CHAPTER 5

An Inventory Model with Server Interruptions and Retrials *

5.1 Introduction

In the previous chapter we assumed that replenishment is instantaneous and no shortage is permitted (shortage cost infinity). In this chapter we assume that replenishment of inventory does not materialize immediately as placement of order, rather takes a random amount of time for order materialization. Thus we have now a finite shortage cost situation. Another salient feature of this chapter is that the QBD we develop here is level dependent. The reason behind this is the state dependent retrial rate of orbital customers. LDQBD’s are much more complex than LIQBD’s since we do not get a repeating pattern for the entries of the infinitesimal generator of the process. However we make it level independent through a process called truncation.

The first study on inventory models with positive lead time, with unsatisfied customers thus created going to an orbit to try again for inventory from there, was by Artalejo et. al [6]. Whereas their approach is algorithmic, Ushakumari [68] produces analytical solution to the same model. Following these, a number of papers on inventory models with retrial of unsatisfied customers emerged. One may refer to Krishnamoorthy and Islam [30,31] Two other papers where an inventory model with retrial of demands is considered

* The results of this chapter was presented as a paper at the 8th International Workshop on Retrial Queues; July 27-29, 2010, Beijing, China. It has also been accepted for publication in the journal ‘Operational Research’ (article No: DOI: 10.1007/s12351-011-0120-8)
are by Sivakumar [61,62]; where the first one considers an (s, S) perishable inventory system in which demands occur from a finite source and those demands that arrive in a stock-out period, are sent to an orbit. The discussion in [62] is on a two-commodity system where customers, encountering both commodities out-of-stock, proceed to an orbit of infinite capacity. In Krishnamoorthy and Jose [34], the authors analyse and compare different (s, S) inventory models with an orbit of infinite capacity. They consider situations where a finite waiting station/no waiting station is provided for fresh/retrial customers. In all these models, the presence of the retrying customers results in an in-flow and out-flow pattern that is distinct from those where only a queue of unsatisfied demands has been considered. In this regard the paper by Krishnamoorthy and Islam [32], where an (s, S) inventory system with a finite pool of unsatisfied demands is studied, has an in-flow and out-flow pattern that is distinct from both the above type of models.

Different kinds of interruptions in service such as server breakdown, server going on vacation, arrival of priority customers, being common real life phenomena, analysis of queueing models with these features partially or completely incorporated, is important. An M/M/1 queueing model with service interruption was first studied by White and Christie [73]. The service interruption of such customers is assumed to be due to arrival of priority customers. The policy adopted is preemptive repeat. However, it may be noted that in the case of exponentially distributed service time, repeat or resumption of interrupted service do not make any difference. Krishnamoorthy et. al [29,37, 38] provide a glimpse of earlier work on queues with service interruption and provide several new results.
5.2. Mathematical Model

In a very recent paper, Krishnamoorthy et. al [41] considered an \((s,S)\) inventory model with positive service time and instantaneous replenishment, where the service process is subject to interruptions. In the present chapter, we extend the above model by assuming that replenishment of items requires a random amount of lead-time. As in [41], the service in the present case is also subject to interruptions. Now since the replenishment is not instantaneous, an increase in waiting time is probable which motivates us to introduce a retrial queueing model here. It may be noted that the interruption process assumed here has the property that, at a time only one interruption is encountered by the server. In our model, whenever the service is restarted after interruption, it is assumed that the entire service is repeated from the beginning and further it is prd. This is in contrast to Nicola, Kulkarni and Trivedi [51] as well as Marie and Trivedi [55].

This chapter is arranged as follows. In section 5.2, we describe the mathematical model under study. In section 5.3, a necessary and sufficient condition for the stability of the system is obtained and steady state distribution is computed. Section 5.4 is devoted to some system performance measures like the expected waiting time of an orbital customer. Finally in section 5.5 we provide some results of the numerical experiments carried out for analyzing different aspects of the system under study. Concluding remarks are given in section 6.

5.2 Mathematical Model

Before proceeding with the modeling of the problem under investigation, we introduce a few notations and assumptions that are used in the sequel:
5.2. Mathematical Model

- Arrival of primary customers -- Poisson process of rate $\lambda$
- Service time -- exponential random variable with parameter $\mu$
- $s$ is the reorder level and $S$ is the maximum number of items that can be stored ($S > 2s$)
- Lead time (the time elapsed from placing an order for replenishment of the item until it is delivered) -- exponentially distributed with parameter $\eta$
- Service interruption process -- Poisson process of rate $\delta_1$
- Interruption duration -- exponential random variable with parameter $\delta_2$
- Retrial rate -- $j\theta$, when there are $j$ customers in the orbit (thus retrial rate turns out to be level dependent)
- Probability of a primary customer joining the system when the server is in interrupted state -- $p$
- Probability of an orbital customer quitting the system after an unsuccessful retrial due to server in interrupted state -- $q$
- Identity matrix of order $n$ -- $I_n$
- Identity matrix of appropriate order -- $I$
- Column vector of 1’s of appropriate order -- $\mathbf{e}$

The model under study is described as follows: Customers arrive to a single server counter according to a Poisson process of rate $\lambda$ where inventory is served. Service times are iid exponentially distributed random variables with parameter $\mu$. Inventory is replenished according to $(s, S)$ policy, with the lead time distribution exponential with parameter $\eta$.

While the server serves a customer, the service may get interrupted with the interruption process governed by a Poisson process of rate $\delta_1$. It is
5.2. Mathematical Model

assumed that while the server is under interruption, no further interruption can befall the server. On completion of an interruption the service restarts, with the duration of an interruption exponentially distributed with parameter $\delta_2$. No waiting space is provided for customers, other than for the one whose service gets interrupted. An arriving customer, finding the server busy, leaves the service area and joins an orbit of infinite capacity from where it retries for service. The duration of the interval between two successive repeated attempts is exponentially distributed with parameter $j\theta$ when the number of customers in the orbit is $j$. While the server is on an interruption, an arriving customer (primary) joins the system with probability $p$ and a retrying customer goes back to the orbit with probability $(1-q)$. With complementary probabilities the customer leaves the system in both cases. When the inventory level is zero no primary arrival or retrial is entertained (primary arrivals have to leave the system without getting admission to orbit and retrial customers stay put in the orbit).

Let $N(t)$ be the number of customers in the orbit and $L(t)$ be the inventory level at time $t$. Also let

$$C(t) = \begin{cases} 
0, & \text{if the server is idle} \\
1, & \text{if the server is busy} \\
2, & \text{if the server is on interruption}
\end{cases}$$

be the server status. Then $\Omega = \{X(t); t \geq 0\} = \{(N(t), C(t), L(t)); t \geq 0\}$ is a Markov chain on the state space $((\mathbb{Z}_+ \cup \{0\}) \times \{0,1,2\} \times \{1,2,3,\ldots,S\}) \cup ((\mathbb{Z}_+ \cup \{0\}) \times \{0\} \times \{0\})$. 

101
5.2. Mathematical Model

The state space of the Markov chain is partitioned into levels $i$ defined as
\[ \{(i,0,j); 0 \leq j \leq S\} \cup \{(i,1,j); 1 \leq j \leq S\} \cup \{(i,2,j); 1 \leq j \leq S\}, \quad i \geq 0 \quad \text{and} \quad Q = S - s. \]

5.2.1 A typical illustration of the transitions of the Markov chain is as given below:

For $i \geq 0$,

- $(i,1,k) \xrightarrow{\lambda} (i+1,1,k)$
- $(i,2,k) \xrightarrow{p\lambda} (i+1,2,k)$
- $(i,0,k) \xrightarrow{i\theta} (i-1,1,k); 1 \leq k \leq S$
- $(i,2,k) \xrightarrow{q\theta} (i-1,2,k)$
- $(i,0,k) \xrightarrow{\lambda} (i,1,k); 1 \leq k \leq S$
- $(i,0,k) \xrightarrow{\eta} (i,0,k+Q); 0 \leq k \leq s$
- $(i,1,k) \xrightarrow{\eta} (i,1,k+Q); 1 \leq k \leq s$
- $(i,2,k) \xrightarrow{\eta} (i,2,k+Q); 1 \leq k \leq s$
- $(i,1,k) \xrightarrow{\mu} (i,0,k-1); 1 \leq k \leq S$
- $(i,1,k) \xrightarrow{\delta_1} (i,2,k)$
- $(i,2,k) \xrightarrow{\delta_2} (i,1,k)$
The Markov chain $\mathcal{Q}$ in which the above transitions occur, is a level-dependent quasi birth and death process (LDQBD) with infinitesimal generator matrix.
5.2. Mathematical Model

\[
\hat{Q} = \begin{bmatrix}
A_{00} & A_{01} & 0 & 0 & 0 \\
A_{10} & A_{11} & A_{0} & 0 & 0 \\
0 & A_{21} & A_{12} & A_{0} & 0 \\
& & \ddots & \ddots & \ddots \\
& & & \ddots & \ddots \\
& & & & A_{11}
\end{bmatrix},
\]

where each entry is a \((3S + 1)x (3S + 1)\) matrix, which is explained in detail below:

\[
A_0 = \begin{bmatrix}
0 & 0 & 0 \\
0 & \lambda I_S & 0 \\
0 & 0 & p \lambda I_S
\end{bmatrix}
\]

represents transitions from level \(i\) to \(i+1\) due to arrival of a customer; note that by assumption, arrival rate is \(p \lambda\), when the server is interrupted.

