}{i!}e^{-\mu \cdot l}\right)\frac{P_k}{1-P_K} \tag{4}$$

The objective of the firm's is to maximize their revenue which includes the following three main parts:

(1) **Expected revenue (net of direct costs)** is represented by $\overline{\lambda}(p-m)$, where m is the unit direct variable cost. Direct variable costs include all costs that are proportional to production volume, such as the costs of direct materials and direct labor.

(2) **Total Congestion costs** is expressed as the mean number of jobs in the system multiplied by the unit holding cost ($L_s \times F$). The congestion costs predominantly is the costs of maintaining WIP including warehousing and storage costs, materials handling expenses, insurance costs, capital opportunity costs, quality expenses arising from deteriorating in-process inventory, and other relevant inventory holding expenses..

(1) **Total Lateness penalty costs** is expressed as (penalty per job per unit lateness) × (throughput rate) × (probability that a job is late) × (expected lateness given that a job is late). Penalty cost

per job per unit lateness (c) reflects direct compensation paid to customers for not meeting the quoted lead-time as well as the cost of expediting orders to meet the due dates.

Further, we can write the optimization of lead-time- and price- dependent demand problem using the M/M/1/K as:

(P0) $$\underset{l,p,\lambda}{\text{Maximize}} \quad \bar{\lambda}(p-m)-(L_s \times F)-\left(c \times \bar{\lambda} \times P(W \geq l) \times W\right) \tag{5}$$

$$\text{Subject to} \quad \lambda \leq a - b_1 p - b_2 l \tag{6}$$

$$1-\sum_{k=0}^{K-1}\left(\sum_{i=0}^{k}\frac{(\mu \cdot l)^i}{i!}e^{-\mu \cdot l}\right)\frac{P_k}{1-P_K} \geq s \tag{7}$$

$$\rho = \frac{\lambda}{\mu} \tag{8}$$

$$P_K = \frac{1-\rho}{1-\rho^{K+1}}\rho^K \tag{9}$$

$$\bar{\lambda} = \lambda(1-P_K) \tag{10}$$

$$\rho \neq 1 \tag{11}$$

$$\lambda, p, l, \bar{\lambda} \geq 0 \tag{12}$$

Where,

		Decision Variables			***Parameters***
p	=	Price of the good/service set by the firm,	a	=	Demand corresponding to zero price and zero quoted lead-time,
l	=	Quoted lead-time,	b_1	=	Price sensitivity of demand,
λ	=	Mean arrival rate (demand),	b_2	=	Lead-time sensitivity of demand,
			μ	=	Mean service rate (Production capacity),
			m	=	Unit direct variable cost,
			s	=	Service level defined by company,
			P_K	=	Probability of balking customer,
			K	=	System capacity.

In this formulation, constraint (6) requires that the mean demand (λ) served by firm does not exceed the demand generated by price (p) and quoted lead-time (l). Constraint (7) expresses the lower bound on the firm service level. Constraint (8) is the formulation of ρ as stated by Hillier and Lieberman (2001). Constraint (9) is probability of balking customer. Constraint (10) is the number of customers that are served and exit the system. Constraint (11) is the condition to use constraint (9) in M/M/1/K (Hillier and Lieberman, 2001). Constraint (12) is the non-negativity constraint.

In the following section (section 4), we start by modeling the system with K = 1 and without penalty cost and holding cost. Then in the section 5, we will investigate whether the new model can be better or not by comparing with existing M/M/1 models. In section 6, we will add back the penalty cost and holding cost in modeling the system with K = 1. Section 7 will be the comparison of the new model (with penalty and holding cost) with the existing M/M/1.

4. M/M/1/1 without Penalty Cost and Holding Cost

In this section, we model the system with K = 1. Since the unit holding cost and the penalty cost are removed, the objective function will be only to maximize the expected revenue. The Probability $(W > l) = e^{-\mu l}$ (proven in Appendix A). Hence, service level constrain can be written as $1-e^{-\mu l} \geq s$. The formulation of this problem is:

(P1) $$\underset{\lambda,l,p,\bar{\lambda}}{\text{Maximize}} \quad \bar{\lambda}(p-m) \tag{13}$$

$$\text{Subject to} \quad \lambda \leq a - b_1 p - b_2 l \tag{14}$$

$$1-e^{-\mu l} \geq s \tag{15}$$

$$\rho = \frac{\lambda}{\mu} \quad (16)$$

$$P_1 = \frac{\rho}{1+\rho} \quad (17)$$

$$\bar{\lambda} = \lambda(1 - P_1) \quad (18)$$

$$\rho \neq 1 \quad (19)$$

$$\lambda, p, l, \bar{\lambda} \geq 0 \quad (20)$$

Eq. (15) can be rewritten as $\mu l \geq \ln(1/(1-s))$. Then, by integrating the equality constrain (eq. (16-19)) into the objective function, we can simplify the formulation as:

$$\underset{\lambda,l,p}{\text{Maximize}} \quad \left(\frac{\lambda\mu}{\mu+\lambda}\right) \times (p - m) \quad (21)$$

$$\text{Subject to} \quad \lambda \leq a - b_1 p - b_2 l \quad (22)$$

$$\mu l \geq \ln(1/(1-s)) \quad (23)$$

$$\lambda, p, l \geq 0 \quad (24)$$

As in Palaka et al. (1998); Pekgün et al. (2008); So and Song (1998), the service level constraint (eq.(23)) is binding at optimality (see proof in appendix C). Hence, $l = \ln(1/(1-s))/\mu$. Denote $\ln(1/(1-s))$ by z, we get:

$$l = \frac{z}{\mu} \quad (25)$$

As stated by Palaka et al. (1998); and Pekgün et al. (2008), the demand constraint (eq.(22)) is also binding at optimality (see proof in appendix B), thus:

$$\lambda = a - b_1 p - b_2 l \Leftrightarrow p = \frac{a - b_2 l - \lambda}{b_1} \quad (26)$$

Substitute the $l = z/\mu$, into eq. (26):

$$p = \frac{a\mu - b_2 z - \lambda\mu}{\mu b_1} \quad (27)$$

Substituting eq. (27) into the objective function, we get a new problem as:

$$\underset{\lambda}{\text{Maximize}} \left(\frac{\lambda\mu}{\mu+\lambda}\right) \times \left(\frac{a\mu - b_2 z - \lambda\mu}{\mu b_1} - m\right) \quad (28)$$

$$\Leftrightarrow \underset{\lambda}{\text{Maximize}} \left(\frac{\lambda\mu}{\mu+\lambda}\right) \times \left(\frac{a\mu - b_2 z - \lambda\mu - m\mu b_1}{\mu b_1}\right) \quad (29)$$

$$\Leftrightarrow \underset{\lambda}{\text{Maximize}} \left(\frac{\lambda a\mu - \lambda b_2 z - \lambda^2 \mu - \lambda m\mu b_1}{\mu b_1 + \lambda b_1}\right) \quad (30)$$

There is a feasibility condition of this problem as explained in proposition 1.

Proposition 1. *The problem of:* $\underset{\lambda}{\text{Maximize}} \left(\frac{\lambda a\mu - \lambda b_2 z - \lambda^2 \mu - \lambda m\mu b_1}{\mu b_1 + \lambda b_1}\right)$*, only feasible and have the objective > 0 if* $\frac{a\mu - b_2 z}{\mu b_1} \geq m$ *(proof in appendix D).*

To find the solution (optimum point) from eq. (30), first we need to find stationary point. Hence we have to find derive the necessary condition where the first order condition is equal to 0.

$$\frac{\partial}{\partial\lambda}\left(\frac{\lambda a\mu-\lambda b_2 z-\lambda^2\mu-\lambda m\mu b_1}{\mu b_1+\lambda b_1}\right)=0 \tag{31}$$

$$\Leftrightarrow \frac{a\mu^2 b_1-\mu b_1 b_2 z-m\mu^2 b_1^{\,2}-2\lambda\mu^2 b_1-\lambda^2\mu b_1}{\left(\mu b_1+\lambda b_1\right)^2}=0 \tag{32}$$

The denominator of eq. (32) is always positive because $\lambda,\mu,b_1\geq 0$. In another words, numerator of eq. (32) should be equal to zero.

$$a\mu^2 b_1-\mu b_1 b_2 z-m\mu^2 b_1^{\,2}-2\lambda\mu^2 b_1-\lambda^2\mu b_1=0 \tag{33}$$

The discriminant (Δ) of eq. (34) should be greater equal to 0 to have roots.

