CHAPTER

11

Managing Economies of Scale in a Supply Chain

Cycle Inventory

LEARNING OBJECTIVES

After reading this chapter, you will be able to

1. Balance the appropriate costs to choose the optimal lot size and cycle inventory in a supply chain.
2. Understand the impact of quantity discounts on lot size and cycle inventory.
3. Devise appropriate discounting schemes for a supply chain.
4. Understand the impact of trade promotions on lot size and cycle inventory.
5. Identify managerial levers that reduce lot size and cycle inventory in a supply chain without increasing cost.

Cycle inventory exists because producing or purchasing in large lots allows a stage of the supply chain to exploit economies of scale and thus lower cost. The presence of fixed costs associated with ordering and transportation, quantity discounts in product pricing, and short-term discounts or promotions encourages different stages of a supply chain to exploit economies of scale and order in large lots. In this chapter, we study how each of these factors affects the lot size and cycle inventories within a supply chain. Our goal is to identify managerial levers that reduce cycle inventory in a supply chain without raising cost.

11.1 THE ROLE OF CYCLE INVENTORY IN A SUPPLY CHAIN

A *lot* or *batch size* is the quantity that a stage of a supply chain either produces or purchases at a time. Consider, for example, a computer store that sells an average of four printers a day. The store manager, however, orders 80 printers from the manufacturer each time he places an order. The lot or batch size in this case is 80 printers. Given daily sales of four printers, it takes an average of 20 days before the store sells the entire lot and purchases a replenishment lot. The computer store holds an inventory of printers because the manager purchases a lot size larger than the store's daily sales. *Cycle inventory* is the average inventory in a supply chain due to either production or purchases in lot sizes that are larger than those demanded by the customer.

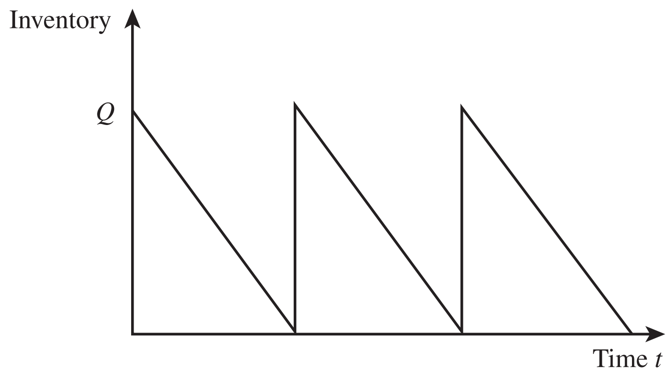


FIGURE 11-1 Inventory Profile of Jeans at Jean-Mart

In the rest of this chapter, we use the following notation:

Q: Quantity in a lot or batch size

D: Demand per unit time

Here, we ignore the impact of demand variability and assume that demand is stable. In Chapter 12, we introduce demand variability and its impact on safety inventory.

Let us consider the cycle inventory of jeans at Jean-Mart, a department store. The demand for jeans is relatively stable at $D = 100$ pairs of jeans per day. The store manager at Jean-Mart currently purchases in lots of $Q = 1{,}000$ pairs. The *inventory profile* of jeans at Jean-Mart is a plot depicting the level of inventory over time, as shown in Figure 11-1.

Because purchases are in lots of $Q = 1{,}000$ units, whereas demand is only $D = 100$ units per day, it takes 10 days for an entire lot to be sold. Over these 10 days, the inventory of jeans at Jean-Mart declines steadily from 1,000 units (when the lot arrives) to 0 (when the last pair is sold). This sequence of a lot arriving and demand depleting inventory until another lot arrives repeats itself every 10 days, as shown in the inventory profile in Figure 11-1.

When demand is steady, cycle inventory and lot size are related as follows:

$$\text{Cycle inventory} = \frac{\text{lot size}}{2} = \frac{Q}{2} \qquad \textbf{(11.1)}$$

For a lot size of 1,000 units, Jean-Mart carries a cycle inventory of $Q/2 = 500$ pairs of jeans. From Equation 11.1, we see that cycle inventory is proportional to the lot size. A supply chain in which stages produce or purchase in larger lots has more cycle inventory than a supply chain in which stages produce and purchase in smaller lots. For example, if a competing department store with the same demand purchases in lot sizes of 200 pairs of jeans, it will carry a cycle inventory of only 100 pairs of jeans.

Lot sizes and cycle inventory also influence the flow time of material within the supply chain. Recall from Little's Law (Equation 3.1) that

$$\text{Average flow time} = \frac{\text{average inventory}}{\text{average flow rate}}$$

For any supply chain, average flow rate equals demand. We thus have

$$\text{Average flow time resulting from cycle inventory} = \frac{\text{cycle inventory}}{\text{demand}} = \frac{Q}{2D}$$

For lot sizes of 1,000 pairs of jeans and daily demand of 100 pairs of jeans, we obtain

$$\text{Average flow time resulting from cycle inventory} = \frac{Q}{2D} = \frac{1{,}000}{2 \times 100} = 5 \text{ days}$$

Cycle inventory at the Jean-Mart store thus adds five days to the average amount of time that jeans spend in the supply chain. The larger the cycle inventory, the longer the lag time between when a product is produced and when it is sold. A lower level of cycle inventory is always desirable, because long time lags leave a firm vulnerable to demand changes in the marketplace. A lower cycle inventory also decreases a firm's working capital requirement. Toyota, for example, keeps a cycle inventory of only a few hours of production between the factory and most suppliers. As a result, Toyota is never left with unneeded parts, and its working capital requirements are less than those of its competitors. Toyota also allocates very little space in the factory to inventory.

Zara and Seven-Eleven Japan are two companies that have built their strategy on the ability to replenish their stores in small lots. Seven-Eleven replenishes its stores in Japan with fresh food three times a day. The small replenishment batch allows Seven-Eleven to provide product that is is always very fresh. Zara replenishes its European stores up to three times a week. Each replenishment batch thus contains only about two days of demand. This ensures that Zara's inventory on hand closely tracks customer demand. In both instances the firms have used small batch replenishment to ensure that their supply closely tracks customer demand trends.

Before we suggest actions that a manager can take to reduce cycle inventory, it is important to understand why stages of a supply chain produce or purchase in large lots and how lot size reduction affects supply chain performance.

Cycle inventory is held to take advantage of economies of scale and reduce cost within a supply chain. For example, apparel is shipped from Asia to North America in full container loads to reduce the transportation cost per unit. Similarly, an integrated steel mill produces hundreds of tons of steel per lot to spread that high cost of setup over a large batch. To understand how the supply chain achieves these economies of scale, we first identify supply chain costs that are influenced by lot size.

The *average price paid per unit purchased* is a key cost in the lot-sizing decision. A buyer may increase the lot size if this action results in a reduction in the price paid per unit purchased. For example, if the jeans manufacturer charges \$20 per pair for orders under 500 pairs of jeans and \$18 per pair for larger orders, the store manager at Jean-Mart gets the lower price by ordering in lots of at least 500 pairs of jeans. The price paid per unit is referred to as the *material cost* and is denoted by C. It is measured in dollars per unit. In many practical situations, material cost displays economies of scale—increasing lot size decreases material cost.

The *fixed ordering cost* includes all costs that do not vary with the size of the order but are incurred each time an order is placed. There may be a fixed administrative cost to place an order, a fixed trucking cost to transport the order, and a fixed labor cost to receive the order. Jean-Mart incurs a cost of \$400 for the truck regardless of the number of pairs of jeans shipped. If the truck can hold up to 2,000 pairs of jeans, a lot size of 100 pairs results in a transportation cost of \$4/pair, whereas a lot size of 1,000 pairs results in a transportation cost of \$0.40/pair. Given the fixed transportation cost per batch, the store manager can reduce transportation cost per unit by increasing the lot size. The fixed ordering cost per lot or batch is denoted by S (commonly thought of as a setup cost) and is measured in dollars per lot. The ordering cost also displays economies of scale—increasing the lot size decreases the fixed ordering cost per unit purchased.

Holding cost is the cost of carrying one unit in inventory for a specified period of time, usually one year. It is a combination of the cost of capital, the cost of physically storing the inventory, and the cost that results from the product becoming obsolete. The holding cost is denoted by H and is measured in dollars per unit per year. It may also be obtained as a fraction, h, of the unit cost of the product. Given a unit cost of C, the holding cost H is given by

$$H = hC \tag{11.2}$$

The total holding cost increases with an increase in lot size and cycle inventory.

To summarize, the following costs must be considered in any lot-sizing decision:

- Average price per unit purchased, \$$C$/unit
- Fixed ordering cost incurred per lot, \$$S$/lot
- Holding cost incurred per unit per year, \$$H$/unit/year = hC

Later in the chapter, we discuss how the various costs may be estimated in practice. However, for the purposes of this discussion, we assume they are already known.

The primary role of cycle inventory is to allow different stages in a supply chain to purchase product in lot sizes that minimize the sum of the material, ordering, and holding costs. If a manager considers the holding cost alone, he or she will reduce the lot size and cycle inventory. Economies of scale in purchasing and ordering, however, motivate a manager to increase the lot size and cycle inventory. A manager must make the trade-off that minimizes total cost when making lot-sizing decisions.

Ideally, cycle inventory decisions should be made considering the total cost across the entire supply chain. In practice, however, it is generally the case that each stage makes its cycle inventory decisions independently. As we discuss later in the chapter, this practice increases the level of cycle inventory as well as the total cost in the supply chain.

Key Point

Cycle inventory exists in a supply chain because different stages exploit economies of scale to lower total cost. The costs considered include material cost, fixed ordering cost, and holding cost.

Any stage of the supply chain exploits economies of scale in its replenishment decisions in the following three typical situations:

1. A fixed cost is incurred each time an order is placed or produced.
2. The supplier offers price discounts based on the quantity purchased per lot.
3. The supplier offers short-term price discounts or holds trade promotions.

In the following sections, we review how purchasing managers can best respond to these situations.

11.2 ESTIMATING CYCLE INVENTORY-RELATED COSTS IN PRACTICE

When setting cycle inventory levels in practice, a common hurdle is estimating the ordering and holding costs. Given the robustness of cycle inventory models, it is better to get a good approximation quickly rather than spend a lot of time trying to estimate costs exactly.

Our goal is to identify incremental costs that change with the lot-sizing decision. We can ignore costs that are unchanged with a change in lot size. For example, if a factory is running at 50 percent of capacity and all labor is full time and not earning overtime, it can be argued that the incremental setup cost for labor is zero. Reducing the lot size in this case will not have any impact on setup cost until either labor is fully utilized (and earning overtime) or machines are fully utilized (with a resulting loss in production capacity).

Inventory Holding Cost

Holding cost is estimated as a percentage of the cost of a product and is the sum of the following major components:

- ***Cost of capital:*** This is the dominant component of holding cost for products that do not become obsolete quickly. The appropriate approach is to evaluate the *weighted-average cost of capital* (WACC), which takes into account the required return on the firm's equity and the cost of its debt (see Brealey and Myers, 2000). These are weighted by the amount of equity and debt financing that the firm has. The formula for the WACC is

$$WACC = \frac{E}{D+E}(R_f + \beta \times MRP) + \frac{D}{D+E} R_b(1-t)$$

where

E = amount of equity
D = amount of debt
R_f = risk-free rate of return (which is usually in the mid-single digits)
β = the firm's beta, a measure of volatility of stock price
MRP = market risk premium (which is around the high single digits)
R_b = rate at which the firm can borrow money (related to its debt rating)
t = tax rate

Most of these numbers can be found in a company's annual report and in any equity research report on the company. The borrowing rate comes from tables listing the rates charged for bonds from firms with the same credit ratings. The risk-free rate is the return on U.S. Treasury bonds, and the market risk premium is the return of the market above the risk-free rate. If access to a company's financial structure is not available, a good approximation can be made by using numbers from public companies in the same industry and of similar size.

- ***Obsolescence (or spoilage) cost:*** The obsolescence cost estimates the rate at which the value of the stored product drops because its market value or quality falls. This cost can range dramatically, from rates of many-thousand percent to virtually zero, depending on the type of product. Perishable products have high obsolescence rates. Even nonperishables can have high obsolescence rates if they have short life cycles. A product with a life cycle of six months has an effective obsolescence cost of 200 percent. At the other end of the spectrum are products such as crude oil that take a long time to spoil or become obsolete. For such products, a low obsolescence rate may be applied.
- ***Handling cost:*** Handling cost should include only incremental receiving and storage costs that vary with the quantity of product received. Quantity-independent handling costs that vary with the number of orders should be included in the order cost. The quantity-dependent handling cost often does not change if quantity varies within a range. If the quantity is within this range (e.g., the range of inventory a crew of four people can unload per period of time), incremental handling cost added to the holding cost is zero. If the quantity handled requires more people, an incremental handling cost is added to the holding cost.
- ***Occupancy cost:*** The occupancy cost reflects the incremental change in space cost due to changing cycle inventory. If the firm is being charged based on the actual number of units held in storage, we have the direct occupancy cost. Firms often lease or purchase a fixed amount of space. As long as a marginal change in cycle inventory does not change the space requirements, the incremental occupancy cost is zero. Occupancy costs often take the form of a step function, with a sudden increase in cost when capacity is fully utilized and new space must be acquired.
- ***Miscellaneous costs:*** The final component of holding cost deals with a number of other relatively small costs. These costs include theft, security, damage, tax, and additional insurance charges that are incurred. Once again, it is important to estimate the incremental change in these costs on changing cycle inventory.

Ordering Cost

The ordering cost includes all incremental costs associated with placing or receiving an extra order that are incurred regardless of the size of the order. Components of ordering cost include the following:

- ***Buyer time:*** Buyer time is the incremental time of the buyer placing the extra order. This cost should be included only if the buyer is utilized fully. The incremental cost of getting an idle buyer to place an order is zero and does not add to the ordering cost. Electronic ordering can significantly reduce the buyer time to place an order.

- ***Transportation costs:*** A fixed transportation cost is often incurred regardless of the size of the order. For instance, if a truck is sent to deliver every order, it costs the same amount to send a half-empty truck as it does a full truck. Less-than-truckload pricing also includes a fixed component that is independent of the quantity shipped and a variable component that increases with the quantity shipped. The fixed component should be included in the ordering cost.
- ***Receiving costs:*** Some receiving costs are incurred regardless of the size of the order. These include any administration work such as purchase order matching and any effort associated with updating inventory records. Receiving costs that are quantity dependent should not be included here.
- ***Other costs:*** Each situation can have costs unique to it that should be considered if they are incurred for each order regardless of the quantity of that order.

The ordering cost is estimated as the sum of all its component costs.

11.3 ECONOMIES OF SCALE TO EXPLOIT FIXED COSTS

To better understand the trade-offs discussed in this section, consider a situation that often arises in daily life—the purchase of groceries and other household products. These may be purchased at a nearby convenience store or at a Costco (a large warehouse club selling consumer goods), which is generally located much farther away. The fixed cost of going shopping is the time it takes to go to either location. This fixed cost is much lower for the nearby convenience store. Prices, however, are higher at the local convenience store. Taking the fixed cost into account, we tend to tailor our lot size decision accordingly. When we need only a small quantity, we go to the nearby convenience store because the benefit of a low fixed cost outweighs the cost of higher prices at the convenience store. When we are buying a large quantity, however, we go to Costco, where the lower prices over the larger quantity purchased more than make up for the higher fixed cost.

In this section, we focus on the situation in which a fixed cost associated with placing, receiving, and transporting an order is incurred for each order. A purchasing manager wants to minimize the total cost of satisfying demand and must therefore make the appropriate cost trade-offs when making the lot-sizing decision. We start by considering the lot-sizing decision for a single product.

Lot Sizing for a Single Product (Economic Order Quantity)

As Best Buy sells its current inventory of HP computers, the purchasing manager places a replenishment order for a new lot of Q computers. Including the cost of transportation, Best Buy incurs a fixed cost of $\$S$ per order. The purchasing manager must decide on the number of computers to order from HP in a lot. For this decision, we assume the following inputs:

$D =$ Annual demand of the product

$S =$ Fixed cost incurred per order

$C =$ Cost per unit of product

$h =$ Holding cost per year as a fraction of product cost

Assume that HP does not offer any discounts, and each unit costs $\$C$ no matter how large an order is. The holding cost is thus given by $H = hC$ (using Equation 11.2). The model is developed using the following basic assumptions:

1. Demand is steady at D units per unit time.
2. No shortages are allowed—that is, all demand must be supplied from stock.
3. Replenishment lead time is fixed (initially assumed to be zero).

The purchasing manager makes the lot-sizing decision to minimize the total cost for the store. He or she must consider three costs when deciding on the lot size:

- Annual material cost
- Annual ordering cost
- Annual holding cost

Because purchase price is independent of lot size, we have

$$\text{Annual material cost} = CD$$

The number of orders must suffice to meet the annual demand D. Given a lot size of Q, we thus have

$$\text{Number of orders per year} = \frac{D}{Q} \quad \textbf{(11.3)}$$

Because an order cost of S is incurred for each order placed, we infer that

$$\text{Annual ordering cost} = \left(\frac{D}{Q}\right)S \quad \textbf{(11.4)}$$

Given a lot size of Q, we have an average inventory of $Q/2$. The annual holding cost is thus the cost of holding $Q/2$ units in inventory for one year and is given as

$$\text{Annual holding cost} = \left(\frac{Q}{2}\right)H = \left(\frac{Q}{2}\right)hC$$

The total annual cost, TC, is the sum of all three costs and is given as

$$\text{Total annual cost, } TC = CD + \left(\frac{D}{Q}\right)S + \left(\frac{Q}{2}\right)hC$$

Figure 11-2 shows the variation in different costs as the lot size is changed. Observe that the annual holding cost increases with an increase in lot size. In contrast, the annual ordering cost declines with an increase in lot size. The material cost is independent of lot size because we have assumed the price to be fixed. The total annual cost thus first declines and then increases with an increase in lot size.

From the perspective of the manager at Best Buy, the optimal lot size is one that minimizes the total cost to Best Buy. It is obtained by taking the first derivative of the total cost with respect to Q and setting it equal to 0 (see Appendix 11A at the end of this chapter). The optimal lot size

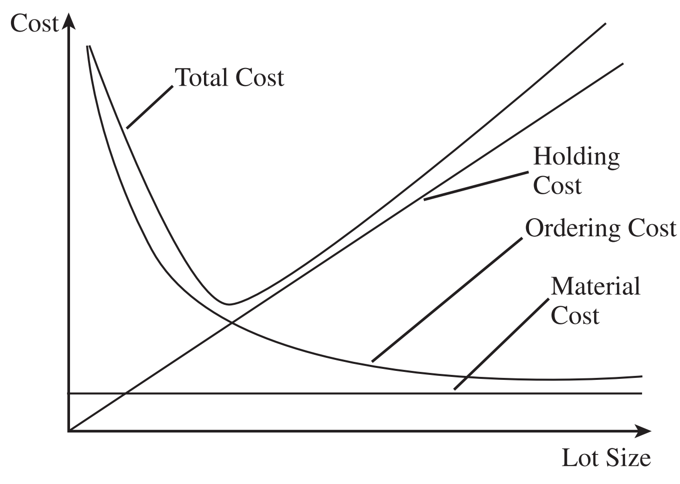


FIGURE 11-2 Effect of Lot Size on Costs at Best Buy

is referred to as the *economic order quantity* (EOQ). It is denoted by Q^* and is given by the following equation:

$$\text{Optimal lot size, } Q^* = \sqrt{\frac{2DS}{hC}} \qquad \textbf{(11.5)}$$

For this formula, it is important to use the same time units for the holding cost rate h and the demand D. With each lot or batch of size Q^*, the cycle inventory in the system is given by $Q^*/2$. The flow time spent by each unit in the system is given by $Q^*/(2D)$. As the optimal lot size increases, so does the cycle inventory and the flow time. The optimal ordering frequency is given by n^*, where

$$n^* = \frac{D}{Q^*} = \sqrt{\frac{DhC}{2S}} \qquad \textbf{(11.6)}$$

In Example 11-1 (see spreadsheet *Chapter11-examples1-6*, worksheet *Example 11-1*), we illustrate the EOQ formula and the procedure to make lot-sizing decisions.

EXAMPLE 11-1 Economic Order Quantity

Demand for the Deskpro computer at Best Buy is 1,000 units per month. Best Buy incurs a fixed order placement, transportation, and receiving cost of \$4,000 each time an order is placed. Each computer costs Best Buy \$500 and the retailer has a holding cost of 20 percent. Evaluate the number of computers that the store manager should order in each replenishment lot.

Analysis:

In this case, the store manager has the following inputs:

Annual demand, $D = 1{,}000 \times 12 = 12{,}000$ units

Order cost per lot, $S = \$4{,}000$

Unit cost per computer, $C = \$500$

Holding cost per year as a fraction of unit cost, $h = 0.2$

Using the EOQ formula (Equation 11.5), the optimal lot size is

$$\text{Optimal order size} = Q^* = \sqrt{\frac{2 \times 12{,}000 \times 4{,}000}{0.2 \times 500}} = 980$$

To minimize the total cost at Best Buy, the store manager orders a lot size of 980 computers for each replenishment order. The cycle inventory is the average resulting inventory and (using Equation 11.1) is given by

$$\text{Cycle inventory} = \frac{Q^*}{2} = \frac{980}{2} = 490$$

For a lot size of $Q^* = 980$, the store manager evaluates

$$\text{Number of orders per year} = \frac{D}{Q^*} = \frac{12{,}000}{980} = 12.24$$

$$\text{Annual ordering and holding cost} = \frac{D}{Q^*}S + \left(\frac{Q^*}{2}\right)hC = \$97{,}980$$

$$\text{Average flow time} = \frac{Q^*}{2D} = \frac{490}{12{,}000} = 0.041 \text{ year} = 0.49 \text{ month}$$

Each computer thus spends 0.49 month, on average, at Best Buy before it is sold because it was purchased in a batch of 980.

A few key insights can be gained from Example 11-1 (see worksheet *Example11-1*). Using a lot size of 1,100 (instead of 980) increases annual costs to \$98,636 (from \$97,980). Even though the order size is more than 10 percent larger than the optimal order size Q^*, total cost increases by only 0.67 percent. This issue can be relevant in practice. Best Buy may find that the economic order quantity for flash drives is 6.5 cases. The manufacturer may be reluctant to ship half a case and may want to charge extra for this service. Our discussion illustrates that Best Buy is perhaps better off with lot sizes of six or seven cases, because this change has a small impact on its inventory-related costs but can save on any fee that the manufacturer charges for shipping half a case.

Key Point

Total ordering and holding costs are relatively stable around the economic order quantity. A firm is often better served by ordering a convenient lot size close to the EOQ rather than the precise EOQ.

If demand at Best Buy increases to 4,000 computers a month (demand has increased by a factor of 4), the EOQ formula shows that the optimal lot size doubles and the number of orders placed per year also doubles. In contrast, average flow time decreases by a factor of 2. In other words, as demand increases, cycle inventory measured in terms of days (or months) of demand should reduce if the lot-sizing decision is made optimally. This observation can be stated as the following Key Point:

Key Point

If demand increases by a factor of k, the optimal lot size increases by a factor of $\sqrt{k}$. The number of orders placed per year should also increase by a factor of $\sqrt{k}$. Flow time attributed to cycle inventory should decrease by a factor of $\sqrt{k}$.

Let us return to the situation in which monthly demand for the Deskpro model is 1,000 computers. Now assume that the manager would like to reduce the lot size to $Q = 200$ units to reduce flow time. If this lot size is decreased without any other change, we have

$$\text{Annual inventory-related costs} = \left(\frac{D}{Q}\right)S + \left(\frac{Q}{2}\right)hC = \$250{,}000$$

This is significantly higher than the total cost of \$97,980 that Best Buy incurred when ordering in lots of 980 units, as in Example 11-1. Thus, there are clear financial reasons that the store manager would be unwilling to reduce the lot size to 200. To make it feasible to reduce the lot size, the manager should work to reduce the fixed order cost S. If the fixed cost associated with each lot is reduced to \$1,000 (from the current value of \$4,000), the optimal lot size reduces to 490 (from the current value of 980). We illustrate the relationship between desired lot size and order cost in Example 11-2 (see worksheet *Example 11-2*).

EXAMPLE 11-2 Relationship Between Desired Lot Size and Ordering Cost

The store manager at Best Buy would like to reduce the optimal lot size from 980 to 200. For this lot size reduction to be optimal, the store manager wants to evaluate how much the ordering cost per lot should be reduced.

Analysis:

In this case, we have

Desired lot size, $Q^* = 200$

Annual demand, $D = 1{,}000 \times 12 = 12{,}000$ units

Unit cost per computer, $C = \$500$

Holding cost per year as a fraction of inventory value, $h = 0.2$

Using the EOQ formula (Equation 11.5), the desired order cost is

$$S = \frac{hC(Q^*)^2}{2D} = \frac{0.2 \times 500 \times 200^2}{2 \times 12{,}000} = \$166.7$$

Thus, the store manager at Best Buy would have to reduce the order cost per lot from \$4,000 to \$166.7 for a lot size of 200 to be optimal.

The observation in Example 11-2 may be stated as in the following Key Point:

Key Point

To reduce the optimal lot size by a factor of k, the fixed order cost S must be reduced by a factor of k^2.

Production Lot Sizing

In the EOQ formula, we have implicitly assumed that the entire lot arrives at the same time. While this may be a reasonable assumption for a retailer receiving a replenishment lot, it is not reasonable in a production environment in which production occurs at a specified rate, say, P. In a production environment, inventory thus builds up at a rate of $P - D$ when production is on, and inventory is depleted at a rate of D when production is off.

With D, h, C, and S as defined earlier, the EOQ formula can be modified to obtain the economic production quantity (EPQ) as follows:

$$Q^P = \sqrt{\frac{2DS}{(1 - D/P)hC}}$$

The annual setup cost in this case is given by

$$\left(\frac{D}{Q^P}\right)S$$

The annual holding cost is given by

$$(1 - D/P)\left(\frac{Q^P}{2}\right)hC$$

Observe that the economic production quantity is the EOQ multiplied by a correction factor that approaches 1 as the production rate becomes much faster than the demand.

For the remainder of this chapter, we restrict our attention to the case in which the entire replenishment lot arrives at the same time, a scenario that applies in most supply chain settings.

Lot Sizing with Capacity Constraint

In our discussion so far we have assumed that the economic order quantity for a retailer will fit on the truck. In reality the truck has a limited capacity, say K. If the economic order quantity Q is more than the K, the retailer will have to pay for more than one truck. In this case, the optimal order quantity is obtained by comparing the cost of ordering K units (a full truck) and Q units ($\lceil Q/K \rceil$ trucks). If the setup cost S arises primarily from the cost of a truck, it is never optimal to order more than one truck. In this case, the optimal order size is the minimum of the EOQ and the truck capacity (K).

11.4 AGGREGATING MULTIPLE PRODUCTS IN A SINGLE ORDER

As we have discussed earlier, a key to reducing lot size is the reduction of the fixed cost incurred per lot. One major source of fixed costs is transportation. In several companies, the array of products sold is divided into families or groups, with each group managed independently by a separate product manager. This results in separate orders and deliveries for each product family, thus increasing the overall cycle inventory. Aggregating orders and deliveries across product families is an effective mechanism to lower cycle inventories. We illustrate the idea of aggregating shipments using the following example.

Consider the data from Example 11-1. Assume that Best Buy purchases four computer models, and the demand for each of the four models is 1,000 units per month. In this case, if each product manager orders separately, he or she would order a lot size of 980 units (as in Example 11-1). Across the four models, the total cycle inventory would thus be $4 \times 980/2 = 1{,}960$ units.

Now consider the case in which a store manager at Best Buy realizes that all four model shipments originate from the same source. She asks the product managers to coordinate their purchasing to ensure that all four products arrive on the same truck. In this case, the optimal combined lot size across all four models turns out to be 1,960 units (use $S = \$4{,}000$, $D = 4 \times 12{,}000 = 48{,}000$, $hC = \$500 \times 0.2 = \$10$ in Equation 11.5). This is equivalent to a lot size of 490 units for each model. As a result of aggregating orders and spreading the fixed transportation cost across multiple products originating from the same supplier, it becomes financially optimal for the store manager at Best Buy to reduce the lot size for each individual product. This action significantly reduces the cycle inventory, as well as the cost to Best Buy.

Another way to achieve this result is to have a single delivery coming from multiple suppliers (allowing fixed transportation cost to be spread across multiple suppliers) or to have a single truck delivering to multiple retailers (allowing fixed transportation cost to be spread across multiple retailers). Firms that import product to the United States from Asia have worked hard to aggregate their shipments across suppliers (often by building hubs in Asia that all suppliers deliver to), allowing them to maintain transportation economies of scale while getting smaller and more frequent deliveries from each supplier. The benefits of aggregation can be stated as in the following Key Point:

Key Point

Aggregating replenishment across products, retailers, or suppliers in a single order allows for a reduction in lot size for individual products because fixed ordering and transportation costs are now spread across multiple products, retailers, or suppliers.

Walmart and other retailers, such as Seven-Eleven Japan, have facilitated aggregation across multiple supply and delivery points without storing intermediate inventories through the use of cross-docking. Each supplier sends full truckloads to the DC, containing an aggregate delivery destined for multiple retail stores. At the DC, each inbound truck is unloaded, product is cross-docked, and outbound trucks are loaded. Each outbound truck now contains product aggregated from several suppliers destined for one retail store.

When considering fixed costs, one cannot ignore the receiving or loading costs. As more products are included in a single order, the product variety on a truck increases. The receiving warehouse now has to update inventory records for more items per truck. In addition, the task of putting inventory into storage now becomes more expensive because each distinct item must be stocked in a separate location. Thus, when attempting to reduce lot sizes, it is important to focus on reducing costs that increase with variety. Advance shipping notices (ASNs) are files that contain precise records of the contents of the truck that are sent electronically by the supplier to the customer. These electronic notices facilitate updating of inventory records as well as the decision regarding storage locations, helping reduce the fixed cost of receiving. RFID technology is also likely to help reduce the fixed costs associated with receiving that are related to product variety. The reduced fixed cost of receiving makes it optimal to reduce the lot size ordered for each product, thus reducing cycle inventory.

We next analyze how optimal lot sizes may be determined when there are fixed costs associated with each lot as well as the variety in the lot.

Lot Sizing with Multiple Products or Customers

In general, the ordering, transportation, and receiving costs of an order grow with the variety of products or pickup points. For example, it is cheaper for Walmart to receive a truck containing a single product than it is to receive a truck containing many different products, because the inventory update and restocking effort is less for a single product. A portion of the fixed cost of an order can be related to transportation (this depends only on the load and is independent of product variety on the truck). A portion of the fixed cost is related to loading and receiving (this cost increases with variety on the truck). We now discuss how optimal lot sizes may be determined in such a setting.

Our objective is to arrive at lot sizes and an ordering policy that minimize the total cost. We assume the following inputs:

D_i: Annual demand for product i

S: Order cost incurred each time an order is placed, independent of the variety of products included in the order

s_i: Additional order cost incurred if product i is included in the order

Let us consider the case in which Best Buy purchases multiple models of a product. The store manager may consider three approaches to the lot-sizing decision:

1. Each product manager orders his or her model independently.
2. The product managers jointly order every product in each lot.
3. Product managers order jointly but not every order contains every product; that is, each order contains a selected subset of the products.