\[
A_{2j} = \begin{bmatrix}
0 & B_{2j} & 0 \\
0 & 0 & 0 \\
0 & 0 & q_j \theta I_S
\end{bmatrix}
\]

represents transitions from level \(j\) to \(j-1\), where

\[
B_{2j} = \begin{bmatrix}
0 \\
0 \\
0
\end{bmatrix}
\]

represents transitions from level \(j\) to \(j-1\) due to retrial of an orbital customer becoming successful and \(q_j \theta I_S\) represents those due to an orbital customer leaving the system after an unsuccessful retrial finding the interrupted server.

\[
A_{ij} = \begin{bmatrix}
D_1 & D_2 & 0 \\
D_4 & D_5 & D_6 \\
0 & D_8 & D_9
\end{bmatrix}
\]

represents transitions within the level \(j\), where the sub-matrices are explained as follows.
5.2. Mathematical Model

\[
D_1 = \begin{bmatrix}
D_{11} & 0 & \eta I_{(s+1)} \\
0 & -\lambda \theta J(s-2s-1) & 0 \\
0 & 0 & -(\lambda \theta) I_{(s+1)}
\end{bmatrix},
\]

with

\[
D_{11} = \begin{bmatrix}
-\eta & 0 \\
0 & -(\eta + \lambda \theta) J(s+1)
\end{bmatrix}
\]

represents transitions within level \( j \), where the server status remains as idle; note that the only transition here is that due to replenishment of the items represented by \( \eta I_{(s+1)} \).

\[
D_2 = \begin{bmatrix}
0 \\
\lambda I_S
\end{bmatrix}_{(S+1) \times S}
\]

represents transitions within level \( j \), where the server status changes from idle to busy due to arrival of a customer.

\[
D_3 = \begin{bmatrix}
-(\eta + \lambda + \mu + \delta J)J_s & 0 & C \\
0 & 0 & -(\lambda + \mu + \delta J)J_{(s-2s-1)} \\
0 & 0 & -(\lambda + \mu + \delta J)I_{(s+1)}
\end{bmatrix}
\]

and

\[
D_3 = \begin{bmatrix}
-(\eta + p\lambda + \delta_2 + jq\theta)I_s & 0 & C \\
0 & 0 & -(p\lambda + \delta_2 + jq\theta)I_{(s-2s-1)} \\
0 & 0 & -(p\lambda + \delta_2 + jq\theta)I_{(s+1)}
\end{bmatrix}
\]

with \( C = [0 \, \eta I_{s}^T]_{s \times (s+1)} \), represents transitions within level \( j \), where the server status remains as busy and interrupted respectively. Note that the only transition here is that due to replenishment of the items represented by the matrix \( C \).
5.3 Analysis of the Model

\[ D_6 = \delta I_s \] represents transitions within level j, due to server interruption

and \( D_8 = \delta I_s \) those due to server status changing from interrupted to busy.

5.3 Analysis of the Model

In this section, using Matrix Analytic Methods (for details on Matrix Analytic Methods, see Neuts [56]), we perform the steady state analysis. First let us look at the stability of the system.

5.3.1 Stability Condition

For investigating the stability condition of the system under study, first we apply Neuts-Rao [57] truncation to the LIQBD. To this end suppose that \( A_i = A_{iN} \) and \( A_s = A_{sN} \) for all \( i \geq N \). The generator matrix of the truncated system \( \Omega_N \) will look as under:

\[
\tilde{Q}_N = \begin{bmatrix}
A_{10} & A_{0} & 0 & 0 & 0 \\
A_{21} & A_{11} & A_{0} & 0 & 0 \\
0 & A_{22} & A_{12} & A_{0} & 0 & 0 \\
& & & \ddots & \ddots & \ddots \\
A_{2N} & A_{1N} & A_{0} & & & \\
0 & A_{2N} & A_{1N} & A_{0} & 0 & \\
& & & & \ddots & \ddots & \ddots 
\end{bmatrix}
\]

Define \( A_N = A_0 + A_{1N} + A_{2N} \) and let \( \pi_N = (\pi_N(0,0), \pi_N(0,1), \pi_N(0,2), \ldots, \pi_N(0,S), \pi_N(1,1), \pi_N(1,2), \ldots, \pi_N(1,S), \pi_N(2,1), \pi_N(2,2), \ldots, \pi_N(2,S)) \) be the steady state vector of \( A_N \).
5.3. Analysis of the Model

From the well-known results of Matrix Analytic Methods (see Neuts [56]), it follows that the truncated system, which is a level-independent quasi birth death process, is stable if and only if \( \pi_N A_{2N} e > \pi_N A_0 e \), that is, if and only if

\[
N \theta \left[ \pi_N (0,1) + \pi_N (0,2) + \ldots + \pi_N (0,S) \right] + qN \theta \left[ \pi_N (2,1) + \pi_N (2,2) + \ldots + \pi_N (2,S) \right] > \lambda \left[ \pi_N (1,1) + \pi_N (1,2) + \ldots + \pi_N (1,S) \right] + p \lambda \left[ \pi_N (2,1) + \pi_N (2,2) + \ldots + \pi_N (2,S) \right].
\]

This reduces to the system being stable if and only if

\[
\left\{ \frac{N \theta \mu}{\lambda} + \frac{qN \theta \delta}{\delta_2} \right\} \left[ \pi_N (1,1) + \pi_N (1,2) + \ldots + \pi_N (1,S) \right] > \left( \lambda + \frac{p \lambda \delta}{\delta_2} \right) \left[ \pi_N (1,1) + \pi_N (1,2) + \ldots + \pi_N (1,S) \right],
\]

which on further simplification yields that the system is stable if and only if

\[
\lambda + \frac{p \lambda \delta}{\delta_2} < \frac{N \theta \mu}{\lambda + N \theta} + \frac{qN \theta \delta}{\delta_2}.
\]

Because of the second factor on the right hand side of the above inequality, we see that the system is stable whenever the probability \( q \) is greater than zero.

Now, when \( q = 0 \), taking the limit in the above inequality as \( N \to \infty \), it reduces to:

\[
\lambda + \frac{p \lambda \delta}{\delta_2} < \mu.
\]

Thus we have the following theorem for stability of the system under study:

**Theorem 5.1**

When the probability \( q \) that a retrying customer leaves the orbit after an unsuccessful retrial, is greater than zero, the Markov Chain \( \Omega \) is stable.
irrespective of the other system parameters and when \( q = 0 \), it is stable if and only if \( \lambda + \frac{p\lambda \delta}{\delta^2} < \mu \).

5.3.2 Computation of Steady State Probability Vector

We find the steady state vector of \( \Omega \), by approximating it with the steady state vector of the truncated system, \( \Omega_N \) with generator matrix \( Q_N \). Let \( \pi^{(N)} = (\pi_0, \pi_1, \pi_2, \ldots) \), be the steady state vector of \( \Omega_N \) where each \( \pi_i \) is a row vector consisting of \( 3S+1 \) elements represented as

\[
\pi_i = (\pi(i,0,0), \pi(i,0,1), \pi(i,0,2), \ldots, \pi(i,0,S), \pi(i,1,1), \pi(i,1,2), \ldots, \pi(i,1,S), \pi(i,2,1), \pi(i,2,2), \ldots, \pi(i,2,S))
\]

Then from known results of Matrix Analytic Methods (see Neuts [56]), it follows that

\[\pi_{N+r} = \pi_{N-r} (R_N)^{r+1}, \text{ for } r \geq 0,\]

where \( R_N \) is the minimal non-negative solution of the matrix quadratic equation,

\[
(R_N)^2 A_{2N} + R_N A_{1N} + A_0 = 0,
\]

\[
\pi_{N-i} = \pi_{N-i-1} R_{N-i}, \text{ for } 1 \leq i \leq N-1, \text{ where}
\]

\[
R_{N-i} = -A_0 \left( A_{1N-i-1} R_{N-i-1} A_{2N-i+1} \right)^{-1}.
\]

Now for computing \( \pi_0 \), we have the equation \( \pi_0 \left( A_{10} + R_1 A_{21} \right) = 0 \). First we take \( \pi_0 \) as the steady state vector of the generator matrix \( A_{10} + R_1 A_{21} \). Then \( \pi_i \), for \( 1 \leq i \leq N-1 \), can be found using the recursive formulae;

\[\pi_i = \pi_{i-1} R_i \]

The steady state probability distribution of the truncated system is then obtained by dividing each \( \pi_i \) with the normalizing constant
5.4. System Performance Measures

5.4.1 Waiting Time Analysis of an Orbital Customer

Since no queue is formed in the orbit, customers, independently of each other, try to access the server. Therefore computation of the waiting time distribution becomes extremely complex though it has been achieved in some special cases (see books [2, 16] for details). Hence we restrict ourselves to the computation of the moments of the waiting time. Though we can find the expected waiting time using Little’s Law, the second moment and variance of the waiting time are not easy to find. These moments are found by approximating the waiting time in the system under study by those in a corresponding system with finite orbit capacity.

