Regents Park Publishers

Tutorial

DEN 423

Waiting Lines

Waiting Lines and Queuing Theory Models

Learning Objectives

After completing this Module, students will be able to:

1. Describe the trade-off curves for cost-of-waiting time and cost-of-service
2. Understand the three parts of a queuing system: the calling population, the queue itself, and the service facility
3. Describe the basic queuing system configurations
4. Understand the assumptions of the common models dealt with in this Module
5. Analyze a variety of operating characteristics of waiting lines

Overview

Introduction

- *Queuing theory* is the study of *waiting lines*
- It is one of the oldest and most widely used quantitative analysis techniques
- Waiting lines are an everyday occurrence for most people
- Queues form in business process as well
- The three basic components of a queuing process are arrivals, service facilities, and the actual waiting line
- Analytical models of waiting lines can help managers evaluate the cost and effectiveness of service systems

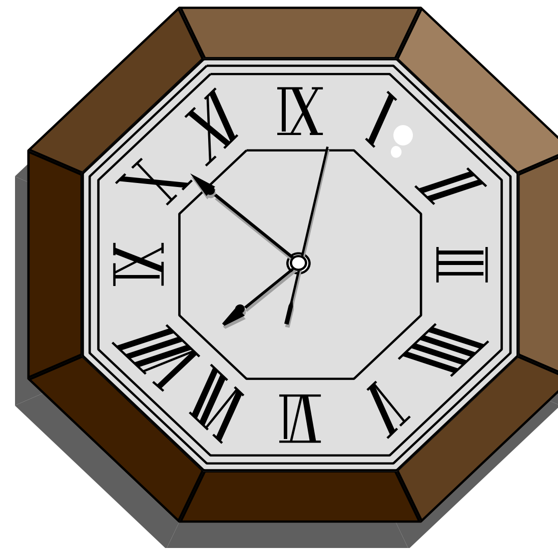
Lines and Waiting

"Every day I get in the queue, that waits for the bus that takes me to you …"

Pete Townshend, *Magic Bus*

Where the Time Goes

In a life time, the average person will spend:



SIX MONTHS :	Waiting at stoplights
EIGHT MONTHS:	Opening junk mail
ONE YEAR:	Looking for stuff
TWO YEARS:	Reading E-mail
FOUR YEARS:	Doing housework
FIVE YEARS:	Waiting in line
SIX YEARS:	Eating

Cultural Attitudes

- "Americans hate to wait. So business is trying a trick or two to make lines seem shorter…" *The New York Times, September 25, 1988*

- "An Englishman, even when he is by himself, will form an orderly queue of one…" *George Mikes, "How to be an Alien"*

Cultural Attitudes

- **"In the Soviet Union, waiting lines were used as a rationing device…"** *Hedrick Smith, "The Russians"*

Cultural Attitudes

- “In the United States, waiting lines were used as a rationing device too…
  - *Doctors’ Appointments at VA*

Waiting Realities

- *Inevitability of Waiting*: Waiting results from variations in arrival rates and service rates

- *Economics of Waiting*: High utilization purchased at the price of customer waiting. Make waiting productive (salad bar) or profitable (drinking bar).

Laws of Service

- *Maister's First Law*:
  Customers compare expectations with perceptions.
- *Maister's Second Law*:
  Is hard to play catch-up ball.
- *Skinner's Law*:
  The other line always moves faster.
- *Jenkin's Corollary*:
  However, when you switch to another other line, the line you left moves faster.

Service Providers have to Care about Customers' Experience

This Experience has to be designed, architected, implemented, and measured.