The first approach does not use any aggregation and results in high cost. The second approach aggregates all products in each order. The weakness of the second approach is that low-demand products are aggregated with high-demand products in every order. Complete aggregation results in high costs if the product-specific order cost for the low-demand products is large. In such a situation, it may be better to order the low-demand products less frequently than the high-demand products. This practice results in a reduction of the product-specific order cost associated with the low-demand product. As a result, the third approach is likely to yield the lowest cost. However, it is more complex to coordinate.

We consider the example of Best Buy purchasing computers and illustrate the effect of each of the three approaches on supply chain costs.

LOTS ARE ORDERED AND DELIVERED INDEPENDENTLY FOR EACH PRODUCT In this approach, each product is ordered independently of the others. This scenario is equivalent to applying the EOQ formula to each product when evaluating lot sizes, as illustrated in Example 11-3 (see worksheet *Example 11-3* in spreadsheet *Chapter11-examples1-6*).

EXAMPLE 11-3 Multiple Products with Lots Ordered and Delivered Independently

Best Buy sells three models of computers, the Litepro, the Medpro, and the Heavypro. Annual demands for the three products are $D_L = 12{,}000$ for the Litepro, $D_M = 1{,}200$ units for the Medpro, and $D_H = 120$ units for the Heavypro. Each model costs Best Buy $500. A fixed transportation cost of $4,000 is incurred each time an order is delivered. For each model ordered and delivered on the same truck, an additional fixed cost of $1,000 per model is incurred for receiving and storage. Best Buy incurs a holding cost of 20 percent. Evaluate the lot sizes that the Best Buy manager should order if lots for each product are ordered and delivered independently. Also evaluate the annual cost of such a policy.

Analysis:

In this example, we have the following information:

Demand, $D_L = 12{,}000$/year, $D_M = 1{,}200$/year, $D_H = 120$/year

Common order cost, $S = \$4{,}000$

Product-specific order cost, $s_L = \$1{,}000$, $s_M = \$1{,}000$, $s_H = \$1{,}000$

Holding cost, $h = 0.2$

Unit cost, $C_L = \$500$, $C_M = \$500$, $C_H = \$500$

Because each model is ordered and delivered independently, a separate truck delivers each model. Thus, a fixed ordering cost of $5,000 ($4,000 + $1,000) is incurred for each product delivery. The optimal ordering policies and resulting costs for the three products (when the three products are ordered independently) are evaluated using the EOQ formula (Equation 11.5) and are shown in Table 11-1.

The Litepro model is ordered 11 times a year, the Medpro model is ordered 3.5 times a year, and the Heavypro model is ordered 1.1 times each year. The annual ordering and holding cost Best Buy incurs if the three models are ordered independently turns out to be $155,140.

TABLE 11-1 Lot Sizes and Costs for Independent Ordering

	Litepro	Medpro	Heavypro
Demand per year	12,000	1,200	120
Fixed cost/order	$5,000	$5,000	$5,000
Optimal order size	1,095	346	110
Cycle inventory	548	173	55
Annual holding cost	$54,772	$17,321	$5,477
Order frequency	11.0/year	3.5/year	1.1/year
Annual ordering cost	$54,772	$17,321	$5,477
Average flow time	2.4 weeks	7.5 weeks	23.7 weeks
Annual cost	$109,544	$34,642	$10,954

Note: Although these figures are correct, some may differ from calculations due to rounding.

Independent ordering is simple to execute but ignores the opportunity to aggregate orders. Thus, the product managers at Best Buy could potentially lower costs by combining orders on a single truck. We next consider the scenario in which all three products are ordered and delivered on the same truck each time an order is placed.

LOTS ARE ORDERED AND DELIVERED JOINTLY FOR ALL THREE MODELS Given that all three models are included each time an order is placed, the combined fixed order cost per order is given by

$$S^* = S + s_L + s_M + s_H$$

The next step is to identify the optimal ordering frequency. Let n be the number of orders placed per year. We then have

$$\text{Annual order cost} = S^* n$$

$$\text{Annual holding cost} = \frac{D_L h C_L}{2n} + \frac{D_M h C_M}{2n} + \frac{D_H h C_H}{2n}$$

The total annual cost is thus given by

$$\text{Total annual cost} = \frac{D_L h C_L}{2n} + \frac{D_M h C_M}{2n} + \frac{D_H h C_H}{2n} + S^* n$$

The optimal order frequency minimizes the total annual cost and is obtained by taking the first derivative of the total cost with respect to n and setting it equal to 0. This results in the optimal order frequency n^*, where

$$n^* = \sqrt{\frac{D_L h C_L + D_M h C_M + D_H h C_H}{2S^*}} \qquad \textbf{(11.7)}$$

Equation 11.7 can be generalized to the case in which there are k items consolidated on a single order, as follows:

$$n^* = \sqrt{\frac{\sum_{i=1}^{k} D_i h C_i}{2S^*}} \qquad \textbf{(11.8)}$$

Truck capacity can also be included in this setting by comparing the total load for the optimal n^* with truck capacity. If the optimal load exceeds truck capacity, n^* is increased until the load equals truck capacity. By applying Equation 11.8 for different values of k, we can also find the optimal number of items or suppliers to be aggregated in a single delivery.

In Example 11-4, we consider the case in which the product managers at Best Buy jointly order all three models each time they place an order (see worksheet *Example 11-4*).

EXAMPLE 11-4 Products Ordered and Delivered Jointly

Consider the Best Buy data in Example 11-3. The three product managers have decided to aggregate and order all three models each time they place an order. Evaluate the optimal lot size for each model.

Analysis:

Because all three models are included in each order, the combined order cost is

$$S^* = S + s_L + s_M + s_H = \$7{,}000 \text{ per order}$$

The optimal order frequency is obtained using Equation 11.7 and is given by

$$n^* = \sqrt{\frac{(12{,}000 \times 100) + (1{,}200 \times 100) + (120 \times 100)}{2 \times 7{,}000}} = 9.75$$

Thus, if each model is to be included in every order and delivery, the product managers at Best Buy should place 9.75 orders each year. In this case, the ordering policies and costs are as shown in Table 11-2.

TABLE 11-2 Lot Sizes and Costs for Joint Ordering at Best Buy

	Litepro	Medpro	Heavypro
Demand per year (*D*)	12,000	1,200	120
Order frequency (n^*)	9.75/year	9.75/year	9.75/year
Optimal order size (D/n^*)	1,230	123	12.3
Cycle inventory	615	61.5	6.15
Annual holding cost	$61,512	$6,151	$615
Average flow time	2.67 weeks	2.67 weeks	2.67 weeks

Because 9.75 orders are placed each year and each order costs a total of $7,000, we have

$$\text{Annual order cost} = 9.75 \times 7{,}000 = \$68{,}250$$

The annual ordering and holding cost, across the three sizes, of the aforementioned policy is given by

$$\text{Annual ordering and holding cost} = \$61{,}512 + \$6{,}151 + \$615 + \$68{,}250 = \$136{,}528$$

Observe that the product managers at Best Buy lower the annual cost from $155,140 to $136,528 by ordering all products jointly. This represents a decrease of about 12 percent.

In Example 11-5, we consider optimal aggregation of orders or deliveries in the presence of capacity constraints (see worksheet *Example 11-5*).

EXAMPLE 11-5 Aggregation with Capacity Constraint

W.W. Grainger sources from hundreds of suppliers and is considering the aggregation of inbound shipments to lower costs. Truckload shipping costs $500 per truck along with $100 per pickup. Average annual demand from each supplier is 10,000 units. Each unit costs $50 and Grainger incurs a holding cost of 20 percent. What is the optimal order frequency and order size if Grainger decides to aggregate four suppliers per truck? What is the optimal order size and frequency if each truck has a capacity of 2,500 units?

Analysis:

In this case, W.W. Grainger has the following inputs:

Demand per product, $D_i = 10{,}000$

Holding cost, $h = 0.2$

Unit cost per product, $C_i = \$50$

Common order cost, $S = \$500$

Supplier-specific order cost, $s_i = \$100$

The combined order cost from four suppliers is given by

$$S^* = S + s_1 + s_2 + s_3 + s_4 = \$900 \text{ per order}$$

From Equation 11.8, the optimal order frequency is

$$n^* = \sqrt{\frac{\sum_{i=1}^{4} D_i h C_i}{2S^*}} = \sqrt{\frac{4 \times 10{,}000 \times 0.2 \times 50}{2 \times 900}} = 14.91$$

It is thus optimal for Grainger to order 14.91 times per year. The annual ordering cost per supplier is

$$\text{Annual order cost} = 14.91 \times \frac{900}{4} = \$3{,}355$$

The quantity ordered from each supplier is $Q = 10{,}000/14.91 = 671$ units per order. The annual holding cost per supplier is

$$\text{Annual holding cost per supplier} = \frac{hC_iQ}{2} = 0.2 \times 50 \times \frac{671}{2} = \$3{,}355$$

This policy, however, requires a total capacity per truck of $4 \times 671 = 2{,}684$ units. Given a truck capacity of 2,500 units, the order frequency must be increased to ensure that the order quantity from each supplier is 2,500/4 = 625. Thus, W.W. Grainger should increase the order frequency to 10,000/625 = 16. The limited truck capacity results in an optimal order frequency of 16 orders per year instead of 14.91 orders per year when truck capacity was ignored. The limited truck capacity will increase the annual order cost per supplier to \$3,600 and decrease the annual holding cost per supplier to \$3,125.

The main advantage of ordering all products jointly is that the system is easy to administer and implement. The disadvantage is that it is not selective enough in combining the particular models that should be ordered together. If product-specific order costs are high and products vary significantly in terms of their sales, it is possible to lower costs by being selective about the products being aggregated in a joint order.

Next, we consider a policy under which the product managers do not necessarily order all models each time an order is placed, but still coordinate their orders.

LOTS ARE ORDERED AND DELIVERED JOINTLY FOR A SELECTED SUBSET OF THE PRODUCTS We first illustrate how being selective in aggregating orders into a single order can lower costs. Consider Example 11-4, in which the manager decides to aggregate all three computer models in every order. The optimal policy from Example 11-4 is to order 9.75 times a year. The disadvantage of this policy is that the Heavypro, with annual demand of only 120 units, is also ordered 9.75 times. Given that a model-specific cost of \$1,000 is incurred with each order, we are essentially adding 1,000/(120/9.75) = \$81.25 in order cost to each Heavypro. If we were to include the Heavypro in every fourth order (instead of every order), though, we would save $9{,}750 \times (3/4) = \$7{,}312.50$ in product-specific ordering cost (save three out of four product-specific orders) while incurring an additional $500 \times 0.2 \times [(120/9.75)/2] \times 3 = \$1{,}846.15$ in holding cost (because the lot size of Heavypro would increase from 120/9.75 to $[120 \times (4/9.75)]$. Such a policy would thus decrease the annual cost relative to complete aggregation by more than \$5,466. This example points to the value of being more selective when aggregating orders.

We now discuss a procedure that is more selective in combining products to be ordered jointly. The procedure we discuss here does not necessarily provide the optimal solution. It does, however, yield an ordering policy whose cost is close to optimal. The approach of the procedure is to first identify the "most frequently" ordered product that is included in every order. The base fixed cost S is then entirely allocated to this product. For each of the "less frequently" ordered products i, the ordering frequency is determined using only the product-specific ordering cost s_i. The frequencies are then adjusted so that each product i is included every m_i orders, where m_i is an integer. We now detail the procedure used.

We first describe the procedure in general and then apply it to the specific example. Assume that the products are indexed by i, where i varies from 1 to l (assuming a total of l products). Each product i has an annual demand D_i, a unit cost C_i, and a product-specific order cost s_i. The common order cost is S.

Step 1: As a first step, identify the most frequently ordered product, assuming each product is ordered independently. In this case, a fixed cost of $S + s_i$ is allocated to each product. For each product i (using Equation 11.6), evaluate the ordering frequency:

$$\bar{n}_i = \sqrt{\frac{hC_iD_i}{2(S + s_i)}}$$

This is the frequency at which product i would be ordered if it were the only product being ordered (in which case a fixed cost of $S + s_i$ would be incurred per order). Let $\bar{n}$ be the frequency of the most frequently ordered product, i^*; that is, $\bar{n}_{i^*}$ is the maximum among all $\bar{n}_i$ $(\bar{n} = \bar{n}_{i^*} = \max \{\bar{n}_i, i = 1, \ldots, l\})$. The most frequently ordered product is i^*, which is included each time an order is placed.

Step 2: For all products $i \neq i^*$, evaluate the ordering frequency:

$$\bar{\bar{n}}_i = \sqrt{\frac{hC_iD_i}{2s_i}}$$

$\bar{\bar{n}}_i$ represents the desired order frequency if product i incurs the product-specific fixed cost s_i only each time it is ordered.

Step 3: Our goal is to include each product $i \neq i^*$ with the most frequently ordered product i^* after an integer number of orders. For all $i \neq i^*$, evaluate the frequency of product i relative to the most frequently ordered product i^* to be m_i, where

$$m_i = \lceil \bar{n}/\bar{\bar{n}}_i \rceil$$

In this case, $\lceil \ \rceil$ is the operation that rounds a fraction up to the closest integer. Product i is included with the most frequently ordered product i^* every m_i orders. Given that the most frequently ordered product i^* is included in every order, $m_{i^*} = 1$.

Step 4: Having decided the ordering frequency of each product i, recalculate the ordering frequency of the most frequently ordered product i^* to be n, where

$$n = \sqrt{\frac{\sum_{i=1}^{l} hC_im_iD_i}{2(S + \sum_{i=1}^{l} s_i/m_i)}} \qquad \textbf{(11.9)}$$

Note that n is a better ordering frequency for the most frequently ordered product i^* than $\bar{n}$ because it takes into account the fact that each of the other products i is included with i^* every m_i orders.

Step 5: For each product, evaluate an order frequency of $n_i = n/m_i$ and the total cost of such an ordering policy. The total annual cost is given by

$$TC = nS + \sum_{i=1}^{l} n_is_i + \sum_{i=1}^{l} \left(\frac{D_i}{2n_i}\right)hC_i$$

This procedure results in *tailored aggregation*, with higher-demand products ordered more frequently and lower-demand products ordered less frequently. Example 11-6 (see worksheet *Example 11-6*) considers tailored aggregation for the Best Buy ordering decision in Example 11-3.

EXAMPLE 11-6 Lot Sizes Ordered and Delivered Jointly for a Selected Subset That Varies by Order

Consider the Best Buy data in Example 11-3. Product managers have decided to order jointly, but to be selective about which models they include in each order. Evaluate the ordering policy and costs using the procedure discussed previously.

TABLE 11-3 Lot Sizes and Costs for Ordering Policy Using Heuristic

	Litepro	Medpro	Heavypro
Demand per year (*D*)	12,000	1,200	120
Order frequency (*n*)	11.47/year	5.74/year	2.29/year
Order size (*D*/*n*)	1,046	209	52
Cycle inventory	523	104.5	26
Annual holding cost	$52,307	$10,461	$2,615
Average flow time	2.27 weeks	4.53 weeks	11.35 weeks

Analysis:

Recall that $S = \$4{,}000$, $s_L = \$1{,}000$, $s_M = \$1{,}000$, $s_H = \$1{,}000$. Applying Step 1, we obtain

$$\overline{n_L} = \sqrt{\frac{hC_LD_L}{2(S + s_L)}} = 11.0,\ \overline{n_M} = \sqrt{\frac{hC_MD_M}{2(S + s_M)}} = 3.5,\ \overline{n_H} = \sqrt{\frac{hC_HD_H}{2(S + s_H)}} = 1.1$$

Clearly, Litepro is the most frequently ordered model. Thus, we set $\overline{n} = 11.0$.

We now apply Step 2 to evaluate the frequency with which Medpro and Heavypro are included with Litepro in the order. We first obtain

$$\overline{\overline{n_M}} = \sqrt{\frac{hC_MD_M}{2s_M}} = 7.7 \quad \text{and} \quad \overline{\overline{n_H}} = \sqrt{\frac{hC_HD_H}{2s_H}} = 2.4$$

Next, we apply Step 3 to evaluate

$$m_M = \left\lceil \frac{\overline{n}}{\overline{\overline{n_M}}} \right\rceil = \left\lceil \frac{11.0}{7.7} \right\rceil = 2 \quad \text{and} \quad m_H = \left\lceil \frac{\overline{n}}{\overline{\overline{n_H}}} \right\rceil = \left\lceil \frac{11.0}{2.4} \right\rceil = 5$$

Thus, Medpro is included in every second order and Heavypro is included in every fifth order (Litepro, the most frequently ordered model, is included in every order). Now that we have decided on the ordering frequency of each model, apply Step 4 (Equation 11.9) to recalculate the ordering frequency of the most frequently ordered model as

$$n = \sqrt{\frac{hC_Lm_LD_L + hC_Mm_MD_M + hC_Hm_MD_H}{2(S + s_L/m_L + s_M/m_M + s_H/m_H)}} = 11.47$$

Thus, the Litepro is ordered 11.47 times per year. Next, we apply Step 5 to obtain an ordering frequency for each product:

$$n_L = 11.47/\text{year},\ n_M = 11.47/2 = 5.74/\text{year, and } n_H = 11.47/5 = 2.29/\text{year}$$

The ordering policies and resulting costs for the three products are shown in Table 11-3.

The annual holding cost of this policy is $65,383.50. The annual order cost is given by

$$nS + n_Ls_L + n_Ms_M + n_Hs_H = \$65{,}383.50$$

The total annual cost is thus equal to $130,767. Tailored aggregation results in a cost reduction of $5,761 (about 4 percent) compared with the joint ordering of all models. The cost reduction results because each model-specific fixed cost of $1,000 is not incurred with every order.

From the Best Buy examples, it follows that aggregation can provide significant cost savings and reduction in cycle inventory in the supply chain. When product-specific order costs (s_i) are small relative to the fixed cost S, complete aggregation, whereby every product is included in every order, is very effective. Tailored aggregation provides little additional value in this setting and may not be worth the additional complexity. If Examples 11-3, 11-4, and 11-6 are repeated with s_i = \$300 (change cells D5:D7 in worksheet *Example 11-3* to 300), we find that tailored aggregation decreases costs by only about 1 percent relative to complete aggregation, whereas complete aggregation decreases costs by more than 25 percent relative to no aggregation. As product-specific order costs increase, however, tailored aggregation becomes more effective. If Examples 11-3, 11-4, and 11-6 are repeated with s_i = \$3,000 (change cells D5:D7 in worksheet *Example 11-3* to 3,000), we find that complete aggregation actually increases costs relative to no aggregation. Tailored aggregation, however, decreases costs by about 10 percent relative to no aggregation. In general, complete aggregation should be used when product-specific order costs are small, and tailored aggregation should be used when product-specific order costs are large.

Next, we consider lot sizes when material cost displays economies of scale.

Key Point

A key to reducing cycle inventory is the reduction of lot size. A key to reducing lot size without increasing costs is reducing the fixed cost associated with each lot. This may be achieved by reducing the fixed cost itself or by aggregating lots across multiple products, customers, or suppliers. When aggregating across multiple products, customers, or suppliers, simple aggregation is effective when product-specific order costs are small, and tailored aggregation is best if product-specific order costs are large.

11.5 ECONOMIES OF SCALE TO EXPLOIT QUANTITY DISCOUNTS

We now consider pricing schedules that encourage buyers to purchase in large lots. There are many instances in business-to-business transactions in which the pricing schedule displays economies of scale, with prices decreasing as lot size increases. A discount is *lot-size based* if the pricing schedule offers discounts based on the quantity ordered in a single lot. A discount is *volume based* if the discount is based on the total quantity purchased over a given period, regardless of the number of lots purchased over that period. In this section, we will see that lot-size–based quantity discounts tend to increase the lot size and cycle inventory in a supply chain. Two commonly used lot-size–based discount schemes are:

- All unit quantity discounts
- Marginal unit quantity discount or multi-block tariffs

To investigate the impact of such quantity discounts on the supply chain, we must answer the following two basic questions:

1. Given a pricing schedule with quantity discounts, what is the optimal purchasing decision for a buyer seeking to maximize profits? How does this decision affect the supply chain in terms of lot sizes, cycle inventories, and flow times?
2. Under what conditions should a supplier offer quantity discounts? What are appropriate pricing schedules that a supplier seeking to maximize profits should offer?

We start by studying the optimal response of a retailer (the buyer) when faced with either of the two lot-size–based discount schemes offered by a manufacturer (the supplier). The retailer's objective is to select lot sizes to minimize the total annual material, order, and holding costs. Next, we evaluate the optimal lot size in the case of all unit quantity discounts.

All Unit Quantity Discounts

In all unit quantity discounts, the pricing schedule contains specified break points $q_0, q_1, \ldots, q_r$, where $q_0 = 0$. If an order placed is at least as large as q_i but smaller than q_{i+1}, each unit is obtained at a cost of C_i. In general, the unit cost decreases as the quantity ordered increases; that

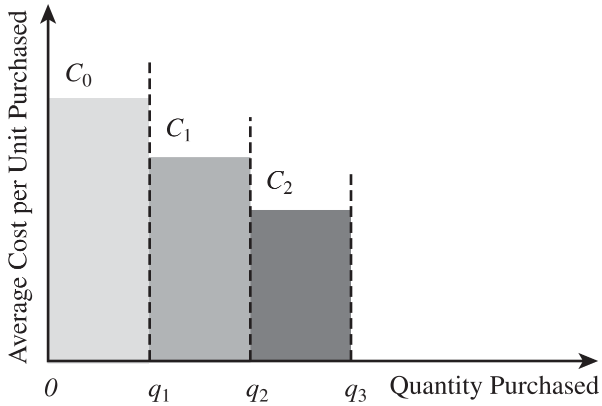


FIGURE 11-3 Average Unit Cost with All Unit Quantity Discounts

is, $C_0 \geq C_1 \geq \cdots \geq C_r$. For all unit discounts, the average unit cost varies with the quantity ordered, as shown in Figure 11-3. The retailer's objective is to decide on lot sizes to maximize profits or, equivalently, to minimize the sum of material, order, and holding costs. The solution procedure evaluates the optimal lot size for each price and picks the lot size that minimizes the overall cost.

Step 1: Evaluate the economic order quantity for each price $C_i, 0 \leq i \leq r$ as follows:

$$Q_i = \sqrt{\frac{2DS}{hC_i}} \tag{11.10}$$

Step 2: We next select the order quantity Q_i^* for each price C_i. There are three possible cases for Q_i:

1. $q_i \leq Q_i < q_{i+1}$
2. $Q_i < q_i$
3. $Q_i \geq q_{i+1}$

Case 3 can be ignored for Q_i because it is considered for Q_{i+1}. Thus, we need to consider only the first two cases. If $q_i \leq Q_i < q_{i+1}$, then set $Q_i^* = Q_i$. If $Q_i < q_i$, then a lot size of Q_i does not result in a discount. In this case, set $Q_i^* = q_i$ to qualify for the discounted price of C_i per unit.

Step 3: For each i, calculate the total annual cost of ordering Q_i^* units (this includes order cost, holding cost, and material cost) as follows:

$$\text{Total annual cost, } TC_i = \left(\frac{D}{Q_i^*}\right)S + \frac{Q_i^*}{2}hC_i + DC_i \tag{11.11}$$

Step 4: Over all i select order quantity Q_i^* with the lowest total cost TC_i.

Goyal (1995) has shown that this procedure can be shortened further by identifying a cutoff price C^* above which the optimal solution cannot occur. Recall that C_r is the lowest unit cost above the final threshold quantity q_r. The cutoff is obtained as follows:

$$C^* = \frac{1}{D}\left(DC_r + \frac{DS}{q_r} + \frac{h}{2}q_rC_r - \sqrt{2hDSC_r}\right)$$

In Example 11-7, we evaluate the optimal lot size given an all unit quantity discount (see worksheets *Example 11-7* and *Example 11-7 check* in spreadsheet *Chapter11-examples7-8*).

EXAMPLE 11-7 All Unit Quantity Discounts

Drugs Online (DO) is an online retailer of prescription drugs and health supplements. Vitamins represent a significant percentage of its sales. Demand for vitamins is 10,000 bottles per month. DO incurs a fixed order placement, transportation, and receiving cost of $100 each time an order for vitamins is placed with the manufacturer. DO incurs a holding cost of 20 percent. The manufacturer uses the following all unit discount pricing schedule. Evaluate the number of bottles that the DO manager should order in each lot.

Order Quantity	Unit Price
0–4,999	$3.00
5,000–9,999	$2.96
10,000 or more	$2.92

Analysis:

In this case, the manager has the following inputs:

$q_0 = 0$, $q_1 = 5{,}000$, $q_2 = 10{,}000$
$C_0 = \$3.00$, $C_1 = \$2.96$, $C_2 = \$2.92$
$D = 120{,}000/\text{year}$, $S = \$100/\text{lot}$, $h = 0.2$

Using Step 1 and Equation 11.10 we obtain

$$Q_0 = \sqrt{\frac{2DS}{hC_0}} = 6{,}325;\ Q_1 = \sqrt{\frac{2DS}{hC_1}} = 6{,}367;\ Q_2 = \sqrt{\frac{2DS}{hC_2}} = 6{,}411$$

In Step 2, we ignore $i = 0$ because $Q_0 = 6{,}325 > q_1 = 5{,}000$. For $i = 1, 2$, we obtain

$$Q_1^* = Q_1 = 6{,}367;\ Q_2^* = q_2 = 10{,}000$$

In Step 3, we obtain the total costs using Equation 11.11 as follows:

$$TC_1 = \left(\frac{D}{Q_1^*}\right)S + \left(\frac{Q_1^*}{2}\right)hC_1 + DC_1 = \$358{,}969;\ TC_2 = \$354{,}520$$

Observe that the lowest total cost is for $i = 2$. Thus, it is optimal for DO to order $Q_2^* = 10{,}000$ bottles per lot and obtain the discount price of $2.92 per bottle.

If the manufacturer in Example 11-7 sold all bottles for $3, it would be optimal for DO to order in lots of 6,325 bottles. The quantity discount is an incentive for DO to order in larger lots of 10,000 bottles, raising both the cycle inventory and the flow time. The impact of the discount is further magnified if DO works hard to reduce its fixed ordering cost to $S = \$4$ (from the current $100). Then, the optimal lot size in the absence of a discount is 1,265 bottles. In the presence of the all unit quantity discount, the optimal lot size is still 10,000 bottles. In this case, the presence of quantity discounts leads to an eightfold increase in average inventory and flow time at DO.

Given that all unit quantity discounts increase the average inventory and flow time in a supply chain, it is important to identify how these discounts add value in a supply chain. Before we consider this question, we discuss marginal unit quantity discounts.

Marginal Unit Quantity Discounts

Marginal (or incremental) unit quantity discounts are also referred to as *multi-block tariffs.* In this case, the pricing schedule contains specified break points $q_0, q_1, \ldots, q_r$. It is not the *average*

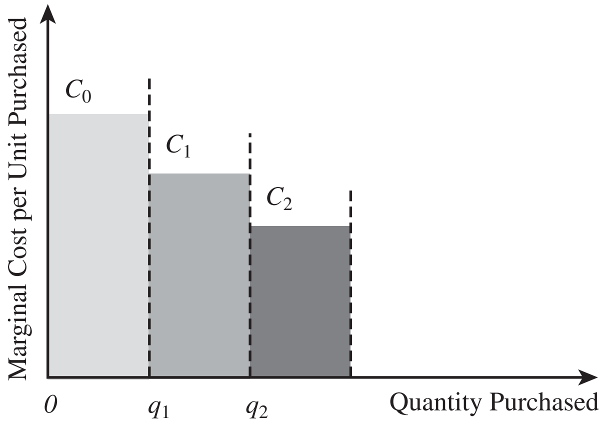


FIGURE 11-4 Marginal Unit Cost with Marginal Unit Quantity Discount

cost of a unit but the *marginal cost* of a unit that decreases at a breakpoint (in contrast to the all unit discount scheme). If an order of size q is placed, the first $q_1 - q_0$ units are priced at C_0, the next $q_2 - q_1$ are priced at C_1, and, in general, $q_{i+1} - q_i$ units are priced at C_i. The marginal cost per unit varies with the quantity purchased, as shown in Figure 11-4.

Faced with such a pricing schedule, the retailer's objective is to decide on a lot size that maximizes profits or, equivalently, minimizes material, order, and holding costs.

The solution procedure discussed here evaluates the optimal lot size for each marginal price C_i (this forces a lot size between q_i and q_{i+1}) and then settles on the lot size that minimizes the overall cost. A more streamlined procedure has been provided by Hu and Munson (2002).

For each value of i, $0 \leq i \leq r$, let V_i be the cost of ordering q_i units. Define $V_0 = 0$ and V_i for $0 \leq i \leq r$ as follows:

$$V_i = C_0(q_1 - q_0) + C_1(q_2 - q_1) + \cdots + C_{i-1}(q_i - q_{i-1}) \quad \textbf{(11.12)}$$

For each value of i, $0 \leq i \leq r - 1$, consider an order of size Q in the range q_i to q_{i+1} units; that is, $q_{i+1} \geq Q \geq q_i$. The material cost of each order of size Q is given by $V_i + (Q - q_i)C_i$. The various costs associated with such an order are as follows:

$$\text{Annual order cost} = \left(\frac{D}{Q}\right)S$$

$$\text{Annual holding cost} = [V_i + (Q - q_i)C_i]h/2$$

$$\text{Annual material cost} = \frac{D}{Q}[V_i + (Q - q_i)C_i]$$

The total annual cost is the sum of the three costs and is given by

$$\text{Total annual cost} = \left(\frac{D}{Q}\right)S + [V_i + (Q - q_i)C_i]h/2 + \frac{D}{Q}[V_i + (Q - q_i)C_i]$$

The optimal lot size for price C_i is obtained by taking the first derivative of the total cost with respect to the lot size and setting it equal to 0. This results in the following optimal lot size:

$$\text{Optimal 1ot size for price } C_i \text{ is } Q_i = \sqrt{\frac{2D(S + V_i - q_iC_i)}{hC_i}} \quad \textbf{(11.13)}$$

Observe that the optimal lot size is obtained using a formula very much like the EOQ formula (Equation 11.5), except that the presence of the quantity discount has the effect of raising the fixed cost per order by $V_i - q_iC_i$ (from S to $S + V_i - q_iC_i$). The overall optimal lot size is obtained as follows:

Step 1: Evaluate the optimal lot size using Equation 11.13 for each price C_i.

Step 2: We next select the order quantity Q_i^* for each price C_i. There are three possible cases for Q_i:

1. If $q_i \le Q_i \le q_{i+1}$ then set $Q_i^* = Q_i$
2. If $Q_i < q_i$ then set $Q_i^* = q_i$
3. If $Q_i > q_{i+1}$ then set $Q_i^* = q_{i+1}$

Step 3: Calculate the total annual cost of ordering Q_i^* units as follows:

$$TC_i = \left(\frac{D}{Q_i^*}\right)S + [V_i + (Q_i^* - q_i)C_i]h/2 + \frac{D}{Q_i^*}[V_i + (Q_i^* - q_i)C_i] \qquad \textbf{(11.14)}$$

Step 4: Over all i, select order size Q_i^* with the lowest cost TC_i.