$$\begin{aligned}\Delta &= \left(-2\mu^2 b_1\right)^2-4\left(-\mu b_1\right)\left(a\mu^2 b_1-\mu b_1 b_2 z-m\mu^2 b_1^{\,2}\right)\\ &=4\mu^4 b_1^{\,2}-4\left(-a\mu^3 b_1^{\,2}+\mu^2 b_1^{\,2} b_2 z+m\mu^3 b_1^{\,3}\right)\\ &=4\mu^4 b_1^{\,2}+4a\mu^3 b_1^{\,2}-4\mu^2 b_1^{\,2} b_2 z-4m\mu^3 b_1^{\,3}\end{aligned}$$

It is proven in appendix E that the discriminant is greater equal to 0 ($\Delta\geq 0$). Hence eq. (34) has two roots which are:

$$\begin{aligned}\lambda &= \frac{-\left(-2\mu^2 b_1\right)\pm\sqrt{\Delta}}{2\left(-\mu b_1\right)}\\ &=\frac{-\left(-2\mu b_1\right)\pm\sqrt{4\mu^4 b_1^{\,2}+4a\mu^3 b_1^{\,2}-4\mu^2 b_1^{\,2} b_2 z-4m\mu^3 b_1^{\,3}}}{2\left(-b_1\right)}\\ &=\frac{2\mu^2 b_1}{-2\mu b_1}\pm\frac{\sqrt{\left(4\mu^2 b_1^{\,2}\right)\times\left(\mu^2+a\mu-b_2 z-m\mu b_1\right)}}{-2\mu b_1}\\ &=-\mu\pm\left(-\sqrt{\mu^2+a\mu-b_2 z-m\mu b_1}\right)\end{aligned}$$

$$\lambda_1=-\mu-\sqrt{\mu^2+a\mu-b_2 z-m\mu b_1}\ \text{ or }\ \lambda_2=-\mu+\sqrt{\mu^2+a\mu-b_2 z-m\mu b_1} \tag{34}$$

If we take a closer look, the $a\mu-b_2 z-m\mu b_1$ is equivalent to λ with $p = m$. Hence, inside the feasibility region $\mu^2+a\mu-b_2 z-m\mu b_1$ will be bigger than 0, thus λ_1 will give us negative value, where λ should be greater than 0. Hence, we only use λ_2. The λ_2 also correspond to the feasibility condition (see appendix F). This roots (λ_2) is the stationary point of the objective function and could be also the optimum point if the objective is a concave function. In appendix G, it is proven that objective function is concave. Hence, the λ_2 is also the optimum point that can be obtained as in proposition 2.

Proposition 2. *For the problem of* $\left(\lambda^*\mu/\left(\mu+\lambda^*\right)\right)\times\left(p^*-m\right)$, *the optimum demand can be found by using equation* $\lambda^*=-\mu+\sqrt{\mu^2+a\mu-b_2 z-m\mu b_1}$ *with* $z=\ln\left(1/\left(1-s\right)\right)$, *the optimum lead-time* $l^*=\ln\left(1/\left(1-s\right)\right)/\mu$ *and optimum price* $p^*=\left(a-b_2 l^*-\lambda^*\right)/b_1$.

5. Comparison with Existing M/M/1

In this section, we compare our modeling (M/M/1/K) with the existing M/M/1 taken from Pekgün et al. (2008). First, we compare the result based on the variation of market potential (a) and lead-time sensitivity (b_2). We use the other parameters as: b_1 = 4; μ = 10; s = 0.95; m = 5.

Table 1 Comparison based on a and b2

b2	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
20	-	40.87%	17.94%	8.29%	3.42%	0.66%

b_2	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
19	-	37.10%	15.27%	6.12%	1.57%	-0.99%
18	-	33.39%	12.61%	3.96%	-0.30%	-2.65%
17	-	29.73%	9.96%	1.79%	-2.18%	-4.33%
16	-	26.13%	7.31%	-0.40%	-4.07%	-6.03%
15	-	22.57%	4.66%	-2.59%	-5.99%	-7.74%
14	-	19.05%	2.01%	-4.80%	-7.92%	-9.48%
13	-	15.57%	-0.65%	-7.03%	-9.87%	-11.24%
12	-	12.11%	-3.32%	-9.29%	-11.86%	-13.04%
11	-	8.68%	-6.02%	-11.57%	-13.88%	-14.87%
10	-	5.26%	-8.74%	-13.90%	-15.94%	-16.74%
9	-	1.85%	-11.49%	-16.28%	-18.06%	-18.67%
8	-	-1.56%	-14.30%	-18.71%	-20.23%	-20.65%
7	-	-4.98%	-17.17%	-21.22%	-22.48%	-22.71%
6	-	-8.43%	-20.13%	-23.83%	-24.83%	-24.86%
5	-	-11.92%	-23.20%	-26.56%	-27.30%	-27.14%
4	-	-15.48%	-26.42%	-29.47%	-29.94%	-29.58%
3	-	-19.14%	-29.88%	-32.63%	-32.83%	-32.25%
2	-	-22.93%	-33.72%	-36.20%	-36.11%	-35.31%
1	-	-26.96%	-38.27%	-40.58%	-40.18%	-39.10%
0	-	-31.37%	-46.41%	-50.00%	-49.07%	-47.47%
	20	30	40	50	60	70
	a					

Note:

-	Problem is infeasible for both M/M/1 and M/M/1/1
-xx.xx%	The M/M/1 is better xx.xx% than M/M/1/1
xx.xx%	The M/M/1/1 is better xx.xx% than M/M/1

As seen in table 1, in a big market potential, the M/M/1 is a better policy. Hence it is better to accept all clients than rejecting them. However, as the customers become more sensitive to lead-time (b_2 increase), the M/M/1 became worse than M/M/1/K. It can be concluded that if the costumer is very sensitive to lead-time, it better to reject some of them to get a better revenue. In small market potential, it is better to reject some clients than accept them all.

Second, we compare based on the variation of market potential (a) and price sensitivity (b_1). The other parameters are: $b_2 = 6$; $\mu = 10$; $s = 0.95$; $m = 5$.

Table 2 Comparison based on a and b_1

b_1	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
20	-	-	-	-	-	-
19	-	-	-	-	-	-
18	-	-	-	-	-	-
17	-	-	-	-	-	-
16	-	-	-	-	-	-
15	-	-	-	-	-	-

b_1	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
14	-	-	-	-	-	-
13	-	-	-	-	-	4.54%
12	-	-	-	-	-	-8.43%
11	-	-	-	-	4.54%	-15.87%
10	-	-	-	-	-8.43%	-20.13%
9	-	-	-	4.54%	-15.87%	-22.51%
8	-	-	-	-8.43%	-20.13%	-23.83%
7	-	-	4.54%	-15.87%	-22.51%	-24.51%
6	-	-	-8.43%	-20.13%	-23.83%	-24.83%
5	-	4.54%	-15.87%	-22.51%	-24.51%	-24.91%
4	-	-8.43%	-20.13%	-23.83%	-24.83%	-24.86%
3	4.54%	-15.87%	-22.51%	-24.51%	-24.91%	-24.72%
2	-8.43%	-20.13%	-23.83%	-24.83%	-24.86%	-24.53%
1	-15.87%	-22.51%	-24.51%	-24.91%	-24.72%	-24.30%
0	-	-	-	-	-	-
	20	30	40	50	60	70
	a					

For a cases that the clients are more sensitive to the price, the M/M/1 is a better policy. Hence it is better to accept all clients regardless the size of the market potential.

Third, we compare based on the variation of market potential (a) and service level (s). The other parameters are: $b_1 = 4$, $b_2 = 6$, $\mu = 10$, $m = 5$.

Table 3 Comparison based on a and s

s	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
0.99	-	2.61%	-10.87%	-15.74%	-17.58%	-18.23%
0.98	-	-2.12%	-14.77%	-19.12%	-20.59%	-20.98%
0.97	-	-4.90%	-17.10%	-21.16%	-22.43%	-22.66%
0.96	-	-6.89%	-18.79%	-22.65%	-23.76%	-23.89%
0.95	-	-8.43%	-20.13%	-23.83%	-24.83%	-24.86%
0.94	-	-9.70%	-21.23%	-24.81%	-25.71%	-25.68%
0.93	-	-10.78%	-22.18%	-25.65%	-26.48%	-26.38%
0.92	-	-11.72%	-23.01%	-26.40%	-27.15%	-27.00%
0.91	-	-12.55%	-23.75%	-27.06%	-27.75%	-27.56%
0.90	-	-13.30%	-24.43%	-27.67%	-28.30%	-28.06%
0.89	-	-13.98%	-25.04%	-28.22%	-28.81%	-28.53%
0.88	-	-14.60%	-25.61%	-28.73%	-29.27%	-28.96%
0.87	-	-15.17%	-26.14%	-29.21%	-29.71%	-29.36%
0.86	-	-15.71%	-26.63%	-29.66%	-30.11%	-29.74%
0.85	-	-16.20%	-27.10%	-30.08%	-30.50%	-30.09%
0.84	-	-16.67%	-27.53%	-30.48%	-30.86%	-30.43%
0.83	-	-17.12%	-27.95%	-30.86%	-31.21%	-30.75%

s	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
0.82	-	-17.53%	-28.35%	-31.22%	-31.54%	-31.05%
0.81	-	-17.93%	-28.72%	-31.56%	-31.86%	-31.35%
0.80	-	-18.31%	-29.08%	-31.90%	-32.16%	-31.63%
	20	30	40	50	60	70
	a					

For this cases, the M/M/1 is a better policy for almost all service level and market potential. There is only one case where M/M/1/K is better. It is when the market level is very small and the service level is near to 1.