Let $E(W_L)$ be the expected waiting time of an orbital customer in the system under study and $E(W_L^{(N)})$ be that in the corresponding system with finite orbit capacity $N$. Then $E(W_L) = \lim_{N \to \infty} E(W_L^{(N)})$.

For the system with finite orbit capacity $N$, $W_L^{(N)}$ can be found as the time until absorption in a Markov chain $\{\hat{X}(t), t \geq 0\}$, where $\{\hat{X}(t), t \geq 0\} = \{(\hat{N}(t), C(t), L(t)), t \geq 0\}$, if the tagged customer is in the orbit and $\hat{X}(t) = \Delta$, if either the tagged customer gets service or quits the system. In the above, $\hat{N}(t)$ denotes the number of customers in the orbit including the tagged customer, $C(t)$ and $L(t)$ are as defined in section 2. Since
the orbit capacity is $N$, we have $1 \leq \hat{N}(t) \leq N$. The state space of the process \[\{\hat{X}(t), t \geq 0\}\] is
\[\{\Delta\} \cup \{(i, 0, k) | 1 \leq i \leq N, 0 \leq k \leq S\} \cup \{(i, j, k) | 1 \leq i \leq N, 1 \leq j \leq 2, 1 \leq k \leq S\},\]
where $\Delta$ is an absorbing state. The generator matrix of this process is
\[
\hat{Q}_w(N) = \begin{bmatrix} T & T^0 \\ 0 & 0 \end{bmatrix},
\]
where $T^0$ is an $N(3S+1) \times 1$ matrix given by
\[
T^0((i-1)(3S+1)+j,1) = \theta, \quad j = 2 \text{ to } S+1; \quad i=1 \text{ to } N,
\]
\[
T^0((i-1)(3S+1) + j,1) = q\theta, \quad j = 2S+2 \text{ to } 3S+1; \quad i=1 \text{ to } N
\]
and the matrix $T$ is given by
\[
T = \begin{bmatrix} A_{11} & A_0 \\ \hat{A}_{21} & A_{12} & A_0 \\ 0 & \hat{A}_{22} & A_{13} & A_0 \\ \vdots & \vdots & \ddots & \vdots \\ \hat{A}_{2(N-2)} & \hat{A}_{1(N-1)} & A_0 \\ \hat{A}_{2(N-1)} & \hat{A}_{1N} \end{bmatrix}, \text{ where}
\]
\[
\hat{A}_{2j} = \begin{bmatrix} 0 & \hat{B}_{2j} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (j-1)q\theta I_S \end{bmatrix}
\]
with
\[
\hat{B}_{2j} = \begin{bmatrix} 0 & 0 \\ (j-1)\theta I_S & 0 \end{bmatrix}_{(S+1) \times (S+1)} \quad \text{and} \quad \hat{A}_{1N} = A_{1N} + A_0
\]

110
and all other matrices are as defined in the generator matrix Q. Thus
\[ E(W_L^{(N)}) = -\alpha T^{-1}e \] (see Neuts [56], page 46),
where \( \alpha = \pi_L = (\pi_{L0}, \pi_{L1}, \pi_{L2}, \ldots, \pi_{LN}) \); \( \pi_{Li} = \pi_i \) with entries corresponding to server is idle states taken as zero. It has been verified numerically that for large \( N \), \( E(W_L^{(N)}) \) converges according to Little’s theorem.

In a similar manner, we can find the second moment of the waiting time of an orbital customer as
\[ E(W_L^{2}) = \lim_{N \to \infty} E\left(W_L^{(N)}\right)^2, \]
where \( E\left(W_L^{(N)}\right)^2 \) is the corresponding second moment in the truncated system and is given by
\[ E\left(W_L^{(N)}\right)^2 = 2\alpha T^{-2}e \] (see Neuts [56], page 46).

Finally, the variance of the waiting time of an orbital customer is given by
\[ V(W_L) = E(W_L^{2}) - (E(W_L))^2. \]

The conditional probability that a customer leaves the system without taking service given that he arrives while the server is busy is given by
\[ P_{WS} = -\alpha T^{-1}\hat{T} \] (see Neuts [56], page 46),
where \( \hat{T} \) is an \( N(3S+1) \times 1 \) matrix whose non zero entries are given by
\[ \hat{T}_{(i-1)(3S+1)+j,1} = q\theta, \quad j = 2S+2 \text{ to } 3S+1; \quad i=1 \text{ to } N \]
5.4. System Performance Measures

5.4.2 Other Performance Measures

The following system performance measures are calculated numerically.

1. The probability that server is busy is given by
   \[ P_\beta = \sum_{i=0}^{\infty} \sum_{j=1}^{S} \pi(i, 1, j). \]

2. The probability that server is on interruption is given by
   \[ P_\alpha = \sum_{i=0}^{\infty} \sum_{j=1}^{S} \pi(i, 2, j). \]

3. The probability that server is idle is given by
   \[ P_\gamma = 1 - P_\alpha - P_\beta. \]

4. The expected number of customers in the orbit is given by
   \[ E(\sigma) = \sum_{i=0}^{\infty} \sum_{j=0}^{S} i \pi(i, 0, j) + \sum_{i=0}^{\infty} \sum_{j=1}^{S} i \{ \pi(i, 1, j) + \pi(i, 2, j) \}. \]

5. The expected inventory level is given by
   \[ E(\omega) = \sum_{i=0}^{\infty} \sum_{j=0}^{S} j \pi(i, 0, j) + \sum_{i=0}^{\infty} \sum_{j=1}^{S} j \{ \pi(i, 1, j) + \pi(i, 2, j) \}. \]

6. The effective rate of successful retrials is given by
   \[ E(\sigma_0) = \sum_{i=0}^{\infty} \sum_{j=0}^{S} i \theta \pi(i, 0, j). \]

7. The effective replenishment rate is given by
   \[ EFRR = \sum_{i=0}^{\infty} \sum_{j=0}^{S} \eta \pi(i, 0, j) + \sum_{i=0}^{\infty} \sum_{j=1}^{S} \eta \{ \pi(i, 1, j) + \pi(i, 2, j) \}. \]

8. The probability that inventory level is zero is given by
   \[ P(L=0) = \sum_{i=0}^{\infty} \pi(i, 0, 0). \]

9. The probability that inventory level is greater than \( s \) is given by
   \[ P(L>s) = \sum_{i=0}^{\infty} \sum_{j=s+1}^{S} \{ \pi(i, 0, j) + \pi(i, 1, j) + \pi(i, 2, j) \}. \]
5.4. System Performance Measures

10. The effective interruption rate is given by $E_{\text{INTR}} = \delta \sum_{i=0}^{\infty} \sum_{j=1}^{S} \pi(i,1,j)$.

11. The effective repair rate is given by $E_{\text{REP}} = \delta \sum_{i=0}^{\infty} \sum_{j=1}^{S} \pi(i,2,j)$.

12. The effective loss rate of orbital customers after seeing an interrupted server on retrial is given by $E_{\text{LOSS}} = \sum_{i=0}^{\infty} \sum_{j=1}^{S} qi\theta \pi(i,2,j)$.

13. The effective rate at which arriving customers are lost on seeing an interrupted server $E_{\text{A LOSS}} = (1-p)\lambda \sum_{i=0}^{\infty} \sum_{j=1}^{S} \pi(i,2,j)$.

14. The effective rate at which customers are lost finding the inventory level as zero $E_{\text{O LOSS}} = \sum_{i=0}^{\infty} \bar{\lambda} \pi(i,0,0)$.

15. The effective rate at which orders are placed $E_{\text{OR}} = \sum_{i=0}^{\infty} \mu \pi(i,1,s+1)$.

16. The expected rate at which customers are lost $E_{\text{LOSS}} = E_{\text{LOSS}} + E_{\text{A LOSS}} + E_{\text{O LOSS}}$.

5.4.3. The Cost Function

To investigate whether an optimal value exists for the re-order level $s$, we studied the following cost function.

$COST = C_{\text{INTR}} * E_{\text{INTR}} + C_{\text{LOSS}} * E_{\text{LOSS}} + C_{\text{N}} * E(\sigma) + C_{\text{I}} * E(\omega)$,

where $C_{\text{INTR}}$ is the cost per interruption per unit time, $C_{\text{LOSS}}$ is the unit time cost assigned when a customer is lost, $C_{\text{N}}$ is the holding cost per customer per unit time and $C_{\text{I}}$ is the inventory holding cost.
5.5 Numerical Illustration

In this section, we provide numerical illustration of the system performance as the underlying parameters vary.