In Example 11-8, we evaluate the optimal lot size given a marginal unit quantity discount (see worksheets *Example 11-8* and *Example 11-8 check* in spreadsheet *Chapter11-examples7-8*).

EXAMPLE 11-8 Marginal Unit Quantity Discounts

Let us return to DO from Example 11-7. Assume that the manufacturer uses the following marginal unit discount pricing schedule:

Order Quantity	Marginal Unit Price
0–5,000	\$3.00
5,000–10,000	\$2.96
Over 10,000	\$2.92

This implies that if an order is placed for 7,000 bottles, the first 5,000 are at a unit cost of \$3.00, with the remaining 2,000 at a unit cost of \$2.96. Evaluate the number of bottles that DO should order in each lot.

Analysis:

In this case, we have

$$q_0 = O, q_1 = 5{,}000, q_2 = 10{,}000$$
$$C_0 = \$3.00, C_1 = \$2.96, C_2 = \$2.92$$
$$D = 120{,}000/\text{year},\ S = \$100/\text{lot},\ h = 0.2$$

Using Equation 11.12, we obtain

$$V_0 = 0;\ V_1 = 3 \times (5{,}000 - 0) = \$15{,}000$$
$$V_2 = 3 \times (5{,}000 - 0) + 2.96 \times (10{,}000 - 5{,}000) = \$29{,}800$$

Using Step 1 and Equation 11.13, we obtain

$$Q_0 = \sqrt{\frac{2D(S + V_0 - q_0C_0)}{hC_0}} = 6{,}325$$
$$Q_1 = \sqrt{\frac{2D(S + V_1 - q_1C_1)}{hC_1}} = 11{,}028$$
$$Q_2 = \sqrt{\frac{2D(S + V_2 - q_2C_2)}{hC_2}} = 16{,}961$$

In Step 2, we set $Q_0^* = q_1 = 5{,}000$ because $Q_0 = 6{,}325 > q_1 = 5{,}000$. Similarly, we obtain $Q_1^* = q_2 = 10{,}000$ (because $Q_1 = 11{,}028 > q_2 = 10{,}000$) and $Q_2^* = Q_2 = 16{,}961$.

In Step 3, we obtain the total cost for $i = 0, 1, 2$ using Equation 11.14 to be

$$TC_0 = \left(\frac{D}{Q_0^*}\right)S + [V_0 + (Q_0^* - q_0)C_0]\,h/2 + \frac{D}{Q_0^*}[V_0 + (Q_0^* - q_0)C_0] = \$363{,}900$$

$$TC_1 = \left(\frac{D}{Q_1^*}\right)S + [V_1 + (Q_1^* - q_1)C_1]\,h/2 + \frac{D}{Q_1^*}[V_1 + (Q_1^* - q_1)C_1] = \$361{,}780$$

$$TC_2 = \left(\frac{D}{Q_2^*}\right)S + [V_2 + (Q_2^* - q_2)C_2]\,h/2 + \frac{D}{Q_2^*}[V_2 + (Q_2^* - q_2)C_2] = \$360{,}365$$

Observe that the lowest cost is for $i = 2$. Thus, it is optimal for DO to order in lots of $Q_2^* = 16{,}961$ bottles. This is much larger than the optimal lot size of 6,325 when the manufacturer does not offer any discount.

If the fixed cost of ordering is $4, the optimal lot size for DO is 15,755 with the discount compared to a lot size of 1,265 without the discount. This discussion demonstrates that there can be significant order sizes—and, thus, significant cycle inventory—in the absence of any formal fixed ordering costs as long as quantity discounts are offered. Thus, quantity discounts lead to a significant buildup of cycle inventory in a supply chain. In many supply chains, quantity discounts contribute more to cycle inventory than fixed ordering costs. This forces us once again to question the value of quantity discounts in a supply chain.

Why Do Suppliers Offer Quantity Discounts?

We have seen that the presence of lot-size–based quantity discounts tends to increase the level of cycle inventory in the supply chain. We now develop reasons why suppliers may offer lot-size–based quantity discounts in a supply chain. In each case, we look for circumstances under which a lot-size–based quantity discount increases supply chain profits. Quantity discounts can increase the supply chain profit for the following two main reasons:

1. Improved coordination to increase total supply chain profits
2. Extraction of surplus by supplier through price discrimination

Munson and Rosenblatt (1998) also provide other factors, such as marketing, that motivate sellers to offer quantity discounts. We now discuss each of the two situations in greater detail.

COORDINATION TO INCREASE TOTAL SUPPLY CHAIN PROFITS A supply chain is *coordinated* if the decisions the retailer and supplier make maximize total supply chain profits. In reality, each stage in a supply chain may have a separate owner and thus attempt to maximize its own profits. For example, each stage of a supply chain is likely to make lot-sizing decisions with an objective of minimizing its own overall costs. The result of this independent decision making can be a lack of coordination in a supply chain because actions that maximize retailer profits may not maximize supply chain profits. In this section, we discuss how a manufacturer may use appropriate quantity discounts to ensure that total supply chain profits are maximized even if the retailer is acting to maximize its own profits.

Quantity discounts for commodity products. Economists have argued that for commodity products such as milk, a competitive market exists and prices are driven down to the products' marginal cost. In this case, the market sets the price and the firm's objective is to lower costs in order to increase profits. Consider, for example, the online retailer DO, discussed earlier. It can be argued that it sells a commodity product. In this supply chain, both the manufacturer and DO incur costs related to each order placed by DO. Assume that the manufacturer has a fixed cost S_M, a unit cost C_M, and a holding cost h_M. The manufacturer incurs fixed costs related to order setup and fulfillment (S_M) and holding costs ($h_M C_M$) as it builds up inventory to replenish the order. Assume that the retailer has a fixed cost S_R, a unit cost C_R, and a holding cost h_R. Thus, DO incurs fixed costs (S_R) for each order it places and holding costs ($h_R C_R$) for the inventory it holds as it slowly sells an order. Even though both parties incur costs associated with the lot-sizing decision made by DO, the retailer makes its lot-sizing decisions based solely on minimizing its local costs. This results in lot-sizing decisions that are locally optimal but do not maximize the supply chain surplus. We illustrate this idea in Example 11-9 (see spreadsheet *Chapter11-quantity discounts* worksheet *Example 11-9*).

EXAMPLE 11-9 The Impact of Locally Optimal Lot Sizes on a Supply Chain

Demand for vitamins is 10,000 bottles per month. DO incurs a fixed order placement, transportation, and receiving cost of $100 each time it places an order for vitamins with the manufacturer. DO incurs a holding cost of 20 percent. The manufacturer charges $3 for each bottle of vitamins purchased. Evaluate the optimal lot size for DO.

Each time DO places an order, the manufacturer must process, pack, and ship the order. The manufacturer has a line packing bottles at a steady rate that matches demand. The manufacturer incurs a fixed-order filling cost of $250, production cost of $2 per bottle, and a holding cost of 20 percent. What is the annual fulfillment and holding cost incurred by the manufacturer as a result of DO's ordering policy?

Analysis:

In this case, we have

$$D = 120{,}000/\text{year},\ S_R = \$100/\text{lot},\ h_R = 0.2,\ C_R = \$3$$
$$S_M = \$250/\text{lot},\ h_M = 0.2,\ C_M = \$2$$

Using the EOQ formula (Equation 11.5), we obtain the optimal lot size and annual cost for DO to be:

$$Q_R = \sqrt{\frac{2DS_R}{h_R C_R}} = \sqrt{\frac{2 \times 120{,}000 \times 100}{0.2 \times 3}} = 6{,}325$$

$$\text{Annual cost for DO} = \left(\frac{D}{Q_R}\right)S_R + \left(\frac{Q_R}{2}\right)h_R C_R = \$3{,}795$$

If DO orders in lots sizes of $Q_R = 6{,}325$, the annual cost incurred by the manufacturer is obtained to be:

$$\text{Annual cost for manufacturer} = \left(\frac{D}{Q_R}\right)S_M + \left(\frac{Q_R}{2}\right)h_M C_M = \$6{,}008$$

The annual supply chain cost (manufacturer + DO) is thus $6,008 + $3,795 = $9,803.

In Example 11-9, DO picks the lot size of 6,325 with an objective of minimizing only its own costs. From a supply chain perspective, the optimal lot size should account for the fact that both DO and the manufacturer incur costs associated with each replenishment lot. If we assume that the manufacturer produces at a rate that matches demand (as assumed in Example 11-9), the total supply chain cost of using a lot size Q is obtained as follows:

$$\text{Annual cost for DO and manufacturer} = \left(\frac{D}{Q}\right)S_R + \left(\frac{Q}{2}\right)h_R C_R + \left(\frac{D}{Q}\right)S_M + \left(\frac{Q}{2}\right)h_M C_M$$

The optimal lot size (Q^*) for the supply chain is obtained by taking the first derivative of the total cost with respect to Q and setting it equal to 0 as follows (see worksheet *Example 11-9*):

$$Q^* = \sqrt{\frac{2D(S_R + S_M)}{h_R C_R + h_M C_M}} = 9{,}165$$

If DO orders in lots of $Q^* = 9{,}165$ units, the total costs for DO and the manufacturer are as follows:

$$\text{Annual cost for DO} = \left(\frac{D}{Q^*}\right)S_R + \left(\frac{Q^*}{2}\right)h_R C_R = \$4{,}059$$

$$\text{Annual cost for manufacturer} = \left(\frac{D}{Q^*}\right)S_M + \left(\frac{Q^*}{2}\right)h_M C_M = \$5{,}106$$

Observe that if DO orders a lot size of 9,165 units, the supply chain cost decreases to $9,165 (from $9,803 when DO ordered its own optimal lot size of 6,325). There is thus an opportunity for the supply chain to save $638. The challenge, however, is that ordering in lots of 9,165 bottles raises the cost for DO by $264 per year from $3,795 to $4,059 (even though it reduces overall supply chain costs). The manufacturer's costs, in contrast, go down by $902 from $6,008 to $5,106 per year. Thus, the manufacturer must offer DO a suitable incentive for DO to raise its lot size. A lot-size–based quantity discount is an appropriate incentive in this case. Example 11-10 (see worksheet *Example 11-10*) provides details of how the manufacturer can design a suitable quantity discount that gets DO to order in lots of 9,165 units even though DO is optimizing its own profits (and not total supply chain profits).

EXAMPLE 11-10 Designing a Suitable Lot-Size–Based Quantity Discount

Consider the data from Example 11-9. Design a suitable quantity discount that gets DO to order in lots of 9,165 units when it aims to minimize only its own total costs.

Analysis:

Recall that ordering in lots of 9,165 units instead of 6,325 increases annual ordering and holding costs for DO by $264. Thus, the manufacturer needs to offer an incentive of at least $264 per year to DO in terms of decreased material cost if DO orders in lots of 9,165 units. Decreasing material cost by $264/year from sales of 120,000 units implies that material cost must be decreased from $3/unit to $3 – 264/120,000 = $2.9978/unit if DO orders in lots of 9,165.

Thus, the appropriate quantity discount is for the manufacturer to charge $3 if DO orders in lots that are smaller than 9,165 units and discount the price to $2.9978 for orders of 9,165 or more.

Observe that offering a lot-size–based discount in this case decreases total supply chain cost. It does, however, increase the lot size the retailer purchases and thus increases cycle inventory in the supply chain.

Key Point

For commodity products for which price is set by the market, manufacturers with large fixed costs per lot can use lot-size–based quantity discounts to maximize total supply chain profits. Lot-size–based discounts, however, increase cycle inventory in the supply chain.

Our discussion on coordination for commodity products highlights the important link between the lot-size–based quantity discount offered and the order costs incurred by the manufacturer. As the manufacturer works on lowering order or setup cost, the discount it offers to retailers should change. For a low enough setup or order cost, the manufacturer gains little from using a lot-size–based quantity discount. In Example 11-9, discussed earlier, if the manufacturer lowers its fixed cost per order from \$250 to \$100, the total supply chain costs are close to the minimum without quantity discounts even if DO is trying to minimize its cost. Thus, if its fixed order costs are lowered to \$100, it makes sense for the manufacturer to eliminate all quantity discounts. In most companies, however, marketing and sales design quantity discounts, whereas operations works on reducing setup or order cost. As a result, changes in pricing do not always occur in response to setup cost reduction in manufacturing. It is important that the two functions coordinate these activities.

Quantity discounts for products for which the firm has market power. Now, consider the scenario in which the manufacturer has invented a new vitamin pill, Vitaherb, which is derived from herbal ingredients and has other properties highly valued in the market. Few competitors have a similar product, so it can be argued that the price at which the retailer DO sells Vitaherb influences demand. Assume that the annual demand faced by DO is given by the demand curve 360,000 – 60,000*p*, where *p* is the price at which DO sells Vitaherb. The manufacturer incurs a production cost of $C_M = \$2$ per bottle of Vitaherb sold. The manufacturer must decide on the price C_R to charge DO, and DO in turn must decide on the price *p* to charge the customer. The profit at DO ($Prof_R$) and the manufacturer ($Prof_M$) as a result of this policy is given by

$$Prof_R = (p - C_R)(360{,}000 - 60{,}000\,p);\ Prof_M = (C_R - C_M)(360{,}000 - 60{,}000\,p)$$

DO picks the price *p* to maximize $Prof_R$. Taking the first derivative with respect to *p* and setting it to 0, we obtain the following relationship between *p* and C_R

$$p = 3 + \frac{C_R}{2} \qquad \textbf{(11.15)}$$

Given that the manufacturer is aware that DO is aiming to optimize its own profits, the manufacturer is able to use the relationship between *p* and C_R to obtains its own profits to be

$$Prof_M = (C_R - C_M)\left(360{,}000 - 60{,}000\left(3 + \frac{C_R}{2}\right)\right) = (C_R - 2)(180{,}000 - 30{,}000\,C_R)$$

The manufacturer picks its price C_R to maximize $Prof_M$. Taking the first derivative of $Prof_M$ with respect C_R to and setting it to 0 we obtain $C_R = \$4$. Substituting back into Equation 11.15, we obtain $p = \$5$. Thus, when DO and the manufacturer make their pricing decisions independently, it is optimal for the manufacturer to charge a wholesale price of $C_R = \$4$ and for DO to charge a retail price of $p = \$5$. The total market demand in this case is 360,000 – 60,000*p* = 60,000 bottles. DO makes a profit of $Prof_R = (5 - 4)(360{,}000 - [60{,}000 \times 5]) = \$60{,}000$ and the manufacturer makes a profit of $Prof_M = (4 - 2)(360{,}000 - [60{,}000 \times 5]) = \$120{,}000$ (see worksheet *2-stage*).

Now, consider the case in which the two stages coordinate their pricing decisions with a goal of maximizing the supply chain profit $Prof_{SC}$, which is given by

$$Prof_{SC} = (p - C_M)(360{,}000 - 60{,}000p)$$

The optimal retail price is obtained by setting the first derivative of $Prof_{SC}$ with respect to *p* to 0. We thus obtain the coordinated retail price to be

$$p = 3 + \frac{C_M}{2} = 3 + \frac{2}{2} = \$4$$

If the two stages coordinate pricing and DO prices at $p = \$4$, market demand is 360,000 – 60,000p = 120,000 bottles. The total supply chain profit if the two stages coordinate is $Prof_{SC} = (\$4 - \$2) \times 120{,}000 = \$240{,}000$. As a result of each stage settings its price independently, the supply chain thus loses $60,000 in profit. This phenomenon is referred to as *double marginalization.* Double marginalization leads to a loss in profit because the supply chain margin is divided between two stages, but each stage makes its pricing decision considering only its own local profits.

Key Point

The supply chain profit is lower if each stage of the supply chain makes its pricing decisions independently, with the objective of maximizing its own profit. A coordinated solution results in higher profit.

Given that independent pricing decisions lower supply chain profits, it is important to consider pricing schemes that may help recover some of these profits even when each stage of the supply chain continues to act independently. We propose two pricing schemes that the manufacturer may use to achieve the coordinated solution and maximize supply chain profits even though DO acts in a way that maximizes its own profit.

1. *Two-part tariff:* In this case, the manufacturer charges its entire profit as an up-front franchise fee *ff* (which could be anywhere between the noncoordinated manufacturer profit $Prof_M$ and the difference between the coordinated supply chain profit and the noncoordinated retailer profit, $Prof_{SC} - Prof_R$) and then sells to the retailer at cost; that is, the manufacturer sets its wholesale price $C_R = C_M$. This pricing scheme is referred to as a *two-part tariff* because the manufacturer sets both the franchise fee and the wholesale price. The retail pricing decision is thus based on maximizing its profits $(p - C_M)(360{,}000 - 60{,}000p) - ff$. Under the two-part tariff, the franchise fee *ff* is paid up front and is thus a fixed cost that does not change with the retail price p. The retailer DO is thus effectively maximizing the coordinated supply chain profits $Prof_{SC} = (p - C_M)(360{,}000 - 60{,}000p)$. Taking the first derivative with respect to p and setting it equal to 0, the optimal coordinated retail price p is evaluated to be

$$p = 3 + \frac{C_M}{2}$$

In the case of DO, recall that total supply chain profit when the two stages coordinate is $Prof_{SC} =$ $240,000 with DO charging the customer $4 per bottle of Vitaherb. The profit made by DO when the two stages do not coordinate is $Prof_R = \$60{,}000$. One option available to the manufacturer is to construct a two-part tariff by which DO is charged an upfront fee of $ff = Prof_{SC} - Prof_R =$ $180,000 (see worksheet *2-part-tariff*) and material cost of $C_R = C_M = \$2$ per bottle. DO maximizes its profit if it prices the vitamins at $p = 3 + C_M / 2 = 3 + 2 / 2 = \$4$ per bottle. It has annual sales of $360{,}000 - 60{,}000p = 120{,}000$ and profits of $60,000. The manufacturer makes a profit of $180,000, which it charges up front. Observe that the use of a two-part tariff has increased supply chain profits from $180,000 to $240,000 even though the retailer DO has made a locally optimal pricing decision given the two-part tariff. A similar result can be obtained as long as the manufacturer sets the up-front fee *ff* to be any value between $120,000 and $180,000 with a wholesale price of $C_R = C_M = 2$.

2. *Volume-based quantity discount:* Observe that the two-part tariff is really a volume-based quantity discount whereby the retailer DO pays a lower average unit cost as it purchases larger quantities each year (the franchise fee *ff* is amortized over more units). This observation can be made explicit by designing a volume-based discount scheme that gets the retailer DO to purchase and sell the quantity sold when the two stages coordinate their actions.

Recall that the coordinated solution results in a retail price of $p = 3 + C_M/2 = 3 + 2/2 = 4$. This retail price results in total demand of $d^{coord} = 360{,}000 - (60{,}000 \times 4) = 120{,}000$. The objective of the manufacturer is to design a volume-based discounting scheme that gets the retailer DO to buy (and sell) $d^{coord} = 120{,}000$ units each year. The pricing scheme must be such that retailer gets a profit of at least \$60,000, and the manufacturer gets a profit of at least \$120,000 (these are the profits that DO and the manufacturer made when their actions were not coordinated).

Several such pricing schemes can be designed. One such scheme is for the manufacturer to charge a wholesale price of $C_R = \$4$ per bottle (this is the same wholesale price that is optimal when the two stages are not coordinated) for annual sales below $d^{coord} = 120{,}000$ units, and to charge $C_R = \$3.50$ (any value between \$3.00 and \$3.50 will work) if sales reach 120,000 or more (see worksheet *Volume Discount*). It is then optimal for DO to order 120,000 units in the year and price them at $p = \$4$ per bottle to the customers (to ensure that they are all sold). The total profit earned by DO $(360{,}000 - 60{,}000p) \times (p - C_R) = \$60{,}000$. The total profit earned by the manufacturer is $120{,}000 \times (C_R - \$2) = 180{,}000$ when $C_R = \$3.50$. The total supply chain profit is \$240,000, which is higher than the \$180,000 that the supply chain earned when actions were not coordinated.

If the manufacturer charges \$3.00 (instead of \$3.50) for sales of 120,000 units or more, it is still optimal for DO to order 120,000 units in the year and price them at $p = \$4$ per bottle. The only difference is that the total profit earned by DO now increases to \$120,000, whereas that for the manufacturer now drops to \$120,000. The total supply chain profits remain at \$240,000. The price that the manufacturer is able to charge (between \$3.00 and \$3.50) for sales of 120,000 or more will depend on the relative bargaining power of the two parties.

At this stage, we have seen that even in the absence of inventory-related costs, quantity discounts play a role in supply chain coordination and improved supply chain profits. Unless the manufacturer has large fixed costs associated with each lot, the discount schemes that are optimal are volume based and not lot-size based. It can be shown that even in the presence of large fixed costs for the manufacturer, a two-part tariff or volume-based discount, with the manufacturer passing on some of the fixed cost to the retailer, optimally coordinates the supply chain and maximizes profits given the assumption that customer demand decreases when the retailer increases price.

A key distinction between lot-size–based and volume discounts is that lot-size discounts are based on the quantity purchased per lot, not the rate of purchase. Volume discounts, in contrast, are based on the rate of purchase or volume purchased on average per specified time period (say, a month, quarter, or year). Lot-size–based discounts tend to raise the cycle inventory in the supply chain by encouraging retailers to increase the size of each lot. Volume-based discounts, in contrast, are compatible with small lots that reduce cycle inventory. Lot-size–based discounts make sense only when the manufacturer incurs high fixed cost per order. In all other instances, it is better to have volume-based discounts.

Key Point

For products for which the firm has market power, two-part tariffs or volume-based quantity discounts can be used to achieve coordination in the supply chain and maximize supply chain profits.

Key Point

For products for which a firm has market power, lot-size–based discounts are not optimal for the supply chain even in the presence of inventory costs. In such a setting, either a two-part tariff or a volume-based discount, with the supplier passing on some of its fixed cost to the retailer, is needed for the supply chain to be coordinated and maximize profits.

One can make the point that even with volume-based discounts, retailers will tend to increase the size of the lot toward the end of the evaluation period. For example, assume that the manufacturer offers DO a 2 percent discount if the number of bottles of Vitaherb purchased over a quarter exceeds 40,000. This policy will not affect the lot sizes DO orders early during the quarter, and DO will order in small lots to match the quantity ordered with demand. Consider a situation, however, in which DO has sold only 30,000 bottles with a week left before the end of the quarter. To get the quantity discount, DO may order 10,000 bottles over the last week even though it expects to sell only 3,000. In this case, cycle inventory in the supply chain goes up in spite of the fact that there is no lot-size–based quantity discount. The situation in which orders peak toward the end of a financial horizon is referred to as the *hockey stick phenomenon* because demand increases dramatically toward the end of a period, similar to the way a hockey stick bends upward toward its end. This phenomenon has been observed in many industries. One possible solution is to base the volume discounts on a rolling horizon. For example, each week the manufacturer may offer DO the volume discount based on sales over the past 12 weeks. Such a rolling horizon dampens the hockey stick phenomenon by making each week the last week in some 12-week horizon.

Thus far, we have discussed only the scenario in which the supply chain has a single retailer. One may ask whether our insights are robust and also apply if the supply chain has multiple retailers, each with different demand curves, all supplied by a single manufacturer. As one would expect, the form of the discount scheme to be offered becomes more complicated in these settings (typically, instead of having only one breakpoint at which the volume-based discount is offered, there are multiple breakpoints). The basic form of the optimal pricing scheme, however, does not change. The optimal discount continues to be volume based, with the average price charged to the retailers decreasing as the rate of purchase (volume purchased per unit time) increases.

PRICE DISCRIMINATION TO MAXIMIZE SUPPLIER PROFITS *Price discrimination* is the practice in which a firm charges different prices to maximize profits. An example of price discrimination is airlines: Passengers traveling on the same plane often pay different prices for their seats.

As discussed in Chapter 16, setting a fixed price for all units does not maximize profits for the manufacturer. In principle, the manufacturer can obtain the entire area under the demand curve above its marginal cost by pricing each unit differently based on customers' marginal willingness to pay at each quantity. Quantity discounts are one mechanism for price discrimination because customers pay different prices based on the quantity purchased.

Next we discuss trade promotions and their impact on lot sizes and cycle inventory in the supply chain.

Key Point

Price discrimination to maximize profits at the manufacturer may also be a reason to offer quantity discounts within a supply chain.

11.6 SHORT-TERM DISCOUNTING: TRADE PROMOTIONS

Manufacturers use *trade promotions* to offer a discounted price to retailers and set a time period over which the discount is effective. For example, a manufacturer of canned soup may offer a price discount of 10 percent for the shipping period December 15 to January 25. For all purchases within the specified time horizon, retailers get a 10 percent discount. In some cases, the manufacturer may require specific actions from the retailer, such as displays, advertising, promotion, and so on, to qualify for the trade promotion. Trade promotions are quite common in the consumer packaged-goods industry, with manufacturers promoting different products at different times of the year.

The goal of trade promotions is to influence retailers to act in a way that helps the manufacturer achieve its objectives. The following are a few of the key goals (from the manufacturer's perspective) of a trade promotion (see Blattberg and Neslin [1990] for more details):

1. Induce retailers to use price discounts, displays, or advertising to spur sales.
2. Shift inventory from the manufacturer to the retailer and the customer.
3. Defend a brand against competition.

Although these may be the manufacturer's objectives, it is not clear that they are always achieved as the result of a trade promotion. Our goal in this section is to investigate the impact of a trade promotion on the behavior of the retailer and the performance of the entire supply chain. The key to understanding this impact is to focus on how a retailer reacts to a trade promotion that a manufacturer offers. In response to a trade promotion, the retailer has the following options:

1. Pass through some or all of the promotion to customers to spur sales.
2. Pass through very little of the promotion to customers but purchase in greater quantity during the promotion period to exploit the temporary reduction in price.

The first action lowers the price of the product for the end customer, leading to increased purchases and, thus, increased sales for the entire supply chain. The second action does not increase purchases by the customer, but increases the amount purchased and held at the retailer. As a result, the cycle inventory and flow time within the supply chain increase.

A *forward buy* occurs when a retailer purchases in the promotional period for sales in future periods. A forward buy helps reduce the retailer's future cost of goods for product sold after the promotion ends. Although a forward buy is often the retailer's appropriate response to a price promotion, it can decrease supply chain profits because it results in higher demand variability, with a resulting increase in inventory and flow times within the supply chain.

Our objective in this section is to understand a retailer's optimal response when faced with a trade promotion. We identify the factors affecting the forward buy and quantify the size of a forward buy by the retailer. We also identify factors that influence the amount of the promotion that a retailer passes on to the customer.

We first illustrate the impact of a trade promotion on forward buying behavior of the retailer. Consider a Cub Foods supermarket selling chicken noodle soup manufactured by the Campbell Soup Company. Customer demand for chicken noodle soup is D cans per year. Campbell charges $\$C$ per can. Cub Foods incurs a holding cost of h (per dollar of inventory held for a year). Using the EOQ formula (Equation 11.5), Cub Foods normally orders in the following lot sizes:

$$Q^* = \sqrt{\frac{2DS}{hC}}$$

Campbell announces that it is offering a discount of $\$d$ per can for the coming four-week period. Cub Foods must decide how much to order at the discounted price compared with the lot size of Q^* that it normally orders. Let Q^d be the lot size ordered at the discounted price.

The costs the retailer must consider when making this decision are material cost, holding cost, and order cost. Increasing the lot size Q^d lowers the material cost for Cub Foods because it purchases more cans (for sale now and in the future) at the discounted price. Increasing the lot size Q^d increases the holding cost because inventories increase. Increasing the lot size Q^d lowers the order cost for Cub Foods because some orders that would otherwise have been placed are now not necessary. Cub Foods' goal is to make the trade-off that minimizes the total cost.

The inventory pattern when a lot size of Q^d is followed by lot sizes of Q^* is shown in Figure 11-5. The objective is to identify Q^d that minimizes the total cost (material cost + ordering cost + holding cost) over the time interval during which the quantity Q^d (ordered during the promotion period) is consumed.

The precise analysis in this case is complex, so we present a result that holds under some restrictions (see Silver, Pyke, and Petersen [1998] for a more detailed discussion). The first key

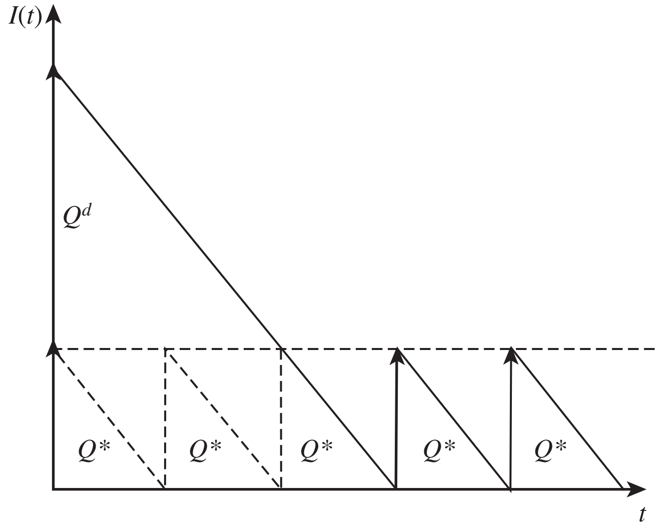


FIGURE 11-5 Inventory Profile for Forward Buying

assumption is that the discount is offered once, with no future discounts. The second key assumption is that the retailer takes no action (such as passing on part of the trade promotion) to influence customer demand. The customer demand thus remains unchanged. The third key assumption is that we analyze a period over which the demand is an integer multiple of Q^*. With these assumptions, the optimal order quantity at the discounted price is given by

$$Q^d = \frac{dD}{(C - d)h} + \frac{CQ^*}{C - d} \tag{11.16}$$

In practice, retailers are often aware of the timing of the next promotion. If the demand until the next anticipated trade promotion is Q_1, it is optimal for the retailer to order $min\{Q^d, Q_1\}$ Observe that the quantity Q^d ordered as a result of the promotion is larger than the regular order quantity Q^*. The forward buy in this case is given by

$$\text{Forward buy} = Q^d - Q^*$$

Even for relatively small discounts, the order size increases by a large quantity, as illustrated in Example 11-11 (see spreadsheet *Chapter11-examples11-12*).

EXAMPLE 11-11 Impact of Trade Promotions on Lot Sizes

DO is a retailer that sells Vitaherb, a popular vitamin diet supplement. Demand for Vitaherb is 120,000 bottles per year. The manufacturer currently charges \$3 for each bottle, and DO incurs a holding cost of 20 percent. DO currently orders in lots of $Q^* = 6{,}325$ bottles. The manufacturer has offered a discount of \$0.15 for all bottles purchased by retailers over the coming month. How many bottles of Vitaherb should DO order given the promotion?