Characteristics of Queuing Systems

Characteristics of a Queuing System

- There are three parts to a queuing system
  1. The arrivals or inputs to the system (sometimes referred to as the *calling population*
  2. The queue or waiting line itself
  3. The service facility
- These components have their own characteristics that must be examined before mathematical models can be developed

Characteristics of a Queuing System

- Arrival Characteristics have three major characteristics, *size, pattern*, and *behavior*
  - Size of the calling population
    - Can be either unlimited (essentially *infinite*) or limited (*finite*)
  - Pattern of arrivals
    - Can arrive according to a known pattern or can arrive *randomly*
    - Random arrivals generally follow a *Poisson distribution*

Characteristics of a Queuing System

- The Poisson distribution is

$$P(X) = \frac{e^{-\lambda}\lambda^{X}}{X!} \text{ for } X = 0, 1, 2, 3, 4, \ldots$$

where

$$\begin{aligned} P(X) &= \text{probability of } X \text{ arrivals} \\ X &= \text{number of arrivals per unit of time} \\ \lambda &= \text{average arrival rate} \\ e &= 2.7183 \end{aligned}$$

Excel: POISSON. DIST

Characteristics of a Queuing System

- If $\lambda = 2$, we can find the values for $X = 0$, 1, and 2

$$P(X) = \frac{e^{-\lambda}\lambda^X}{X!}$$

$$P(0) = \frac{e^{-2}2^0}{0!} = \frac{0.1353(1)}{1} = 0.1353 = 14\%$$

$$P(1) = \frac{e^{-2}2^1}{1!} = \frac{e^{-2}2}{1} = \frac{0.1353(2)}{1} = 0.2706 = 27\%$$

$$P(2) = \frac{e^{-2}2^2}{2!} = \frac{e^{-2}4}{2(1)} = \frac{0.1353(4)}{2} = 0.2706 = 27\%$$

Characteristics of a Queuing System

- Behavior of arrivals
  - Most queuing models assume customers are patient and will wait in the queue until they are served and do not switch lines
  - *Balking* refers to customers who refuse to join the queue
  - *Reneging* customers enter the queue but become impatient and leave without receiving their service
  - That these behaviors exist is a strong argument for the use of queuing theory to managing waiting lines

Characteristics of a Queuing System

- Waiting Line Characteristics
  - Waiting lines can be either *limited* or *unlimited*
  - Queue discipline refers to the rule by which customers in the line receive service
  - The most common rule is *first-in, first-out* (*FIFO*)
  - Other rules are possible and may be based on other important characteristics
  - Other rules can be applied to select which customers enter which queue, but may apply FIFO once they are in the queue