Fourth, we compare based on the variation of market potential (a) and production capacity (μ). The other parameters are: $b_1 = 4$, $b_2 = 6$, $s = 0.95$, $m = 5$.

Table 4 Comparison based on a and μ

μ	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
20	-	-13.12%	-24.91%	-32.01%	-35.86%	-37.62%
19	-	-13.26%	-25.20%	-32.11%	-35.65%	-37.16%
18	-	-13.35%	-25.42%	-32.09%	-35.31%	-36.58%
17	-	-13.38%	-25.55%	-31.94%	-34.82%	-35.86%
16	-	-13.34%	-25.57%	-31.62%	-34.17%	-34.99%
15	-	-13.18%	-25.43%	-31.11%	-33.32%	-33.95%
14	-	-12.87%	-25.09%	-30.36%	-32.26%	-32.71%
13	-	-12.34%	-24.48%	-29.32%	-30.93%	-31.23%
12	-	-11.52%	-23.53%	-27.95%	-29.30%	-29.47%
11	-	-10.28%	-22.12%	-26.15%	-27.30%	-27.37%
10	-	-8.43%	-20.13%	-23.83%	-24.83%	-24.86%
9	-	-5.72%	-17.35%	-20.83%	-21.77%	-21.82%
8	-	-1.70%	-13.51%	-16.94%	-17.94%	-18.09%
7	-	4.32%	-8.15%	-11.82%	-13.05%	-13.42%
6	-	13.62%	-0.53%	-4.90%	-6.63%	-7.40%
5	-	28.77%	10.79%	4.86%	2.15%	0.67%
4	-	56.01%	28.86%	19.59%	14.97%	12.21%
3	-	116.15%	61.96%	44.62%	35.87%	30.51%
2	-	359.61%	144.13%	98.95%	78.23%	66.03%
1	-	-	962.02%	359.37%	240.29%	186.72%
0	-	-	-	-	-	-
	20	30	40	50	60	70
	a					

It is shown that regardless the size of the market potential, the M/M/1 is better when the firm has a big production capacity (mean service rate). In the other hand, when the firm has a small production capacity, it is better to use the M/M/1/K.

It is shown that the use of the client rejection policy can be better in some cases even when the penalty and holding cost are removed. Therefore, in the next section, we are going to study the case where the penalty and holding cost are added.

6. M/M/1/1 with Penalty Cost and Holding Cost

In this section we take into account the penalty and holding cost, we consider the system capacity equal to 1 (K = 1). With the addition of penalty cost and holding cost, the objective will be based on three part: expected revenue, total congestion costs, and total lateness penalty costs. The formulation of this three part have been discussed in section 3. The service level of this case is $1-e^{-\mu l} \geq s$ as the Probability$(W > l) = e^{-\mu l}$ (proven in Appendix A). Thus the formulation of this problem is:

$$\text{(P2)}\ \underset{\lambda,p,l,\bar{\lambda}}{Maximize} \quad \bar{\lambda}(p-m) - \frac{F\lambda}{\mu+\lambda} - \frac{c\bar{\lambda}}{\mu} e^{-\mu l} \tag{35}$$

$$\text{Subject to} \quad \lambda \leq a - b_1 p - b_2 l \tag{36}$$

$$1 - e^{-\mu l} \geq s \tag{37}$$

$$\rho = \frac{\lambda}{\mu} \tag{38}$$

$$P_1 = \frac{\rho}{1+\rho} \tag{39}$$

$$\bar{\lambda} = \lambda(1 - P_1) \tag{40}$$

$$\rho \neq 1 \tag{41}$$

$$\bar{\lambda}, \lambda, l, p \geq 0 \tag{42}$$

Integrating the equality constraints (eq. (38 – 41)) to the objective function and rewrite $1-e^{-\mu l} \geq s$ to $\mu l \geq \ln(1/(1-s))$, we get a new problem as:

$$\underset{\lambda,l,p}{\text{Maximize}} \quad \frac{\lambda\left(p\mu - m\mu - F - ce^{-\mu l}\right)}{\mu+\lambda} \tag{43}$$

$$\text{Subject to} \quad \lambda \leq a - b_1 p - b_2 l \tag{44}$$

$$\mu l \geq \ln(1/(1-s)) \tag{45}$$

$$\lambda, p, l \geq 0 \tag{46}$$

As proven in Palaka et al. (1998); and Pekgün et al. (2008), the demand constraint (eq. (44)) is binding at optimality (see demonstration in appendix H). Thus by removing the price (*p*), the problem became:

$$\underset{\lambda,l}{Maximize} \quad \frac{\lambda\left[\left(\mu(a - b_2 l - \lambda)/b_1\right) - m\mu - F - ce^{-\mu l}\right]}{\mu+\lambda} \tag{47}$$

$$\text{Subject to} \quad \mu l \geq \ln(1/(1-s)) \tag{48}$$

$$l, p \geq 0 \tag{49}$$

As in Palaka et al. (1998), the critical value for the service level is defined as $s_c = 1 - b_2/b_1 c$. And this cases are considered as two mutually exclusive cases:

(1) the industry standard service level is strictly less than the critical service level $s < s_c$, and

(2) the industry standard service level is greater than the critical service level $s \geq s_c$.

Those two cases are proven in appendix I. There also is a feasibility condition of this problem as explained in proposition 3.

Proposition 3. *The problem of:* $\underset{\lambda,l,p}{\text{Maximize}} \ \frac{\lambda\left(p\mu - m\mu - F - ce^{-\mu l}\right)}{\mu+\lambda}$ *only feasible if* $\frac{a - b_2 l}{b_1} \geq m$ *and have the objective > 0 if* $a\mu - \mu b_2 l - \mu m b_1 - F b_1 - c b_1 e^{-\mu l} \geq 0$ *(proof in appendix J).*

The lead-time (*l*) in proposition 3 can be found as we proceed in finding the lead-time of the problem. The lead-time can be found based on the two mutually exclusive cases of the service level:

$$1-e^{-\mu l} = Max\{s_c, s\} \quad \Leftrightarrow \quad l^* = \frac{1}{\mu}\ln\left(\frac{1}{1-Max\{s_c, s\}}\right)$$

$$\Leftrightarrow \quad l^* = \frac{1}{\mu}\ln x \text{ where, } x = Max\{1/(1-s), b_1 c/b_2\}.$$

Next, to find the demand (λ), we apply Lagrangian Multiplier method as seen in appendix I. This Lagrangian multiplier method give us a stationary point to problem that must satisfy:

$$\frac{\partial L}{\partial \lambda} = 0, \frac{\partial L}{\partial l} = 0, \ \gamma\frac{\partial L}{\partial \gamma} = 0, \gamma \ge 0 \text{ and } \mu l \ge \ln\left(\frac{1}{1-s}\right)$$

where,

$$L(\lambda, l, \gamma) = \frac{\lambda\left(p\mu - m\mu - F - ce^{-\mu l}\right)}{\mu+\lambda} + \gamma\left\{\mu l - \ln\left(\frac{1}{1-s}\right)\right\} \tag{50}$$

with $\lambda = a - b_1 p - b_2 l$, hence $p = \dfrac{a - b_2 l - \lambda}{b_1}$

$$L(\lambda, l, \gamma) = \frac{\lambda\left[\left(\mu\left(a - b_2 l - \lambda\right)/b_1\right) - m\mu - F - ce^{-\mu l}\right]}{\mu+\lambda} + \gamma\left\{\mu l - \ln\left(\frac{1}{1-s}\right)\right\} \tag{51}$$

If we derive eq. (51) based λ, we can find the demand of this problem.

$$\frac{\partial L}{\partial \lambda} = 0 \quad \Leftrightarrow \quad \frac{\lambda\left[F + ce^{-\mu l} + m\mu + \left(\mu\left(\lambda - a + b_2 l\right)/b_1\right)\right]}{(\mu+\lambda)^2}$$

$$- \frac{F + ce^{-\mu l} + m\mu + \left(\mu\left(\lambda - a + b_2 l\right)/b_1\right)}{\mu+\lambda} - \frac{\mu\lambda}{b_1(\mu+\lambda)} = 0$$

$$\Leftrightarrow \quad \frac{\mu\left(a\mu - \mu b_2 l - 2\mu\lambda - \mu m b_1 - F b_1 - b_1 c e^{-\mu l} - \lambda^2\right)}{b_1(\mu+\lambda)^2} = 0 \tag{52}$$