Analysis:

In the absence of any promotion, DO orders in lot sizes of $Q^* = 6{,}325$ bottles. Given a monthly demand of $D = 10{,}000$ bottles, DO normally orders every 0.6325 months. In the absence of the trade promotion we have the following:

$$\text{Cycle inventory at DO} = Q^*/2 = 6{,}325/2 = 3{,}162.50 \text{ bottles}$$
$$\text{Average flow time} = Q^*/2D = 6{,}325/(2D) = 0.3162 \text{ months}$$

The optimal lot size during the promotion is obtained using Equation 11.15 and is given by

$$Q^d = \frac{dD}{(C-d)h} + \frac{CQ^*}{C-d} = \frac{0.15 \times 120{,}000}{(3.00 - 0.15) \times 0.20} + \frac{3 \times 6{,}325}{3.00 - 0.15} = 38{,}236$$

During the promotion, DO should place an order for a lot size of 38,236. In other words, DO places an order for 3.8236 months' worth of demand. In the presence of the trade promotion we have

$$\text{Cycle inventory at DO} = Q^d/2 = 38{,}236/2 = 19{,}118 \text{ bottles}$$
$$\text{Average flow time} = Q^d/(2D) = 38{,}236/(20{,}000) = 1.9118 \text{ months}$$

In this case, the forward buy is given by

$$\text{Forward buy} = Q^d - Q^* = 38{,}236 - 6{,}325 = 31{,}911 \text{ bottles}$$

As a result of this forward buy, DO will not place any order for the next 3.8236 months (without a forward buy, DO would have placed another 31,912/6,325 = 5.05 orders for 6,325 bottles each during this period). Observe that a 5 percent discount causes the lot size to increase by more than 500 percent.

As the example illustrates, forward buying as a result of trade promotions leads to a significant increase in the quantity ordered by the retailer. The large order is then followed by a period of small orders to compensate for the inventory built up at the retailer. The fluctuation in orders as a result of trade promotions is one of the major contributors to the bullwhip effect discussed in Chapter 10. The retailer can justify the forward buying during a trade promotion because it decreases its total cost. In contrast, the manufacturer can justify this action only as a competitive necessity (to counter a competitor's promotion) or if it has either inadvertently built up a lot of excess inventory or the forward buy allows the manufacturer to smooth demand by shifting it from peak- to low-demand periods. In practice, manufacturers often build up inventory in anticipation of planned promotions. During the trade promotion, this inventory shifts to the retailer, primarily as a forward buy. If the forward buy during trade promotions is a significant fraction of total sales, manufacturers end up reducing the revenues they earn from sales because most of the product is sold at a discount. The increase in inventory and the decrease in revenues often lead to a reduction in manufacturer as well as total supply chain profits as a result of trade promotions (see Blattberg and Neslin [1990] for more details).

Key Point

Trade promotions lead to a significant increase in lot size and cycle inventory because of forward buying by the retailer. This generally results in reduced supply chain profits unless the trade promotion reduces demand fluctuations.

Now, let us consider the extent to which the retailer may find it optimal to pass through some of the discount to the end customer to spur sales. As Example 11-12 shows, it is not optimal for the retailer to pass through the entire discount to the customer. In other words, it is optimal for the retailer to capture part of the promotion and pass through only part of it to the customer.

EXAMPLE 11-12 How Much of a Discount Should the Retailer Pass Through?

Assume that DO faces a demand curve for Vitaherb of $300{,}000 - 60{,}000p$. The normal price charged by the manufacturer to the retailer is $C_R = \$3$ per bottle. Ignoring all inventory-related costs, evaluate the optimal response of DO to a discount of \$0.15 per unit.

Analysis:

The profits for DO, the retailer, are given as follows:

$$Prof_R = (300{,}000 - 60{,}000\,p)\,p - (300{,}000 - 60{,}000\,p)C_R$$

The retailer prices to maximize profits, and the optimal retail price is obtained by setting the first derivative of retailer profits with respect to p to 0. This implies that

$$300{,}000 - 120{,}000p + 60{,}000\,C_R = 0$$

or

$$p = (300{,}000 + 60{,}000\,C_R)/120{,}000 \qquad \textbf{(11.17)}$$

Substituting $C_R = \$3$ into Equation 11.17, we obtain a retail price of $p = \$4$. As a result, the customer demand at the retailer in the absence of the promotion is

$$D_R = 30{,}000 - 60{,}000\,p = 60{,}000$$

During the promotion, the manufacturer offers a discount of \$0.15, resulting in a price to the retailer of $C_R = \$2.85$. Substituting into Equation 11.17, the optimal price set by DO is

$$p = (300{,}000 + 60{,}000 \times 2.85)/120{,}000 = \$3.925$$

Observe that the retailer's optimal response is to pass through only \$0.075 of the \$0.15 discount to the customer. The retailer does not pass through the entire discount. At the discounted price, DO experiences a demand of

$$D_R = 300{,}000 - 60{,}000\,p = 64{,}500$$

This represents an increase of 7.5 percent in demand relative to the base case. It is optimal here for DO to pass on half the trade promotion discount to the customers. This action results in a 7.5 percent increase in customer demand.

From Examples 11-11 and 11-12, observe that the increase in customer demand resulting from a trade promotion (7.5 percent of demand in Example 11-12) is small relative to the increased purchase by the retailer due to forward buying (500 percent from Example 11-11). The impact of the increase in customer demand may be further dampened by customer behavior. For many products, such as detergent and toothpaste, most of the increase in customer purchases is a forward buy by the customer; customers are unlikely to start brushing their teeth more frequently simply because they have purchased a lot of toothpaste. For such products, a trade promotion does not truly increase demand.

Key Point

Faced with a short-term discount, it is optimal for retailers to pass through only a fraction of the discount to the customer, keeping the rest for themselves. Simultaneously, it is optimal for retailers to increase the purchase lot size and forward buy for future periods. Thus, trade promotions often lead to an increase of cycle inventory in a supply chain without a significant increase in customer demand.

Manufacturers have always struggled with the fact that retailers pass along only a small fraction of a trade discount to the customer. In a study conducted by Kurt Salmon and Associates (1993), almost a quarter of all distributor inventories in the dry-grocery supply chain could be attributed to forward buying.

Our previous discussion supports the claim that trade promotions generally increase cycle inventory in a supply chain and hurt performance. This realization has led many firms, including the world's largest retailer, Walmart, and several manufacturers, such as P&G, to adopt "everyday low pricing" (EDLP). Here, the price is fixed over time and no short-term discounts are offered. This eliminates any incentive for forward buying. As a result, all stages of the supply chain purchase in quantities that match demand.

In general, the discount passed through by the retailer to the consumer is influenced by the retailer deal elasticity, which is the increase in retail sales per unit discount in price. The higher the deal elasticity, the more of the discount the retailer is likely to pass through to the consumer. Thus, trade promotions by the manufacturer may make sense for products with a high deal elasticity that ensures high pass-through by the retailer, and high holding costs that ensure low forward buying. Blattberg and Neslin (1990) identify paper goods as products with high deal elasticity and holding cost. They also identify trade promotions as being more effective with strong brands relative to weak brands.

Trade promotions may also make sense as a competitive response. In a category such as cola, some customers are loyal to their brand, whereas others switch depending on the brand being offered at the lowest price. Consider a situation in which one of the competitors—say, Pepsi—offers retailers a trade promotion. Retailers increase their purchases of Pepsi and pass through some of the discount to the customer. Price-sensitive customers increase their purchase of Pepsi. If a competitor such as Coca-Cola does not respond, it loses some market share in the form of price-sensitive customers. A case can be made that a trade promotion by Coca-Cola is justified in such a setting as a competitive response. Observe that with both competitors offering trade promotions, there is no real increase in demand for either unless customer consumption grows. Inventory in the supply chain, however, does increase for both brands. This is, then, a situation in which trade promotions are a competitive necessity, but they increase supply chain inventory, leading to reduced profits for all competitors.

Trade promotions should be designed so retailers limit their forward buying and pass along more of the discount to end customers. The manufacturer's objective is to increase market share and sales without allowing the retailer to forward buy significant amounts. This outcome can be achieved by offering discounts to the retailer that are based on actual sales to customers rather than the amount purchased by the retailer. The discount price thus applies to items sold to customers (sell-through) during the promotion, not the quantity purchased by the retailer (sell-in). This eliminates all incentive for forward buying.

Given the information technology in place, many manufacturers today offer scanner-based promotions by which the retailer receives credit for the promotion discount for every unit sold. Another option is to limit the allocation to a retailer based on past sales. This is also an effort to limit the amount that the retailer can forward buy. It is unlikely, however, that retailers will accept such schemes for weak brands.

11.7 MANAGING MULTIECHELON CYCLE INVENTORY

A *multiechelon* supply chain has multiple stages and possibly many players at each stage. The lack of coordination in lot sizing decisions across the supply chain results in high costs and more cycle inventory than required. The goal in a multiechelon system is to decrease total costs by coordinating orders across the supply chain.

Consider a simple multiechelon system with one manufacturer supplying one retailer. Assume that production is instantaneous, so the manufacturer can produce a lot when needed. If the two stages are not synchronized, the manufacturer may produce a new lot of size Q right after shipping a lot of size Q to the retailer. Inventory at the two stages in this case is as shown in Figure 11-6. In this case, the retailer carries an average inventory of $Q/2$ and the manufacturer carries an average inventory of about Q.

Overall supply chain inventory can be lowered if the manufacturer synchronizes its production to be ready just in time to be shipped to the retailer. In this case, the manufacturer carries no inventory and the retailer carries an average inventory of $Q/2$. Synchronization of production and replenishment allows the supply chain to lower total cycle inventory from about $3Q/2$ to $Q/2$.

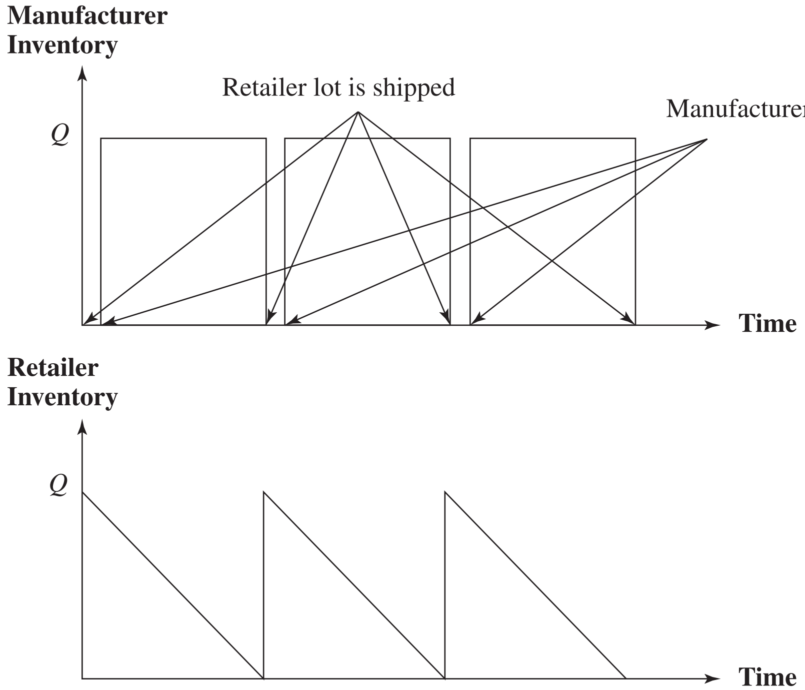


FIGURE 11-6 Inventory Profile at Retailer and Manufacturer with No Synchronization

For a simple multiechelon supply chain with only one player at each stage, ordering policies in which the lot size at each stage is an integer multiple of the lot size at its immediate customer have been shown to be quite close to optimal. When lot sizes are integer multiples, coordination of ordering across stages allows for a portion of the delivery to a stage to be cross-docked on to the next stage. The extent of cross-docking depends on the ratio of the fixed cost of ordering S and the holding cost H at each stage. The closer this ratio is between two stages, the higher is the optimal percentage of cross-docked product. Munson, Hu, and Rosenblatt (2003) provide optimal order quantities in a multiechelon setting with a single manufacturer supplying a single retailer.

If one party (distributor) in a supply chain supplies multiple parties (retailers) at the next stage of the supply chain, it is important to distinguish retailers with high demand from those with low demand. In this setting, Roundy (1985) has shown that a near-optimal policy results if retailers are grouped such that all retailers in one group order together and, for any retailer, either the ordering frequency is an integer multiple of the ordering frequency at the distributor or the ordering frequency at the distributor is an integer multiple of the frequency at the retailer. An integer replenishment policy has every player ordering periodically, with the length of the reorder interval for each player an integer multiple of some base period. An example of such a policy is shown in Figure 11-7. Under this policy, the distributor places a replenishment order every two weeks. Some retailers place replenishment orders every week, and others place replenishment orders every two or four weeks. Observe that for retailers ordering more frequently than the distributor, the retailers' ordering frequency is an integer multiple of the distributor's frequency. For retailers ordering less frequently than the distributor, the distributor's ordering frequency is an integer multiple of the retailers' frequency.

If an integer replenishment policy is synchronized across the two stages, the distributor can cross-dock part of its supply on to the next stage. All shipments to retailers ordering no more frequently than the distributor (every two or four weeks) are cross-docked, as shown in Figure 11-7. For retailers ordering more frequently (every week) than the distributor, half the orders are cross-docked, with the other half shipped from inventory, as shown in Figure 11-7.

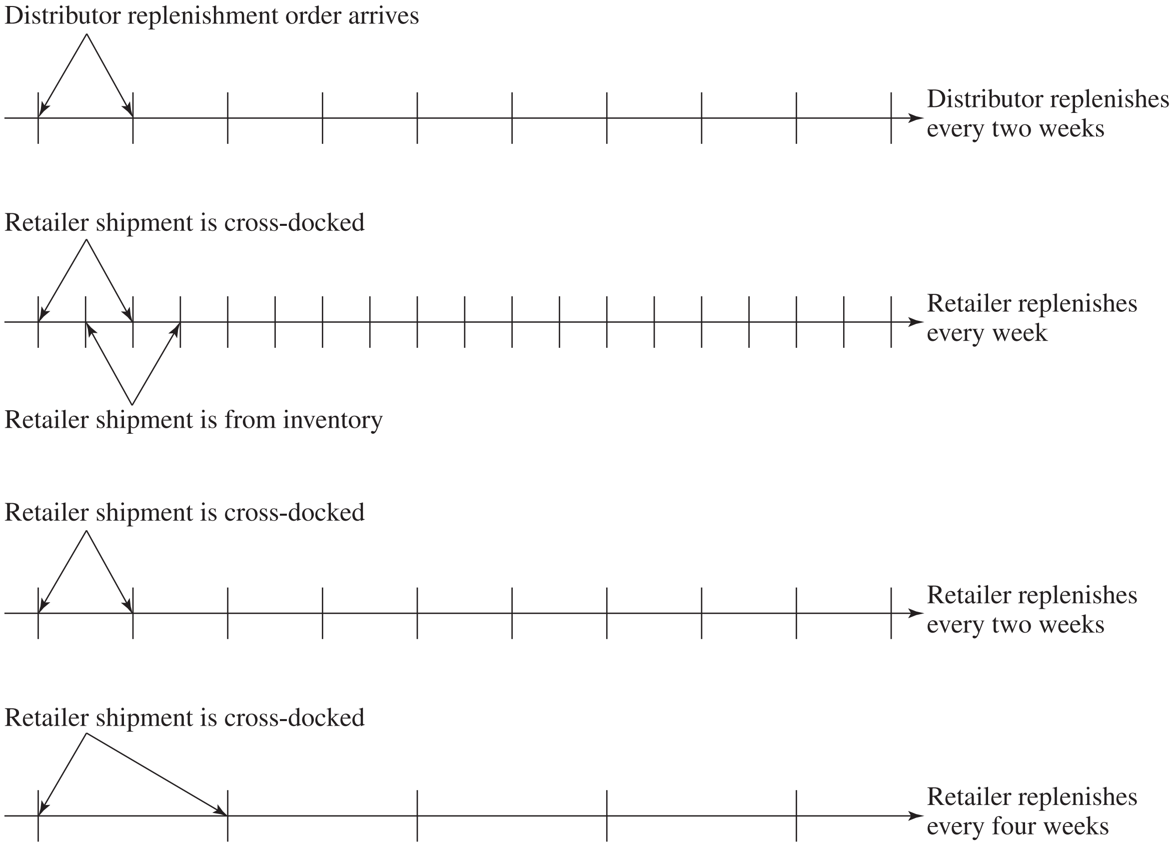


FIGURE 11-7 Illustration of an Integer Replenishment Policy

Integer replenishment policies for the supply chain shown in Figure 11-8 can be summarized as follows:

- Divide all parties within a stage into groups such that all parties within a group order from the same supplier and have the same reorder interval.
- Set reorder intervals across stages such that the receipt of a replenishment order at any stage is synchronized with the shipment of a replenishment order to at least one customer. The synchronized portion can be cross-docked.
- For customers with a longer reorder interval than the supplier, make the customer's reorder interval an integer multiple of the supplier's interval and synchronize replenishment at the two stages to facilitate cross-docking. In other words, a supplier should cross-dock all orders from customers that reorder less frequently than the supplier.

Key Point

Integer replenishment policies can be synchronized in multiechelon supply chains to keep cycle inventory and order costs low. Under such policies, the reorder interval at any stage is an integer multiple of a base reorder interval. Synchronized integer replenishment policies facilitate a high level of cross-docking across the supply chain.

- For customers with a shorter reorder interval than the supplier, make the supplier's reorder interval an integer multiple of the customer's interval and synchronize replenishment at the

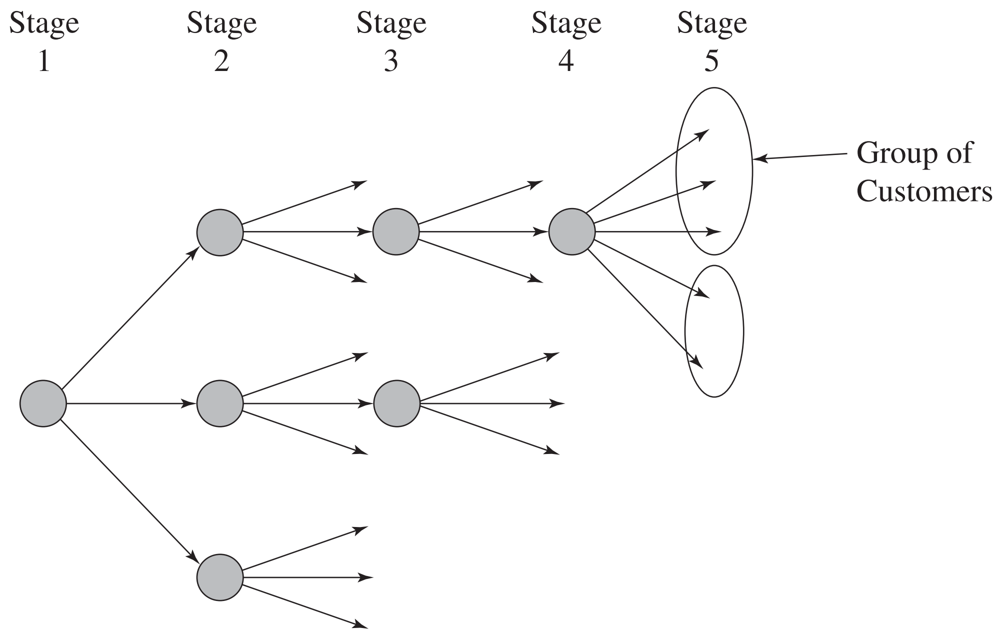


FIGURE 11-8 A Multiechelon Distribution Supply Chain

two stages to facilitate cross-docking. In other words, a supplier should cross-dock one out of every *k* shipments to a customer that orders more frequently than the supplier, where *k* is an integer.

- The relative frequency of reordering depends on the setup cost, holding cost, and demand at different parties.

Although the integer policies discussed above synchronize replenishment within the supply chain and decrease cycle inventories, they increase safety inventories, because of the lack of flexibility with the timing of a reorder, as discussed in Chapter 12. Thus, these polices make the most sense for supply chains in which cycle inventories are large and demand is relatively predictable.

11.8 SUMMARY OF LEARNING OBJECTIVES

1. ***Balance the appropriate costs to choose the optimal lot size and cycle inventory in a supply chain.*** Cycle inventory generally equals half the lot size. Therefore, as the lot size grows, so does the cycle inventory. In deciding on the optimal amount of cycle inventory, the supply chain goal is to minimize the total cost—the order cost, holding cost, and material cost. As cycle inventory increases, so does the holding cost. However, the order cost and, in some instances, the material cost decrease with an increase in lot size and cycle inventory. The EOQ balances the three costs to obtain the optimal lot size. The higher the order and transportation cost, the higher the lot size and cycle inventory.

2. ***Understand the impact of quantity discounts on lot size and cycle inventory.*** Lot-size–based quantity discounts increase the lot size and cycle inventory within the supply chain because they encourage buyers to purchase in larger quantities to take advantage of the decrease in price.

3. ***Devise appropriate discounting schemes for a supply chain.*** Quantity discounts are justified to increase total supply chain profits when independent lot-sizing decisions in a supply chain lead to suboptimal solutions from an overall supply chain perspective. If suppliers have large fixed costs, suitable lot-size–based quantity discounts can be justified because they help coordinate the supply chain. Volume-based discounts are more effective than lot-size–based discounts in increasing supply chain profits without increasing lot size and cycle inventory.

4. *Understand the impact of trade promotions on lot size and cycle inventory.* Trade promotions increase inventory and total supply chain costs through forward buying, which shifts future demand to the present and creates a spike in demand followed by a dip. The increased variability raises inventories and costs.

5. *Identify managerial levers that reduce lot size and cycle inventory in a supply chain without increasing cost.* The key managerial levers for reducing lot size, and thus cycle inventory, in the supply chain without increasing cost are the following:

- Reduce fixed ordering and transportation costs incurred per order.
- Implement volume-based discounting schemes rather than individual lot-size–based discounting schemes.
- Eliminate or reduce trade promotions and encourage EDLP. Base trade promotions on sell-through rather than sell-in to the retailer.

Discussion Questions

1. Consider a supermarket deciding on the size of its replenishment order from Procter & Gamble. What costs should it take into account when making this decision?
2. Discuss how various costs for the supermarket in Question 1 change as it decreases the lot size ordered from Procter & Gamble.
3. As demand at the supermarket chain in Question 1 grows, how would you expect the cycle inventory measured in days of inventory to change? Explain.
4. The manager at the supermarket in Question 1 wants to decrease the lot size without increasing the costs he incurs. What actions can he take to achieve this objective?
5. "Though owned by different parties, members of the supply chain should coordinate with each other instead of pursuing their own objectives." Do you agree or disagree? Why?
6. Should a commodity product manufacturer provide lot size-based or volume-based quantity discounts in order to maximize total supply chain profits?
7. What is the difference between lot-size–based and volume-based quantity discounts?
8. Do you think integer replenishment policies should be synchronized in multi-echelon supply chains? Why?
9. Why is it appropriate to include only the incremental cost when estimating the holding and order cost for a firm?

Exercises

1. Harley Davidson has its engine assembly plant in Milwaukee and its motorcycle assembly plant in Pennsylvania. Engines are transported between the two plants using trucks. Each truck trip costs $1,500. The motorcycle plant assembles and sells 300 motorcycles each day. Each engine costs $450 and Harley incurs a holding cost of 20 percent per year. How many engines should Harley load onto each truck? What is the cycle inventory of engines at Harley?
2. Harley has decided to implement just-in-time (JIT) at the motorcycle assembly plant. As part of this initiative it has reduced the number of engines loaded on each truck to 100. If each truck trip still costs $1,500, how does this decision impact annual costs at Harley? What should the cost of each truck be if a load of 100 engines is to be optimal for Harley?
3. A North Face retail store in Chicago sells 500 jackets each month. Each jacket costs the store $100 and the company has an annual holding cost of 25 percent. The fixed cost of a replenishment order (including transportation) is $100. The store currently places a replenishment order every month for 500 jackets. What is the annual holding and ordering cost? On average, how long does a jacket spend in inventory? If the retail store wants to minimize ordering and holding cost, what order size do you recommend? How much would the optimal order reduce holding and ordering cost relative to the current policy?
4. Target purchases home goods made by a supplier in China. Target's stores in the United States sell 200,000 units of home goods each month. Each unit costs $10 and the company has an annual holding cost of 20 percent. Placing a replenishment order incurs clerical costs of $500/order. The shipping company charges $5,000 as a fixed cost per shipment along with a variable cost of $0.10 per unit shipped. What is the optimal order size for Target? What is the annual holding cost of the optimal policy? How many orders per year does Target place? What is the annual fixed transportation cost? What is the annual variable transportation cost? What is the annual clerical cost?
5. Amazon sells 20,000 units of consumer electronics from Samsung every month. Each unit costs $100 and Amazon has

a holding cost of 20 percent. The fixed clerical and transportation cost for each order Amazon places with Samsung is $4,000. What is the optimal size of the order that Amazon should place with Samsung? With the goal of reducing inventories, Amazon would like to reduce the size of each order it places with Samsung to 2,500 units (allowing it to get four replenishment orders every month). How much should it reduce the fixed cost per order for an order of 2,500 units to be optimal?

6. Amazon sells 10,000 Lenovo PCs every month. Each PC costs $500 and Amazon has a holding cost of 20 percent. For what fixed cost per order would an order size of 10,000 units be optimal? For what fixed cost per order would an order size of 2,500 units be optimal?

7. A steel rolling mill can produce I-beams at the rate of 20 tons per week. Customer demand for the beams is 5 tons per week. To produce I-beams, the mill must go through a setup that requires changing to the appropriate rolling patterns. Each setup costs the mill $10,000 in labor and lost production. I-beams cost the mill $2,000 per ton and the mill has a holding cost of 25 percent. What is the optimal production batch size for I-beams? What is the annual setup cost of the optimal policy? What is the annual holding cost?

8. A steel rolling mill can produce I-beams at the rate of 20 tons per week. Customer demand for the beams is 5 tons per week. I-beams cost the mill $2,000 per ton and the mill has a holding cost of 25 percent. To produce I-beams, the mill must go through a setup that requires changing to the appropriate rolling patterns. The mill would like to produce I-beams in batches of 40 tons (resulting in a production batch every eight weeks). For what changeover cost would this batch size be optimal?

9. An electronics company has two contract manufacturers in Asia: Foxconn assembles its tablets and smart phones and Flextronics assembles its laptops. Monthly demand for tablets and smartphones is 10,000 units, whereas that for laptops is 4,000. Tablets cost the company $100, laptops cost $400, and the company has a holding cost of 25 percent. Currently the company has to place separate orders with Foxconn and Flextronics and receives separate shipments. The fixed cost of each shipment is $10,000. What is the optimal order size and order frequency with each of Foxconn and Flextronics?

The company is thinking of combining all assembly with the same contract manufacturer. This will allow for a single shipment of all products from Asia. If the fixed cost of each shipment remains $10,000, what is the optimal order frequency and order size from the combined orders? How much reduction in cycle inventory can the company expect as a result of combining orders and shipments?

10. Harley purchases components from three suppliers. Components purchased from Supplier A are priced at $5 each and used at the rate of 20,000 units per month. Components purchased from Supplier B are priced at $4 each and are used at the rate of 2,500 units per month. Components purchased from Supplier C are priced at $5 each and used at the rate of 900 units per month. Currently, Harley purchases a separate truckload from each supplier. As part of its JIT drive, Harley has decided to aggregate purchases from the three suppliers. The trucking company charges a fixed cost of $400 for the truck with an additional charge of $100 for each stop. Thus, if Harley asks for a pickup from only one supplier, the trucking company charges $500; from two suppliers, it charges $600; and from three suppliers, it charges $700. Suggest a replenishment strategy for Harley that minimizes annual cost. Assume a holding cost of 20 percent per year. Compare the cost of your strategy with Harley's current strategy of ordering separately from each supplier. What is the cycle inventory of each component at Harley?

11. Ford and GM carry spare parts for their dealers at a third-party warehouse in Michigan's Upper Peninsula. Demand for Ford spare parts is 100 units per month, whereas demand for GM parts is 120 per month. Each spare part costs $100 and both companies have a holding cost of 20 percent. Currently, each company uses a separate truck to ship these parts. Each truck has a fixed cost of $500. What is the optimal order size and frequency for Ford? For GM? What is the annual ordering and holding cost for each company?

A third-party logistics provider has offered to combine shipments for each of the two companies on a single truck. This will increase the cost of each truck to $600. If the two companies agree to the joint shipment, what is the optimal order frequency and size? What is the annual ordering and holding cost for the two companies combined? Should Ford and GM accept the third party's proposal? How should they divide the fixed cost per truck among themselves?

12. Prefab, a furniture manufacturer, uses 20,000 square feet of plywood per month. Its trucking company charges Prefab $400 per shipment, independent of the quantity purchased. The manufacturer offers an all unit quantity discount with a price of $1 per square foot for orders under 20,000 square feet, $0.98 per square foot for orders between 20,000 square feet and 40,000 square feet, and $0.96 per square foot for orders larger than 40,000 square feet. Prefab incurs a holding cost of 20 percent. What is the optimal lot size for Prefab? What is the annual cost of such a policy? What is the cycle inventory of plywood at Prefab? How does it compare with the cycle inventory if the manufacturer does not offer a quantity discount but sells all plywood at $0.96 per square foot?

Now consider the case in which the manufacturer offers a marginal unit quantity discount for the plywood. The first 20,000 square feet of any order are sold at $1 per square foot, the next 20,000 square feet are sold at $0.98 per square foot, and any quantity larger than 40,000 square feet is sold for $0.96 per square foot. What is the optimal lot size for Prefab given this pricing structure? How much cycle inventory of plywood will Prefab carry given the ordering policy?

13. Demand for fasteners at W.W. Grainger is 20,000 boxes per month. The holding cost at Grainger is 20 percent per year. Each order incurs a fixed cost of $400. The supplier offers an all unit discount pricing scheme with a price of $5 per box for orders under 30,000 and a price of $4.90 for all orders of 30,000 or more. How many boxes should Grainger order per replenishment?

14. Now, consider Exercise 13 with a marginal unit quantity discount. Demand for fasteners at W.W. Grainger is 20,000 boxes per month. The holding cost at Grainger is 20 percent per year. Each order incurs a fixed cost of $400. The supplier offers a marginal unit discount pricing scheme with a price of $5 per box for the first 30,000 and a price of $4.90 per unit for each unit above 30,000 in an order. How many boxes should Grainger order per replenishment?

15. Demand for phones at Amazon is 5,000 per month. The holding cost at Amazon is 25 percent and the company incurs a fixed cost of $500 for each order placed. The supplier offers an all unit quantity discount with a price of $200 per phone for all orders under 10,000, a price of $195 for all orders of 10,000 or more but under 20,000 and a price of $190 for all orders of 20,000 or more. How many phones should Amazon order per replenishment?

16. Demand for phones at Amazon is 5,000 per month. The holding cost at Amazon is 25 percent and the company incurs a fixed cost of $500 for each order placed. The supplier offers a marginal unit quantity discount with a price of $200 per phone for the first 10,000 phones in an order, a price of $195 for the next 10,000 phones in the order, and a price of $190 for the quantity above 20,000 in the order. How many phones should Amazon order per replenishment?

17. Dominick's supermarket chain sells Nut Flakes, a popular cereal manufactured by the Tastee cereal company. Demand for Nut Flakes is 1,000 boxes per week. Dominick's has a holding cost of 25 percent and incurs a fixed trucking cost of $200 for each replenishment order it places with Tastee. Given that Tastee normally charges $2 per box of Nut Flakes, how much should Dominick's order in each replenishment lot?

Tastee runs a trade promotion for a month, lowering the price of Nut Flakes to $1.80. How much should Dominick's order, given the short-term price reduction?