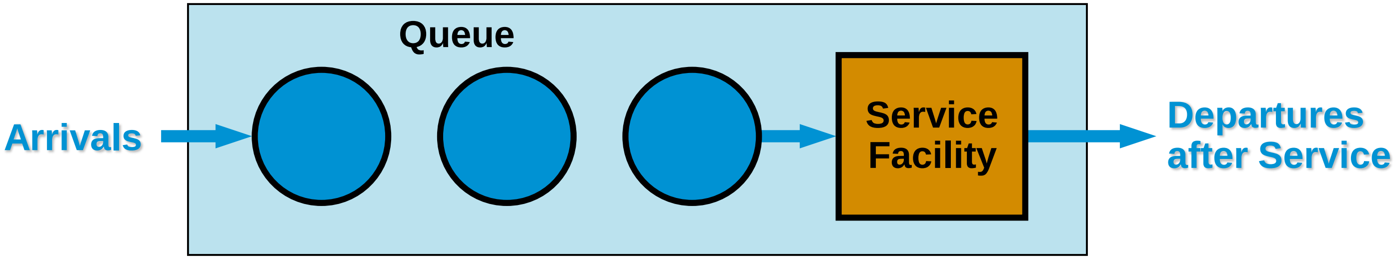
Queuing Systems

Characteristics of a Queuing System

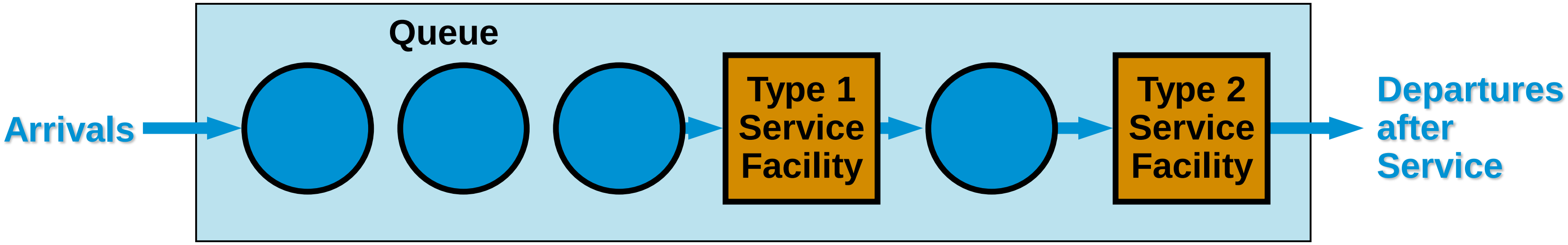
- Service Facility Characteristics
  - Basic queuing system configurations
    - Service systems are classified in terms of the number of channels, or servers, and the number of phases, or service stops
    - A *single-channel system* with one server is quite common
    - *Multichannel systems* exist when multiple servers are fed by one common waiting line
    - In a *single-phase system* the customer receives service form just one server
    - If a customer has to go through more than one server, it is a *multiphase system*