It is shown that the denominator of eq. (52) is always positive because $\lambda, \mu, b_1 \ge 0$. In another words, numerator of eq. (52) should be equal to zero.

$$\mu\left(a\mu - \mu b_2 l - 2\mu\lambda - \mu m b_1 - F b_1 - b_1 c e^{-\mu l} - \lambda^2\right) = 0 \tag{53}$$

$$a\mu - \mu b_2 l - 2\mu\lambda - \mu m b_1 - F b_1 - b_1 c e^{-\mu l} - \lambda^2 = 0 \tag{54}$$

The discriminant (Δ) of eq. (34) should be greater equal to 0 to have roots.

$$\begin{aligned}\Delta &= (-2\mu)^2 - 4(-1)\left(a\mu - \mu b_2 l - \mu m b_1 - F b_1 - b_1 c e^{-\mu l}\right)\\ &= 4\mu^2 + 4a\mu - 4\mu b_2 l - 4\mu m b_1 - 4F b_1 - 4 b_1 c e^{-\mu l}\end{aligned} \tag{55}$$

In appendix K, it is proven that the discriminant (Δ) is bigger equal to zero. Hence eq. (54) has two roots which are:

$$\begin{aligned}\lambda &= \frac{-(-2\mu) \pm \sqrt{\Delta}}{2(-1)}\\ &= \frac{2\mu \pm \sqrt{4\mu^2 + 4a\mu - 4\mu b_2 l - 4\mu m b_1 - 4F b_1 - 4 b_1 c e^{-\mu l}}}{-2}\\ &= -\mu \pm \sqrt{\mu^2 + a\mu - \mu b_2 l - \mu m b_1 - F b_1 - b_1 c e^{-\mu l}}\\ \lambda_1 &= -\mu - \sqrt{\mu^2 + a\mu - \mu b_2 l - \mu b_1 m - F b_1 - b_1 c e^{-\mu l}} \quad \text{or}\\ \lambda_2 &= -\mu + \sqrt{\mu^2 + a\mu - \mu b_2 l - \mu b_1 m - F b_1 - b_1 c e^{-\mu l}}\end{aligned} \tag{56}$$

The $\mu^2 - F b_1 + a\mu - \mu b_2 l - \mu b_1 m - b_1 c e^{-\mu l}$ is proven to be bigger equal to zero (detail see the prove of discriminant in appendix K). It causes the λ_1 will always have value lower than zero. On the other side, the λ_2 can have value greater than zero. The λ_2 will have a positive value if the proposition 3 is respected. This also correspond to the feasibility and non-negativity of objective function proposition (see proof in appendix L).

The lead-time (l) and demand (λ), that are found by applying Lagrangian Multiplier method, are only candidates for the optimality. It can be the optimum point if the objective function is concave. However, it is proven in appendix M that the objective is concave function. Thus, this points is also the optimum points. As a summary, the optimum point of this problem can be found based on the proposition 4.

Proposition 4. *For the problem of* $\lambda\left(p\mu - m\mu - F - ce^{-\mu l}\right)/(\mu+\lambda)$, *the optimum demand can be found by using equation* $\lambda^* = -\mu + \sqrt{\mu^2 + a\mu - \mu b_2 l - \mu b_1 m - F b_1 - b_1 c e^{-\mu l}}$ *, the optimum lead-time* $l^* = \ln x/\mu$ *with* $x = Max\{1/(1-s), b_1 c/b_2\}$ *and optimum price* $p^* = \left(a - b_2 l^* - \lambda^*\right)/b_1$.

Next based on the result found in this section, we will investigate this rejection policy with penalty and holding cost by comparing with the existing M/M/1 model.

7. Comparison with Existing M/M/1

In this section, we compare our modeling (M/M/1/K) with the existing M/M/1 taken from Palaka et al. (1998). First, we compare the result based on the variation of market potential (a) and lead-time sensitivity (b_2). We use the other parameters as: b_1 = 4; μ = 10; s = 0.95; m = 5; F = 2; c = 10.

Table 5 Comparison based on a and b_2

b_2	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
20	-	53.96%	26.95%	15.50%	9.58%	6.11%
19	-	49.95%	24.23%	13.33%	7.74%	4.49%
18	-	46.02%	21.53%	11.17%	5.90%	2.86%
17	-	42.16%	18.84%	9.02%	4.05%	1.22%
16	-	38.37%	16.17%	6.86%	2.19%	-0.43%
15	-	34.64%	13.51%	4.69%	0.33%	-2.09%
14	-	30.96%	10.85%	2.52%	-1.54%	-3.76%
13	-	27.34%	8.20%	0.34%	-3.43%	-5.45%
12	-	23.77%	5.56%	-1.85%	-5.34%	-7.16%
11	-	20.24%	2.91%	-4.05%	-7.26%	-8.89%
10	-	16.74%	0.25%	-6.27%	-9.21%	-10.65%
9	-	13.28%	-2.42%	-8.52%	-11.19%	-12.43%
8	-	9.84%	-5.10%	-10.80%	-13.19%	-14.25%
7	-	6.42%	-7.81%	-13.11%	-15.24%	-16.10%
6	-	3.01%	-10.56%	-15.47%	-17.33%	-18.01%
5	-	-0.40%	-13.35%	-17.88%	-19.49%	-19.97%
4	-	-3.82%	-16.19%	-20.36%	-21.71%	-22.00%
3	-	-7.26%	-19.12%	-22.93%	-24.02%	-24.12%
2	-	-10.74%	-22.14%	-25.62%	-26.45%	-26.35%
1	-	-14.64%	-25.65%	-28.76%	-29.30%	-28.98%
0	-	-20.37%	-31.10%	-33.75%	-33.86%	-33.21%
	20	30	40	50	60	70
	a					

As seen in table 5, the M/M/1/K is better when dealing with customers that are sensitive to lead-time. The lower the market potential is, the better the M/M/1/K.

Second, we compare based on the variation of market potential (a) and price sensitivity (b_1). The other parameters are: $b_2 = 6$; $\mu = 10$; $s = 0.95$; $m = 5$; $F = 2$; $c = 10$.

Table 6 Comparison based on a and b_1

b_1	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
20	-	-	-	-	-	-
19	-	-	-	-	-	-
18	-	-	-	-	-	-
17	-	-	-	-	-	-
16	-	-	-	-	-	-
15	-	-	-	-	-	-
14	-	-	-	-	-	-
13	-	-	-	-	-	-
12	-	-	-	-	-	26,18%
11	-	-	-	-	42,97%	11,99%
10	-	-	-	-	20,26%	2,93%
9	-	-	-	35,50%	7,00%	-3,27%
8	-	-	-	14,45%	-1,51%	-7,76%
7	-	-	28,29%	2,01%	-7,34%	-11,16%
6	-	-	8,70%	-6,00%	-11,56%	-13,87%
5	-	21,30%	-2,99%	-11,49%	-14,77%	-16,10%
4	-	3,01%	-10,56%	-15,47%	-17,33%	-18,01%
3	14,50%	-8,05%	-15,75%	-18,51%	-19,47%	-19,70%
2	-2,69%	-15,24%	-19,53%	-20,96%	-21,32%	-21,23%
1	-13,22%	-20,20%	-22,45%	-23,04%	-22,98%	-22,67%
0	-	-	-	-	-	-
	20	30	40	50	60	70
	A					

For clients that are sensitive to price, the M/M/1/K could be better when the size of the market potential is small.

Third, we compare based on the variation of market potential (a) and service level (s). The other parameters are: $b_1 = 4$; $b_2 = 6$; $\mu = 10$; $m = 5$; $F = 2$; $c = 10$.