18. Flanger is an industrial distributor that sources from hundreds of suppliers. The two modes of transportation available for inbound shipping are LTL (less than truckload) and TL (truckload). LTL shipping costs $1 per unit, whereas TL shipping costs $500 per truck. Each truck can carry up to 1,000 units. Flanger wants a rule assigning products to shipping mode (TL or LTL) based on annual demand. Each unit costs $50, and Flanger uses a holding cost of 22 percent. Flanger incurs a fixed cost of $150 for each order placed with a supplier.

a. Determine a threshold for annual demand above which TL is preferred and below which LTL is preferred.
b. How does the threshold change [relative to part (a)] if unit cost is $150 (instead of $50) with all other data unchanged? Which mode becomes preferable as unit cost grows?
c. How does the threshold change [relative to part (a)] if the LTL cost comes down to $0.8 per unit (instead of $1 per unit)?

19. SuperPart, an auto parts distributor, has a large warehouse in the Chicago region and is deciding on a policy for the use of TL or LTL transportation for inbound shipping. LTL shipping costs $1 per unit. TL shipping costs $850 per truck plus $150 per pickup. Thus, a truck used to pick up from three suppliers costs $850 + (3 \times 100) = \$1{,}150$. A truck can carry up to 2,000 units. SuperPart incurs a fixed cost of $150 for each order placed with a supplier. Thus, an order with three distinct suppliers incurs an ordering cost of $450. Each unit costs $50, and SuperPart uses a holding cost of 23 percent. Assume that product from each supplier has an annual demand of 3,000 units. SuperPart has thousands of suppliers and the company must decide on the number of suppliers to group per truck if using TL.

a. What is the optimal order size and annual cost if LTL shipping is used? What is the time between orders?
b. What is the optimal order size and annual cost if TL shipping is used with a separate truck for each supplier? What is the time between orders?
c. What is the optimal order size and annual cost per product if TL shipping is used but two suppliers are grouped together per truck?
d. What is the optimal number of suppliers that should be grouped together? What is the optimal order size and annual cost per product in this case? What is the time between orders?
e. Which shipping policy would you recommend if each product has an annual demand of 3,000? Which shipping policy would you recommend for products with an annual demand of 1,500? Which shipping policy would you recommend for products with an annual demand of 18,000?

20. PlasFib is a manufacturer of synthetic fibers used for making furniture upholstery. PlasFib manufactures fiber in 50 colors on one line. When changing over from one color to the next, part of the line must be cleaned, leading to a loss of material. Each changeover costs $120 in lost material and changeover labor. Assume that each changeover requires the line to shut down for 0.5 hour. When it is running, the line produces fiber at the rate of 100 pounds per hour.

The fibers sold by PlasFib are divided into three categories. There are 5 fast-moving colors that average sales of 31,000 pounds per color per year. There are 10 medium-moving colors that average sales of 12,000 pounds per color per year. The remaining are slow-moving products and average sales of 3,200 pounds per year each. Each pound of fiber costs $6 and PlasFib has a holding cost of 15 percent.

a. What is the batch size that PlasFib should produce for each fast-, medium-, and slow-moving color? How many days of demand does this translate into?
b. What is the annual setup and holding cost of the policies you suggested in part (a)?
c. How many hours of plant operation will the above schedule require in a year (including a half-hour of setup per batch)?

21. TopOil, a refiner in Indiana, serves three customers near Nashville, Tennessee, and maintains consignment inventory (owned by TopOil) at each location. Currently, TopOil uses TL transportation to deliver separately to each customer. Each truck costs $800 plus $250 per stop. Thus, delivering to each customer separately costs $1,050 per truck. TopOil is considering aggregating deliveries to Nashville on a single truck. Demand at the large customer is 60 tons a year, demand at the medium customer is 24 tons per year, and demand at the small customer is 8 tons per year. Product cost for TopOil is $10,000 per ton, and it uses a holding cost of 25 percent. Truck capacity is 12 tons.

a. What is the annual transportation and holding cost if TopOil ships a full truckload each time a customer is running out of stock? How many days of inventory is carried at each customer under this policy?
b. What is the optimal delivery policy to each customer if TopOil ships separately to each of them? What is the annual transportation and holding cost? How many days of inventory is carried at each customer under this policy?
c. What is the optimal delivery policy to each customer if TopOil aggregates shipments to each of the three customers on every truck that goes to Nashville? What is the annual transportation and holding cost? How many days of inventory are carried at each customer under this policy?
d. Can you come up with a tailored policy that has lower costs than the policies in (b) or (c)? What are the costs and inventories for your suggested policy?

22. Crunchy, a cereal manufacturer, has dedicated a plant for one major retail chain. Sales at the retail chain average about 20,000 boxes a month and production at the plant keeps pace with this average demand. Each box of cereal costs Crunchy $3 and is sold to the retailer at a wholesale price of $5. Both Crunchy and the retailer use a holding cost of 20 percent. For each order placed, the retailer incurs an ordering cost of $200. Crunchy incurs the cost of transportation and loading that totals $1,000 per order shipped.

a. Given that it is trying to minimize its ordering and holding costs, what lot size will the retailer ask for in each order? What are the annual ordering and holding costs for the retailer as a result of this policy? What are the annual ordering and holding costs for Crunchy as a result of this policy? What is the total inventory cost across both parties as a result of this policy?
b. What lot size minimizes the inventory costs (ordering, delivery, and holding) across both Crunchy and the retailer? How much reduction in cost relative to (a) results from this policy?
c. Design an all unit quantity discount that results in the retailer ordering the quantity in (b).
d. How much of the $1,000 delivery cost should Crunchy pass along to the retailer for each lot to get the retailer to order the quantity in (b)?

23. A steel service center sources products from an integrated steel mill at a cost of $2,000 per ton. Demand for steel at the service center is 50 tons per month. The service center has a holding cost of 25 percent and incurs a fixed cost of $2,000 for each order. How many tons of steel should the service center order per replenishment? What is the annual ordering and holding cost incurred by the service center?

The integrated steel mill incurs a fixed cost of $4,000 for each order placed by the steel service center. Steel costs the mill $1,000 per ton and the mill has a holding cost of 20 percent. Assuming that the mill builds up its steel (for the service center) at the rate of 50 tons per month, what are the annual fixed cost and holding cost incurred by the mill as a result of the service center's ordering policy? What is the annual cost incurred by both the service center and the steel mill?

If the steel mill and the service center could work in a coordinated manner, what is the optimal order size that minimizes their joint fixed and holding costs? What annual savings could the supply chain expect as a result of coordination? Design an all unit quantity discount that the integrated steel mill could use to get the service center to order the coordinated amount without increasing annual costs at the service center.

24. The Orange company has introduced a new music device called the J-Pod. The J-Pod is sold through Good Buy, a major electronics retailer. Good Buy has estimated that demand for the J-Pod will depend on the final retail price p according to the demand curve

$$\text{Demand } D = 2{,}000{,}000 - 2{,}000p$$

The production cost for Orange is $100 per J-Pod.

a. What wholesale price should Orange charge for the J-Pod? At this wholesale price, what retail price should Good Buy set? What are the profits for Orange and Good Buy at equilibrium?
b. If Orange decides to discount the wholesale price by $40, how much of a discount should Good Buy offer to customers if it wants to maximize its own profits? What fraction of the discount offered by Orange does Good Buy pass along to the customer?

25. The Orange company prices J-Pods at $550 per unit. Good Buy sells the J-Pods at $775. Annual demand at this retail price turns out to be 450,000 units. Good Buy incurs ordering, receiving, and transportation costs of $10,000 for each lot of J-Pods ordered. The holding cost used by the retailer is 20 percent.

a. What is the optimal lot size that Good Buy should order?
b. The Orange company has discounted J-Pods by $40 for the short term (about the next two weeks). Good Buy has decided not to change the retail price but may change the lot size ordered from Orange. How should Good Buy adjust its lot size given this discount? How much does the lot size increase because of the discount?

Bibliography

Blattberg, Robert C., and Scott A. Neslin. *Sales Promotion: Concepts, Methods, and Strategies*. Upper Saddle River, NJ: Prentice Hall, 1990.

Brealey, Richard A., and Stewart C. Myers. *Principles of Corporate Finance*. Boston, MA: Irwin McGraw-Hill, 2000.

Buzzell, Robert, John Quelch, and Walter Salmon. "The Costly Bargain of Trade Promotions." *Harvard Business Review* (March–April 1990): 141–149.

Crowther, John F. "Rationale for Quantity Discounts." *Harvard Business Review* (March–April 1964): 121–127.

Dolan, Robert J. "Quantity Discounts: Managerial Issues and Research Opportunities." *Marketing Science* (1987): 6, 1–24.

Federgruen, Awi, and Yu-Sheng Zheng. "Optimal Power-of-Two Replenishment Strategies in Capacitated General Production/Distribution Networks." *Management Science* (1993): 39, 710–727.

Goyal, Suresh K. "A Simple Procedure for Price Break Models." *Production Planning & Control* (1995): 6, 584–585.

Hu, Jianli, and Charles L. Munson. "Dynamic Demand Lot-sizing Rules for Incremental Quantity Discounts." *Journal of the Operational Research Society* (2002): 53, 855–863.

Kurt Salmon Associates, Inc. *Efficient Consumer Response*. Washington, DC: Food Marketing Institute, 1993.

Lee, Hau L., and Corey Billington. "Managing Supply Chain Inventories: Pitfalls and Opportunities." *Sloan Management Review* (Spring 1992): 65–73.

Maxwell, William L., and John A. Muckstadt. "Establishing Consistent and Realistic Reorder Intervals in Production-Distribution Systems." *Operations Research* (1985): 33, 1316–1341.

Munson, Charles L., Jianli Hu, and Meir J. Rosenblatt. "Teaching the Costs of Uncoordinated Supply Chains." *Interfaces* (2003): 33, 24–39.

Munson, Charles L., and Meir J. Rosenblatt. "Theories and Realities of Quantity Discounts." *Production and Operations Management* (1998): 7, 352–369.

Roundy, Robin. "98%-Effective Integer-Ratio Lot-Sizing for One-Warehouse Multi-Retailer Systems." *Management Science* (1985): 31, 1416–1429.

Roundy, Robin. "A 98%-Effective Lot-Sizing Rule for a Multi-Product, Multi-Stage Production Inventory System." *Mathematics of Operations Research* (1986): 11, 699–727.

Silver, Edward A., David Pyke, and Rein Petersen. *Inventory Management and Production Planning and Scheduling*. New York: Wiley, 1998.

Zipkin, Paul H. *Foundations of Inventory Management*. Boston, MA: Irwin McGraw-Hill, 2000.

CASE STUDY

Delivery Strategy at MoonChem

John Kresge, vice president of supply chain, was very concerned as he left the meeting at MoonChem, a manufacturer of specialty chemicals. The year-end meeting evaluated financial performance and discussed the fact that the firm was achieving only two inventory turns a year. A more careful look revealed that more than half the inventory MoonChem owned was in consignment with its customers. This was very surprising, given that only 20 percent of its customers carried consignment inventory. John was responsible for inventory as well as transportation costs. He decided to take a careful look at the management of consignment inventory and come up with an appropriate plan.

MoonChem Operations

MoonChem, a manufacturer of specialty chemicals, had eight manufacturing plants and 40 distribution centers. The plants manufactured the base chemicals, and the distribution centers mixed them to produce hundreds of end products that fit customer specifications. In the specialty chemicals market, MoonChem decided to differentiate itself in the Midwest region by providing consignment inventory to its customers. The company wanted to take this strategy national if it proved effective. MoonChem kept the chemicals required by each customer in the Midwest region on consignment at the customers' sites. Customers used the chemicals as needed, and MoonChem managed replenishment to ensure availability. In most instances, consumption of chemicals by customers was stable. MoonChem owned the consignment inventories and was paid for the chemicals as they were used.

Distribution at MoonChem

MoonChem used Golden trucking, a full-truckload carrier, for all its shipments. Each truck had a capacity of 40,000 pounds; Golden charged a fixed rate given the origin and destination, regardless of the quantity shipped on the truck. MoonChem sent full truckloads to each customer to replenish its consignment inventory.

The Illinois Pilot Study

John decided to take a careful look at his distribution operations. He focused on Illinois, which was supplied from the Chicago distribution center. He broke up Illinois into a collection of zip codes that were contiguous, as shown in Figure 11-9. He restricted attention to the Peoria region, which was classified as zip code 615. A careful study of the Peoria region revealed two large customers, six medium-sized customers, and twelve small customers. The annual consumption at each type of customer was as shown in Table 11-4. Golden charged $400 for each shipment from Chicago to Peoria, and MoonChem's policy was to send a full truckload to each customer as needed.

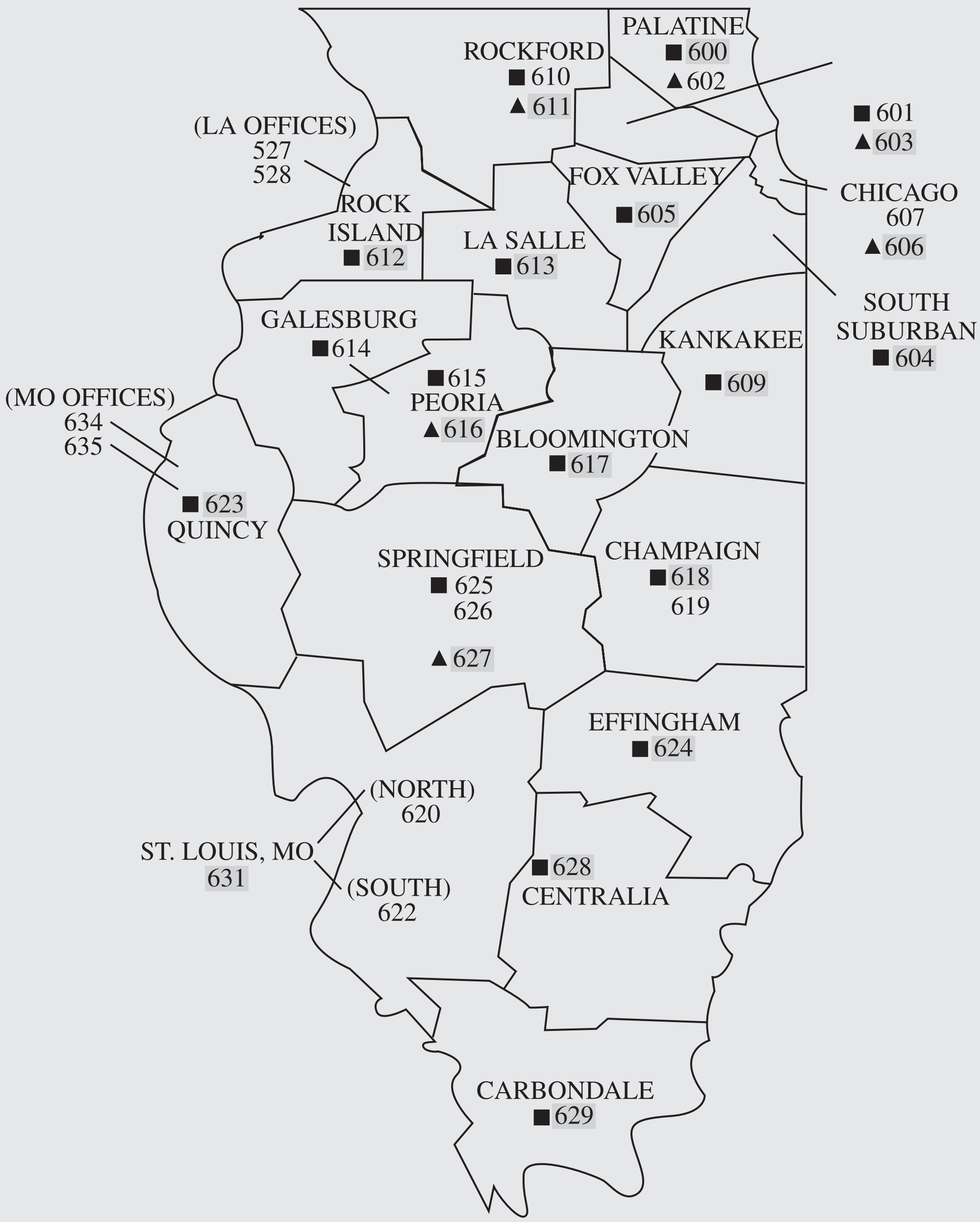


FIGURE 11-9 Illinois Zip Code Map

TABLE 11-4 Customer Profile for MoonChem in Peoria Region

Customer Type	Number of Customers	Consumption (Pounds per Month)
Small	12	1,000
Medium	6	6,000
Large	2	12,000

John checked with Golden to find out what it would take to include shipments for multiple customers on a single load. Golden informed him that it would charge $350 per truck and add $50 for each drop-off for which Golden was responsible. Thus, if Golden carried a truck that had to make one delivery, the total charge would be $400. However, if a truck had to make four deliveries, the total charge would be $550.

Each pound of chemical in consignment cost MoonChem $1, and MoonChem had a holding cost of 25 percent. John wanted to analyze a few different options for distribution available in the Peoria region to decide on the optimal distribution policy. One was to aggregate all 20 customers into each truck going to Peoria. The other was to separate the 20 customers into two groups with one large, three medium, and six small customers in each group. Each group would then be aggregated into a single truck going to Peoria. The detailed study of the Peoria region would provide the blueprint for the distribution strategy that MoonChem planned to roll out nationally.

Questions

1. What is the annual cost of MoonChem's strategy of sending full truckloads to each customer in the Peoria region to replenish consignment inventory?
2. Consider different delivery options and evaluate the cost of each. What delivery option do you recommend for MoonChem?
3. How does your recommendation impact consignment inventory for MoonChem?

CASE STUDY

NAN

NZ All Natural Ice Cream (NAN) began in 1984 as a small shop in Christchurch, which is the largest city in the South Island of New Zealand. With NAN specializing in producing ice cream with all natural ingredients without any artificial colors or flavors, it is highly regarded by the public concerned about health, and NAN's sales volume has increased significantly over the years. NAN continues to diversify its product range to include nondairy, low-fat sorbets; low-fat frozen yogurts; and premium ice cream with a wide range of flavors. In 2010, it dominated the whole New Zealand premium grocery ice cream category with 7 out of 10 top-selling SKUs. It has more than 100 employees, and its products have been exported to more than 30 countries.

Since Auckland, which is in the North Island of New Zealand, is geographically closer to the primary fresh milk and cream supply of the Waikato (heartland of various free-range dairy farms), NAN decides to set up its own production plants in Auckland to keep production near the supply of raw materials so that the manufacturing process could be smoother and more secured. Currently, three plants and five distribution centers are dispersed over the country. The plants generally manufacture the base products, and the distribution centers produce hundreds of end products that fit customer specifications. To further expand its domestic market, NAN considered providing consignment inventory to its North Island's customers and implementing this strategy throughout the country if it proved effective. NAN would keep the ice cream products required by each customer in the North Island on consignment at the customers' sites. Customers would sell the ice cream as demanded, and NAN would replenish regularly to ensure sufficient product supply. According to previous sales data, the consumption of ice cream by customers tended to be stable. NAN owned the consignment inventories and got paid when the ice cream was consumed.

NAN assigned all its inland replenishment function to a 3PL company named SuperTruck through open bidding. SuperTruck is famous for its fast response and prompt delivery through expertise in fleet management. All of its trucks have refrigerated units built either directly on the frame or transported by trailer and are powered by diesel-operated generators. SuperTruck has to fulfill every order NAN places within 24 hours or reimburse NAN for all damages in case of a failure to do so. Each truck has a capacity of 40,000 liters, and SuperTruck has been contracted to charge a fixed flat rate given the origin and destination regardless of the quantity loaded. NAN usually sends full truckloads (TL) to each customer to replenish its consignment inventory.

Katy Leung has recently been employed by NAN as North Island regional supply chain manager to oversee logistics operations in the region. Katy is responsible for inventory as well as transportation costs and decides to review the current operations for possible cost savings. In particular, Katy wonders if NAN should include multiple shipments for different customers on a single

TABLE 11-5 Customer Profile for NAN in Wellington

Customer Type	Number of Customers	Consumption (Liters/month)
Small	12	1,000
Medium	6	6,000
Large	2	12,000

truckload. Table 11-4 shows the customer profile in Wellington—one of the distribution centers.

SuperTruck charges \$400 for each truckload from the Wellington DC to each customer based on NAN's policy to send a full truckload as needed. In response to Katy's recent enquiry, SuperTruck indicated that it would charge \$350 per truck plus \$50 for each drop-off location. That is, if a truck made only one delivery as existing practice, the total charge would be \$350 + \$50 = \$400. However, if it had to make three extra deliveries (total of 4 locations), the total charge would then be \$350 + \$200 = \$550.

Each liter of ice cream product in consignment costs NAN \$1, while the holding cost is maintained at 25 percent. Consider the different distribution possibilities in the Wellington example and recommend the best option. This will help shape the future marketing and distribution strategy that NAN plans to roll out throughout all of New Zealand.

Questions

1. Calculate the annual cost of NAN's strategy of sending only full truckloads to each customer in Wellington to replenish consignment inventory.
2. Evaluate the costs for different delivery options and recommend the best option that NAN should adopt.
3. How does your recommendation affect NAN's consignment inventory?

APPENDIX 11A

Economic Order Quantity

Objective

Derive the economic order quantity (EOQ) formula.

Analysis:

Given an annual demand D, order cost S, unit cost C, and holding cost h, our goal is to estimate the lot size Q that minimizes the total annual cost. For a lot size of Q, the total annual cost is given by

$$\text{Total annual cost, } TC = (D/Q)S + (Q/2)\,hC + CD$$

To minimize the total cost, we take the first derivative with respect to the lot size Q and set it to zero. Taking the first derivative with respect to Q, we have

$$\frac{d(TC)}{dQ} = -\frac{DS}{Q^2} + \frac{hC}{2}$$

Setting the first derivative to be zero, the EOQ is given by

$$Q^2 = \frac{2DS}{hC} \quad or \quad Q = \sqrt{\frac{2DS}{hC}}$$

CHAPTER

12

Managing Uncertainty in a Supply Chain

Safety Inventory

LEARNING OBJECTIVES

After reading this chapter, you will be able to

1. Describe different measures of product availability.
2. Understand the role of safety inventory in a supply chain.
3. Identify factors that influence the required level of safety inventory.
4. Use available managerial levers to lower safety inventory without hurting product availability.

In this chapter, we discuss how safety inventory can help a supply chain improve product availability in the presence of supply and demand variability. We discuss various measures of product availability and how managers can set safety inventory levels to provide the desired product availability. We also explore what managers can do to reduce the amount of safety inventory required while maintaining or even improving product availability.

12.1 THE ROLE OF SAFETY INVENTORY IN A SUPPLY CHAIN

Safety inventory is inventory carried to satisfy demand that exceeds the amount forecast. Safety inventory is required because demand is uncertain, and a product shortage may result if actual demand exceeds the forecast demand. Consider, for example, Bloomingdale's, a high-end department store. Bloomingdale's sells purses purchased from Gucci, an Italian manufacturer. Given the high transportation cost from Italy, the store manager at Bloomingdale's orders in lots of 600 purses. Demand for purses at Bloomingdale's averages 100 a week. Gucci takes three weeks to deliver the purses to Bloomingdale's in response to an order. If there is no demand uncertainty and exactly 100 purses are sold each week, the store manager at Bloomingdale's can place an order when the store has exactly 300 purses remaining. In the absence of demand uncertainty, such a policy ensures that the new lot arrives just as the last purse is being sold at the store.

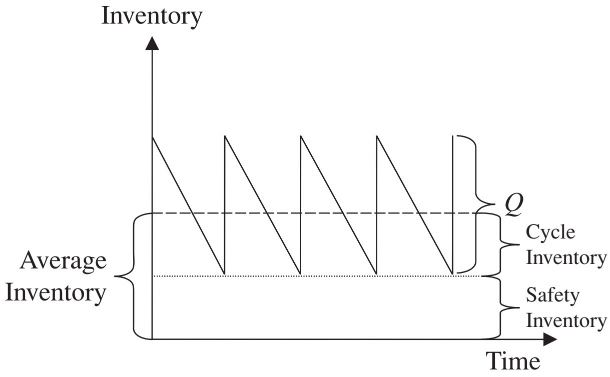


FIGURE 12-1 Inventory Profile with Safety Inventory

However, given demand fluctuations and forecast errors, actual demand over the three weeks may be higher or lower than the 300 purses that were forecast. If the actual demand at Bloomingdale's is higher than 300, some customers will be unable to purchase purses, resulting in a potential loss of margin for Bloomingdale's. The store manager thus decides to place an order with Gucci when the store still has 400 purses. This policy improves product availability for the customer because the store now runs out of purses only if the demand over the three weeks exceeds 400. Given an average weekly demand of 100 purses, the store will have an average of 100 purses remaining when the replenishment lot arrives. Safety inventory is the average inventory remaining when the replenishment lot arrives. Thus, Bloomingdale's carries a safety inventory of 100 purses.

Given a lot size of $Q = 600$ purses, the cycle inventory, the focus of the previous chapter, is $Q/2 = 300$ purses. The inventory profile at Bloomingdale's in the presence of safety inventory is shown in Figure 12-1, which illustrates that the average inventory at Bloomingdale's is the sum of the cycle and safety inventories.

This example illustrates a trade-off that a supply chain manager must consider when planning safety inventory. On one hand, raising the level of safety inventory increases product availability, and thus the margin captured from customer purchases. On the other hand, raising the level of safety inventory increases inventory holding costs. This issue is particularly significant in industries in which product life cycles are short and demand is volatile. Carrying excessive inventory can help counter demand volatility but can really hurt if new products come onto the market and demand for the product in inventory dries up. The inventory on hand then becomes worthless.

In today's business environment, it has become easier for customers to search across stores for product availability. If Amazon is out of a book, for example, a customer can easily check to see whether barnesandnoble.com has the title available. The increased ease of searching puts pressure on firms to improve product availability. Simultaneously, product variety has grown with increased customization. As a result, markets have become increasingly heterogeneous and demand for individual products is unstable and difficult to forecast. Both the increased variety and the greater pressure for availability push firms to raise the level of safety inventory they hold. Given the product variety and high demand uncertainty in most high-tech supply chains, a significant fraction of the inventory carried is safety inventory.

As product variety has grown, however, product life cycles have shrunk. Thus, it is more likely that a product that is "hot" today will be obsolete tomorrow, which increases the cost to firms of carrying too much inventory. Thus, a key to the success of any supply chain is to figure out ways to decrease the level of safety inventory carried without hurting the level of product availability.

The importance of reduced safety inventories is emphasized by the experience of Nordstrom, Macy's, and Saks during the 2008–2009 recession. Nordstrom outperformed the other two chains by moving its inventories about twice as fast as its competitors. In 2008 (2009), Nordstrom carried an average of about 2 (2) months, Macy's carried about 4 (4.15) months, and Saks carried about 4.24 (4.67) months of inventory. A key to Nordstrom's success has been its ability to provide a high level of product availability to customers while carrying low levels of safety inventory in its supply chain. This fact has also played an important role in the success of Zara, Walmart, and Seven-Eleven Japan.

For any supply chain, three key questions need to be considered when planning safety inventory:

1. What is the appropriate level of product availability?
2. How much safety inventory is needed for the desired level of product availability?
3. What actions can be taken to reduce safety inventory without hurting product availability?

The first question is discussed in detail in Chapter 13. The remainder of this chapter focuses on answering the second and third questions, assuming a desired level of product availability. Next, we consider factors that influence the appropriate level of safety inventory.

12.2 FACTORS AFFECTING THE LEVEL OF SAFETY INVENTORY

The appropriate level of safety inventory is determined by the following two factors:

- The uncertainty of both demand and supply
- The desired level of product availability

As the uncertainty of supply or demand grows, the required level of safety inventories increases. Demand for milk at a supermarket is quite predictable. As a result, supermarkets can operate with low levels of safety inventory relative to demand. In contrast, demand for spices at the same supermarket is much harder to predict. Thus the supermarket needs to carry high levels of safety inventory for spices relative to demand. Whereas most of the milk inventory at a supermarket is cycle inventory (with very little being safety inventory), most of the spice inventory is safety inventory carried to deal with uncertainty of demand.

As the desired level of product availability increases, the required level of safety inventory also increases. If the supermarket targets a higher level of product availability for a certain spice, it must carry a higher level of safety inventory for that spice.

Next, we discuss some measures of demand uncertainty.

Measuring Demand Uncertainty

As discussed in Chapter 7, demand has a systematic as well as a random component. The random component is a measure of demand uncertainty. The goal of forecasting is to predict the systematic component and estimate the random component. The random component is usually estimated as the standard deviation of forecast error. We illustrate our ideas using uncertain demand for a smartphone at B&M Office Supplies as the context. We assume that periodic demand for the phone at B&M is normally distributed with the following inputs:

D: Average demand per period

σ_D: Standard deviation of demand (forecast error) per period

Even though standard deviation of demand is not necessarily the same as forecast error, we treat the two to be interchangeable in our discussion. Safety inventory calculations should really be based on forecast error.

Lead time is the gap between the time an order is placed and when it is received. In our discussion, we denote the lead time by L. In the B&M example, L is the time between when B&M orders phones and when they are delivered. In this case, B&M is exposed to the uncertainty of demand during the lead time. Whether B&M is able to satisfy all demand from inventory depends on the demand for phones experienced during the lead time and the inventory B&M has when a replenishment order is placed. Thus, B&M must estimate the uncertainty of demand during the lead time, not just in a single period. We now evaluate the distribution of demand over L periods, given the distribution of demand during each period.

EVALUATING DEMAND DISTRIBUTION OVER *L* PERIODS Assume that demand for each period i, $i = 1, \ldots, L$, is normally distributed with a mean D_i and standard deviation σ_i. Let ρ_{ij} be the correlation coefficient of demand between periods i and j. In this case, the total demand during L periods is normally distributed with a mean of D_L and a standard deviation of σ_L, where the following is true:

$$D_L = \sum_{i=1}^{L} D_i, \quad \sigma_L = \sqrt{\sum_{i=1}^{L} \sigma_i^2 + 2\sum_{i>j} \rho_{ij}\sigma_i\sigma_j} \qquad \textbf{(12.1)}$$

Demand in two periods is *perfectly positively correlated* if $\rho_{ij} = 1$. Demand in two periods is *perfectly negatively correlated* if $\rho_{ij} = -1$. Demand in two periods is *independent* if $\rho_{ij} = 0$. If demand during each of L periods is independent and normally distributed with a mean of D and a standard deviation of σ_D, Equation 12.1 can be used to show that total demand during the L periods is normally distributed with a mean D_L and a standard deviation of σ_L, where the following is true:

$$D_L = D \times L, \quad \sigma_L = \sqrt{L}\sigma_D \qquad \textbf{(12.2)}$$

Another important measure of uncertainty is the *coefficient of variation* (cv), which is the ratio of the standard deviation to the mean. Given demand with a mean of μ and a standard deviation of σ, we have

$$cv = \sigma / \mu$$

The coefficient of variation measures the size of the uncertainty relative to demand. It captures the fact that a product with a mean demand of 100 and a standard deviation of 100 has greater demand uncertainty than a product with a mean demand of 1,000 and a standard deviation of 100. Considering the standard deviation alone cannot capture this difference.

Next, we discuss some measures of product availability.