Characteristics of a Queuing System

- Four basic queuing system configurations



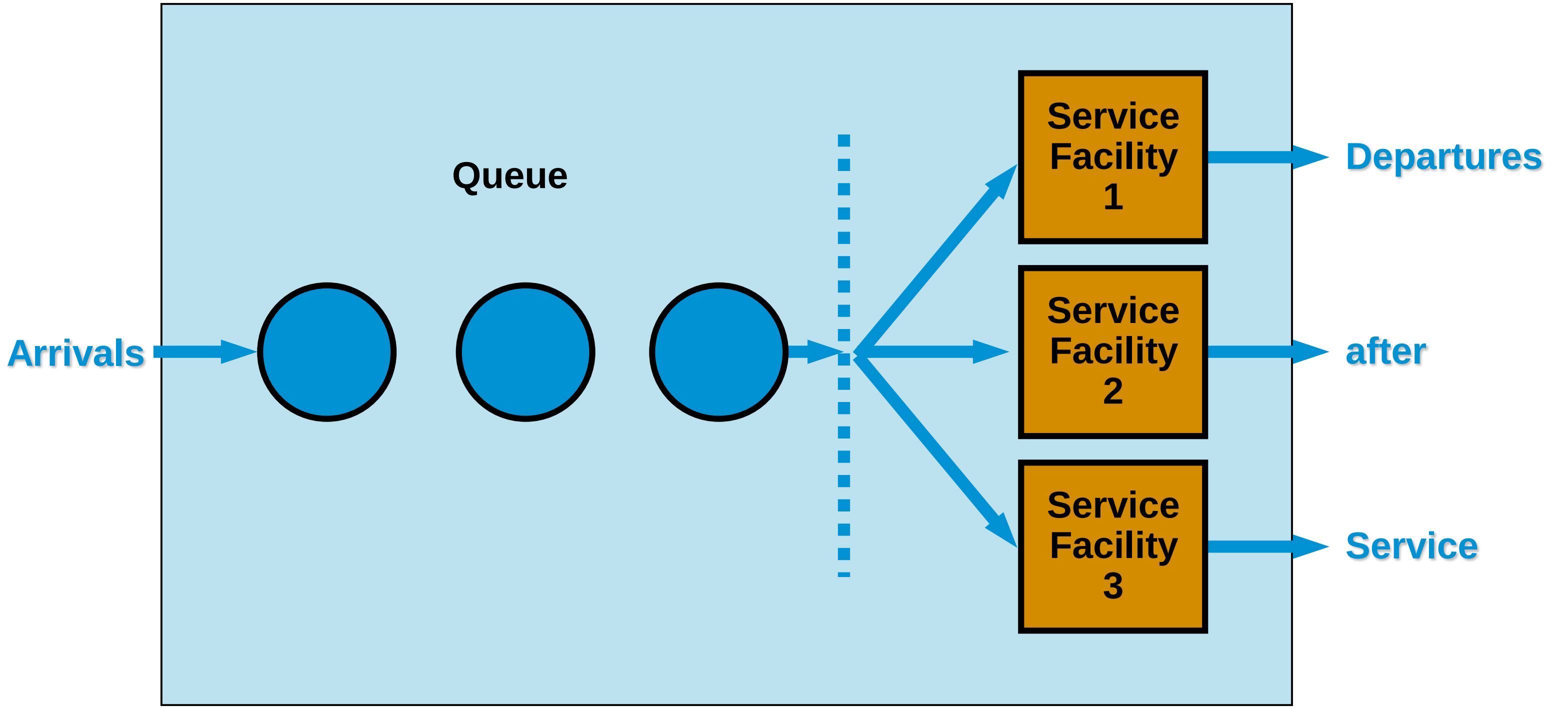
Single-Channel, Single-Phase System



Single-Channel, Multiphase System

Characteristics of a Queuing System

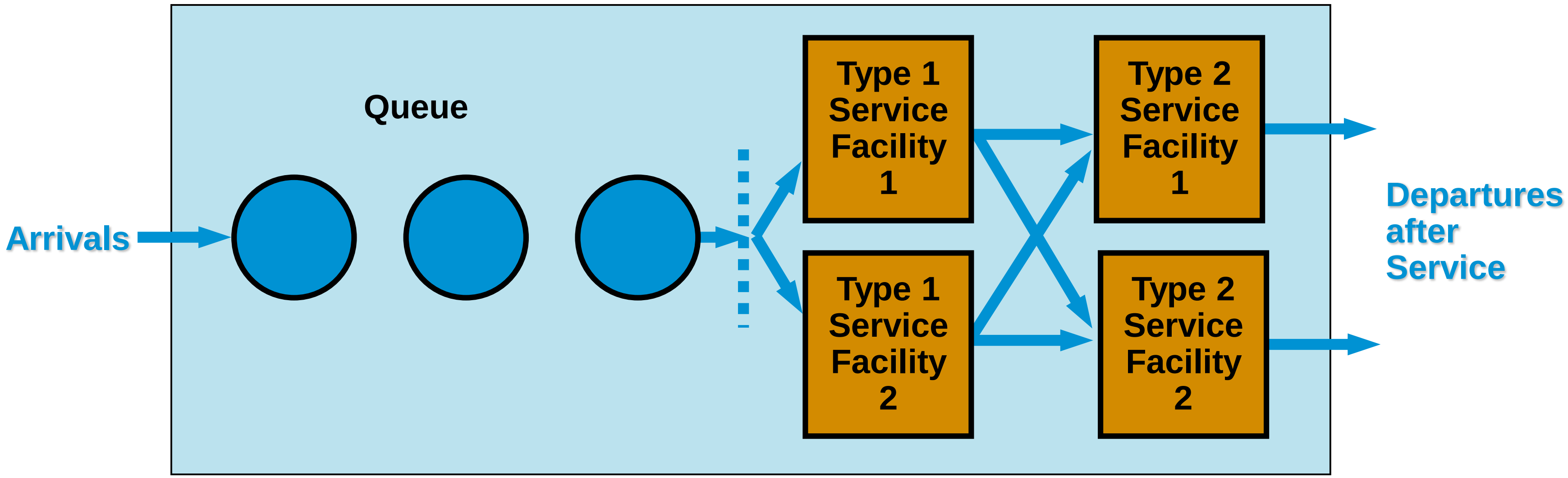
- Four basic queuing system configurations