Table 7 Comparison based on a and s

s	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
0,99	-	12,21%	-3,25%	-9,22%	-11,81%	-12,99%
0,98	-	7,90%	-6,63%	-12,10%	-14,35%	-15,29%
0,97	-	5,58%	-8,48%	-13,68%	-15,75%	-16,56%
0,96	-	4,07%	-9,69%	-14,72%	-16,67%	-17,41%
0,95	-	3,01%	-10,56%	-15,47%	-17,33%	-18,01%
0,94	-	2,22%	-11,20%	-16,02%	-17,83%	-18,46%

s	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
0,93	-	1,62%	-11,68%	-16,44%	-18,20%	-18,80%
0,92	-	1,16%	-12,06%	-16,76%	-18,49%	-19,06%
0,91	-	0,81%	-12,34%	-17,01%	-18,71%	-19,26%
0,90	-	0,55%	-12,56%	-17,20%	-18,88%	-19,42%
0,89	-	0,35%	-12,72%	-17,34%	-19,00%	-19,53%
0,88	-	0,21%	-12,84%	-17,44%	-19,09%	-19,61%
0,87	-	0,12%	-12,91%	-17,50%	-19,15%	-19,66%
0,86	-	0,07%	-12,95%	-17,54%	-19,18%	-19,69%
0,85	-	0,06%	-12,97%	-17,55%	-19,19%	-19,70%
0,84	-	0,06%	-12,97%	-17,55%	-19,19%	-19,70%
0,83	-	0,06%	-12,97%	-17,55%	-19,19%	-19,70%
0,82	-	0,06%	-12,97%	-17,55%	-19,19%	-19,70%
0,81	-	0,06%	-12,97%	-17,55%	-19,19%	-19,70%
0,80	-	0,06%	-12,97%	-17,55%	-19,19%	-19,70%
	20	30	40	50	60	70
	a					

For this cases, the M/M/1 is a better policy for almost all service level and market potential. There is only one case where M/M/1/K is better. It is when the market level is very small.

Fourth, we compare based on the variation of market potential (a) and production rate (μ). The other parameters are: $b_1 = 4$; $b_2 = 6$; $s = 0.95$; $m = 5$; $F = 2$; $c = 10$.

Table 8 Comparison based on a and μ

μ	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
20	-	-9,82%	-21,64%	-28,60%	-32,38%	-34,20%
19	-	-9,62%	-21,61%	-28,40%	-31,93%	-33,55%
18	-	-9,32%	-21,47%	-28,06%	-31,33%	-32,75%
17	-	-8,91%	-21,19%	-27,56%	-30,56%	-31,80%
16	-	-8,33%	-20,74%	-26,86%	-29,61%	-30,68%
15	-	-7,55%	-20,07%	-25,93%	-28,43%	-29,35%
14	-	-6,49%	-19,12%	-24,71%	-26,99%	-27,78%
13	-	-5,07%	-17,82%	-23,14%	-25,23%	-25,93%
12	-	-3,15%	-16,06%	-21,16%	-23,10%	-23,74%
11	-	-0,55%	-13,70%	-18,65%	-20,51%	-21,14%
10	-	3,01%	-10,56%	-15,47%	-17,33%	-18,01%
9	-	7,95%	-6,36%	-11,41%	-13,40%	-14,21%
8	-	14,97%	-0,67%	-6,15%	-8,45%	-9,51%
7	-	25,30%	7,18%	0,79%	-2,08%	-3,57%
6	-	41,28%	18,43%	10,30%	6,40%	4,19%
5	-	68,27%	35,52%	24,01%	18,26%	14,82%
4	-	121,78%	64,17%	45,51%	36,16%	30,47%
3	-	277,90%	122,19%	84,63%	66,97%	56,50%

μ	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
2	-	-	312,26%	183,40%	136,67%	111,73%
1	-	-	-	1496,14%	550,83%	363,51%
0	-	-	-	-	-	-
	20	30	40	50	60	70
	a					

It is shown that regardless the size of the market potential, the M/M/1 is better when the firm has a big production capacity (mean service rate). In the other hand, when the firm has a small production capacity, it is better to use the M/M/1/K.

Fifth, we compare based on the variation of market potential (*a*) and holding cost (*F*). The other parameters are: $b_1 = 4$; $b_2 = 6$; $s = 0.95$; $\mu = 10$; $m = 5$; $c = 10$.

Table 9 Comparison based on a and F

F	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
20	-	99,96%	55,78%	37,64%	28,05%	22,22%
19	-	93,10%	51,75%	34,63%	25,57%	20,07%
18	-	86,50%	47,78%	31,64%	23,10%	17,94%
17	-	80,15%	43,88%	28,67%	20,64%	15,80%
16	-	74,01%	40,04%	25,72%	18,18%	13,66%
15	-	68,09%	36,26%	22,80%	15,73%	11,52%
14	-	62,34%	32,53%	19,88%	13,29%	9,37%
13	-	56,77%	28,84%	16,99%	10,84%	7,22%
12	-	51,36%	25,19%	14,10%	8,39%	5,06%
11	-	46,09%	21,57%	11,21%	5,93%	2,89%
10	-	40,94%	17,99%	8,33%	3,46%	0,70%
9	-	35,92%	14,43%	5,44%	0,98%	-1,51%
8	-	31,00%	10,88%	2,55%	-1,52%	-3,74%
7	-	26,18%	7,35%	-0,36%	-4,05%	-6,00%
6	-	21,44%	3,81%	-3,30%	-6,60%	-8,30%
5	-	16,76%	0,27%	-6,26%	-9,20%	-10,63%
4	-	12,14%	-3,30%	-9,27%	-11,84%	-13,02%
3	-	7,56%	-6,90%	-12,33%	-14,55%	-15,48%
2	-	3,01%	-10,56%	-15,47%	-17,33%	-18,01%
1	-	-1,55%	-14,29%	-18,70%	-20,22%	-20,64%
0	-	-6,13%	-18,14%	-22,07%	-23,25%	-23,41%
	20	30	40	50	60	70
	a					

As the holding cost became more expensive, it is better to reject some clients. It is shown in table 9. there is a limit of holding cost where we should accept or reject clients for the optimal firm policy.

Sixth, we compare based on the variation of market potential (*a*) and penalty cost (*c*). The other parameters are: $b_1 = 4$; $b_2 = 6$; $s = 0.95$; $\mu = 10$; $m = 5$; $F = 2$.

Table 10 Comparison based on a and c

c	M/M/1					
	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1	M/M/1/1
20	-	5,28%	-8,72%	-13,89%	-15,93%	-16,73%
19	-	5,05%	-8,90%	-14,04%	-16,07%	-16,86%
18	-	4,83%	-9,09%	-14,20%	-16,21%	-16,98%
17	-	4,60%	-9,27%	-14,36%	-16,35%	-17,11%
16	-	4,37%	-9,45%	-14,52%	-16,49%	-17,24%
15	-	4,14%	-9,64%	-14,67%	-16,63%	-17,37%
14	-	3,92%	-9,82%	-14,83%	-16,77%	-17,49%
13	-	3,69%	-10,00%	-14,99%	-16,91%	-17,62%
12	-	3,46%	-10,19%	-15,15%	-17,05%	-17,75%
11	-	3,23%	-10,37%	-15,31%	-17,19%	-17,88%
10	-	3,01%	-10,56%	-15,47%	-17,33%	-18,01%
9	-	2,78%	-10,74%	-15,63%	-17,48%	-18,14%
8	-	2,55%	-10,93%	-15,79%	-17,62%	-18,27%
7	-	2,32%	-11,11%	-15,95%	-17,76%	-18,40%
6	-	2,09%	-11,30%	-16,11%	-17,90%	-18,53%
5	-	1,87%	-11,48%	-16,27%	-18,05%	-18,66%
4	-	1,64%	-11,67%	-16,43%	-18,19%	-18,79%
3	-	1,41%	-11,85%	-16,59%	-18,33%	-18,92%
2	-	1,18%	-12,04%	-16,75%	-18,48%	-19,05%
1	-	0,96%	-12,23%	-16,91%	-18,62%	-19,18%
0	-	0,73%	-12,41%	-17,07%	-18,76%	-19,31%
	20	30	40	50	60	70
	a					

For a small size market potential, it is better to reject some clients regardless the penalty cost. For big market, accept all clients would be the best policy for the firms.

8. Conclusion

In this paper, we model a make-to-order firm under lead-time- and price-dependent demand as M/M/1/K. The system has capacity K and the customers are lead-time- and price-sensitive to demand. We compare our model with the existing model taken from Pekgün et al. (2008) and Palaka et al. (1998). In the case where the penalty and holding cost aren't considered, the M/M/1/K is better when the customers are lead-time sensitive, the market potential is small, and the firm's production capacity (mean service rate) is small. In the case where the penalty and holding cost are considered, the M/M/1/K is better when the customers are lead-time sensitive, the market potential is small, the firm's production capacity (mean service rate) is small and the holding cost is high.

An extension of this research which is $K > 1$ is still undergoing research. Another extension is modeling the system in M/D/1.

Reference

Baker, W., Marn, M., Zawada, C., 2001. Price smarter on the Net. Harv. Bus. Rev. 79, 122–127.

Geary, S., Zonnenberg, J.P., 2000. What It Means To Be Best In Class. Supply Chain Manag. Rev. 4, 43–48.

Gross, D., Shortle, J.F., Thompson, J.M., Harris, C., 2008. Fundamentals of queueing theory, Fourth Edi. ed. John Wiley & Sons, Inc., Hoboken, New Jersey.