Measuring Product Availability

Product availability reflects a firm's ability to fill a customer order out of available inventory. A *stockout* results if a customer order arrives when product is not available. There are several ways to measure product availability. Some of the important measures are listed next.

1. *Product fill rate (fr)* is the fraction of product demand that is satisfied from product in inventory. Fill rate is equivalent to the probability that product demand is supplied from available inventory. Fill rate should be measured over specified amounts of demand rather than over time. Thus, it is more appropriate to measure fill rate over every million units of demand rather than every month. Assume that B&M provides smartphones to 90 percent of its customers from inventory, with the remaining 10 percent lost to a neighboring competitor because of a lack of available inventory. In this case, B&M achieves a fill rate of 90 percent.

2. *Order fill rate* is the fraction of orders that are filled from available inventory. Order fill rate should also be measured over a specified number of orders rather than over time. In a multiproduct scenario, an order is filled from inventory only if all products in the order can be supplied from the available inventory. In the case of B&M, a customer may order a phone along with a laptop. The order is filled from inventory only if both the phone and the laptop are available through the store. Order fill rates tend to be lower than product fill rates because all products must be in stock for an order to be filled.

3. *Cycle service level (CSL)* is the fraction of replenishment cycles that end with all the customer demand being met. A *replenishment cycle* is the interval between two successive replenishment deliveries. The CSL is equal to the probability of not having a stockout in a replenishment cycle. CSL should be measured over a specified number of replenishment cycles. If B&M orders replenishment lots of 600 phones, the interval between the arrival of two successive replenishment lots is a replenishment cycle. If the manager at B&M manages inventory such that the store does not run out of inventory in 6 out of 10 replenishment cycles, the store achieves a CSL of 0.6 or 60 percent. Observe that a CSL of 0.6 typically results in a much higher fill rate. In the 60 percent of cycles in which B&M does not run out of inventory, all customer demand is satisfied from available inventory. In the 40 percent of cycles in which a stockout does occur, most of the customer demand is satisfied from inventory. Only the small fraction toward the end of the cycle that arrives after B&M is out of inventory is lost. As a result, the fill rate is much higher than 0.6.

The distinction between product fill rate and order fill rate is usually not significant in a single-product situation. When a firm is selling multiple products, however, this difference may be significant. For example, if most orders include 10 or more products that are to be shipped, an out-of-stock situation of one product results in the order not being filled from stock. The firm in this case may have a poor order fill rate even though it has good product fill rates. Tracking order fill rates is important when customers place a high value on the entire order being filled at one time.

Next, we describe two replenishment policies that are often used in practice.

Replenishment Policies

A replenishment policy consists of decisions regarding when to reorder and how much to reorder. These decisions determine the cycle and safety inventories along with the fill rate fr and the cycle service level CSL. Replenishment policies may take any of several forms. We restrict attention to two types:

1. *Continuous review:* Inventory is continuously tracked, and an order for a lot size Q is placed when the inventory declines to the reorder point (ROP). As an example, consider the store manager at B&M who continuously tracks the inventory of phones. She orders 600 phones when the inventory drops below ROP $= 400$. In this case, the size of the order does not change from one order to the next. The time between orders may fluctuate, given variable demand.

2. *Periodic review:* Inventory status is checked at regular periodic intervals, and an order is placed to raise the inventory level to a specified threshold. As an example, consider the purchase of flash drives at B&M. The store manager does not track flash drive inventory continuously. Every Thursday, employees check flash drive inventory, and the manager orders enough so that the total of the available inventory and the size of the order equals 1,000 flash drives. In this case, the time between orders is fixed. The size of each order, however, can fluctuate given variable demand.

These inventory policies are not comprehensive, but they suffice to illustrate the key managerial issues concerning safety inventories.

12.3 DETERMINING THE APPROPRIATE LEVEL OF SAFETY INVENTORY

We now discuss the relationship between safety inventory and the CSL and fr. In this section, we restrict our attention to the continuous review policy. The periodic review policy is discussed in detail in Section 12.6. The continuous review policy consists of a lot size Q ordered when the inventory on hand declines to the ROP. Assume that weekly demand is normally distributed, with mean D and standard deviation σ_D. Assume replenishment lead time of L weeks.

Linking Safety Inventory and Cycle Service Level

We first show how cycle service levels can be evaluated given a replenishment policy (and thus the corresponding safety inventory). We then show how to determine the required safety inventory given a desired cycle service level.

EVALUATING SAFETY INVENTORY GIVEN A REPLENISHMENT POLICY In the case of B&M, safety inventory corresponds to the average number of phones on hand when a replenishment order arrives. Given the lead time of L weeks and a mean weekly demand of D, using Equation 12.2, we have

$$\text{Expected demand during lead time} = D \times L$$

Given that the store manager places a replenishment order when ROP phones are on hand, we have

$$\text{Safety inventory, } ss = ROP - D \times L \tag{12.3}$$

This is because, on average, $D \times L$ phones will sell over the L weeks between when the order is placed and when the lot arrives. The average safety inventory when the replenishment lot arrives is thus $ROP - D \times L$. The evaluation of safety inventory for a given inventory policy is described in Example 12-1 (see spreadsheet *Chapter 12-examples* worksheet *Example 12-1*).

EXAMPLE 12-1 Evaluating Safety Inventory Given an Inventory Policy

Assume that weekly demand for phones at B&M Office Supplies is normally distributed, with a mean of 2,500 and a standard deviation of 500. The manufacturer takes two weeks to fill an order placed by the B&M manager. The store manager currently orders 10,000 phones when the inventory on hand drops to 6,000. Evaluate the safety inventory and the average inventory carried by B&M. Also evaluate the average time a phone spends at B&M.

Analysis:

Under this replenishment policy, we have

Average demand per week, $D = 2{,}500$

Standard deviation of weekly demand, $\sigma_D = 500$

Average lead time for replenishment, $L = 2$ weeks

Reorder point, $ROP = 6{,}000$

Average lot size, $Q = 10{,}000$

Using Equation 12.3, we thus have

$$\text{Safety inventory, } ss = ROP - D \times L = 6{,}000 - 5{,}000 = 1{,}000$$

B&M thus carries a safety inventory of 1,000 phones. From Chapter 11, recall that

$$\text{Cycle inventory} = Q/2 = 10{,}000/2 = 5{,}000$$

We thus have

$$\text{Average inventory} = \text{cycle inventory} + \text{safety inventory} = 5{,}000 + 1{,}000 = 6{,}000$$

B&M thus carries an average of 6,000 phones in inventory. Using Little's law (Equation 3.1), we have

$$\text{Average flow time} = \text{average inventory/throughput} = 6{,}000/2{,}500 = 2.4 \text{ weeks}$$

Each phone thus spends an average of 2.4 weeks at B&M.

Next, we discuss how to evaluate the CSL given a replenishment policy.

EVALUATING CYCLE SERVICE LEVEL GIVEN A REPLENISHMENT POLICY Given a replenishment policy, our goal is to evaluate the CSL, the probability of not stocking out in a replenishment cycle. We return to B&M's continuous review replenishment policy of ordering Q units when the inventory on hand drops to the ROP. The lead time is L weeks and weekly demand is normally distributed, with a mean of D and a standard deviation of σ_D. Observe that a stockout occurs in a cycle if demand during the lead time is larger than the ROP. Thus, we have

$$CSL = \text{Prob}(\text{demand during lead time of } L \text{ weeks} \leq ROP)$$

To evaluate this probability, we need to obtain the distribution of demand during the lead time. From Equation 12.2, we know that demand during lead time is normally distributed, with a mean of D_L and a standard deviation of σ_L. Using the notation for the normal distribution from Appendix 12A and the equivalent Excel function from Equation 12.22 in Appendix 12B, the CSL is

$$CSL = F(ROP, D_L, \sigma_L) = NORMDIST(ROP, D_L, \sigma_L, 1) \tag{12.4}$$

We now illustrate this evaluation in Example 12-2 (see worksheet *Example 12-2*).

EXAMPLE 12-2 Evaluating Cycle Service Level Given a Replenishment Policy

Weekly demand for phones at B&M is normally distributed, with a mean of 2,500 and a standard deviation of 500. The replenishment lead time is two weeks. Assume that the demand is independent from one week to the next. Evaluate the CSL resulting from a policy of ordering 10,000 phones when there are 6,000 phones in inventory.

Analysis:

In this case, we have

$$Q = 10{,}000,\ ROP = 6{,}000,\ L = 2 \text{ weeks}$$
$$D = 2{,}500/\text{week},\ \sigma_D = 500$$

Observe that B&M runs the risk of stocking out during the lead time of two weeks between when an order is placed and when the replenishment arrives. Thus, whether or not a stockout occurs depends on the demand during the lead time of two weeks.

Because demand across time is independent, we use Equation 12.2 to obtain demand during the lead time to be normally distributed with a mean of D_L and a standard deviation of σ_L, where

$$D_L = D \times L = 2 \times 2{,}500 = 5{,}000, \quad \sigma_L = \sqrt{L}\sigma_D = \sqrt{2} \times 500 = 707$$

Using Equation 12.4, the CSL is evaluated as

$$CSL = F(ROP, D_L, \sigma_L) = NORMDIST(ROP, D_L, \sigma_L, 1)$$
$$= NORMDIST(6000,\ 5000,\ 707, 1) = 0.92$$

A CSL of 0.92 implies that in 92 percent of the replenishment cycles, B&M supplies all demand from available inventory. In the remaining 8 percent of the cycles, stockouts occur and some demand is not satisfied because of the lack of inventory.

We now discuss how the appropriate level of safety inventory may be obtained given a desired CSL.

Evaluating Safety Inventory Given Desired Cycle Service Level

In many practical settings, firms have a desired level of product availability and want to design replenishment polices that achieve this level. For example, Walmart has a desired level of product availability for each product sold in a store. The store manager must design a replenishment policy with the appropriate level of safety inventory to meet this goal. The desired level of product availability may be determined by trading off the cost of holding inventory with the cost of a stockout. This trade-off is discussed in detail in Chapter 13. In other instances, the desired level of product availability (in terms of CSL or fill rate) is stated explicitly in contracts, and management must design replenishment policies that achieve the desired target.

EVALUATING REQUIRED SAFETY INVENTORY GIVEN DESIRED CYCLE SERVICE LEVEL Our goal is to obtain the appropriate level of safety inventory given the desired CSL. We assume that a continuous review replenishment policy is followed. Consider the store manager at Walmart responsible for designing replenishment policies for all products in the store. He has targeted a CSL for the basic box of Lego building blocks. Given a lead time of L, the store manager wants to identify a suitable reorder point ROP and safety inventory that achieves the desired service level. Assume that demand for Legos at Walmart is normally distributed and independent from one week to the next. We assume the following inputs:

Desired cycle service level $= CSL$
Mean demand during lead time $= D_L$
Standard deviation of demand during lead time $= \sigma_L$

From Equation 12.3, recall that $ROP = D_L + ss$. The store manager needs to identify safety inventory ss such that the following is true:

$$\text{Probability}(\text{demand during lead time} \leq D_L + ss) = CSL$$

Given that demand is normally distributed, the store manager must identify safety inventory ss such that the following is true (using Equation 12.4):

$$F(D_L + ss, D_L, \sigma_L) = CSL$$

Given the definition of the inverse normal in Appendix 12A and the equivalent Excel function from Appendix 12B, we obtain

$$D_L + ss = F^{-1}(CSL, D_L, \sigma_L) = NORMINV(CSL, D_L, \sigma_L)$$
$$\text{or } ss = F^{-1}(CSL, D_L, \sigma_L) - D_L = NORMINV(CSL, D_L, \sigma_L) - D_L$$

Using the definition of the standard normal distribution and its inverse from Appendix 12A, and the equivalent Excel function from Appendix 12B, it can also be shown that the following is true:

$$ss = F_S^{-1}(CSL) \times \sigma_L = F_S^{-1}(CSL) \times \sqrt{L}\sigma_D = NORMSINV(CSL) \times \sqrt{L}\sigma_D \quad \textbf{(12.5)}$$

In Example 12-3 (see worksheet *Example 12-3*), we illustrate the evaluation of safety inventory given a desired CSL.

EXAMPLE 12-3 **Evaluating Safety Inventory Given a Desired Cycle Service Level**

Weekly demand for Legos at a Walmart store is normally distributed, with a mean of 2,500 boxes and a standard deviation of 500. The replenishment lead time is two weeks. Assuming a continuous-review replenishment policy, evaluate the safety inventory that the store should carry to achieve a CSL of 90 percent.

Analysis:

In this case we have

$$D = 2{,}500/\text{week},\ \sigma_D = 500,\ CSL = 0.9,\ L = 2 \text{ weeks}$$

Because demand across time is independent, we use Equation 12.2 to find demand during the lead time to be normally distributed with a mean of D_L and a standard deviation of σ_L, where

$$D_L = D \times L = 2 \times 2{,}500 = 5{,}000; \quad \sigma_L = \sqrt{L}\sigma_D = \sqrt{2} \times 500 = 707$$

Using Equation 12.5, we obtain

$$ss = F_s^{-1}(CSL) \times \sigma_L = NORMSINV(CSL) \times \sigma_L = NORMSINV(0.90) \times 707 = 906$$

Thus, the required safety inventory to achieve a CSL of 90 percent is 906 boxes.

Linking Safety Inventory and Fill Rate

We now show how fill rates can be evaluated given a replenishment policy (and thus the corresponding safety inventory). We then show how to determine the required safety inventory given a desired fill rate.

EVALUATING FILL RATE GIVEN A REPLENISHMENT POLICY Recall that fill rate measures the proportion of customer demand that is satisfied from available inventory. Fill rate is generally a more relevant measure than cycle service level because it allows the retailer to estimate the fraction of demand that is turned into sales. The two measures are closely related, as raising the cycle service level also raises the fill rate for a firm. Our discussion focuses on evaluating fill rate for a continuous review policy under which Q units are ordered when the quantity on hand drops to the ROP.

To evaluate the fill rate, it is important to understand the process by which a stockout occurs during a replenishment cycle. A stockout occurs if the demand during the lead time exceeds the ROP. We thus need to evaluate the average amount of demand in excess of the ROP in each replenishment cycle.

The *expected shortage per replenishment cycle* (ESC) is the average units of demand that are not satisfied from inventory in stock per replenishment cycle. Given a lot size of Q (which is also the average demand in a replenishment cycle), the fraction of demand lost is thus ESC/Q.

The product fill rate fr is thus given by

$$fr = 1 - ESC/Q = (Q - ESC)Q \qquad \textbf{(12.6)}$$

A shortage occurs in a replenishment cycle only if the demand during the lead time exceeds the ROP. Let $f(x)$ be the density function of the demand distribution during the lead time. The ESC is given by

$$ESC = \int_{x=ROP}^{\infty} (x - ROP) f(x)\,dx \qquad \textbf{(12.7)}$$

When demand during the lead time is normally distributed with mean D_L and standard deviation σ_L, given a safety inventory ss, Equation 12.7 can be simplified to

$$ESC = -ss\left[1 - F_s\left(\frac{ss}{\sigma_L}\right)\right] + \sigma_L f_s\left(\frac{ss}{\sigma_L}\right) \qquad \textbf{(12.8)}$$

where F_s is the standard normal cumulative distribution function and f_s is the standard normal density function. The standard normal distribution has a mean of 0 and a standard deviation of 1. A detailed description of the normal distribution is given in Appendix 12A. Details of the simplification in Equation 12.8 are described in Appendix 12C. Using Excel functions

	A	B	C	D	E
1	*Inputs*				
2	*Q*	*D*	σ_D	*L*	*ss*
3	10,000	2,500	500	2	1,000
4	*Distribution of demand during lead time*				
5	D_L	σ_L			
6	5,000	707			
7	*Cycle Service Level and Fill Rate*				
8	*CSL*	*ESC*	*fr*		
9	0.92	25.13	0.9975		

Cell	Cell Formula	Equation
A6	=B3*D3	12.2
B6	=SQRT(D3)*C3	12.2
A9	=NORMDIST(A6+E3, A6, B6, 1)	12.4
B9	=-E3*(1-NORMDIST(E3/B6, 0, 1, 1)) + B6*NORMDIST(E3/B6, 0, 1, 0)	12.8
C9	=(A3-B9)/A3	12.5

FIGURE 12-2 Excel Solution of Example 12-4

(Equations 12.25 and 12.26) discussed in Appendix 12B, ESC may be evaluated (using Equation 12.8) as

$$ESC = -ss[1 - NORMDIST(ss/\sigma_L, 0, 1, 1)] + \sigma_L\, NORMDIST(ss/\sigma_L, 0, 1, 0) \qquad \textbf{(12.9)}$$

Given the ESC, we can use Equation 12.6 to evaluate the fill rate *fr*. Next, we illustrate this evaluation in Example 12-4 (see worksheet *Example 12-4* and Figure 12-2).

EXAMPLE 12-4 Evaluating Fill Rate Given a Replenishment Policy

From Example 12-2, recall that weekly demand for phones at B&M is normally distributed, with a mean of 2,500 and a standard deviation of 500. The replenishment lead time is two weeks. Assume that the demand is independent from one week to the next. Evaluate the fill rate resulting from the policy of ordering 10,000 phones when there are 6,000 phones in inventory.

Analysis:

From the analysis of Example 12-2, we have

Lot size, $Q = 10{,}000$

Average demand during lead time, $D_L = 5{,}000$

Standard deviation of demand during lead time, $\sigma_L = 707$

Using Equation 12.3, we obtain

$$\text{Safety inventory, } ss = ROP - D_L = 6{,}000 - 5{,}000 = 1{,}000$$

From Equation 12.9, we thus have

$$ESC = -1{,}000[1 - NORMDIST(1{,}000/707, 0, 1, 1)] + 707\, NORMDIST(1{,}000/707, 0, 1, 0) = 25$$

Thus, on average, in each replenishment cycle, 25 phones are demanded by customers but not available in inventory. Using Equation 12.6, we thus obtain the following fill rate:

$$fr = (Q - ESC)/Q = (10{,}000 - 25)/10{,}000 = 0.9975$$

In other words, 99.75 percent of the demand is filled from inventory in stock. This is much higher than the CSL of 92 percent that resulted in Example 12-2 for the same replenishment policy.

A few key observations should be made. First, observe that the fill rate (0.9975) in Example 12-4 is significantly higher than the CSL (0.92) in Example 12-2 for the same replenishment policy. Next, by rerunning the examples with a different lot size (in worksheet *Example 12-4*), we can observe the impact of lot-size changes on the service level. Increasing the lot size of phones from 10,000 to 20,000 has no impact on the CSL (which stays at 0.92). The fill rate, however, now increases to 0.9987. This happens because an increase in lot size results in fewer replenishment cycles. In the case of B&M, an increase in lot size from 10,000 to 20,000 results in replenishment occurring once every eight weeks instead of once every four weeks. With a 92 percent CSL, a lot size of 10,000 results in, on average, one cycle with a stockout per year. With a lot size of 20,000, we have, on average, one stockout every two years. Thus, the fill rate is higher.

Key Point

Both fill rate and cycle service level increase as the safety inventory is increased. For the same safety inventory, an increase in lot size increases the fill rate but not the cycle service level.

EVALUATING REQUIRED SAFETY INVENTORY GIVEN DESIRED FILL RATE For a continuous review replenishment policy, we now evaluate the required safety inventory given a desired fill rate *fr*. Consider the store manager at Walmart targeting a fill rate *fr* for Lego building blocks. The current replenishment lot size is Q. The first step is to obtain the ESC using Equation 12.6.

The next step is to obtain a safety inventory *ss* that solves Equation 12.8 (and its Excel equivalent, Equation 12.9) given the ESC evaluated earlier. It is not possible to give a formula that provides the answer. The appropriate safety inventory that solves Equation 12.9 can be obtained easily using Excel and trying different values of *ss*. In Excel, the safety inventory may also be obtained directly using the tool *GOALSEEK,* as illustrated in Example 12-5 (use worksheet *Example 12-5*).

EXAMPLE 12-5 Evaluating Safety Inventory Given Desired Fill Rate

Weekly demand for Legos at a Walmart store is normally distributed, with a mean of 2,500 boxes and a standard deviation of 500. The replenishment lead time is two weeks. The store manager currently orders replenishment lots of 10,000 boxes from Lego. Assuming a continuous-review replenishment policy, evaluate the safety inventory the store should carry to achieve a fill rate of 97.5 percent.

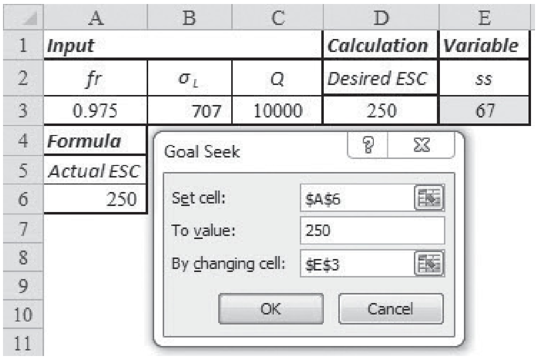
Analysis:

In this case, we have

Desired fill rate, $fr = 0.975$

Lot size, $Q = 10{,}000$ boxes

Standard deviation of demand during lead time, $\sigma_L = \sqrt{2} \times 500 = 707$



Cell	Cell Formula	Equation
A6	-E3*(1-NORMSDIST(E3/B3, 0, 1,1)) + B3*NORMDIST(E3/B3, 0, 1, 0)	12.10

FIGURE 12-3 Spreadsheet to Solve for *ss* Using *GOALSEEK*

From Equation 12.6, we thus obtain an ESC as

$$ESC = (1 - fr)\,Q = (1 - 0.975)10{,}000 = 250$$

Now we need to solve Equation 12.8 for the safety inventory *ss*, where

$$ESC = 250 = -ss\left[1 - F_s\left(\frac{ss}{\sigma_L}\right)\right] + \sigma_L f_s\left(\frac{ss}{\sigma_L}\right) = -ss\left[1 - F_s\left(\frac{ss}{707}\right)\right] + 707 f_s\left(\frac{ss}{707}\right)$$

Using Equation 12.9, this equation may be restated with Excel functions as follows:

$$250 = -ss[1 - NORMDIST(ss/707{,}0{,}1{,}1)] + 707NORMDIST(ss/707{,}0{,}1{,}0) \quad \textbf{(12.10)}$$

Equation 12.10 may be solved in Excel by trying different values of *ss* until the equation is satisfied. A more elegant approach for solving Equation 12.10 is to use the Excel tool *GOALSEEK*, as follows.

In the worksheet *Example 12-5,* invoke *GOALSEEK* using Data | What-If Analysis | Goal Seek. In the *GOALSEEK* dialog box, enter the data as shown in Figure 12-3 and click the OK button. In this case, cell D3 is changed until the value of the formula in cell A6 equals 250.

Using *GOALSEEK*, we obtain a safety inventory of $ss = 67$ boxes, as shown in Figure 12-3. Thus, the store manager at Walmart should target a safety inventory of 67 boxes to achieve the desired fill rate of 97.5 percent.

Impact of Desired Product Availability and Uncertainty on Safety Inventory

The two key factors that affect the required level of safety inventory are the desired level of product availability and uncertainty. We now discuss the impact that each factor has on the safety inventory.

As the desired product availability goes up, the required safety inventory also increases because the supply chain must now be able to accommodate uncommonly high demand or uncommonly low supply. For the Walmart situation in Example 12-5, we evaluate the required safety inventory for varying levels of fill rate as shown in Table 12-1.

Observe that raising the fill rate from 97.5 percent to 98.0 percent requires an additional 116 units of safety inventory, whereas raising the fill rate from 99.0 percent to 99.5 percent requires an additional 268 units of safety inventory. Thus, the marginal increase in safety inventory grows as product availability rises. This phenomenon highlights the importance of selecting

TABLE 12-1 Required Safety Inventory for Different Values of Fill Rate

Fill Rate	Safety Inventory
97.5%	67
98.0%	183
98.5%	321
99.0%	499
99.5%	767

suitable product availability levels. It is important for a supply chain manager to be aware of the products that require a high level of availability and hold high safety inventories only for those products. It is not appropriate to select a high level of product availability and require it arbitrarily for all products.

Key Point

The required safety inventory grows rapidly with an increase in the desired product availability.

From Equation 12.5, we see that the required safety inventory *ss* is also influenced by the standard deviation of demand during the lead time, σ_L. The standard deviation of demand during the lead time is influenced by the duration of the lead time L and the standard deviation of periodic demand σ_D, as shown in Equation 12.2. The relationship between safety inventory and σ_D is linear, in that a 10 percent increase in σ_D results in a 10 percent increase in safety inventory. Safety inventory also increases with an increase in lead time L. The safety inventory, however, is proportional to the square root of the lead time (if demand is independent over time) and thus grows more slowly than the lead time itself.

Key Point

The required safety inventory increases with an increase in the lead time and the uncertainty of periodic demand.

A goal of any supply chain manager is to reduce the level of safety inventory required in a way that does not adversely affect product availability. The previous discussion highlights two key managerial levers that may be used to achieve this goal:

1. ***Reduce the supplier lead time L:*** If lead time decreases by a factor of k, the required safety inventory decreases by a factor of $\sqrt{k}$. The only caveat here is that reducing the supplier lead time requires significant effort from the supplier, whereas reduction in safety inventory occurs at the retailer. Thus, it is important for the retailer to share some of the resulting benefits, as discussed in Chapter 10. Walmart, Seven-Eleven Japan, and many other retailers apply tremendous pressure on their suppliers to reduce the replenishment lead time. Apparel retailer Zara has built its entire strategy around using local flexible production to reduce replenishment lead times. In each case, the benefit has manifested itself in the form of reduced safety inventory while maintaining the desired level of product availability.

2. ***Reduce the underlying uncertainty of demand (represented by σ_D):*** If uncertainty represented by σ_D is reduced by a factor of k, the required safety inventory also decreases by a factor of k. A reduction in uncertainty may be achieved by better market intelligence, increased supply chain visibility, and the use of more sophisticated forecasting methods. Seven-Eleven Japan provides its store managers with detailed data about prior demand along with weather and other factors that may influence demand. This market intelligence allows the store managers to make better forecasts, reducing uncertainty. In most supply chains, however, the key to reducing the underlying forecast uncertainty is to link all forecasts throughout the supply chain to customer demand data. A lot of the demand uncertainty exists only because each stage of the supply chain plans and forecasts independently. This distorts demand throughout the supply chain, increasing uncertainty. Improved visibility and coordination, as discussed in Chapter 10, can often reduce the demand uncertainty significantly. Zara plans its production and replenishment based on actual sales at its retail stores to ensure that no unnecessary uncertainties are introduced. Both Walmart and Seven-Eleven Japan share demand information with their suppliers, reducing uncertainty and thus safety inventory within the supply chain.

We illustrate the benefits of reducing lead time and demand uncertainty in Example 12-6 (see worksheet *Example 12-6*).

EXAMPLE 12-6 Benefits of Reducing Lead Time and Demand Uncertainty

Weekly demand for white shirts at a Target store is normally distributed, with a mean of 2,500 and a standard deviation of 800. The replenishment lead time from the current supplier is nine weeks. The store manager aims for a cycle service level of 95 percent. What savings in safety inventory can the store expect if the supplier reduces lead time to one week? What savings in safety inventory can the store expect if reduced demand uncertainty results in a standard deviation of demand of 400?

Analysis:

For the base case, we have

$$D = 2{,}500/\text{week},\ \sigma_D = 800,\ CSL = 0.95$$

From Equation 12.5, we thus obtain the base case safety inventory to be

$$ss = NORMSINV(CSL) \times \sqrt{L}\sigma_D = NORMSINV(.95) \times \sqrt{9} \times 800 = 3{,}948$$

If the suppler reduces the lead time L to one week, the required safety inventory is given by

$$ss = NORMSINV(CSL) \times \sqrt{L}\sigma_D = NORMSINV(.95) \times \sqrt{1} \times 800 = 1{,}316$$

Thus, reducing the lead time from nine weeks to one week reduces the required safety inventory by 2,632 shirts.

We now consider the benefits of reducing forecast error. If Target reduces the standard deviation from 800 to 400 (for the nine-week lead time), the required safety inventory is obtained as follows:

$$ss = NORMSINV(CSL) \times \sqrt{L}\sigma_D = NORMSINV(.95) \times \sqrt{9} \times 400 = 1{,}974$$

Thus, reducing the standard deviation (equal to forecast error) of demand from 800 to 400 reduces the required safety inventory by 1,974 shirts.

12.4 IMPACT OF SUPPLY UNCERTAINTY ON SAFETY INVENTORY

In our discussion to this point, we have focused on situations with demand uncertainty in the form of a forecast error. In many practical situations, supply uncertainty also plays a significant role. The impact of supply uncertainty is well illustrated by the impact of the grounding of MSC

Napoli on the south coast of Britain in January 2007. The container ship was carrying more than 1,000 tons of nickel, a key ingredient of stainless steel. Given that 1,000 tons was almost 20 percent of the 5,052 tons of nickel then stored in warehouses globally, this delay in bringing nickel to market resulted in significant shortages and raised the price of nickel by about 20 percent in the first 3.5 weeks of January 2007. Supply uncertainty arises because of many factors, including production delays, transportation delays, and quality problems. Supply chains must account for supply uncertainty when planning safety inventories.

In this section, we incorporate supply uncertainty by assuming that lead time is uncertain and identify the impact of lead time uncertainty on safety inventories. Assume that the customer demand per period for tablets at Amazon and the replenishment lead time from the supplier are normally distributed. We are provided the following inputs:

D: Average demand per period

σ_D: Standard deviation of demand per period

L: Average lead time for replenishment

s_L: Standard deviation of lead time

We consider the safety inventory requirements given that Amazon follows a continuous review policy to manage tablet inventory. Amazon experiences a stockout of product if demand during the lead time exceeds the ROP—that is, the quantity on hand when Amazon places a replenishment order. Thus, we need to identify the distribution of customer demand during the lead time. Given that both lead time and periodic demand are uncertain, demand during the lead time is normally distributed with a mean of D_L and a standard deviation σ_L, where

$$D_L = D \times L; \quad \sigma_L = \sqrt{L\sigma_D^2 + D^2 s_L^2} \tag{12.11}$$

Given the distribution of demand during the lead time in Equation 12.11 and a desired CSL, Amazon can obtain the required safety inventory using Equation 12.5. If product availability is specified as a fill rate, Amazon can obtain the required safety inventory using the procedure outlined in Example 12-5. In Example 12-7, we illustrate the impact of lead time uncertainty on the required level of safety inventory at Amazon (see worksheet *Example 12-7*).

EXAMPLE 12-7 Impact of Lead Time Uncertainty on Safety Inventory

Daily demand for tablets at Amazon is normally distributed, with a mean of 2,500 and a standard deviation of 500. The tablet supplier takes an average of $L = 7$ days to replenish inventory at Amazon. Amazon is targeting a CSL of 90 percent (providing a fill rate close to 100 percent) for its tablet inventory. Evaluate the safety inventory of tablets that Amazon must carry if the standard deviation of the lead time is seven days. Amazon is working with the supplier to reduce the standard deviation to zero. Evaluate the reduction in safety inventory that Amazon can expect as a result of this initiative.