5.5.1 Effect of the Retrial Rate $\theta$

Table 1(a) shows that as the retrial rate $\theta$ increases, the loss rate of retrying customers $ER_{LOSS}$ increases; the main reason for this is the high value for the system quitting probability $q (= 0.6)$. This increase in the loss rate leads to a decrease in the expected number of customers $E(\sigma)$ and hence a decrease in the server busy probability $P_\beta$ and server interruption probability $P_\alpha$. Note that the decrease in the server interruption probability may be occurring because of the possible decrease in the number of services due to customer loss from the orbit. Hence the decrease in $P_\alpha$, considering the corresponding decrease in $P_\beta$, should not be viewed as a gain to the system under study. Also the decrease in the interruption probability $P_\alpha$ may be taken as the reason for the decrease in the loss rate $EA_{LOSS}$ of arriving customers. From the Table, one can infer that the idle probability of the server is increasing with $\theta$; but this does not imply an increase in the number of successful retrials $E(s\tau)$. This decrease in $E(s\tau)$ with increase in $\theta$ may be due to the decrease in the number of customers in the system. There is a slight increase in the expected inventory level and a narrow decrease in the effective replenishment rate $EFRR$; the reason for this could be the decrease in server busy probability. Because of the increase in the expected inventory level, the loss rate due to zero inventory $EOL_{LOSS}$ must be increasing; but the Table displays a constant $EOL_{LOSS}$, which indicates that the change may be too small. Table 1(b) shows a decrease in the
expected waiting time of an orbital customer with increase in retrial rate \( \theta \); but one can see in the same Table that the conditional probability that a customer may quit the system without receiving any service is increasing. So the decrease in the waiting time does not favor the orbital customers.

**Table 5.1 (a): Effect of retrial rate \( \theta \) on various performance measures**

\[ \lambda = 2, \mu = 4, \eta = 1, \delta_1 = 2, \delta_2 = 2.5, p = 0.5, q = 0.6, s = 10, S = 31 \]

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( P_\beta )</th>
<th>( P_\alpha )</th>
<th>EFRR</th>
<th>( EA_{LOSS} )</th>
<th>( ER_{LOSS} )</th>
<th>( EO_{LOSS} )</th>
<th>( E(\sigma) )</th>
<th>( E(\tau) )</th>
<th>( E(\omega) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.0</td>
<td>0.3332</td>
<td>0.2666</td>
<td>0.0635</td>
<td>0.2666</td>
<td>0.4798</td>
<td>0.0004</td>
<td>0.7449</td>
<td>0.5334</td>
<td>19.6672</td>
</tr>
<tr>
<td>3.2</td>
<td>0.3325</td>
<td>0.2660</td>
<td>0.0633</td>
<td>0.2660</td>
<td>0.5108</td>
<td>0.0004</td>
<td>0.7065</td>
<td>0.5284</td>
<td>19.6698</td>
</tr>
<tr>
<td>3.4</td>
<td>0.3319</td>
<td>0.2655</td>
<td>0.0632</td>
<td>0.2655</td>
<td>0.5417</td>
<td>0.0004</td>
<td>0.6727</td>
<td>0.5236</td>
<td>19.6723</td>
</tr>
<tr>
<td>3.6</td>
<td>0.3313</td>
<td>0.2651</td>
<td>0.0631</td>
<td>0.2651</td>
<td>0.5725</td>
<td>0.0004</td>
<td>0.6426</td>
<td>0.5191</td>
<td>19.6747</td>
</tr>
<tr>
<td>3.8</td>
<td>0.3308</td>
<td>0.2646</td>
<td>0.0630</td>
<td>0.2646</td>
<td>0.6033</td>
<td>0.0004</td>
<td>0.6157</td>
<td>0.5148</td>
<td>19.6770</td>
</tr>
<tr>
<td>4.0</td>
<td>0.3302</td>
<td>0.2642</td>
<td>0.0629</td>
<td>0.2642</td>
<td>0.6340</td>
<td>0.0004</td>
<td>0.5915</td>
<td>0.5108</td>
<td>19.6791</td>
</tr>
<tr>
<td>4.2</td>
<td>0.3297</td>
<td>0.2638</td>
<td>0.0628</td>
<td>0.2638</td>
<td>0.6647</td>
<td>0.0004</td>
<td>0.5696</td>
<td>0.5070</td>
<td>19.6811</td>
</tr>
<tr>
<td>4.4</td>
<td>0.3292</td>
<td>0.2634</td>
<td>0.0627</td>
<td>0.2634</td>
<td>0.6954</td>
<td>0.0004</td>
<td>0.5496</td>
<td>0.5034</td>
<td>19.6830</td>
</tr>
</tbody>
</table>

**Table 5.1 (b): Effect of retrial rate \( \theta \) on waiting time**

\[ \lambda = 2, \mu = 4, \eta = 1, \delta_1 = 2, \delta_2 = 2.5, p = 0.5, q = 0.6, s = 4, S = 10 \]

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( E(W_L) )</th>
<th>( V(W_L) )</th>
<th>( P_{WS} )</th>
<th>( E(\sigma) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.0</td>
<td>0.2985</td>
<td>0.2042</td>
<td>0.1728</td>
<td>0.7517</td>
</tr>
<tr>
<td>3.2</td>
<td>0.2858</td>
<td>0.1883</td>
<td>0.1762</td>
<td>0.7133</td>
</tr>
<tr>
<td>3.4</td>
<td>0.2745</td>
<td>0.1748</td>
<td>0.1794</td>
<td>0.6794</td>
</tr>
<tr>
<td>3.6</td>
<td>0.2644</td>
<td>0.1631</td>
<td>0.1825</td>
<td>0.6493</td>
</tr>
<tr>
<td>3.8</td>
<td>0.2552</td>
<td>0.1529</td>
<td>0.1854</td>
<td>0.6223</td>
</tr>
<tr>
<td>4.0</td>
<td>0.2469</td>
<td>0.144</td>
<td>0.1882</td>
<td>0.5981</td>
</tr>
<tr>
<td>4.2</td>
<td>0.2393</td>
<td>0.1361</td>
<td>0.1909</td>
<td>0.5761</td>
</tr>
<tr>
<td>4.4</td>
<td>0.2323</td>
<td>0.1291</td>
<td>0.1935</td>
<td>0.5561</td>
</tr>
</tbody>
</table>
5.5.2 Effect of the Interruption Rate $\delta_1$

It follows from Table 2(a) that, as the interruption rate $\delta_1$ increases, the probability that the server being interrupted $P_\alpha$ increases and the server busy probability $P_\beta$ decreases; but the server busy probability is high compared to the server interruption probability and the reason for this may be the high repair rate compared to that of interruption rate. Note that as the interruption rate increases, this gap between $P_\alpha$ and $P_\beta$ diminishes with $P_\alpha$ dominating $P_\beta$. As frequent interruptions can cause lengthier services, loss rates $E_{A_{\text{LOSS}}}$ and $E_{R_{\text{LOSS}}}$ increases with increase in $\delta_1$. Note that the expected inventory level is increasing with interruption rate, which may be due to the decrease in the server busy probability and so less inventory may be served. Same reasoning can be made for the decrease in the effective replenishment rate $E_{\text{FRR}}$. The increase in the expected inventory level points to a decrease in the probability that the inventory level in the system is 0, and hence a decrease in the loss rate $E_{O_{\text{LOSS}}}$. The increase in the loss rate explains the decrease in the expected number of orbital customers. From the Table, one can infer that the idle probability is decreasing and this in turn leads to a decrease in the expected number of successful retrials $E(s\tau)$; another reason for this could be the increased loss rate of orbital customers. In Table 2(b), one can see that the expected waiting time of a customer in the orbit is decreasing with increase in the interruption rate and at the same time the conditional $P_{W_{\text{WS}}}$ probability that an orbital customer leaves the system without opting for service increases, which together points to the fact that the decrease in the waiting time is not in
favor of the customer. This happens despite a low number of customers in the orbit, which indicates the harm, which interruptions can cause.