Hillier, F., Lieberman, G., 2001. Introduction To Operations Research.
Ho, T.H., Zheng, Y.-S., 2004. Setting Customer Expectation in Service Delivery: An Integrated Marketing-Operations Perspective. Manage. Sci. 50, 479–488. doi:10.1287/mnsc.1040.0170
Huang, J., Leng, M., Parlar, M., 2013. Demand Functions in Decision Modeling : A Comprehensive Survey and Research Directions. Decis. Sci. 44, 557–609. doi:10.1111/deci.12021
Kapuscinski, R., Tayur, S., 2007. Reliable Due-Date Setting in a Capacitated MTO System with Two Customer Classes. Oper. Res. 55, 56–74. doi:10.1287/opre.1060.0339
Kleinrock, L., 1975. QUEUEING SYSTEMS, VOLUME 1: THEORY, VOLUME 1: ed, John Wiley & Sons, Inc. John Wiley & Sons, Inc., New York. doi:10.1002/net.3230060210
Lawton, C., 2007. How H-P Reclaimed Its PC Lead Over Dell 1–5.
Liu, L., Parlar, M., Zhu, S.X., 2007. Pricing and Lead Time Decisions in Decentralized Supply Chains. Manage. Sci. 53, 713–725. doi:10.1287/mnsc.1060.0653
Palaka, K., Erlebacher, S., Kropp, D.H., 1998. Lead-time setting, capacity utilization, and pricing decisions under lead-time dependent demand. IIE Trans. 30, 151–163. doi:10.1080/07408179808966447
Pekgün, P., Griffin, P.M.P., Keskinocak, P., 2008. Coordination of marketing and production for price and leadtime decisions. IIE Trans. 40, 12–30. doi:10.1080/07408170701245346
Ray, S., Jewkes, E.M., 2004. Customer lead time management when both demand and price are lead time sensitive. Eur. J. Oper. Res. 153, 769–781. doi:10.1016/S0377-2217(02)00655-0
Slotnick, S.A., 2014. Lead-time quotation when customers are sensitive to reputation. Int. J. Prod. Res. 52, 713–726. doi:10.1080/00207543.2013.828176
So, K.C., Song, J.-S., 1998. Price, delivery time guarantees and capacity selection. Eur. J. Oper. Res. 111, 28–49. doi:10.1016/S0377-2217(97)00314-7
Sztrik, J., 2011. Basic Queueing Theory. University of Debrecen: Faculty of Informatics, Debrecen.
Thomopoulos, N.T., 2012. Fundamentals of Queuing Systems: Statistical Methods for Analyzing Queuing Models. Springer Science & Business Media.
Zhao, X., Stecke, K.E., Prasad, A., 2012. Lead Time and Price Quotation Mode Selection: Uniform or Differentiated? Prod. Oper. Manag. 21, 177–193. doi:10.1111/j.1937-5956.2011.01248.x

Appendix

A. The probability *(W > l)*)

In Palaka et al. (1998), So and Song (1998), and Pekgün et al. (2008), Service level constrain is defined as the probability that the firm is able to meet the quoted lead-time (P(W $\le l$) ≥ s). And Sztrik (2011) formulate P(W $\le l$) for the M/M/1/K as

$$P(W \le l) = 1 - \sum_{k=0}^{K-1}\left(\sum_{i=0}^{k}\frac{(\mu \cdot l)^i}{i!}e^{-\mu \cdot l}\right)\frac{P_k}{1-P_K}$$

For the problem of the M/M/1/1, the K equal to 1 ($K = 1$). Hence P(W $\le l$) for the M/M/1/1

$$1 - \sum_{k=0}^{K-1}\left(\sum_{i=0}^{k}\frac{(\mu \cdot l)^i}{i!}e^{-\mu \cdot l}\right)\frac{P_k}{1-P_K} = 1 - \sum_{k=0}^{0}\left(\sum_{i=0}^{k}\frac{(\mu \cdot l)^i}{i!}e^{-\mu \cdot l}\right)\frac{P_k}{1-P_K}$$

$$= 1 - \left(e^{-\mu l}\right)\frac{P_0}{1-P_1}$$

With,

$$P_0 = \frac{1-\rho}{1-\rho^{K+1}} = \frac{1}{1+\rho} \text{ and } P_K = P_1 = \frac{1-\rho}{1-\rho^{K+1}}\rho^K = \frac{\rho}{1+\rho}$$

Hence,

$$1-\left(e^{-\mu\cdot l}\right)\frac{P_0}{1-P_1}=1-e^{-\mu l}\left(\left(\frac{1}{1+\rho}\right)\Big/\left(1-\frac{\rho}{1+\rho}\right)\right)$$

$$=1-e^{-\mu l}\left(\left(\frac{1}{(1+\rho)}\right)\times\left(\frac{(1+\rho)}{1}\right)\right)$$

$$=1-e^{-\mu l}$$

Therefore, the service level constrain for the M/M/1/1 is $1-e^{-\mu l}\geq s$

B. Proof demand constraint is binding

As in Palaka et al. (1998) and Pekgün et al. (2008), suppose that the optimal solution is given by price, p^*, quoted lead-time l^*, and demand rate λ^*, and that $\lambda^* < \Lambda(p^*,1^*.)$. Since the revenues are non-decreasing in p, one could increase the price to p' (while holding the demand rate and quoted lead-time constant) until $l^* = \Lambda(p',1')$. This change will increase revenues without increasing the queueing costs, direct variable costs, and lateness penalties. Therefore, (p^*, l^*, and λ^*) cannot be an optimal solution.

To demonstrate:

$$\underset{\lambda,l,p,\bar{\lambda}}{\text{Maximize}} \quad \bar{\lambda}(p-m)$$

$$\text{Subject to} \quad \lambda\leq a-b_1p-b_2l$$

$$1-\sum_{k=0}^{K-1}\left(\sum_{i=0}^{k}\frac{(\mu\cdot l)^i}{i!}e^{-\mu\cdot l}\right)\frac{P_k}{1-P_K}\geq s$$

Assume: $\lambda^*\leq a-b_1p^*-b_2l^*$, we will have $F^*=\bar{\lambda}^*(p^*-m)$. If we increase the price to p', then we will have $F'=\bar{\lambda}^*(p'-m)$ which is bigger than the F^*. Hence, there is a contradiction which is mean that the constraint is binding.

C. Proof Service level constraint is binding

The service constraint must be binding at optimality as in Palaka et al. (1998) and Pekgün et al. (2008). For the M/M/1/1, applying the Lagrangian multiplier method, we see that a stationary point to problem (P1) must satisfy:

$$\frac{\partial L}{\partial\lambda}=0,\ \frac{\partial L}{\partial l}=0,\ \gamma\frac{\partial L}{\partial\gamma}=0,\ \gamma\geq 0 \text{ and } \mu l\geq\ln\left(\frac{1}{1-s}\right)$$

Where,

$$L(\lambda,l,\gamma)=\left(\frac{\lambda\mu}{\mu+\lambda}\right)\times(p-m)+\gamma\left\{\mu l-\ln\left(\frac{1}{1-s}\right)\right\}$$

With $\lambda=a-b_1p-b_2l$, hence $p=\dfrac{a-b_2l-\lambda}{b_1}$

$$L(\lambda,l,\gamma)=\left(\frac{\lambda\mu}{\mu+\lambda}\right)\times\left(\frac{a-b_2l-\lambda-mb_1}{b_1}\right)+\gamma\left\{\mu l-\ln\left(\frac{1}{1-s}\right)\right\}$$

$$\frac{\partial L}{\partial l}=0 \quad\Leftrightarrow\quad \gamma\mu-\frac{\mu\lambda b_2}{b_1(\mu+\lambda)}=0$$

$$\gamma\mu=\frac{\mu\lambda b_2}{b_1(\mu+\lambda)}$$

$$\gamma=\frac{\lambda b_2}{b_1(\mu+\lambda)}$$

$$\gamma \frac{\partial L}{\partial \gamma} = 0 \qquad \Leftrightarrow \qquad \gamma \left\{ \mu l - \ln\left(\frac{1}{1-s} \right) \right\} = 0$$

We know that $\gamma = \lambda b_2 / \left(b_1(\mu + \lambda)\right)$ with $\lambda, \mu \geq 0$. Hence, $\gamma \neq 0$. Thus, $\mu l - \ln\left(1/(1-s)\right) = 0$ which means service level constraint is binding.