Multichannel, Single-Phase System

Characteristics of a Queuing System

- Four basic queuing system configurations



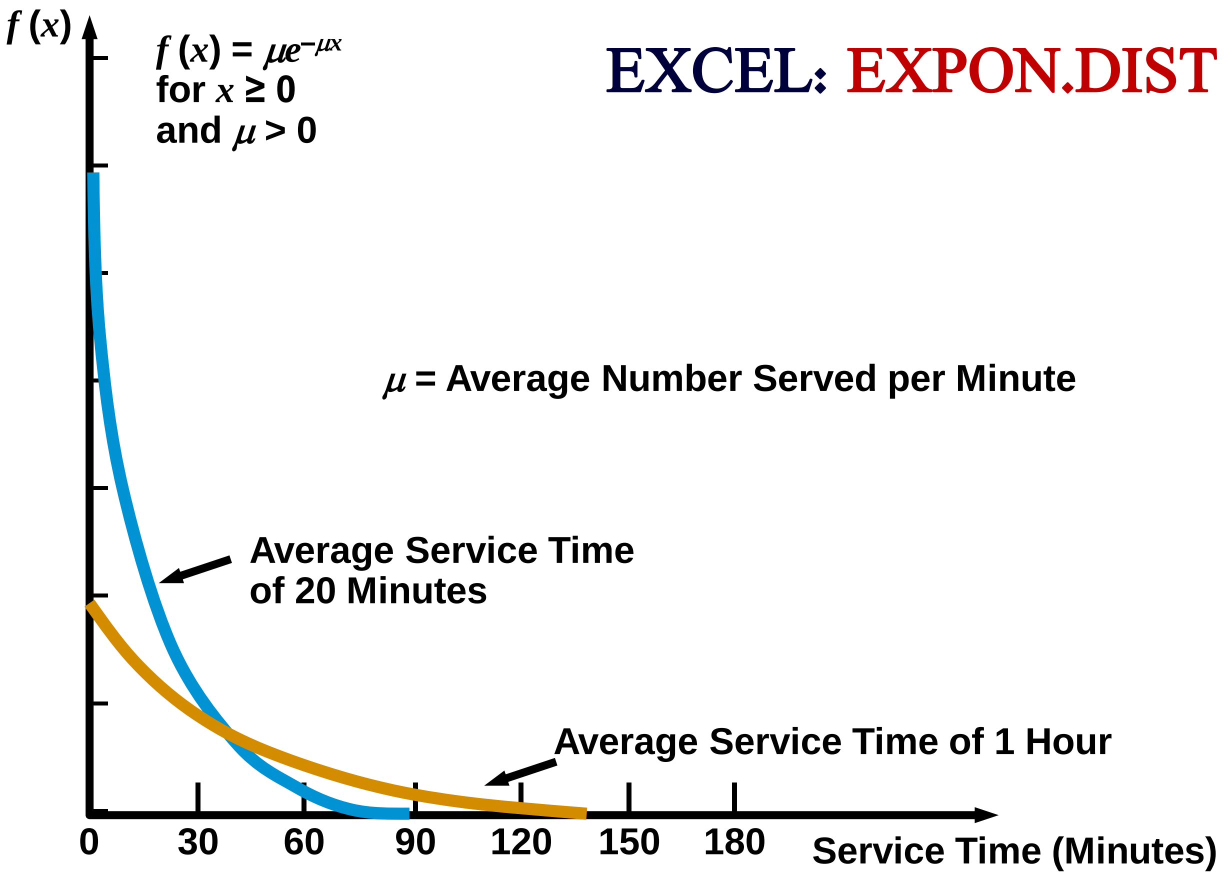
Multichannel, Multiphase System

Characteristics of a Queuing System

- Service time distribution
  - Service patterns can be either constant or random
  - Constant service times are often machine controlled
  - More often, service times are randomly distributed according to a ***negative exponential probability distribution***
  - Models are based on the assumption of particular probability distributions
  - Analysts should take to ensure observations fit the assumed distributions when applying these models

Characteristics of a Queuing System

- Two examples of exponential distribution for service times



Kendall Notation

Identifying Models Using Kendall Notation

- D. G. Kendall (Scottish Statistician) developed a notation for queuing models that specifies the pattern of arrival, the service time distribution, and the number of channels
- It is of the form

Arrival distribution	/	Service time distribution	/	Number of service channels open

- Specific letters are used to represent probability distributions

  M = Poisson distribution for number of occurrences

  D = constant (deterministic) rate

  G = general distribution with known mean and variance

Identifying Models Using Kendall Notation

- So a single channel model with Poisson arrivals and exponential service times would be represented by

$$M/M/1$$

- If a second channel is added we would have

$$M/M/2$$

- A three channel system with Poisson arrivals and constant service time would be

$$M/D/3$$

- A four channel system with Poisson arrivals and normally distributed service times would be

$$M/G/4$$

Operating Characteristics of Waiting Lines

(M/M/1) Model

Single-Channel Model, Poisson Arrivals, Exponential Service Times (M/M/1)