Table 5.2(a): Effect of the interruption rate $\delta_1$ on various performance measures $\lambda=2, \mu=4, \theta=3, \eta=1, \delta_2=2.5, p=0.5, q=0.6, s=10, S=31$

<table>
<thead>
<tr>
<th>$\delta_1$</th>
<th>$P_\beta$</th>
<th>$P_\alpha$</th>
<th>$E_{\text{INTR}}$</th>
<th>EFRR</th>
<th>$E_{\text{LOSS}}$</th>
<th>$E_{\text{RLOSS}}$</th>
<th>ELOSS</th>
<th>$(\sigma)$</th>
<th>$(\sigma t)$</th>
<th>$(\omega)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.3941</td>
<td>0.1577</td>
<td>0.3941</td>
<td>0.0751</td>
<td>0.1577</td>
<td>0.4414</td>
<td>0.0010</td>
<td>0.8498</td>
<td>0.6831</td>
<td>19.4234</td>
</tr>
<tr>
<td>1.2</td>
<td>0.3798</td>
<td>0.1823</td>
<td>0.4557</td>
<td>0.0723</td>
<td>0.1823</td>
<td>0.4962</td>
<td>0.0008</td>
<td>0.8214</td>
<td>0.6457</td>
<td>19.4808</td>
</tr>
<tr>
<td>1.4</td>
<td>0.3667</td>
<td>0.2053</td>
<td>0.5134</td>
<td>0.0698</td>
<td>0.2053</td>
<td>0.5453</td>
<td>0.0007</td>
<td>0.7976</td>
<td>0.6128</td>
<td>19.5332</td>
</tr>
<tr>
<td>1.6</td>
<td>0.3547</td>
<td>0.2270</td>
<td>0.5676</td>
<td>0.0676</td>
<td>0.2270</td>
<td>0.5896</td>
<td>0.0006</td>
<td>0.7774</td>
<td>0.5836</td>
<td>19.5813</td>
</tr>
<tr>
<td>1.8</td>
<td>0.3435</td>
<td>0.2473</td>
<td>0.6184</td>
<td>0.0654</td>
<td>0.2474</td>
<td>0.6300</td>
<td>0.0005</td>
<td>0.7600</td>
<td>0.5572</td>
<td>19.6258</td>
</tr>
<tr>
<td>2.0</td>
<td>0.3332</td>
<td>0.2666</td>
<td>0.6664</td>
<td>0.0635</td>
<td>0.2666</td>
<td>0.6670</td>
<td>0.0004</td>
<td>0.7449</td>
<td>0.5335</td>
<td>19.6672</td>
</tr>
<tr>
<td>2.2</td>
<td>0.3236</td>
<td>0.2847</td>
<td>0.7118</td>
<td>0.0616</td>
<td>0.2847</td>
<td>0.7011</td>
<td>0.0004</td>
<td>0.7316</td>
<td>0.5118</td>
<td>19.7057</td>
</tr>
<tr>
<td>3.0</td>
<td>0.2966</td>
<td>0.3487</td>
<td>0.8717</td>
<td>0.0553</td>
<td>0.3487</td>
<td>0.4885</td>
<td>0.0002</td>
<td>0.6917</td>
<td>0.4413</td>
<td>19.8377</td>
</tr>
<tr>
<td>6.0</td>
<td>0.2124</td>
<td>0.5098</td>
<td>1.2748</td>
<td>0.0405</td>
<td>0.5098</td>
<td>0.6405</td>
<td>0.0000</td>
<td>0.6239</td>
<td>0.2942</td>
<td>20.1503</td>
</tr>
</tbody>
</table>

Table 5.2(b): Effect of the interruption rate $\delta_1$ on waiting time $\lambda=2, \mu=4, \theta=3, \eta=1, \delta_2=2.5, p=0.5, q=0.6, s=4, S=10$

<table>
<thead>
<tr>
<th>$\delta_1$</th>
<th>$E(W_L)$</th>
<th>$V(W_L)$</th>
<th>$(\sigma)$</th>
<th>$P_{WS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.3052</td>
<td>0.2509</td>
<td>0.8635</td>
<td>0.1027</td>
</tr>
<tr>
<td>1.2</td>
<td>0.3025</td>
<td>0.2374</td>
<td>0.8332</td>
<td>0.1185</td>
</tr>
<tr>
<td>1.4</td>
<td>0.3007</td>
<td>0.2266</td>
<td>0.8078</td>
<td>0.1333</td>
</tr>
<tr>
<td>1.6</td>
<td>0.2995</td>
<td>0.2179</td>
<td>0.7863</td>
<td>0.1473</td>
</tr>
<tr>
<td>1.8</td>
<td>0.2988</td>
<td>0.2104</td>
<td>0.7677</td>
<td>0.1604</td>
</tr>
<tr>
<td>2.0</td>
<td>0.2985</td>
<td>0.2042</td>
<td>0.7516</td>
<td>0.1728</td>
</tr>
</tbody>
</table>
5.5 Numerical Illustration

Figure 5.1(a), (b). Impact of the interruption rate $\delta_1$ on the expected number of customers in the system $E(\sigma)$ and on the expected inventory level $E(\omega)$ with parameters $\lambda=2$, $\mu=4$, $\eta=1$, $\delta_1=2$, $\delta_2=2.5$, $p=0.5$, $q=0.6$, $s=10$, $S=31$: The Figures show an opposite behavior of the measures $E(\sigma)$ and $E(\omega)$ with increase in the interruption rate; where the expected number of customers is decreasing due to loss and the expected inventory increases because of a possible drop in the number of service completions. The Figures reflect the harm brought to the system by interruption.

5.5.3 Effect of the Repair Rate $\delta_2$

In Table 3(a), one sees that, the increase in the repair rate leads to a decrease in the server interruption probability and to an increase in the server busy probability. This is expected; as the repair rate increases, the span of interruption period must be decreasing. The reason for increase in the effective interruption rate $E_{\text{INTR}}$ is the increase in the server busy probability. Here note that the interruption rate $\delta_1$ and the repair rate $\delta_2$ are the only two parameters
whose changes make the server busy probability $P_\beta$ and the server interruption probability $P_\alpha$ to vary in opposite directions; that is, if one probability increases, the other probability decreases. Now as these probabilities vary in opposite directions, the effective interruption rate varies in the same direction as server busy probability. Coming back to the repair rate $\delta_2$, from the Table 3(a), one observes that as the repair rate increases, the loss rate of the retrying customers $ER_{\text{LOSS}}$ decreases and hence the expected orbit size $E(\sigma)$ increases. This is expected, since fast repairs allow the server to render service to more customers. Now more number of services leads to a decrease in the inventory level, an increase in the effective replenishment rate $EFRR$ and a narrow increase in the loss rate of customers due to zero inventory $EO_{\text{LOSS}}$. From Table 3(a), one can infer that the server idle probability is increasing, which is also due to an increase in the number of fast service completions and after each service completion, the server becomes idle as we are considering a retrial queue. The increase in the server idle probability then leads to a decrease in the loss rate of arriving customers $EA_{\text{LOSS}}$ and to an increase in the expected number of successful retrials. In Table 3(b), one can see that the expected waiting time of an orbital customer is decreasing with increase in $\delta_2$, which is also due to fast service completions. Note also that the conditional probability that an orbital customer may leave the system without receiving service is decreasing when the repair becomes faster.
Table 5.3(a): Effect of the repair rate $\delta_2$ on various performance measures

$\lambda=2, \mu=4, \theta=3, \eta=1, \delta_1=2.5, p=0.5, q=0.6, s=10, S=31$

<table>
<thead>
<tr>
<th>$\delta_2$</th>
<th>$P_\beta$</th>
<th>$P_\alpha$</th>
<th>$E_{\text{INTR}}$</th>
<th>EFRR</th>
<th>$E_{\Delta \text{LOSS}}$</th>
<th>$E_{\text{RLOSS}}$</th>
<th>$E_{\text{OLoss}}$</th>
<th>$E(\sigma)$</th>
<th>$E(\tau r)$</th>
<th>$E(\omega)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0.2807</td>
<td>0.3743</td>
<td>0.5614</td>
<td>0.0535</td>
<td>0.3743</td>
<td>0.6737</td>
<td>0.0001</td>
<td>0.6973</td>
<td>0.4335</td>
<td>19.8771</td>
</tr>
<tr>
<td>1.7</td>
<td>0.2942</td>
<td>0.3461</td>
<td>0.5884</td>
<td>0.056</td>
<td>0.3461</td>
<td>0.6230</td>
<td>0.0002</td>
<td>0.7082</td>
<td>0.4580</td>
<td>19.8232</td>
</tr>
<tr>
<td>1.9</td>
<td>0.3059</td>
<td>0.3220</td>
<td>0.6117</td>
<td>0.0583</td>
<td>0.3220</td>
<td>0.5795</td>
<td>0.0002</td>
<td>0.7183</td>
<td>0.4799</td>
<td>19.7765</td>
</tr>
<tr>
<td>2.1</td>
<td>0.3161</td>
<td>0.3010</td>
<td>0.6322</td>
<td>0.0602</td>
<td>0.3010</td>
<td>0.5419</td>
<td>0.0002</td>
<td>0.7277</td>
<td>0.4995</td>
<td>19.7356</td>
</tr>
<tr>
<td>2.3</td>
<td>0.3251</td>
<td>0.2827</td>
<td>0.6503</td>
<td>0.0619</td>
<td>0.2827</td>
<td>0.5089</td>
<td>0.0002</td>
<td>0.7365</td>
<td>0.5172</td>
<td>19.6994</td>
</tr>
<tr>
<td>2.5</td>
<td>0.3332</td>
<td>0.2666</td>
<td>0.6664</td>
<td>0.0635</td>
<td>0.2666</td>
<td>0.4798</td>
<td>0.0003</td>
<td>0.7449</td>
<td>0.5334</td>
<td>19.6672</td>
</tr>
<tr>
<td>2.7</td>
<td>0.3405</td>
<td>0.2522</td>
<td>0.6809</td>
<td>0.0648</td>
<td>0.2522</td>
<td>0.4539</td>
<td>0.0003</td>
<td>0.7528</td>
<td>0.5483</td>
<td>19.6382</td>
</tr>
</tbody>
</table>