Analysis:

In this case, we have

Average demand per period, $D = 2{,}500$

Standard deviation of demand per period, $\sigma_D = 500$

Average lead time for replenishment, $L = 7$ days

Standard deviation of lead time, $s_L = 7$ days

TABLE 12-2 Required Safety Inventory as a Function of Lead Time Uncertainty

s_L	σ_L	*ss* (units)	*ss* (days)
6	15,058	19,298	7.72
5	12,570	16,109	6.44
4	10,087	12,927	5.17
3	7,616	9,760	3.90
2	5,172	6,628	2.65
1	2,828	3,625	1.45
0	1,323	1,695	0.68

We first evaluate the distribution of demand during the lead time. Using Equation 12.11, we have

$$\text{Mean demand during lead time, } D_L = D \times L = 2{,}500 \times 7 = 17{,}500$$

$$\text{Standard deviation of demand during lead time, } \sigma_L = \sqrt{L\sigma_D^2 + D^2 s_L^2}$$
$$= \sqrt{7 \times 500^2 + 2{,}500^2 \times 7^2} = 17{,}550$$

The required safety inventory is obtained using Equations 12.5 and 12.27, as follows:

$$ss = NORMSINV(CSL) \times \sigma_L = NORMSINV(0.90) \times 17{,}550 = 22{,}491 \text{ tablets}$$

If the standard deviation of lead time is seven days, Amazon must carry a safety inventory of 22,491 tablets. This is equivalent to about nine days of demand for tablets.

In Table 12-2, we provide the required safety inventory as Amazon works with the supplier to reduce the standard deviation of lead time (s_L) from six down to zero. From Table 12-2, observe that the reduction in lead time uncertainty allows Amazon to reduce its safety inventory of tablets by a significant amount. As the standard deviation of lead time declines from seven days to zero, the amount of safety inventory declines from about nine days of demand to less than a day of demand.

The preceding example emphasizes the impact of lead time variability on safety inventory requirements (and thus material flow time) and the large potential benefits from reducing lead time variability or improving on-time deliveries. Often, safety inventory calculations in practice do not include any measure of supply uncertainty, resulting in levels that may be lower than required. This hurts product availability.

Key Point

A reduction in supply uncertainty can help to dramatically reduce the required safety inventory without hurting product availability.

In practice, variability of supply lead time is caused by practices at both the supplier and the party receiving the order. Suppliers sometimes have poor planning tools that do not allow them to schedule production in a way that can be executed. Today, most supply chain planning software suites have good production planning tools that allow suppliers to promise lead times that can be met. This helps reduce lead time variability. The lack of visibility for a supplier into future customer plans is also a significant factor that increases supply chain uncertainty. W.W. Grainger was able to get its suppliers to reduce both lead time and lead time variability by sharing its future plans with them. This allowed suppliers to schedule Grainger orders into production without waiting for the orders to actually arrive. The quantity produced was finalized closer to actual production. In other instances, the behavior of the party placing the order often increases lead time variability. In one instance, a distributor placed orders to all suppliers on the same day of the week. As a result, all deliveries arrived on the same day of the week. The surge in deliveries made it impossible for all of them to be recorded into inventory on the day they arrived. This led to a perception that supply lead times were long and variable. Just by leveling out the orders over the week, the lead time and the lead time variability were significantly reduced, allowing the distributor to reduce its safety inventory.

Next, we discuss how aggregation can help reduce the amount of safety inventory in the supply chain.

12.5 IMPACT OF AGGREGATION ON SAFETY INVENTORY

In practice, supply chains have varying degrees of inventory aggregation. For example, Barnes & Noble sells books from retail stores with inventory geographically distributed across the country. Amazon, in contrast, ships all its books from a few facilities. Seven-Eleven Japan has many small convenience stores densely distributed across Japan. In contrast, supermarkets tend to be much larger, with fewer outlets that are not as densely distributed. Redbox rents its movies from tens of thousands of kiosks distributed across the United States. In contrast, Netflix centralizes its DVD inventory at fewer than fifty distribution centers.

Our goal is to understand how aggregation in each of these cases affects forecast accuracy and safety inventories. Consider k regions, with demand in each region normally distributed with the following characteristics:

D_i: Mean periodic demand in region i, $i = 1, \ldots, k$

σ_i: Standard deviation of periodic demand in region i, $i = 1, \ldots, k$

ρ_{ij}: Correlation of periodic demand for regions i, j, $1 \le i \ne j \le k$

There are two ways to serve demand in the k regions. One is to have local inventories in each region and the other is to aggregate all inventories into one centralized facility. Our goal is to compare safety inventories in the two cases. With a replenishment lead time of L and a desired cycle service level CSL, the total safety inventory in the decentralized case is (using Equation 12.5):

$$\text{Total safety inventory in decentralized option} = \sum_{i=1}^{k} F_S^{-1}(CSL) \times \sqrt{L} \times \sigma_i \qquad \textbf{(12.12)}$$

If all inventories are aggregated in a central location, we need to evaluate the distribution of aggregated demand. The aggregate demand is normally distributed, with a mean of D^C, standard deviation of σ_D^C, and a variance of var(D^C), as follows:

$$D^C = \textstyle\sum_{i=1}^{k} D_i; \quad \text{var}(D^C) = \sum_{i=1}^{k} \sigma_i^2 + 2\sum_{i>j} \rho_{ij}\sigma_i\sigma_j; \quad \sigma_D^C = \sqrt{\text{var}(D^C)} \qquad \textbf{(12.13)}$$

Observe that Equation 12.13 is like Equation 12.1 except that we are aggregating across k regions rather than L periods. If all k regions have demand that is identically distributed, with mean D and standard deviation σ_D, and have the same correlation ρ, Equation 12.13 can be simplified as

$$D^C = kD; \sigma_D^C = \sqrt{k\sigma^2 + k(k-1)\rho\sigma^2} \qquad \textbf{(12.14)}$$

If all k regions have demand that is independent ($\rho_{ij} = 0$) and identically distributed, with mean D and standard deviation σ_D, Equation 12.13 can be simplified as

$$D^C = kD; \sigma_D^C = \sqrt{k}\sigma_D \qquad \textbf{(12.15)}$$

Using Equations 12.5 and 12.13, the required safety inventory at the centralized location is given as

$$\text{Required safety inventory on aggregation} = F_S^{-1}(CSL) \times \sqrt{L} \times \sigma_D^C \qquad \textbf{(12.16)}$$

The holding cost savings on aggregation per unit sold are obtained by dividing the savings in holding cost by the total demand kD. If H is the holding cost per unit, using Equations 12.12 and 12.16, the savings per units are

$$\text{Holding} - \text{cost savings on aggregation per unit sold}$$
$$= \frac{F_S^{-1}(CSL) \times \sqrt{L} \times H}{D^C} \times \left(\sum_{i=1}^{k} \sigma_i - \sigma_D^C\right) \qquad \textbf{(12.17)}$$

From Equation 12.13, it follows that the difference $\left(\sum_{i=1}^{k} \sigma_i - \sigma_D^C\right)$ is influenced by the correlation coefficients ρ_{ij}. This difference is large when the correlation coefficients are close to -1 (negative correlation) and shrinks as they approach $+1$ (positive correlation). Inventory savings on aggregation are always positive as long as the correlation coefficients are less than 1. From Equation 12.17, we thus draw the following conclusions regarding the value of aggregation:

- The safety inventory savings on aggregation increase with the desired cycle service level *CSL*.
- The safety inventory savings on aggregation increase with the replenishment lead time L.
- The safety inventory savings on aggregation increase with the holding cost H.
- The safety inventory savings on aggregation increase with the coefficient of variation (σ_D/D) of demand.
- The safety inventory savings on aggregation decrease as the correlation coefficients increase.

In Example 12-8 (see worksheet *Example 12-8*), we illustrate the inventory savings on aggregation and the impact of the correlation coefficient on these savings.

EXAMPLE 12-8 Impact of Correlation on Value of Aggregation

A BMW dealership has $k = 4$ retail outlets serving the entire Chicago area (disaggregate option). Weekly demand at each outlet is normally distributed, with a mean of $D = 25$ cars and a standard deviation of $\sigma_D = 5$. The lead time for replenishment from the manufacturer is $L = 2$ weeks. Each outlet covers a separate geographic area, and the correlation of demand across any pair of areas is ρ. The dealership is considering the possibility of replacing the four outlets with a single large outlet (aggregate option). Assume that the demand in the central outlet is the sum of the demand across all four areas. The dealership is targeting a CSL of 0.90. Compare the level of safety inventory needed in the two options as the correlation coefficient ρ varies between 0 and 1.

Analysis:

We provide a detailed analysis for the case when demand in each area is independent (i.e., $\rho = 0$). For each retail outlet we have

Standard deviation of weekly demand, $\sigma_D = 5$

Replenishment lead time, $L = 2$ weeks

Using Equation 12.12, the required safety inventory in the decentralized option for $CSL = 0.90$ is

$$\text{Total required safety inventory, } ss = k \times F_s^{-1}(CSL) \times \sqrt{L} \times \sigma_D$$
$$= 4 \times F_s^{-1}(0.9) \times \sqrt{2} \times 5 = 4 \times NORMSINV(0.9) \times \sqrt{2} \times 5 = 36.25 \text{ cars}$$

Now, consider the aggregate option. Using Equation 12.14, the standard deviation of aggregate weekly demand is

$$\text{Standard deviation of weekly demand at central outlet, } \sigma_D^C = \sqrt{4 \times 5^2 + 4 \times 3 \times 5^2 \times \rho}$$

TABLE 12-3 Safety Inventory in the Disaggregate and Aggregate Options

ρ	Disaggregate Safety Inventory	Aggregate Safety Inventory
0	36.25	18.12
0.2	36.25	22.93
0.4	36.25	26.88
0.6	36.25	30.33
0.8	36.25	33.42
1.0	36.25	36.25

For a CSL of 0.90 and $\rho = 0$, safety inventory required for the aggregate option (using Equation 12.16) is given as

$$ss = F_S^{-1}(0.90) \times \sqrt{L} \times \sigma_D^C = NORMSINV(0.90) \times \sqrt{2} \times 10 = 18.12$$

Using Equations 12.12 to 12.16, the required level of safety inventory for the disaggregate as well as the aggregate option can be obtained for different values of ρ as shown in Table 12-3 using worksheet *Example 12-8*. Observe that the safety inventory for the disaggregate option is higher than for the aggregate option except when all demands are perfectly positively correlated. The benefit of aggregation decreases as demand in different areas is more positively correlated.

Example 12-8 and the previous discussion demonstrate that aggregation reduces demand uncertainty—and, thus, the required safety inventory—as long as the demand being aggregated is not perfectly positively correlated. Demand for most products does not show perfect positive correlation across different geographic regions. Products such as heating oil are likely to have demand that is positively correlated across nearby regions. In contrast, products such as milk and sugar are likely to have demand that is much more independent across regions. If demand in different geographic regions is about the same size and independent, aggregation reduces safety inventory by the square root of the number of regions aggregated. In other words, if the number of independent stocking locations decreases by a factor of n, the average safety inventory is expected to decrease by a factor of $\sqrt{n}$. This principle is referred to as the *square-root law* and is illustrated in Figure 12-4.

Most online retailers exploit the benefits of aggregation in terms of reduced inventories. For example, Blue Nile sells diamonds online and serves the entire United States out of one

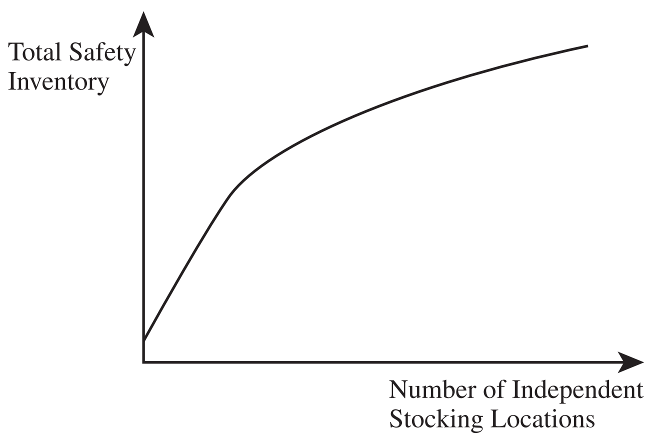


FIGURE 12-4 Square-Root Law

warehouse. As a result, it has lower levels of diamond inventories than jewelry chains such as Tiffany and Zales, which must keep inventory in every retail store.

There are situations, however, in which physical aggregation of inventories in one location may not be optimal. There are two major disadvantages of aggregating all inventories in one location:

1. Increase in response time to customer order
2. Increase in transportation cost to customer

Both disadvantages result because the average distance between the inventory and the customer increases with aggregation. Either the customer must travel farther to reach the product or the product must be shipped over longer distances to reach the customer. A retail chain such as Gap has the option of building many small retail outlets or a few large ones. Gap tends to have many smaller outlets distributed evenly in a region because this strategy reduces the distance that customers travel to reach a store. If Gap had one large centralized outlet, the average distance that customers need to travel would increase, and thus the response time would increase. A desire to decrease customer response time is thus the impetus for the firm to have multiple outlets. Another example is McMaster-Carr, a distributor of MRO supplies. McMaster-Carr uses UPS for shipping product to customers. Because shipping charges are based on distance, having one centralized warehouse increases the average shipping cost as well as the response time to the customer. Thus, McMaster-Carr has five warehouses that allow it to provide next-day delivery to a large fraction of the United States. Next-day delivery by UPS would not be feasible at a reasonable cost if McMaster-Carr had only one warehouse. Even Amazon, which started with one warehouse in Seattle, has added more warehouses in other parts of the United States in an effort to improve response time and reduce transportation cost to the customer. We illustrate the trade-offs of centralization in Example 12-9 (see worksheet *Example 12-9*).

EXAMPLE 12-9 Trade-Offs of Physical Centralization

The Shanghai branch office of an Italian coffee machine company is considering setting up either one distribution center for each of its east, south, west, and north regions or simply one center in Ningbo for the whole of China. The weekly demand for the automatic espresso coffee machine is normally distributed with a mean of 1,000 units and a standard deviation of 300 units. Although the demand for each region is independent, the supply lead time is more or less the same—four weeks. Each machine costs $1,000 and the holding cost is 20 percent. With the next-day delivery promise, the branch office needs to bear an inland trucking cost of $10/unit for all four regional centers. However, if a single national distribution center is decided upon, a more expensive transport fleet is needed and that will cost $13/unit for next-day service. Setting up and operating four regional DCs costs $150,000 per year more than building and operating the single Ningbo national distribution center.

Assume that the Italian company would like a CSL of 0.95. What should the Shanghai branch office decide based on the cost considerations?

Analysis:

Observe that using only one distribution center would decrease facility and inventory costs but increase transportation costs. We therefore have to evaluate the change in each cost category on aggregation.

We start with inventory costs. For each region we have

$$D = \frac{1{,}000}{\text{week}}, \sigma_D = 300, L = 4 \text{ weeks}$$

Given the desired $CSL = 0.95$, the required safety inventory across all four regional distribution centers is obtained using Equation 12.9 to be

$$\begin{aligned}\text{Total required safety inventory, } ss &= 4 \times F_s^{-1}(CSL) \times \sqrt{L} \times \sigma_D \\ &= 4 \times NORMSINV(0.95) \times \sqrt{4} \times 300 = 3{,}948\end{aligned}$$

Now consider the aggregate option. Because demand in all four areas is independent, $\rho = 0$. Using Equation 12.14, the standard deviation of aggregate weekly demand is

$$\text{Standard deviation of weekly demand at national distribution center, } \sigma_D^C = \sqrt{4} \times 300 = 600$$

For a CSL of 0.95, safety inventory required for the aggregate option (using Equation 12.15) is given as

$$ss = F_S^{-1}(0.95) \times \sqrt{L} \times \sigma_D^C = NORMSINV(0.95) \times \sqrt{4} \times 600 = 1{,}974$$

We can now evaluate the changes in inventory, transportation, and facility costs upon aggregation as follows:

$$\text{Decrease in annual inventory holding cost on aggregation} = (3{,}948 - 1{,}974) \times \$1{,}000 \times 0.2 = \$394{,}765$$

$$\text{Decrease in annual facility costs on aggregation} = \$150{,}000$$

$$\text{Increase in annual transportation costs on aggregation} = 52 \times 1{,}000 \times (13 - 10) \times 4 = \$624{,}000$$

Observe that in this case, the annual costs for the Shanghai branch office will be increased by $\$624{,}000 - \$394{,}765 - \$150{,}000 = \$79{,}235$ upon centralization. It is clearly better to run the four centers in the east, south, west, and north regions instead of the Ningbo center.

Example 12-9 and the previous discussion highlight instances in which physical aggregation of inventory at one location may not be optimal. However, aggregating safety inventory has clear benefits. We now discuss various methods by which a supply chain can extract the benefits of aggregation without having to physically centralize all inventories in one location.

Information Centralization

Redbox uses *information centralization* to virtually aggregate its inventories of DVDs despite having tens of thousands of vending machines. The company has set up an online system that allows customers to locate nearby vending machines with the DVD they are searching for in stock. This allows Redbox to provide a much higher level of product availability than would be possible if a customer found out about availability only by visiting a vending machine. The benefit of information centralization derives from the fact that most customers get their DVD from the vending machine closest to their house. In case of a stockout at the closest vending machine, the customer is served from another vending machine, thus improving product availability without adding to inventories.

Retailers such as Gap also use information centralization effectively. If a store does not have the size or color that a customer wants, store employees can use their information system to inform the customer of the closest store with the product in inventory. Customers can then either go to this store or have the product delivered to their house. Gap thus uses information centralization to virtually aggregate inventory across all retail stores even though the inventory is physically separated. This allows Gap to reduce the amount of safety inventory it carries while providing a high level of product availability.

Walmart has an information system in place that allows store managers to search other stores for an excess of items that may be hot sellers at their stores. Walmart provides transportation that allows store managers to exchange products so they arrive at stores where they are in high demand. In this case, Walmart uses information centralization with a responsive transportation system to reduce the amount of safety inventory carried while providing a high level of product availability.

Specialization

Most supply chains provide a variety of products to customers. When inventory is carried at multiple locations, a key decision for a supply chain manager is whether all products should be stocked at every location. Clearly, a product that does not sell in a geographic region should not be carried in inventory by the warehouse or retail store located there. For example, it does not make sense for a Sears retail store in southern Florida to carry a wide variety of snow boots in inventory.

Another important factor that must be considered when making stocking decisions is the reduction in safety inventory that results from aggregation. If aggregation reduces the required safety inventory for a product by a large amount, it is better to carry the product in one central location. If aggregation reduces the required safety inventory for a product by a small amount, it may be best to carry the product in multiple decentralized locations to reduce response time and transportation cost.

The reduction in safety inventory due to aggregation is strongly influenced by the demand's coefficient of variation. For a product with a low coefficient of variation, disaggregate demand can be forecast with accuracy. As a result, the inventory benefit from aggregation is minimal. For a product with a high coefficient of variation of demand, disaggregate demand is difficult to forecast. In this case, aggregation improves forecast accuracy significantly, providing great benefits. We illustrate this idea in Example 12-10 (see worksheet *Example 12-10*).

EXAMPLE 12-10 Impact of Coefficient of Variation on Value of Aggregation

Assume that W.W. Grainger, a supplier of MRO products, has 1,600 stores distributed throughout the United States. Consider two products—large electric motors and industrial cleaner. Large electric motors are high-value items with low demand, whereas the industrial cleaner is a low-value item with high demand. Each motor costs $500 and each can of cleaner costs $30. Weekly demand for motors at each store is normally distributed, with a mean of 20 and a standard deviation of

TABLE 12-4 Value of Aggregation at W.W. Grainger

	Motors	Cleaner
Inventory is stocked in each store		
Mean weekly demand per store	20	1,000
Standard deviation	40	100
Coefficient of variation	2.0	0.1
Safety inventory per store	132	329
Total safety inventory	211,200	526,400
Value of safety inventory	$105,600,000	$15,792,000
Inventory is aggregated at the DC		
Mean weekly aggregate demand	32,000	1,600,000
Standard deviation of aggregate demand	1,600	4,000
Coefficient of variation	0.05	0.0025
Aggregate safety inventory	5,264	13,159
Value of safety inventory	$2,632,000	$394,770
Savings		
Total inventory saving on aggregation	$102,968,000	$15,397,230
Total holding cost saving on aggregation	$25,742,000	$3,849,308
Holding cost saving per unit sold	$15.47	$0.046
Savings as a percentage of product cost	3.09%	0.15%

40. Weekly demand for cleaner at each store is normally distributed, with a mean of 1,000 and a standard deviation of 100. Demand experienced by each store is independent, and supply lead time for both motors and cleaner is four weeks. W.W. Grainger has a holding cost of 25 percent. For each of the two products, evaluate the reduction in safety inventories that will result if they are removed from retail stores and carried only in a centralized DC. Assume a desired CSL of 0.95.

Analysis:

The evaluation of safety inventories and the value of aggregation for each of the two products is shown in Table 12-4. All calculations use the approach discussed earlier and illustrated in Example 12-8. As Table 12-4 shows, the inventory reduction benefit from centralizing motors is much larger than the benefit from centralizing cleaner. From this analysis, W.W. Grainger should stock cleaner at the stores and motors in the DC. Given that cleaner is a high-demand item, customers will be able to pick it up on the same day at the stores. Given that motors are a low-demand item, customers may be willing to wait the extra day that shipping from the DC will entail.

Key Point

The higher the coefficient of variation of an item, the greater is the reduction in safety inventories as a result of centralization.

Items with low demand are referred to as *slow-moving items* and typically have a high coefficient of variation, whereas items with high demand are referred to as *fast-moving items* and typically have a low coefficient of variation. For many supply chains, specializing the distribution network with fast-moving items stocked at decentralized locations close to the customer and slow-moving items stocked at a centralized location can significantly reduce the safety inventory carried without hurting customer response time or adding to transportation costs. The centralized location then specializes in handling slow-moving items.

Of course, other factors also need to be considered when deciding on the allocation of products to stocking locations. For example, an item that is considered an emergency item because the customer needs it urgently may be stocked at stores even if it has a high coefficient of variation. In this case the customer will be willing to pay a premium for having the item available at a store. One also needs to consider the cost of the item. High-value items provide a greater benefit from centralization than do low-value items.

The insights from Example 12-10 and the above discussion are summarized in Figure 12-5. In general, decentralized networks like Costco provide a low-cost supply chain for fast-moving, predictable, low-value products like detergent. Centralized networks like Blue Nile provide a low-cost supply chain for slow-moving, unpredictable, high-value products like diamonds. A decentralized supply chain like Tiffany may carry slow-moving items like diamonds as long as customers are willing to pay a premium for this choice. Similarly, a centralized supply chain like Amazon may carry a fast-moving item like detergent, but only if customers are willing to pay a premium. It can be argued that Amazon's inability to extract a significant enough premium from its customers for the fast-moving items it sells has hurt its profitability.

Item Type	Centralized Inventories	Decentralized Inventories
Fast Moving Predictable {Low value}	Customer willing to pay premium?	Low cost
Slow Moving Unpredictable {High value}	Low cost	Customer willing to pay premium?

FIGURE 12-5 Specialization of Inventory Based on Product Type

It is important for firms with bricks-and-mortar stores to take the idea of specialization into account when incorporating the online channel into an omni-channel strategy. Consider, for example, a bookstore chain such as Barnes & Noble, which carries about 100,000 titles at each retail store. The titles carried can be divided into two broad categories—best sellers with high demand and other books with much lower demand. Barnes & Noble can design an omni-channel strategy under which the retail stores carry primarily best sellers in inventory. They may also carry one, or at most two, copies of each of the other titles, to allow customers to browse. Customers should be able to access all titles that are not in the store via electronic kiosks in the store, which provide access to barnesandnoble.com inventory. This strategy allows customers to access an increased variety of books from Barnes & Noble stores. Customers could place orders for low-volume titles with barnesandnoble.com while purchasing high-volume titles at the store itself. This strategy of specialization would allow Barnes & Noble to aggregate all slow-moving items to be sold by the online channel. All best sellers would be decentralized and carried close to the customer. The supply chain thus reduces inventory costs for slow-moving items at the expense of somewhat higher transportation costs. For the fast-moving items, the supply chain provides a lower transportation cost and better response time by carrying the items at retail stores close to the customer.

Home Depot follows a similar strategy and integrates its online channel with its retail stores. The retail stores carry fast-moving items, and the customer is able to order slow-moving variants online. Home Depot is thus able to increase the variety of products available to customers while keeping supply chain inventories down. Walmart.com has also employed a strategy of carrying slower-moving items online.

Product Substitution

Substitution refers to the use of one product to satisfy demand for a different product. Substitution may occur in two situations:

1. ***Manufacturer-driven substitution:*** The manufacturer or supplier makes the decision to substitute. Typically, the manufacturer substitutes a higher-value product for a lower-value product that is not in inventory. For example, Dell may install a 1.2-terabyte hard drive into a customer order requiring a 1-terabyte drive if the smaller drive is out of stock.
2. ***Customer-driven substitution:*** Customers make the decision to substitute. For example, a customer walking into a Walmart store to buy a gallon of detergent may buy the half-gallon size if the gallon size is not available. The customer substitutes the half-gallon size for the gallon size.

In both cases, exploiting substitution allows the supply chain to satisfy demand using aggregate inventories, which permits the supply chain to reduce safety inventories without hurting product availability. In general, given two products or components, substitution may be one-way (i.e., only one of the products [components] substitutes for the other) or two-way (i.e., either product [component] substitutes for the other). We briefly discuss one-way substitution in the context of manufacturer-driven substitution and two-way substitution in the context of customer-driven substitution.

MANUFACTURER-DRIVEN ONE-WAY SUBSTITUTION Consider a server manufacturer selling direct to customers that offers drives that vary in size from 0.8 to 1.2 terabytes. Customers are charged according to the size of drive that they select, with larger sizes being more expensive. If a customer orders a 1-terabyte drive and the manufacturer is out of drives of this size, there are two possible choices: (1) delay or deny the customer order or (2) substitute a larger drive that is in stock (say, a 1.2-terabyte drive) and fill the customer order on time. The first case is potentially a lost sale or loss of future sales because the customer experiences a delayed delivery. In the second case, the manufacturer installs a higher-cost component, reducing the company's profit margin. These factors, along with the fact that only larger drives can substitute for smaller drives, must be considered when the manufacturer makes inventory decisions for individual drive sizes.

Substitution allows the server manufacturer to aggregate demand across the components, reducing safety inventories required. The value of substitution increases as demand uncertainty increases. Thus, the manufacturer should consider substitution for components displaying high demand uncertainty.

The desired degree of substitution is influenced by the cost differential between the higher-value and lower-value component. If the cost differential is very small, the manufacturer should aggregate most of the demand and carry most of its inventory in the form of the higher-value component. As the cost differential increases, though, the benefit of substitution decreases. In this case, the manufacturer will find it more profitable to carry inventory of each of the two components and decrease the amount of substitution.

The desired level of substitution is also influenced by the correlation of demand between the products. If demand between two components is strongly positively correlated, there is little value in substitution. As demand for the two components becomes less positively correlated (or even negatively correlated), the benefit of substitution increases.

Key Point

Manufacturer-driven substitution increases overall profitability for the manufacturer by allowing some aggregation of demand, which reduces the inventory requirements for the same level of availability.

CUSTOMER-DRIVEN TWO-WAY SUBSTITUTION Consider W.W. Grainger selling two brands of motors, GE and SE, which have similar performance characteristics. Customers are generally willing to purchase either brand, depending on product availability. If W.W. Grainger managers do not recognize customer substitution, they will not encourage it. For a given level of product availability, they will thus have to carry high levels of safety inventory of each brand. If its managers recognize and encourage customer substitution, they can aggregate the safety inventory across the two brands, thereby improving product availability.

W.W. Grainger does a good job of recognizing customer substitution. When a customer calls or goes online to place an order and the product he or she requests is not available, the customer is immediately told the availability of all equivalent products that may be substituted. Most customers ultimately buy a substitute product in this case. W.W. Grainger exploits this substitution by managing safety inventory of all substitutable products jointly. Recognition and exploitation of customer substitution allows W.W. Grainger to provide a high level of product availability with lower levels of safety inventory.

A good understanding of customer-driven substitution is important in the retail industry. It must be exploited when merchandising to ensure that substitute products are placed near each other, allowing a customer to buy one if the other is out of stock. In the online channel, substitution requires a retailer to present the availability of substitute products if the one the customer requests is out of stock. The supply chain is thus able to reduce the required level of safety inventory while providing a high level of product availability.

Key Point

Recognition of customer-driven substitution and joint management of inventories across substitutable products allows a supply chain to reduce the required safety inventory while ensuring a high level of product availability.

The demand uncertainties and the correlation of demand between the substitutable products influence the benefit to a retailer from exploiting substitution. The greater the demand uncertainty, the greater is the benefit of substitution. The less positive the correlation of demand between substitutable products, the greater is the benefit from exploiting substitution.

Component Commonality

In any supply chain, a significant amount of inventory is held in the form of components. A single product such as a server contains hundreds of components. When a supply chain is producing a large variety of products, component inventories can easily become very large. The use of common components in a variety of products is an effective supply chain strategy to exploit aggregation and reduce component inventories.

Dell sells thousands of server configurations to customers. An extreme option for Dell is to design distinct components that are suited to the performance of a particular configuration. Under this option, Dell would use different memory, hard drive, and other components for each distinct finished product. The other option is to design products such that common components are used in different finished products.

Without common components, the uncertainty of demand for any component is the same as the uncertainty of demand for the finished product in which it is used. Given the large number of components in each finished product, demand uncertainty will be high, resulting in high levels of safety inventory. When products with common components are designed, the demand for each component is an aggregation of the demand for all the finished products of which the component is a part. Component demand is thus more predictable than the demand for any one finished product. This fact reduces the component inventories carried in the supply chain. This idea has been a key factor for success in the electronics industry and has also started to play a big role in the auto industry. With increasing product variety, component commonality is a key to reducing supply chain inventories without hurting product availability. We illustrate the basic idea behind component commonality in Example 12-11 (see worksheet *Example 12-11*).

EXAMPLE 12-11 Value of Component Commonality

Assume that Dell is to manufacture 27 servers with three distinct components: processor, memory, and hard drive. Under the disaggregate option, Dell designs specific components for each server, resulting in $3 \times 27 = 81$ distinct components. Under the common-component option, Dell designs servers such that three distinct processors, three distinct memory units, and three distinct hard drives can be combined to create 27 servers. Each component is thus used in nine servers. Monthly demand for each of the 27 servers is independent and normally distributed, with a mean of 5,000 and a standard deviation of 3,000. The replenishment lead time for each component is one month. Dell is targeting a CSL of 95 percent for component inventory. Evaluate the safety inventory requirements with and without the use of component commonality. Also evaluate the change in safety inventory requirements as the number of finished products of which a component is a part varies from one to nine.