- Assumptions of the model
  - Arrivals are served on a FIFO basis
  - No balking or reneging
  - Arrivals are independent of each other but rate is constant over time
  - Arrivals follow a Poisson distribution
  - Service times are variable and independent but the average is known
  - Service times follow a negative exponential distribution
  - Average service rate is greater than the average arrival rate

Single-Channel Model, Poisson Arrivals, Exponential Service Times (M/M/1)

- When these assumptions are met, we can develop a series of equations that define the queue's ***operating characteristics***

- Queuing Equations
  - We let

    $\lambda =$ mean number of arrivals per time period

    $\mu =$ mean number of people or items served per time period

  - The arrival rate and the service rate must be for the same time period

Note

1. Please do not memorize (but need to understand the meaning of) any of the equations shown here.

2. However, you will need to know how to use Excel calculators that will be provided to you.

Single-Channel Model, Poisson Arrivals, Exponential Service Times (M/M/1)

1. The average number of customers or units in the system, L

$$L = \frac{\lambda}{\mu - \lambda}$$

2. The average time a customer spends in the system, W

$$W = \frac{1}{\mu - \lambda}$$

3. The average number of customers in the queue, L_q

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

Single-Channel Model, Poisson Arrivals, Exponential Service Times (M/M/1)

4. The average time a customer spends waiting in the queue, W_q

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

5. The *utilization factor* for the system, ρ, the probability the service facility is being used

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

Single-Channel Model, Poisson Arrivals, Exponential Service Times (M/M/1)

6. The percent idle time, P_0, the probability no one is in the system

$$P_0 = 1 - \frac{\lambda}{\mu}$$

7. The probability that the number of customers in the system is greater than k, $P_{n>k}$

$$P_{n>k} = \left(\frac{\lambda}{\mu}\right)^{k+1}$$

Enhancing the Queuing Environment

- Reducing waiting time is not the only way to reduce waiting cost
- Reducing waiting cost (C_w) will also reduce total waiting cost
- This might be less expensive to achieve than reducing either W or W_q

Multichannel Model, Poisson Arrivals, Exponential Service Times (M/M/m)

- Assumptions of the model
  - Arrivals are served on a FIFO basis
  - No balking or reneging
  - Arrivals are independent of each other but rate is constant over time
  - Arrivals follow a Poisson distribution
  - Service times are variable and independent but the average is known
  - Service times follow a negative exponential distribution
  - Average service rate is greater than the average arrival rate

Multichannel Model, Poisson Arrivals, Exponential Service Times (M/M/m)

- Equations for the multichannel queuing model
- We let

$$m = \text{number of channels open}$$
$$\lambda = \text{average arrival rate}$$
$$\mu = \text{average service rate at each channel}$$

1. The probability that there are zero customers in the system

This value will be provided in a table

$$P_0 = \frac{1}{\left[\sum_{n=0}^{n=m-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n\right] + \frac{1}{m!}\left(\frac{\lambda}{\mu}\right)^m \frac{m\mu}{m\mu - \lambda}} \quad \text{for } m\mu > \lambda$$

Multichannel Model, Poisson Arrivals, Exponential Service Times (M/M/m)

2. The average number of customer in the system

$$L = \frac{\lambda\mu(\lambda/\mu)^m}{(m-1)!(m\mu-\lambda)^2}P_0 + \frac{\lambda}{\mu}$$

3. The average time a unit spends in the waiting line or being served, in the system

$$W = \frac{\mu(\lambda/\mu)^m}{(m-1)!(m\mu-\lambda)^2}P_0 + \frac{1}{\mu} = \frac{L}{\lambda}$$

Multichannel Model, Poisson Arrivals, Exponential Service Times (M/M/m)