Table 5.3(b): Effect of the repair rate $\delta_2$ on waiting time

$\lambda=2, \mu=4, \theta=3, \eta=1, \delta_1=2, p=0.5, q=0.6, s=4, S=10$

<table>
<thead>
<tr>
<th>$\delta_2$</th>
<th>$E(W_L)$</th>
<th>$V(W_L)$</th>
<th>$P_{WS}$</th>
<th>$E(\sigma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0.3061</td>
<td>0.1879</td>
<td>0.2437</td>
<td>0.702</td>
</tr>
<tr>
<td>1.7</td>
<td>0.3037</td>
<td>0.1915</td>
<td>0.225</td>
<td>0.7133</td>
</tr>
<tr>
<td>1.9</td>
<td>0.3019</td>
<td>0.1949</td>
<td>0.209</td>
<td>0.7239</td>
</tr>
<tr>
<td>2.1</td>
<td>0.3005</td>
<td>0.1982</td>
<td>0.1953</td>
<td>0.7337</td>
</tr>
<tr>
<td>2.3</td>
<td>0.2994</td>
<td>0.2012</td>
<td>0.1833</td>
<td>0.7429</td>
</tr>
<tr>
<td>2.5</td>
<td>0.2985</td>
<td>0.2042</td>
<td>0.1728</td>
<td>0.7517</td>
</tr>
<tr>
<td>2.7</td>
<td>0.2978</td>
<td>0.2070</td>
<td>0.1636</td>
<td>0.7599</td>
</tr>
</tbody>
</table>

5.5.4 Effect of the Re-Order Level $s$

Tables 4(a) and (b) describes the effect of the re-order level $s$ on various system performance measures. As the re-order level $s$ increases, expected inventory level increases and hence the probability for a loss due to zero
5.5 Numerical Illustration

inventories in the system decreases. Note that the rate at which the retrying customers and arriving customers are lost due to interruption increases with increase in $s$. The reason for this may be the increase in the server busy probability $P_{\beta}$ and a corresponding increase in the server interruption probability $P_{\alpha}$. The decrease in the expected number of customers in the orbit also has the same reason. The Table shows that as $s$ increases, there is a slight decrease in the rate of successful retrials; which can be attributed to the slight decrease in the server idle probability. Obviously, the effective replenishment rate EFRR has to increase with increase in $s$ but the lower values for EFRR as well as the high values for the expected inventory level together suggests that replenishment is not frequently occurring in the system. Despite of the high service rate (twice as much as the arrival rate), the less frequent replenishments points to the severe effect of interruption on the system behavior. Table 4(b) shows a narrow decrease in the waiting time of a customer in the orbit with an increase in $s$. Note that one has to wait for some time besides a low expected number of customers in the orbit; and that there is a high probability that one may choose to leave the system without opting for service.

Table 4(a) shows that the expected inventory level in the system is high even when the re-order level $s$ is small. This made us to investigate whether an optimal value for $s$ can be found. For this we studied the cost function defined in section 4.2. In the cost function, the measures $E_{\text{INTR}}$ and $E(\omega)$ shows an increase with increase in $s$, while the measures $E_{\text{LOSS}}$ and $E(\sigma)$ shows a decrease. Here note that the increase in the expected inventory is significant as compared to the changes in the other measures and therefore the cost will
ultimately be increasing with $s$, which points to the optimal value zero for $s$. What we want to capture is the decrease in the measures $E_{\text{LOSS}}$ and $E(\sigma)$; and hence get a convex nature for the cost function and an optimal value other than zero for $s$. For doing this, we assume a comparatively large cost for the loss of customers $C_{\text{LOSS}}$, which is also reasonable in many practical situations. As expected, this assumption leads us to an optimal value for $s$ ($s = 5$), as one can infer from Table 4(c). This Table also shows that if the cost $C_{\text{LOSS}}$ is not very high, the cost function is linearly increasing. These results can be more easily verified from Figures 3(a) and (b).

**Table 5.4(a): Effect of the re-order level $s$ on various performance measures**

$\lambda=2$, $\mu=4$, $\theta=3$, $\eta=1$, $\delta_1=2$, $\delta_2=2.5$, $p=0.5$, $q=0.6$, $S=31$

<table>
<thead>
<tr>
<th>$s$</th>
<th>$P_\beta$</th>
<th>$P_\alpha$</th>
<th>$E_{\text{FRR}}$</th>
<th>$E_{A_{\text{LOSS}}}$</th>
<th>$E_{R_{\text{LOSS}}}$</th>
<th>$E_{O_{\text{LOSS}}}$</th>
<th>$E(\sigma)$</th>
<th>$E(s\tau)$</th>
<th>$E(\omega)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.3323</td>
<td>0.2658</td>
<td>0.0511</td>
<td>0.2658</td>
<td>0.4785</td>
<td>0.0058</td>
<td>0.7457</td>
<td>0.540</td>
<td>17.19</td>
</tr>
<tr>
<td>6</td>
<td>0.3327</td>
<td>0.2662</td>
<td>0.0532</td>
<td>0.2662</td>
<td>0.4791</td>
<td>0.0034</td>
<td>0.7454</td>
<td>0.537</td>
<td>17.68</td>
</tr>
<tr>
<td>7</td>
<td>0.3329</td>
<td>0.2664</td>
<td>0.0555</td>
<td>0.2664</td>
<td>0.4794</td>
<td>0.0020</td>
<td>0.7451</td>
<td>0.535</td>
<td>18.17</td>
</tr>
<tr>
<td>8</td>
<td>0.3331</td>
<td>0.2665</td>
<td>0.0579</td>
<td>0.2665</td>
<td>0.4796</td>
<td>0.0012</td>
<td>0.745</td>
<td>0.534</td>
<td>18.67</td>
</tr>
<tr>
<td>9</td>
<td>0.3332</td>
<td>0.2665</td>
<td>0.0606</td>
<td>0.2665</td>
<td>0.4797</td>
<td>0.0007</td>
<td>0.7449</td>
<td>0.534</td>
<td>19.17</td>
</tr>
<tr>
<td>10</td>
<td>0.3332</td>
<td>0.2666</td>
<td>0.0635</td>
<td>0.2666</td>
<td>0.4798</td>
<td>0.0004</td>
<td>0.7449</td>
<td>0.533</td>
<td>19.67</td>
</tr>
<tr>
<td>11</td>
<td>0.3332</td>
<td>0.2666</td>
<td>0.0666</td>
<td>0.2666</td>
<td>0.4799</td>
<td>0.0002</td>
<td>0.7449</td>
<td>0.533</td>
<td>20.17</td>
</tr>
<tr>
<td>12</td>
<td>0.3333</td>
<td>0.2666</td>
<td>0.0702</td>
<td>0.2666</td>
<td>0.4799</td>
<td>0.0001</td>
<td>0.7448</td>
<td>0.533</td>
<td>20.67</td>
</tr>
</tbody>
</table>
Table 5.4(b): Effect of the re-order level $s$ on waiting time
\[ \lambda=2, \mu=4, \theta=3, \eta=1, \delta_1=2, \delta_2=2.5, p=0.5, q=0.6, S=15 \]

<table>
<thead>
<tr>
<th>$s$</th>
<th>$E(W_L)$</th>
<th>$V(W_L)$</th>
<th>$P_{WS}$</th>
<th>$E(\sigma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.2984</td>
<td>0.2022</td>
<td>0.1733</td>
<td>0.7508</td>
</tr>
<tr>
<td>4</td>
<td>0.2981</td>
<td>0.1952</td>
<td>0.1746</td>
<td>0.7486</td>
</tr>
<tr>
<td>5</td>
<td>0.2979</td>
<td>0.1910</td>
<td>0.1754</td>
<td>0.7472</td>
</tr>
<tr>
<td>6</td>
<td>0.2977</td>
<td>0.1884</td>
<td>0.1759</td>
<td>0.7463</td>
</tr>
<tr>
<td>7</td>
<td>0.2977</td>
<td>0.1869</td>
<td>0.1761</td>
<td>0.7458</td>
</tr>
</tbody>
</table>

Table 5.4(c): Effect of the re-order level $s$ on the cost function
\[ \lambda=2, \mu=4, \theta=3, \eta=1, \delta_1=2, \delta_2=2.5, p=0.5, q=0.6, S=25, CN=50, CI=60, CINTR=40 \]