D. Feasibility Condition and Non-Negative Objective Function

Note that: $\lambda = a - b_1 p - b_2 l$ from appendix B and $\lambda \geq 0$, hence:

$$a - b_1 p - b_2 l \geq 0 \Leftrightarrow a - b_2 l \geq b_1 p$$
$$\frac{(a - b_2 l)}{b_1} \geq p$$

Note that: $1 - e^{-(\mu)l} = s$ from appendix C

$$1 - e^{-(\mu)l} = s \quad \Leftrightarrow \mu l = \ln\left(1/(1-s)\right)$$
$$\Leftrightarrow l = \frac{\ln\left(1/(1-s)\right)}{\mu}$$
$$\Leftrightarrow l = \frac{z}{\mu}$$

Equation above imply that

$$\frac{(a - b_2 l)}{b_1} \geq p \Leftrightarrow \frac{a\mu - b_2 z}{\mu b_1} \geq p \Leftrightarrow \text{The feasibility condition}$$

Note that $p \geq m$ to have the positive objective function. It implies that:

$$\frac{a\mu - b_2 z}{\mu b_1} \geq m$$

E. Proof that Discriminant ≥ 0

Suppose that $\Delta \geq 0$

$$4\mu^4 b_1^{\ 2} + 4a\mu^3 b_1^{\ 2} - 4\mu^2 b_1^{\ 2} b_2 z - 4m\mu^3 b_1^{\ 3} \geq 0$$
$$\left(4\mu^2 b_1^{\ 2}\right) \times \left(\mu^2 + a\mu - b_2 z - m\mu b_1\right) \geq 0$$
$$\mu^2 + a\mu - b_2 z - m\mu b_1 \geq 0 \text{ with } z = \mu l$$
$$\mu^2 + a\mu - b_2 \mu l - m\mu b_1 \geq 0$$
$$\mu + a - b_2 l - m b_1 \geq 0$$

The $a - b_2 l - m b_1$ is equivalent to λ with $p = m$ and λ is non-negative. Hence, it is proven that $\mu + a - b_2 l - m b_1 \geq 0$. Thus, the $\Delta \geq 0$

F. The λ_2 correspond to the feasibility condition.

Suppose

$$\lambda_2 \geq 0 \Leftrightarrow \qquad -\mu + \sqrt{\mu^2 + a\mu - b_2 z - m\mu b_1} \geq 0$$
$$\sqrt{\mu^2 + a\mu - b_2 z - m\mu b_1} \geq \mu$$
$$\mu^2 + a\mu - b_2 z - m\mu b_1 \geq \mu^2$$
$$a\mu - b_2 z - m\mu b_1 \geq 0$$
$$a\mu - b_2 z \geq m\mu b_1$$
$$\frac{a\mu - b_2 z}{\mu b_1} \geq m$$

This result corresponds to the feasibility condition & non-negative objective function

G. Proof that the objective function is concave in M/M/1 without penalty and holding cost

This part of the appendix demonstrates that the objective function for the M/M/1/1 without penalty and holding cost is concave. The objective function of this problem is:

$$\underset{\lambda,l,p,\bar{\lambda}}{\text{Maximize}}\ \bar{\lambda}(p-m) \Leftrightarrow \underset{\lambda}{\text{Maximize}}\left(\frac{\lambda a\mu - \lambda b_2 z - \lambda^2\mu - \lambda m\mu b_1}{\mu b_1 + \lambda b_1}\right)$$

Suppose that the $\lambda_1 = 0$, then the objective become:

$$\begin{aligned} f(\lambda,l) = \bar{\lambda}(p-m) \quad &= \frac{\lambda a\mu - \lambda b_2 z - \lambda^2\mu - \lambda m\mu b_1}{\mu b_1 + \lambda b_1} \\ &= \frac{0}{\mu b_1} \\ &= 0 \end{aligned}$$

If the $\lambda_2 = +\infty$, then the objective function is:

$$\begin{aligned} f(\lambda,l) = \bar{\lambda}(p-m) \quad &= \frac{\lambda a\mu - \lambda b_2 z - \lambda^2\mu - \lambda m\mu b_1}{\mu b_1 + \lambda b_1} \\ &= \frac{-\infty}{+\infty} \\ &= -\infty \end{aligned}$$

It can be summarized as shown in table below

			The objective with λ^*
The objective with	$\lambda_1 = 0$	$F(\lambda_1) = 0$	$F(\lambda^*) \geq F(\lambda_1)$
	$\lambda_2 = +\infty$	$F(\lambda_2) = -\infty$	$F(\lambda^*) \geq F(\lambda_2)$

This prove that the objective with $\lambda_1 = 0$ and $\lambda_2 = +\infty$ compared to λ^* is lower. Hence, the objective function is concave.

H. Proof demand constraint is binding (M/M/1/1 with holding and penalty cost)

As in Palaka et al. (1998) and Pekgün et al. (2008), suppose that the optimal solution is given by price, p^*, quoted lead-time l^*, and demand rate λ^*, and that $\lambda^* < \Lambda(p^*,1^*.)$. Since the revenues are non-decreasing in p, one could increase the price to p' (while holding the demand rate and quoted lead-time constant) until $l^* = \Lambda(p',1')$. This change will increase revenues without increasing the queueing costs, direct variable costs, and lateness penalties. Therefore, (p^*, l^*, and λ^*) cannot be an optimal solution.

To demonstrate:

$$\underset{\lambda,l,p,\bar{\lambda}}{\text{Maximize}} \quad \bar{\lambda}(p-m)-\frac{F\lambda}{\mu+\lambda}-\frac{c\bar{\lambda}}{\mu}e^{-\mu l}$$

$$\text{Subject to} \quad \lambda \le a-b_1 p-b_2 l$$

$$1-\sum_{k=0}^{K-1}\left(\sum_{i=0}^{k}\frac{(\mu\cdot l)^i}{i!}e^{-\mu\cdot l}\right)\frac{P_k}{1-P_K}\ge s$$

Assume: $\lambda^* \le a-b_1 p^* - b_2 l^*$, we will have $F^* = \bar{\lambda}^*\left(p^*-m\right)$. If we increase the price to *p'*, then we will have $F' = \bar{\lambda}^*\left(p'-m\right)$ which is bigger than the F^*. Hence, there is a contradiction which is mean that the constraint is binding.

I. Proof Service level constraint (M/M/1/1 with holding and penalty cost)

As in Palaka et al. (1998), applying the Lagrangian multiplier method, we see that a stationary point to problem (P2) must satisfy:

$$\frac{\partial L}{\partial \lambda}=0,\ \frac{\partial L}{\partial l}=0,\ \gamma\frac{\partial L}{\partial \gamma}=0,\ \gamma\ge 0 \text{ and } \mu l \ge \ln\left(\frac{1}{1-s}\right)$$

Where,

$$L(\lambda,l,\gamma)=\frac{\lambda\left(p\mu-m\mu-F-ce^{-\mu l}\right)}{\mu+\lambda}+\gamma\left\{\mu l-\ln\left(\frac{1}{1-s}\right)\right\}$$

With $\lambda = a-b_1 p-b_2 l$, hence $p=\dfrac{a-b_2 l-\lambda}{b_1}$

$$L(\lambda,l,\gamma)=\frac{\lambda\left[\left(\mu(a-b_2 l-\lambda)/b_1\right)-m\mu-F-ce^{-\mu l}\right]}{\mu+\lambda}+\gamma\left\{\mu l-\ln\left(\frac{1}{1-s}\right)\right\}$$

$$\frac{\partial L}{\partial l}=0 \quad \Leftrightarrow \quad \gamma\mu-\frac{\lambda\left[\left(\mu b_2/b_1\right)-\mu c e^{-\mu l}\right]}{\mu+\lambda}=0$$

$$\gamma\mu(\mu+\lambda)-\mu(\lambda b_2/b_1)+\mu\lambda c e^{-\mu l}=0$$

$$\frac{\gamma(\mu+\lambda)}{\lambda c}-\frac{b_2}{cb_1}+e^{-\mu l}=0$$

$$1-\frac{b_2}{cb_1}+\frac{\gamma(\mu+\lambda)}{\lambda c}=1-e^{-\mu l}$$

$$s_c+\frac{\gamma(\mu+\lambda)}{\lambda c}=1-e^{-\mu l}$$

$$\gamma\frac{\partial L}{\partial \gamma}=0 \quad \Leftrightarrow \quad \gamma\left\{\mu l-\ln\left(\frac{1}{1-s}\right)\right\}=0$$

We know that $\gamma(\partial L/\partial\gamma)=0$. This imply that $\gamma=0$ or $\mu l-\ln(1/(1-s))=0$. Suppose that the service level is non-binding, thus $\gamma=0$ and $1-e^{-\mu l}>s$. It implies that:

$$1-\frac{b_2}{cb_1}+\frac{\gamma(\mu+\lambda)}{\lambda c}>s$$

$$1-\frac{b_2}{cb_1}>s$$

$$s_c>s$$

Suppose that the service level is binding, thus $\gamma\ne 0$ and $1-e^{-\mu l}=s$. It implies that:

$$1-\frac{b_2}{cb_1}+\frac{\gamma(\mu+\lambda)}{\lambda c}=s$$
$$s_c+\frac{\gamma(\mu+\lambda)}{\lambda c}=s$$
which imply that $s_c \leq s$.