4. The average number of customers or units in line waiting for service

$$L_q = L - \frac{\lambda}{\mu}$$

5. The average number of customers or units in line waiting for service

$$W_q = W - \frac{1}{\mu} = \frac{L_q}{\lambda}$$

6. The average number of customers or units in line waiting for service

$$\rho = \frac{\lambda}{m\mu}$$

Waiting Costs

Waiting Line Costs

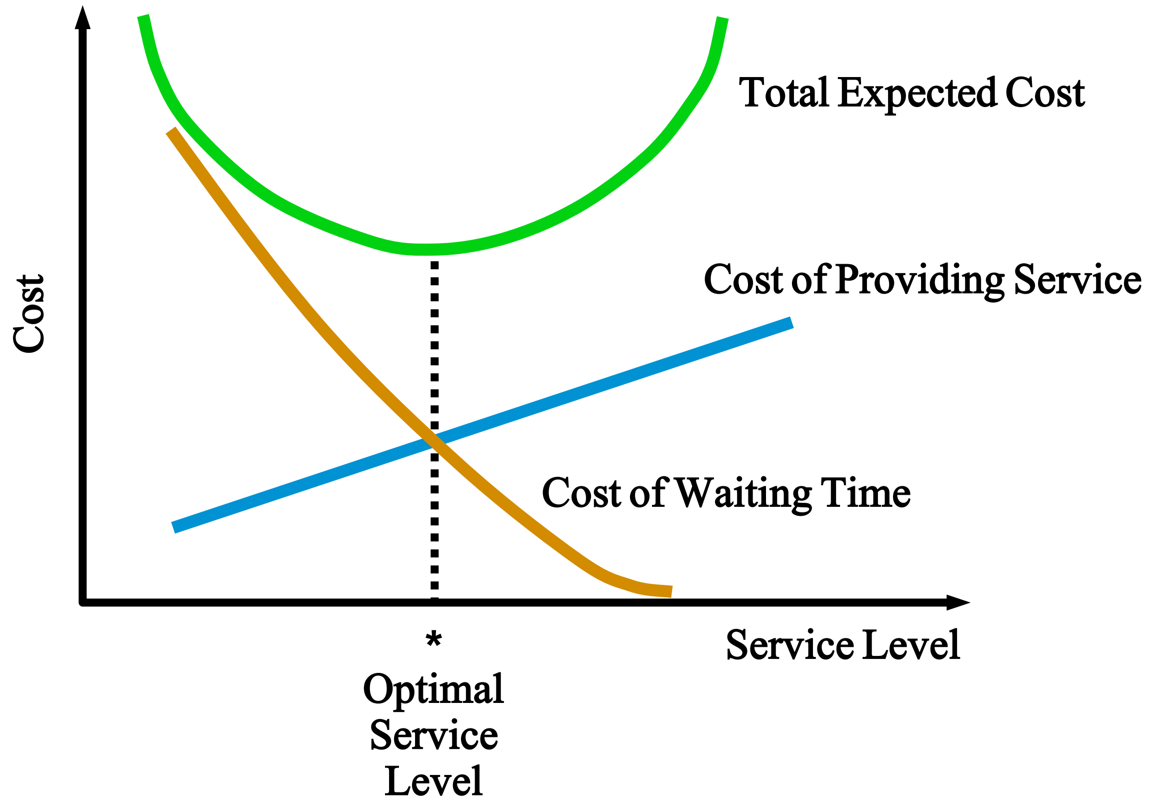
- Most waiting line problems are focused on finding the ideal level of service a firm should provide
- In most cases, this service level is something management can control
- When an organization *does* have control, they often try to find the balance between two extremes
- A *large staff* and *many* service facilities generally results in high levels of service but have high costs

Waiting Line Costs

- Having the ***minimum*** number of service facilities keeps ***service cost*** down but may result in dissatisfied customers
- There is generally a trade-off between cost of providing service and cost of waiting time
- Service facilities are evaluated on their ***total expected cost*** which is the sum of ***service costs*** and ***waiting costs***
- Organizations typically want to find the service level that minimizes the total expected cost

Waiting Costs

- Queuing costs and service level



Psychology of Waiting

Psychology of Waiting

The perception of waiting time is as important as the actual time of waiting

[Dr. Queue]

Mathematics of Waiting Lines

[Queueing Models]

Regents Park Publishers

Tutorial

Waiting Lines

End