<table>
<thead>
<tr>
<th>$S$</th>
<th>COST ($C_{LOSS}=10000$)</th>
<th>COST ($C_{LOSS}=500$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>7546</td>
<td>1075</td>
</tr>
<tr>
<td>4</td>
<td>7510</td>
<td>1094</td>
</tr>
<tr>
<td>5</td>
<td>7500</td>
<td>1117</td>
</tr>
<tr>
<td>6</td>
<td>7504</td>
<td>1140</td>
</tr>
<tr>
<td>7</td>
<td>7516</td>
<td>1164</td>
</tr>
<tr>
<td>8</td>
<td>7534</td>
<td>1189</td>
</tr>
<tr>
<td>9</td>
<td>7555</td>
<td>1213</td>
</tr>
</tbody>
</table>
5.5 Numerical Illustration

Figure 5.2 (a), (b). Impact of the re-order level \( s \) on expected waiting time and on the loss rate of customers with parameters \( \lambda = 2, \ \mu = 4, \ \theta = 3, \ \eta = 1, \ \delta_1 = 2, \ \delta_2 = 2.5, \ p = 0.5, \ q = 0.6, \ S = 15 \): Both the waiting time and the loss rate are decreasing with increase in \( s \). An increase in the expected inventory in the system brought by the increase in \( s \) can be thought of as the reason for the decrease in the total loss rate. A reference to table 4(a) shows that with an increase in \( s \), there is a slight increase in the loss rate due to customers (both orbital and external) seeing an interrupted server. This together with the increase in the expected inventory level can be thought of as the reason behind the decrease in the waiting time. Note that the waiting time of an orbital customer may end with a quit from the system without receiving service.
5.5 Numerical Illustration

**Figure 5.3 (a), (b).** Investigation of an optimal value for the re-order level \( s \) with parameters \( \lambda = 2, \mu = 4, \theta = 3, \eta = 1, \delta_1 = 2, \delta_2 = 2.5, p = 0.5, q = 0.6, S = 25, CN = 50, CI = 60, CINTR = 40: \)

The cost curve in Figure (a), where the cost incurred due to customer loss is high (\(=10000\)), shows a convex nature for the cost function and an optimal value 5 for the re-order level \( s \); while in Figure (b), as the loss cost is not very high (\(=500\)), the cost function is linearly increasing with increase in \( s \).

### 5.5.5 Effect of the Maximum Inventory Level \( S \)

Table 5(a) shows that as the maximum inventory level \( S \) has only very little effect on majority of the system performance measures. The decrease in the effective replenishment rate EFRR and the increase in the
expected inventory level in the system is quiet natural with increase in \( S \). Because the inventory in the system is increasing, the loss due to zero inventories is decreasing. There is a very narrow increase in the server interruption probability, which may be due to the narrow increase in the number of orbital customers.

Table 5.5(a): Effect of the maximum inventory level \( S \) on various performance measures with \( \lambda=2, \mu=4, \theta=3, \eta=1, \delta_2=2.5, \delta_1=2, p=0.5, q=0.6, s=10 \)

<table>
<thead>
<tr>
<th>( S )</th>
<th>( P_\beta )</th>
<th>( P_\alpha )</th>
<th>( \mathcal{E}_{\text{DTR}} )</th>
<th>( \text{EFRR} )</th>
<th>( \mathcal{E}_{\text{LOSS}} )</th>
<th>( \mathcal{E}_{\text{LOSS}} )</th>
<th>( \mathcal{E}_{\text{LOSS}} )</th>
<th>( E(\sigma) )</th>
<th>( E(\sigma) )</th>
<th>( E(\omega) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>0.3332</td>
<td>0.2665</td>
<td>0.6663</td>
<td>0.1025</td>
<td>0.2665</td>
<td>0.4798</td>
<td>0.0007</td>
<td>0.7449</td>
<td>0.5338</td>
<td>15.666</td>
</tr>
<tr>
<td>24</td>
<td>0.3332</td>
<td>0.2665</td>
<td>0.6663</td>
<td>0.0952</td>
<td>0.2665</td>
<td>0.4798</td>
<td>0.0006</td>
<td>0.7449</td>
<td>0.5337</td>
<td>16.166</td>
</tr>
<tr>
<td>25</td>
<td>0.3332</td>
<td>0.2665</td>
<td>0.6664</td>
<td>0.0888</td>
<td>0.2665</td>
<td>0.4798</td>
<td>0.0005</td>
<td>0.7449</td>
<td>0.5336</td>
<td>16.666</td>
</tr>
<tr>
<td>26</td>
<td>0.3332</td>
<td>0.2665</td>
<td>0.6664</td>
<td>0.0833</td>
<td>0.2665</td>
<td>0.4798</td>
<td>0.0005</td>
<td>0.7449</td>
<td>0.5336</td>
<td>17.166</td>
</tr>
<tr>
<td>27</td>
<td>0.3332</td>
<td>0.2665</td>
<td>0.6664</td>
<td>0.0784</td>
<td>0.2665</td>
<td>0.4798</td>
<td>0.0004</td>
<td>0.7449</td>
<td>0.5335</td>
<td>17.666</td>
</tr>
<tr>
<td>28</td>
<td>0.3332</td>
<td>0.2665</td>
<td>0.6664</td>
<td>0.074</td>
<td>0.2666</td>
<td>0.4798</td>
<td>0.0005</td>
<td>0.7449</td>
<td>0.5335</td>
<td>18.166</td>
</tr>
<tr>
<td>29</td>
<td>0.3332</td>
<td>0.2666</td>
<td>0.6664</td>
<td>0.0701</td>
<td>0.2666</td>
<td>0.4798</td>
<td>0.0005</td>
<td>0.7449</td>
<td>0.5335</td>
<td>18.666</td>
</tr>
<tr>
<td>30</td>
<td>0.3332</td>
<td>0.2666</td>
<td>0.6664</td>
<td>0.0666</td>
<td>0.2666</td>
<td>0.4798</td>
<td>0.0004</td>
<td>0.7449</td>
<td>0.5335</td>
<td>19.166</td>
</tr>
</tbody>
</table>

Table 5.5(b): Effect of the maximum inventory level \( S \) on waiting time

\( \lambda=2, \mu=4, \theta=3, \eta=1, \delta_1=2, \delta_2=2.5, p=0.5, q=0.6, s=4 \)

<table>
<thead>
<tr>
<th>( S )</th>
<th>( E(W_L) )</th>
<th>( V(W_L) )</th>
<th>( P_{WS} )</th>
<th>( E(\sigma) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.2985</td>
<td>0.2042</td>
<td>0.1728</td>
<td>0.7517</td>
</tr>
<tr>
<td>11</td>
<td>0.2984</td>
<td>0.2014</td>
<td>0.1734</td>
<td>0.7507</td>
</tr>
<tr>
<td>12</td>
<td>0.2983</td>
<td>0.1993</td>
<td>0.1738</td>
<td>0.7500</td>
</tr>
<tr>
<td>13</td>
<td>0.2982</td>
<td>0.1976</td>
<td>0.1741</td>
<td>0.7494</td>
</tr>
<tr>
<td>14</td>
<td>0.2981</td>
<td>0.1963</td>
<td>0.1744</td>
<td>0.7490</td>
</tr>
<tr>
<td>15</td>
<td>0.2981</td>
<td>0.1952</td>
<td>0.1746</td>
<td>0.7486</td>
</tr>
</tbody>
</table>
5.5.6 Effect of the Joining Probability \( p \)

Table 6(a) and (b) studies the effect of the joining probability \( p \) of an arriving customer on the system behavior. Quiet naturally, the loss rate of arriving customers decreases with increase in the joining probability \( p \). As more customers join the system, the expected number of customers \( E(\sigma) \), increases; the increase in \( E(\sigma) \) then leads to more retrials and therefore an increase in both the successful number of retrials \( E(\sigma_1) \) and in the loss rate after retrials \( ER_{LOSS} \). The increase in the number of orbital customers makes the server busier; but as the number of services increases, the probability of seeing an interrupted server \( P_\alpha \) also increases. The possible increase in the number of services leads to an increase, though narrow, in the effective replenishment rate \( EFRR \). For similar reasons, there is a decrease in the expected inventory level in the system and an increase in the loss rate of customers due to zero inventory \( EO_{LOSS} \). Both these changes, especially that in \( EO_{LOSS} \), are narrow which reflects the interruption factor affecting the system performance. Table 6(b) shows a slight decrease in the expected waiting time of a customer in the orbit; also note that the conditional probability that an orbital customer quits the system without receiving service is decreasing with increase in \( p \). This is expected because when the joining probability \( p \) increases, the server busy probability shows a significant increase compared to the increase in the server interruption probability; and the customer loss, on arrival, takes place with probability \( 1-p \) only when the server is interrupted. Thus increase in \( p \) leads to more service completions and this favors the system performance.
Table 5.6(a): Effect of the joining probability \( p \) on various performance measures

\( \lambda = 2, \mu = 4, \theta = 3, \eta = 1, \delta_1 = 2, \delta_2 = 2.5, q = 0.6, s = 10, S = 31 \)