J. Feasibility Condition and Non-Negative Objective Function (M/M/1/1 with holding and penalty cost)

Note that: $\lambda = a - b_1 p - b_2 l$ from appendix B and $\lambda \geq 0$, hence:

$$a-b_1p-b_2l \geq 0 \iff a-b_2l \geq b_1 p$$
$$\frac{(a-b_2l)}{b_1} \geq p \qquad \text{(Condition 1)}$$

As $p \geq m$ and condition 1 imply that the feasibility condition of this problem is:

$\frac{(a-b_2l)}{b_1} \geq m$, with $l^* = \frac{1}{\mu}\ln x$ and $x = Max\{1/(1-s), b_1c/b_2\}$ ass proven in section 6.

The, to have the non-negative objective function which is correspond to have a feasibility condition:

$$\bar{\lambda}(p-m)-\frac{F\lambda}{\mu+\lambda}-\frac{c\bar{\lambda}}{\mu}e^{-\mu l} \geq 0 \iff \frac{\lambda\left(p\mu-m\mu-F-ce^{-\mu l}\right)}{\mu+\lambda} \geq 0$$

It is obvious that $\mu+\lambda \geq 0$ since the μ and λ is always bigger than zero. Hence, the numerator should be:

$$\lambda\left(p\mu-m\mu-F-ce^{-\mu l}\right) \geq 0$$
$$p\mu-m\mu-F-ce^{-\mu l} \geq 0$$
$$p \geq m+F/\mu+ce^{-\mu l}/\mu \qquad \text{(Condition 2)}$$

Combine this two p condition, we have:

$$\frac{(a-b_2l)}{b_1} \geq p \text{ and } p \geq m+F/\mu+ce^{-\mu l}/\mu$$
$$\iff \frac{(a-b_2l)}{b_1} \geq m+F/\mu+ce^{-\mu l}/\mu$$
$$\iff \frac{(a-b_2l)}{b_1}-m-F/\mu-ce^{-\mu l}/\mu \geq 0$$
$$\iff \mu(a-b_2l)-\mu mb_1-Fb_1-cb_1e^{-\mu l} \geq 0$$
$$\iff a\mu-\mu b_2l-\mu mb_1-Fb_1-cb_1e^{-\mu l} \geq 0$$

Those equation should be bigger equal to zero to have a non-negative objective function.

K. Proof that Discriminant ≥ 0 (M/M/1/1 with holding and penalty cost)

Suppose that:

$$\Delta \geq 0 \iff (-2\mu)^2-4(-1)\left(a\mu-\mu b_2l-\mu mb_1-Fb_1-b_1ce^{-\mu l}\right) \geq 0$$
$$4\mu^2+4a\mu-4\mu b_2l-4\mu mb_1-4Fb_1-4b_1ce^{-\mu l} \geq 0$$
$$\mu^2+a\mu-\mu b_2l-\mu mb_1-Fb_1-b_1ce^{-\mu l} \geq 0$$

The $a\mu-\mu b_2l-\mu mb_1-Fb_1-b_1ce^{-\mu l}$ corresponds to the non-negativity condition of appendix J. This imply that to have a non-negative objective functions, we have to have $a\mu-\mu b_2l-\mu mb_1-Fb_1$ $-b_1ce^{-\mu l} \geq 0$. With the $\mu \geq 0$, we can say that $\mu^2+a\mu-\mu b_2l-\mu mb_1-Fb_1-b_1ce^{-\mu l} \geq 0$ and $\Delta \geq 0$ are proven.

L. The λ_2 correspond to the feasibility condition (M/M/1/1 with holding and penalty cost)

$$\lambda_2 \geq 0 \iff \quad -\mu + \sqrt{\mu^2 + a\mu - \mu b_2 l - \mu b_1 m - Fb_1 - b_1 c e^{-\mu l}} \geq 0$$
$$\sqrt{\mu^2 + a\mu - \mu b_2 l - \mu b_1 m - Fb_1 - b_1 c e^{-\mu l}} \geq \mu$$
$$\mu^2 + a\mu - \mu b_2 l - \mu b_1 m - Fb_1 - b_1 c e^{-\mu l} \geq \mu^2$$
$$a\mu - \mu b_2 l - \mu b_1 m - Fb_1 - b_1 c e^{-\mu l} \geq 0$$

This final equation corresponds to the non-negativity of the objective function in appendix J.

M. Proof that the objective function is concave in M/M/1 with penalty and holding cost

This part of the appendix demonstrates that the objective function for the M/M/1/1 with penalty and holding cost is concave. The objective function of this problem is:

$$\underset{\lambda,p,l,\bar{\lambda}}{Maximize} \quad \bar{\lambda}(p-m) - \frac{F\lambda}{\mu+\lambda} - \frac{c\bar{\lambda}}{\mu} e^{-\mu l}$$

$$\iff \underset{\lambda,l}{Maximize} \quad \frac{\lambda\left[\left(\mu\left(a - b_2 l - \lambda\right)/b_1\right) - m\mu - F - ce^{-\mu l}\right]}{\mu+\lambda}$$

Suppose that the $\lambda_1 = 0$ and $l_1 = 0$, then the objective become:

$$f(\lambda,l) = \bar{\lambda}(p-m) - \frac{F\lambda}{\mu+\lambda} - \frac{c\bar{\lambda}}{\mu} e^{-\mu l} = \frac{\lambda\left[\left(\mu\left(a - b_2 l - \lambda\right)/b_1\right) - m\mu - F - ce^{-\mu l}\right]}{\mu+\lambda}$$
$$= \frac{0\left[\left(a\mu/b_1\right) - m\mu - F - c\right]}{\mu}$$
$$= \frac{0}{\mu b_1}$$
$$= 0$$

Suppose that the $\lambda_2 = +\infty$ and $l_1 = 0$, then the objective become:

$$f(\lambda,l) = \bar{\lambda}(p-m) - \frac{F\lambda}{\mu+\lambda} - \frac{c\bar{\lambda}}{\mu} e^{-\mu l} = \frac{\lambda\left[\left(\mu\left(a - b_2 l - \lambda\right)/b_1\right) - m\mu - F - ce^{-\mu l}\right]}{\mu+\lambda}$$
$$= \frac{\infty\left[\left(-\infty/b_1\right) - m\mu - F - c\right]}{\infty}$$
$$= \frac{-\infty}{\infty}$$
$$= -\infty$$

Suppose that the $\lambda_1 = 0$ and $l_1 = +\infty$, then the objective become:

$$f(\lambda,l) = \bar{\lambda}(p-m) - \frac{F\lambda}{\mu+\lambda} - \frac{c\bar{\lambda}}{\mu} e^{-\mu l} = \frac{\lambda\left[\left(\mu\left(a - b_2 l - \lambda\right)/b_1\right) - m\mu - F - ce^{-\mu l}\right]}{\mu+\lambda}$$
$$= \frac{0\left[\left(-\infty/b_1\right) - m\mu - F - \infty\right]}{\mu}$$
$$= \frac{0}{\mu}$$
$$= 0$$

Suppose that the $\lambda_2 = +\infty$ and $l_2 = +\infty$, then the objective become:

$$f(\lambda,l) = \bar{\lambda}(p-m) - \frac{F\lambda}{\mu+\lambda} - \frac{c\bar{\lambda}}{\mu} e^{-\mu l} = \frac{\lambda\left[\left(\mu\left(a - b_2 l - \lambda\right)/b_1\right) - m\mu - F - ce^{-\mu l}\right]}{\mu+\lambda}$$

$$= \frac{\infty\left[\mu\left(a - \infty - \infty / b_1\right) - m\mu - F - \infty\right]}{\infty}$$

$$= \frac{-\infty}{\infty}$$

$$= -\infty$$

It can be summarized as shown in table below

			The objective with λ^*, l^*
The objective with	$\lambda_1 = 0, l_1 = 0$	$F(\lambda_1, l_1) = 0$	$F(\lambda^*, l^*) \geq F(\lambda_1, l_1)$
	$\lambda_2 = +\infty, l_1 = 0$	$F(\lambda_2, l_1) = -\infty$	$F(\lambda^*, l^*) \geq F(\lambda_2, l_1)$
	$\lambda_1 = 0, l_2 = +\infty$	$F(\lambda_1, l_2) = 0$	$F(\lambda^*, l^*) \geq F(\lambda_1, l_2)$
	$\lambda_2 = +\infty, l_2 = +\infty$	$F(\lambda_2, l_2) = -\infty$	$F(\lambda^*, l^*) \geq F(\lambda_2, l_2)$

This prove that the other compared to λ^* is lower. Hence, the objective function is concave.