<table>
<thead>
<tr>
<th>( p )</th>
<th>( P_\beta )</th>
<th>( P_\alpha )</th>
<th>EFRR</th>
<th>( E_A^{\text{LOSS}} )</th>
<th>( E_R^{\text{LOSS}} )</th>
<th>( E_O^{\text{LOSS}} )</th>
<th>( E(\sigma) )</th>
<th>( E(st) )</th>
<th>( E(\omega) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.3186</td>
<td>0.2549</td>
<td>0.0607</td>
<td>0.5097</td>
<td>0.4588</td>
<td>0.0003</td>
<td>0.5180</td>
<td>0.4220</td>
<td>19.7257</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3244</td>
<td>0.2596</td>
<td>0.0618</td>
<td>0.4153</td>
<td>0.4672</td>
<td>0.0004</td>
<td>0.6052</td>
<td>0.4666</td>
<td>19.7022</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3303</td>
<td>0.2642</td>
<td>0.0629</td>
<td>0.3171</td>
<td>0.4756</td>
<td>0.0004</td>
<td>0.6971</td>
<td>0.5112</td>
<td>19.6788</td>
</tr>
<tr>
<td>0.6</td>
<td>0.3361</td>
<td>0.2689</td>
<td>0.0640</td>
<td>0.2151</td>
<td>0.4840</td>
<td>0.0004</td>
<td>0.7939</td>
<td>0.5556</td>
<td>19.6550</td>
</tr>
<tr>
<td>0.8</td>
<td>0.3419</td>
<td>0.2735</td>
<td>0.0651</td>
<td>0.1094</td>
<td>0.4923</td>
<td>0.0005</td>
<td>0.8955</td>
<td>0.5998</td>
<td>19.6324</td>
</tr>
<tr>
<td>1</td>
<td>0.3476</td>
<td>0.2781</td>
<td>0.0662</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>0.6436</td>
</tr>
</tbody>
</table>

Table 5.6(b): Effect of the joining probability \( p \) on waiting time

\( \lambda = 2, \mu = 4, \theta = 3, \eta = 1, \delta_1 = 2, \delta_2 = 2.5, q = 0.6, s = 4, S = 10 \)

<table>
<thead>
<tr>
<th>( p )</th>
<th>( E(W_L) )</th>
<th>( V(W_L) )</th>
<th>( P_{WS} )</th>
<th>( E(\sigma) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.3067</td>
<td>0.2214</td>
<td>0.182</td>
<td>0.5251</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3015</td>
<td>0.2114</td>
<td>0.1769</td>
<td>0.6122</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2990</td>
<td>0.2058</td>
<td>0.1738</td>
<td>0.7040</td>
</tr>
<tr>
<td>0.6</td>
<td>0.2985</td>
<td>0.2033</td>
<td>0.1722</td>
<td>0.8006</td>
</tr>
<tr>
<td>0.8</td>
<td>0.2995</td>
<td>0.2030</td>
<td>0.1717</td>
<td>0.9020</td>
</tr>
<tr>
<td>1</td>
<td>0.3017</td>
<td>0.2044</td>
<td>0.1721</td>
<td>1.0080</td>
</tr>
</tbody>
</table>

5.5.7 Effect of the System Quitting Probability After an Unsuccessful Retrial \( q \)

After studying the effect of the joining probability \( p \), now in Tables 7(a) and (b), we focus our attention on the effect of the system quitting probability \( q \) of an orbiting customer after an unsuccessful retrial. The increase
in the loss rate after retrials $ER_{\text{LOSS}}$ and a resulting decrease in the expected number of orbital customers $E(\sigma)$ with increase in the loss probability $q$ is obvious. The decrease in the number of customers results in a decrease in the server busy probability $P_\beta$ and also in the server interruption probability $P_\alpha$.

Now this decrease in the probability $P_\alpha$ leads to a decrease in the loss rate of an arriving customer $EA_{\text{LOSS}}$. Note that the decrease in the server busy probability leads to an increase in the expected inventory level in the system and hence to a decrease in the effective replenishment rate $EFRR$ and in the customer loss rate due to zero inventory in the system, $EO_{\text{LOSS}}$. Table (b) shows that an increase in $q$ implies a decrease in the expected waiting time of an orbital customer; again this is not in favor of the customers as the probability that the customer quits the system before taken in to service is increasing.

**Table 5.7(a): Effect of the system quitting probability $q$ on various performance measures with $\lambda=2$, $\mu=4$, $\theta=3$, $\eta=1$, $\delta_2=2.5$, $\delta_1=2$, $p=0.5$, $s=10$, $S=31$**

<table>
<thead>
<tr>
<th>$q$</th>
<th>$P_\beta$</th>
<th>$P_\alpha$</th>
<th>$EFRR$</th>
<th>$EA_{\text{LOSS}}$</th>
<th>$ER_{\text{LOSS}}$</th>
<th>$EO_{\text{LOSS}}$</th>
<th>$E(\sigma)$</th>
<th>$E(\sigma \tau)$</th>
<th>$E(\omega)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.4163</td>
<td>0.3331</td>
<td>0.0793</td>
<td>0.3331</td>
<td>0</td>
<td>0.0170</td>
<td>3.6393</td>
<td>1.7600</td>
<td>19.335</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3593</td>
<td>0.2874</td>
<td>0.0684</td>
<td>0.2874</td>
<td>0.1724</td>
<td>0.0006</td>
<td>1.2689</td>
<td>0.7324</td>
<td>19.563</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3423</td>
<td>0.2739</td>
<td>0.0652</td>
<td>0.2739</td>
<td>0.3286</td>
<td>0.0005</td>
<td>0.9104</td>
<td>0.6031</td>
<td>19.631</td>
</tr>
<tr>
<td>0.6</td>
<td>0.3332</td>
<td>0.2666</td>
<td>0.0635</td>
<td>0.2666</td>
<td>0.4798</td>
<td>0.0004</td>
<td>0.7449</td>
<td>0.5334</td>
<td>19.667</td>
</tr>
<tr>
<td>0.8</td>
<td>0.3274</td>
<td>0.2619</td>
<td>0.0624</td>
<td>0.2619</td>
<td>0.6286</td>
<td>0.0004</td>
<td>0.6475</td>
<td>0.4891</td>
<td>19.690</td>
</tr>
<tr>
<td>1</td>
<td>0.3234</td>
<td>0.2587</td>
<td>0.0616</td>
<td>0.2587</td>
<td>0.776</td>
<td>0.0003</td>
<td>0.5829</td>
<td>0.4583</td>
<td>19.707</td>
</tr>
</tbody>
</table>
5.6 Concluding Remarks

Table 5.7(b): Effect of the system quitting probability $q$ on waiting time

$\lambda=2$, $\mu=4$, $\theta=3$, $\eta=1$, $\delta_1=2$, $\delta_2=2.5$, $p=0.5$, $s=4$, $S=10$

<table>
<thead>
<tr>
<th>$q$</th>
<th>$E(W_L)$</th>
<th>$V(W_L)$</th>
<th>$P_{WS}$</th>
<th>$E(\sigma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.3443</td>
<td>4.375</td>
<td>0</td>
<td>3.6609</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4819</td>
<td>0.4784</td>
<td>0.1131</td>
<td>1.2753</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3549</td>
<td>0.2728</td>
<td>0.1497</td>
<td>0.9171</td>
</tr>
<tr>
<td>0.6</td>
<td>0.2985</td>
<td>0.2042</td>
<td>0.1728</td>
<td>0.7517</td>
</tr>
<tr>
<td>0.8</td>
<td>0.2655</td>
<td>0.1702</td>
<td>0.1900</td>
<td>0.6542</td>
</tr>
<tr>
<td>1</td>
<td>0.2434</td>
<td>0.1497</td>
<td>0.2037</td>
<td>0.5895</td>
</tr>
</tbody>
</table>

5.6 Concluding Remarks

In this chapter we have considered an (s,S) inventory problem with positive service time and lead time. This is the first work in inventory with service interruption---server is subject to interruption while service is in progress. No waiting space is provided for customers, other than for the one whose service gets interrupted. Hence when a service is going on an external arrival has to go to an orbit of infinite capacity. Customers do not join the orbit when inventory level is zero nor when the server is under interruption. Retrial rate is a linear function of the number of customers present in the orbit. Retrial customers, encountering the server in breakdown condition, leave the system for ever with positive probability. This leads to the system being stable always. Also a primary customer, encountering the server in breakdown condition,
chooses to leave the system with positive probability or joins the orbit with complementary probability. All distributions involved are assumed to be exponential.

This system is studied by analyzing a truncated system and then we extended the results to the system with unlimited capacity for the orbit. The system is shown to be stable whenever the probability of leaving system for ever with positive probability as a consequence of the retrial customer encountering the server in break down condition. In the absence of this explicit condition for stability is derived.

A first step to extend the results here is to replace a few of the exponential distributions assumed in the chapter by more general distributions, such as the phase type. If service time duration is assumed to be phase type or at least Erlang of order two or more, then a few of the phases could be provided with protection from interruption. The cost of such protection could be incorporated to investigate the optimal number of phases to be protected. The quality of approximation of the expected waiting time will be improved in a future work.