Analysis:

We first evaluate the disaggregate option, in which components are specific to a server. For each component, we have

$$\text{Standard deviation of monthly demand} = 3{,}000$$

Given a lead time of one month and a total of 81 components across 27 servers, we thus use Equation 12.12 to obtain

$$\text{Total safety inventory required} = 81 \times NORMSINV(0.95) \times \sqrt{1} \times 3{,}000 = 399{,}699 \text{ units}$$

TABLE 12-5 Marginal Benefit of Component Commonality

Number of Finished Products per Component	Safety Inventory	Marginal Reduction in Safety Inventory	Total Reduction in Safety Inventory
1	399,699		
2	282,630	117,069	117,069
3	230,766	51,864	168,933
4	199,849	30,917	199,850
5	178,751	21,098	220,948
6	163,176	15,575	236,523
7	151,072	12,104	248,627
8	141,315	9,757	258,384
9	133,233	8,082	266,466

In the case of component commonality, each component ends up in nine finished products. Therefore, the demand at the component level is the sum of demand across nine products. Using Equations 12.15 and 12.16, the safety inventory required for each component is thus

$$\text{Safety inventory per common component} = NORMSINV(0.95) \times \sqrt{1} \times \sqrt{9} \times 3{,}000$$
$$= 14{,}803.68 \text{ units}$$

With component commonality, there are a total of nine distinct components. The total safety inventory across all nine components is thus

$$\text{Total safety inventory required} = 9 \times 14{,}803.68 = 133{,}233$$

Thus, having each component common to nine products results in a reduction in safety inventory for Dell from 399,699 to 133,233 units.

In Table 12-5, we evaluate the marginal benefit in terms of reduction in safety inventory as a result of increasing component commonality. Starting with the required safety inventory when each component is used in only one finished product, we evaluate the safety inventory as the number of products in which a component is used increases to nine. Observe that component commonality decreases the required safety inventory for Dell. The marginal benefit of commonality, however, declines as a component is used in more and more finished products.

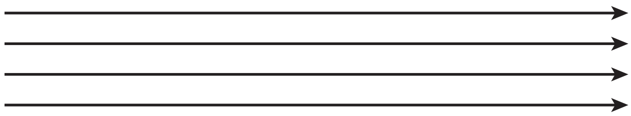
As a component is used in more finished products, it must be more flexible. As a result, the cost of producing the component typically increases with increasing commonality. Given that the marginal benefit of component commonality decreases as we increase commonality, we need to trade off the increase in component cost and the decrease in safety inventory when deciding on the appropriate level of component commonality.

Key Point

Component commonality decreases the safety inventory required. The marginal benefit, however, decreases with increasing commonality.

Postponement

Postponement is the ability of a supply chain to delay product differentiation or customization until closer to the time the product is sold. The goal is to have common components in the supply chain for most of the push phase and move product differentiation as close to the pull phase of



Supply Chain Flows without Postponement

Supply Chain Flows with Component Commonality and Postponement

FIGURE 12-6 Supply Chain Flows without and with Postponement

the supply chain as possible. For example, the final mixing of paint today is done at the retail store after the customer has selected the color he or she wants. Thus, paint variety is produced only when demand is known with certainty. Postponement coupled with component commonality allows paint retailers to carry significantly lower safety inventories than in the past, when mixing was done at the paint factory. In the past, the factory manager had to forecast paint demand by color when planning production. Today, a factory manager needs to forecast only aggregate paint demand because mixing has been postponed until after customer demand is known. As a result, each retail store primarily carries aggregate inventory in the form of base paint that is configured to the appropriate color based on customer demand.

Another classic example of postponement is the production process at Benetton to make colored knit garments. The original process called for the thread to be dyed and then knitted and assembled into garments. The entire process required up to six months. Because the color of the final garment was fixed the moment the thread was dyed, demand for individual colors had to be forecast far in advance (up to six months). Benetton developed a manufacturing technology that allowed it to dye knitted garments to the appropriate color. Now, *greige* thread (the term used for thread that has not yet been dyed) can be purchased, knitted, and assembled into garments before dyeing. The dyeing of the garments is done much closer to the selling season. In fact, part of the dyeing is done after the start of the selling season, when demand is known with great accuracy. In this case, Benetton has postponed the color customization of the knit garments. When thread is purchased, only the aggregate demand across all colors needs to be forecast. Given that this decision is made far in advance, when forecasts are least likely to be accurate, there is great advantage to this aggregation. As Benetton moves closer to the selling season, the forecast uncertainty reduces. At the time Benetton dyes the knit garments, demand is known with a high degree of accuracy. Thus, postponement allows Benetton to exploit aggregation and significantly reduce the level of safety inventory carried. Supply chain flows with and without postponement are illustrated in Figure 12-6.

Without component commonality and postponement, product differentiation occurs early on in the supply chain, and most of the supply chain inventories are disaggregate. Postponement allows the supply chain to delay product differentiation. As a result, most of the inventories in the supply chain are aggregate. Postponement thus allows a supply chain to exploit aggregation to reduce safety inventories without hurting product availability. We illustrate the benefits of postponement in Example 12-12 (see worksheet *Example 12-12*). A more nuanced discussion of the value of postponement is given in Chapter 13.

EXAMPLE 12-12 Value of Postponement

Consider a paint retailer that sells 100 different colors of paint. Assume that weekly demand for each color is independent and is normally distributed with a mean of 30 and a standard deviation of 10. The replenishment lead time from the paint factory is two weeks and the retailer aims for a CSL = 0.95. How much safety stock will the retailer have to hold if paint is mixed at the factory and held in inventory at the retailer as individual colors? How does the safety stock requirement change if the retailer holds base paint (supplied by the paint factory) and mixes colors on demand?

Analysis:

We first evaluate the disaggregate option without postponement, in which the retailer holds safety inventory for each color sold. For each color, we have

$$D = 30/\text{week}, \sigma_D = 10, L = 2 \text{ weeks}$$

Given the desired CSL = 0.95, the required safety inventory across all 100 colors is obtained using Equation 12.12 to be

$$\text{Total required safety inventory, } ss = 100 \times F_S^{-1}(CSL) \times \sqrt{L} \times \sigma_D$$
$$= 100 \times NORMSINV(0.95) \times \sqrt{2} \times 10 = 2{,}326$$

Now, consider the option whereby mixing is postponed until after the customer orders. Safety inventory is held in the form of base paint, whose demand is an aggregate of demand of the 100 colors. Because demand in all 100 colors is independent, $\rho = 0$. Using Equation 12.15, the standard deviation of aggregate weekly demand of base paint is

$$\text{Standard deviation of weekly demand of base paint, } \sigma_D^C = \sqrt{100} \times 10 = 100$$

For a CSL of 0.95, safety inventory required for the aggregate option (using Equation 12.16) is given as

$$ss = F_S^{-1}(0.95) \times \sqrt{L} \times \sigma_D^C = NORMSINV(0.95) \times \sqrt{2} \times 100 = 233$$

Observe that postponement reduces the required safety inventory at the paint retailer from 2,326 units to 233 units.

Postponement can be a powerful concept when customers are willing to wait a little for their orders to arrive. This delay offers the supply chain an opportunity to reduce inventories by postponing product differentiation until after the customer order arrives. It is important that the manufacturing process be designed in a way that enables assembly to be completed quickly. Given that customers are often willing to wait for delivery, several furniture and window manufacturers have postponed some of the assembly processes for their products.

12.6 IMPACT OF REPLENISHMENT POLICIES ON SAFETY INVENTORY

In this section, we describe the evaluation of safety inventories for both continuous and periodic-review replenishment policies. We highlight the fact that periodic review policies require more safety inventory than continuous review policies for the same level of product availability. To simplify the discussion, we focus on the CSL as the measure of product availability. The managerial implications are the same if we use fill rate; the analysis, however, is more cumbersome.

Continuous Review Policies

Given that continuous review policies were discussed in detail in Section 12.2, we reiterate only the main points here. When using a continuous review policy, a manager orders Q units when the inventory drops to the ROP. Clearly, a continuous review policy requires technology that monitors the level of available inventory. This is the case for many firms such as Walmart and Dell, whose inventories are monitored continuously.

Given a desired CSL, our goal is to identify the required safety inventory *ss* and the ROP. We assume that demand is normally distributed, with the following inputs:

D: Average demand per period

σ_D: Standard deviation of demand per period

L: Average lead time for replenishment

The ROP represents the available inventory to meet demand during the lead time L. A stockout occurs if the demand during the lead time is larger than the ROP. If demand across periods is independent, demand during the lead time is normally distributed with the following:

$$\text{Mean demand during lead time, } D_L = D \times L$$
$$\text{Standard deviation of demand during lead time, } \sigma_L = \sqrt{L}\sigma_D$$

Given the desired CSL, the required safety inventory (*ss*) obtained using Equation 12.5 and the ROP obtained using Equation 12.3 are

$$ss = F_S^{-1}(CSL) \times \sigma_L = NORMSINV(CSL) \times \sqrt{L}\sigma_D, ROP = D_L + ss$$

A manager using a continuous review policy has to account only for the uncertainty of demand during the lead time. This is because the continuous monitoring of inventory allows the manager to adjust the timing of the replenishment order, depending on the demand experienced. If demand is very high, inventory reaches the ROP quickly, leading to a quick replenishment order. If demand is very low, inventory drops slowly to the ROP, leading to a delayed replenishment order. The manager, however, has no recourse during the lead time once a replenishment order has been placed. The available safety inventory thus must cover for the uncertainty of demand over this period.

Typically, in continuous review policies, the lot size ordered is kept fixed between replenishment cycles. The optimal lot size may be evaluated using the EOQ formula discussed in Chapter 11.

Periodic Review Policies

In periodic review policies, inventory levels are reviewed after a fixed period of time T and an order is placed such that the level of current inventory plus the replenishment lot size equals a prespecified level called the *order-up-to level* (OUL). The *review interval* is the time T between successive orders. Observe that the size of each order may vary, depending on the demand experienced between successive orders and the resulting inventory at the time of ordering. Periodic review policies are simpler for retailers to implement because they do not require that the retailer have the capability of monitoring inventory continuously. Suppliers may also prefer them because they result in replenishment orders placed at regular intervals.

Let us consider the store manager at Walmart who is responsible for designing a replenishment policy for Lego building blocks. He wants to analyze the impact on safety inventory if he decides to use a periodic review policy. Demand for Legos is normally distributed and independent from one week to the next. We assume the following inputs:

D: Average demand per period

σ_D: Standard deviation of demand per period

L: Average lead time for replenishment

T: Review interval

CSL: Desired cycle service level

To understand the safety inventory requirement, we track the sequence of events over time as the store manager places orders. The store manager places the first order at time 0 such that the lot size ordered and the inventory on hand sum to the OUL. Once an order is placed, the replenishment lot arrives after the lead time L. The next review period is time T, when the store manager places the next order, which then arrives at time $T + L$. The OUL represents the inventory available to meet all demand that arises between periods 0 and $T + L$. The Walmart store will experience a stockout if demand during the time interval between 0 and $T + L$ exceeds the OUL. Thus, the store manager must identify an OUL such that the following is true:

$$\text{Probability}(\text{demand during } T + L \leq OUL) = CSL$$

The next step is to evaluate the distribution of demand during the time interval $T + L$. Using Equation 12.2, demand during the time interval $T + L$ is normally distributed, with

$$\text{Mean demand during } T + L \text{ periods, } D_{T+L} = (T + L)D$$
$$\text{Standard deviation of demand during } T + L \text{ periods, } \sigma_{T+L} = \sqrt{T + L}\,\sigma_D$$

The safety inventory in this case is the quantity in excess of D_{T+L} carried by Walmart over the time interval $T + L$. The OUL and the safety inventory ss are related as follows:

$$OUL = D_{T+L} + ss \qquad \textbf{(12.18)}$$

Given the desired CSL, the safety inventory (ss) required is given by

$$ss = F_S^{-1}(CSL) \times \sigma_{T+L} = NORMSINV(CSL) \times \sigma_{T+L} \qquad \textbf{(12.19)}$$

The average lot size equals the average demand during the review period T and is given as

$$\text{Average lot size, } Q = D_T = D \times T \qquad \textbf{(12.20)}$$

In Figure 12-7, we show the inventory profile for a periodic review policy with lead time $L = 4$ and reorder interval $T = 7$. On day 7, the company places an order that determines available inventory until day 18 (as illustrated in Figure 12-7 by the dashed line from point 1 to point 2). As a result, the safety inventory must be sufficient to buffer demand variability over $T + L = 7 + 4 = 11$ days.

We illustrate the periodic review policy for Walmart in Example 12-13 (see worksheet *Example 12-13*).

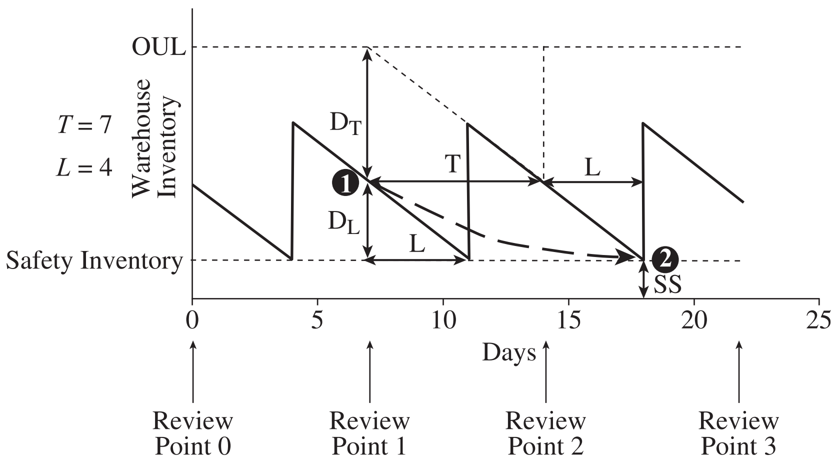


FIGURE 12-7 Inventory Profile for Periodic Review Policy with $L = 4$, $T = 7$

EXAMPLE 12-13 Evaluation Safety Inventory for a Periodic Review Policy

Weekly demand for Legos at a Walmart store is normally distributed, with a mean of 2,500 boxes and a standard deviation of 500. The replenishment lead time is two weeks, and the store manager has decided to review inventory every four weeks. Assuming a periodic-review replenishment policy, evaluate the safety inventory that the store should carry to provide a CSL of 90 percent. Evaluate the OUL for such a policy.

Analysis:

In this case, we have

Average demand per period, $D = 2{,}500$

Standard deviation of demand per period, $\sigma_D = 500$

Average lead time for replenishment, $L = 2$ weeks

Review interval, $T = 4$ weeks

We first obtain the distribution of demand during the time interval $T + L$. Using Equation 12.2, demand during the time interval $T + L$ is normally distributed, with

Mean demand during $T + L$ periods, $D_{T+L} = (T + L)D = (2 + 4)2{,}500 = 15{,}000$

Standard deviation of demand during $T + L$ periods, $\sigma_{T+L} = \sqrt{T + L}\sigma_D$

$$= (\sqrt{4 + 2})500 = 1{,}225$$

From Equation 12.19, the required safety inventory for a $CSL = 0.90$ is given as

$$ss = F_S^{-1}(CSL) \times \sigma_{T+L} = NORMSINV(CSL) \times \sigma_{T+L}$$

$$= NORMSINV(0.90) \times 1{,}225 = 1{,}570 \text{ boxes}$$

Using Equation 12.18, the OUL is given by

$$OUL = D_{T+L} + ss = 15{,}000 + 1{,}570 = 16{,}570$$

The store manager thus orders the difference between 16,570 and current inventory every four weeks.

We can now compare the safety inventory required when using continuous and periodic review policies. With a continuous review policy, the safety inventory is used to cover for demand uncertainty over the lead time L. With a periodic review policy, the safety inventory is used to cover for demand uncertainty over the lead time and the review interval $L + T$. Given that higher uncertainty must be accounted for, periodic review policies require a higher level of safety inventory. This argument can be confirmed by comparing the results in Examples 12-4 and 12-13. For a 90 percent CSL, the store manager requires a safety inventory of 906 boxes when using a continuous review and a safety inventory of 1,570 boxes when using a periodic review.

Key Point

Periodic review replenishment policies require more safety inventory than continuous review policies for the same lead time and level of product availability.

Of course, periodic review policies are somewhat simpler to implement because they do not require continuous tracking of inventory. Given the broad use of bar codes and POS systems as well as the emergence of RFID technology, continuous tracking of all inventories is much more commonplace today than it was a decade ago. In some instances, companies partition their products based on their value. High-value products are managed using continuous review policies, and low-value products are managed using periodic review policies. This makes sense if the cost of perpetual tracking of inventory is more than the savings in safety inventory that result from switching all products to a continuous review policy.

12.7 MANAGING SAFETY INVENTORY IN A MULTIECHELON SUPPLY CHAIN

In our discussion so far, we have assumed that each stage of the supply chain has a well-defined demand and supply distribution that it uses to set its safety inventory levels. In practice, this is not true for multiechelon supply chains. Consider a simple multiechelon supply chain with a supplier feeding a retailer that sells to the final customer. The retailer needs to know demand as well as supply uncertainty to set safety inventory levels. Supply uncertainty, however, is influenced by the level of safety inventory the supplier chooses to carry. If a retailer order arrives when the supplier has enough inventory, the supply lead time is short. In contrast, if the retailer order arrives when the supplier is out of stock, the replenishment lead time for the retailer increases. Thus, if the supplier increases its level of safety inventory, the retailer can reduce the safety inventory it holds. This implies that the level of safety inventory at all stages in a multiechelon supply chain should be related.

All inventory between a stage and the final customer is called the *echelon inventory*. Echelon inventory at a retailer is just the inventory at the retailer or in the pipeline coming to the retailer. Echelon inventory at a distributor, however, includes inventory at the distributor and all retailers served by the distributor. In a multiechelon setting, ROPs and OULs at any stage should be based on echelon inventory and not local inventory. Thus, a distributor should decide its safety inventory levels based on the level of safety inventory carried by all retailers supplied by it. The more safety inventory retailers carry, the less safety inventory the distributor needs to carry. As retailers decrease the level of safety inventory they carry, the distributor must increase its safety inventory to ensure regular replenishment at the retailers.

If all stages in a supply chain attempt to jointly manage their echelon inventory, the issue of how the inventory is divided among various stages becomes important. Carrying inventory upstream in a supply chain allows for more aggregation and thus reduces the amount of inventory required. Carrying inventory upstream, however, increases the probability that the final customer will have to wait because product is not available at a stage close to him or her. Thus, in a multiechelon supply chain, a decision must be made with regard to the level of safety inventory carried at different stages. If inventory is expensive to hold and customers are willing to tolerate a delay, it is better to increase the amount of safety inventory carried upstream, far from the final customer, to exploit the benefits of aggregation. If inventory is inexpensive to hold and customers are time sensitive, it is better to carry more safety inventory downstream, closer to the final customer.

12.8 THE ROLE OF IT IN INVENTORY MANAGEMENT

Besides the basics of formalizing inventory replenishment procedures for thousands of SKUs, the two most significant contributions of IT systems can be improved inventory visibility and better coordination in the supply chain.

An excellent example of the benefits of improved inventory visibility is Nordstrom, a department store chain in the United States. The company was always very good at managing its inventories (IT systems played an important role here) but had historically separated its online inventories and its store inventories. In September 2009, the company started integrating store inventories onto its website. Customers are now able to access inventory no matter where it is available. If they prefer home delivery, Nordstrom can now use store inventory to serve them. If, however, they prefer to pick up the item themselves, Nordstrom allows them to reserve it for pickup. The increased inventory visibility allows Nordstrom to serve its online customers better while also drawing more traffic to stores. In 2010, Walmart also added a similar feature, called "Pick Up Today," which allows customers to place orders online and pick them up a few hours later at a retail store. Customers are alerted (typically through a text message) when the order is ready. Redbox uses inventory visibility at each of its vending machines to guide customers to the closest kiosk that has the desired DVD in stock. In each example, the increased visibility provided by IT systems allows the firm to improve product availability to the customer without increasing inventories.

Another area in which improved visibility could play a significant role is locating in-store or in-warehouse inventory. Often, a store or warehouse has inventory available but in the wrong place. The net result is a loss in product availability despite carrying inventory. Good RFID systems have the potential to address this issue. Although there has been limited success using RFID systems at the item level in stores (there has been some success with high-value apparel), there has been success in areas such as warehousing of aircraft spare parts.

IT systems have also played a significant role in better integrating different stages of the supply chain. A classic example is the continuous replenishment program (CRP) set up between Procter and Gamble (P&G) and Walmart that allowed P&G to replenish diaper inventory at Walmart based on the visibility of available inventories and sales at Walmart. This coordination allowed the two firms to improve service levels while reducing inventories. Over time, the program evolved into collaborative planning, forecasting, and replenishment (discussed in greater detail in Chapter 10), which allows better coordination of planning and replenishment across multiple supply chain partners through improved visibility of inventories and sales. Even though each of these programs uses IT as a foundation, it is important to acknowledge that success requires important organizational changes and leadership commitment as discussed in Chapter 10. Good IT systems are a necessary but not a sufficient condition for success.

It is important to recognize that the value of the IT system in each of the cases discussed here is tightly linked to the accuracy of the inventory information. Inaccurate inventory information leads to flawed decisions and could in the worst case create mistrust among supply chain partners attempting to coordinate the decisions and actions. A study by DeHoratius and Raman (2008) found that about 65 percent of the inventory records checked for a retailer were inaccurate. That is, for 65 percent of the records checked, the inventory on hand did not match the inventory showing in the IT system. Without reasonably accurate inventory records, the value provided by an IT system will be limited.

12.9 ESTIMATING AND MANAGING SAFETY INVENTORY IN PRACTICE

1. ***Account for the fact that supply chain demand is lumpy.*** In practice, a manufacturer or distributor does not order one unit at a time but instead often orders in a large lot. Thus, demand observed by different stages of the supply chain tends to be lumpy. Lumpiness adds to the variability of demand. For example, when using a continuous review policy, lumpiness may lead to inventory dropping far below the ROP before a replenishment order is placed. On average, inventory will drop below the ROP by half the average size of an order. The lumpiness can be accounted for in practice by raising the safety inventory suggested by the models discussed earlier by half the average size of an order.

2. ***Adjust inventory policies if demand is seasonal.*** In practice, demand is often seasonal, with the mean and the standard deviation of demand varying by the time of year. Thus, a given ROP or OUL may correspond to ten days of demand during the low-demand season and only two days of demand during the peak demand season. If the lead time is one week, stockouts are certain to occur during the peak season. In the presence of seasonality, it is not appropriate to select an average demand and standard deviation over the year to evaluate fixed ROPs and OULs. Both the mean and the standard deviation of demand must be adjusted by the time of year to reflect changing demand. Corresponding adjustments in the ROPs, OULs, and safety inventories must be made over the year. It is best to evaluate all inventory parameters such as ROPs and OULs in terms of days of demand. A simple heuristic that keeps days of demand constant over time helps account for seasonality by automatically adjusting the ROP and OUL.

3. *Use simulation to test inventory policies.* Given that demand is most likely not normally distributed and may be seasonal, it is a good idea to use a computer simulation to test and adjust inventory policies before they are implemented. The simulation should use a demand pattern that truly reflects actual demand, including any lumpiness as well as seasonality. The inventory policies obtained using the models discussed in the chapter can then be tested and adjusted if needed to obtain the desired service levels. Surprisingly powerful simulations can be built using Excel, as we discuss in Chapter 13. Identifying problems in a simulation can save a lot of time and money compared to facing these problems once the inventory policy is in place.

4. *Start with a pilot.* Even a simulation cannot identify all problems that may arise when using an inventory policy. Once an inventory policy has been selected and tested using simulation, it is often a good idea to start implementation with a pilot program of products that are representative of the entire set of products in inventory. By starting with a pilot, many of the problems (both in the inventory policies themselves and in the process of applying the policies) can be solved. Getting these problems solved before the policy is rolled out to all the products can save a lot of time and money.

5. *Monitor service levels.* Once an inventory policy has been implemented, it is important that its performance be tracked and monitored. Monitoring is crucial because it allows a supply chain to identify when a policy is not working well and make adjustments before supply chain performance is affected significantly. Monitoring requires not just tracking the inventory levels but also tracking any stockouts that may result. Historically, firms have not tracked stockouts very well, partly because they are difficult to track and partly because of the perception that stockouts affect the customer but not the firm itself. Stockouts can be difficult to measure in a situation such as a supermarket, where the customer simply does not buy the product when it is not on the shelf. However, there are simple ways to estimate stockouts. At a supermarket, the fraction of time that a shelf does not contain a product may be used to estimate the fill rate. Stockouts are in fact easier to estimate online, where the number of clicks on an out-of-stock product can be measured. Given the fraction of clicks that turn into orders and the average size of an order, demand during a stockout can be estimated.

6. *Focus on reducing safety inventories.* Given that safety inventory is often a large fraction of the total inventory in a supply chain, the ability to reduce safety inventory without hurting product availability can significantly increase supply chain profitability. This is particularly important in the high-tech industry, in which product life cycles are short. In this chapter, we discussed a variety of managerial levers that can help reduce safety inventories without hurting availability. Supply chain managers must focus continuously on using these levers to reduce safety inventories.

12.10 SUMMARY OF LEARNING OBJECTIVES

1. *Describe different measures of product availability.* The three basic measures of product availability are product fill rate, order fill rate, and cycle service level. Product fill rate is the fraction of demand for a product that is filled from inventory. Order fill rate is the fraction of orders that are completely filled. Cycle service level is the fraction of replenishment cycles in which no stockouts occur.

2. *Understand the role of safety inventory in a supply chain.* Safety inventory helps a supply chain provide customers with a high level of product availability in spite of supply and demand uncertainty. It is carried just in case demand exceeds the amount forecasted or supply arrives later than expected.

3. ***Identify factors that influence the required level of safety inventory.*** Safety inventory is influenced by demand uncertainty, replenishment lead times, lead time variability, and desired product availability. As any one of them increases, the required safety inventory also increases. The required safety inventory is also influenced by the inventory policy implemented. Continuous review policies require less safety inventory than periodic review policies.

4. ***Use available managerial levers to lower safety inventory without hurting product availability.*** The required level of safety inventory may be reduced and product availability may be improved if a supply chain can reduce demand uncertainty, replenishment lead times, and the variability of lead times. A switch from periodic monitoring to continuous monitoring can also help reduce inventories. Another key managerial lever to reduce the required safety inventories is to exploit aggregation. This may be achieved by physically aggregating inventories, virtually aggregating inventories using information centralization, specializing inventories based on demand volume, exploiting substitution, using component commonality, and postponing product differentiation.

Discussion Questions

1. Is there any impact on safety inventory level if the customer gets more market information?
2. As a supermarket operator, would you hold more safety inventory for a newly introduced detergent or an existing detergent? Why?
3. What are the pros and cons of the various measures of product availability?
4. Describe the two types of ordering policies and the impact that each of them has on safety inventory.
5. Suggest how to reduce the safety inventory level without hurting product availability.
6. Why can a Home Depot, with a few large stores, provide a higher level of product availability with lower inventories than a hardware store chain such as True Value, with many small stores?
7. "It is well-known that inventory aggregation could always reduce demand uncertainty and therefore result in a reduction in required safety inventory level." Do you agree or disagree, why?
8. In the 1980s, paint was sold by color and size in paint stores. Today, paint is mixed at the paint store according to the color required. Discuss what, if any, impact this change has on safety inventories in the supply chain.
9. A new technology allows books to be printed in 10 minutes. Barnes & Noble has decided to purchase these machines for each store. It must decide which books to carry in stock and which books to print on demand using this technology. Do you recommend it for best sellers or for other books? Why?
10. Consider a firm like Zara that has developed production capabilities with very short replenishment lead times. Do you think this capability is more valuable for its online operations or its store operation? Why?
11. What are the different impacts of continuous review policy and periodic review policy on the requirement of safety inventory level?
12. What capabilities can local suppliers in high-cost countries develop if they are to effectively compete against overseas suppliers in low-cost countries? Discuss how each capability impacts the level of inventory in the supply chain.

Exercises

1. Weekly demand for smartphones at an Apple store is normally distributed, with a mean of 500 and a standard deviation of 300. Foxconn, the assembler, takes four weeks to supply an Apple order. Apple is targeting a CSL of 95 percent and monitors its inventory continuously. How much safety inventory of cell phones should the Apple store carry? What should its ROP be?
2. Weekly demand for jeans at a Gap store is normally distributed, with a mean of 100 and a standard deviation of 50. The supply plant takes three weeks to supply a Gap order. The store manager monitors its inventory continuously and reorders jeans when the available inventory drops below 350. How much safety stock does the store carry? What CSL does the store achieve? If the store manager wants to target a CSL of 95 percent, how much safety inventory of jeans should the store carry? What should its ROP be?
3. Weekly demand for Motorola cell phones at a Best Buy store is normally distributed, with a mean of 300 and a standard deviation of 200. Motorola takes two weeks to supply a Best Buy order. Best Buy is targeting a CSL of 95 percent and monitors its inventory continuously. How much safety inventory of cell phones should Best Buy carry? What should its ROP be?
4. Reconsider the Best Buy store in Exercise 3. The store manager has decided to follow a periodic review policy to manage inventory of cell phones. She plans to order every three weeks. Given a desired CSL of 95 percent, how much safety inventory should the store carry? What should its OUL be?

5. Assume that the Best Buy store in Exercise 3 has a policy of ordering cell phones from Motorola in lots of 500. Weekly demand for Motorola cell phones at the store is normally distributed, with a mean of 300 and a standard deviation of 200. Motorola takes two weeks to supply an order. If the store manager is using a continuous review policy and targeting a fill rate of 99 percent, what safety inventory should the store carry? What should its ROP be?

6. Weekly demand for HP printers at a Sam's Club store is normally distributed, with a mean of 250 and a standard deviation of 150. The store manager continuously monitors inventory and currently orders 1,000 printers each time the inventory drops to 600 printers. HP currently takes two weeks to fill an order. How much safety inventory does the store carry? What CSL does Sam's Club achieve as a result of this policy? What fill rate does the store achieve?

7. Return to the Sam's Club store in Exercise 6. Assume that the supply lead time from HP is normally distributed, with a mean of 2 weeks and a standard deviation of 1.5 weeks. How much safety inventory should Sam's Club carry if it wants to provide a CSL of 95 percent? How does the required safety inventory change as the standard deviation of lead time is reduced from 1.5 weeks to zero in intervals of 0.5 weeks?

8. Weekly demand for detergent at a Walmart store is normally distributed, with a mean of 3,000 and a standard deviation of 700. The store manager continuously monitors detergent inventory and places a replenishment order for 10,000 units when there are 7,000 units in inventory. The supplier takes two weeks to supply a Walmart order. What CSL does the store achieve? What fill rate does the store achieve?

9. Weekly demand for paper towels at a Target store is normally distributed, with a mean of 1,000 and a standard deviation of 300. The supplier takes two weeks to supply a Target order, which is for a batch size of 5,000. Target is aiming for a fill rate of 99 percent and monitors its inventory continuously. How much safety inventory of paper towels should Target carry? What should its ROP be?

10. Weekly demand for electric motors at a Japanese motor manufacturer is normally distributed, with a mean of 1,000 and a standard deviation of 1,000. Motors are currently assembled in China and delivered at a cost of 20,000 yen/motor. The supplier takes eight weeks to supply an order. A local Japanese manufacturer has offered to deliver motors with a lead time of one week at a cost of 20,400 yen per motor. The motor manufacturer is targeting a CSL of 99 percent and monitors its inventory continuously. The manufacturer incurs a holding cost of 25 percent. Should the manufacturer accept the local supplier's offer?

11. Weekly demand for private label washing machines at Karstadt, a German department store chain, is normally distributed with a mean of 500 and a standard deviation of 300. Karstadt currently has a supply source in China that delivers machines at a cost of 200 euro. The lead time required by the supplier is normally distributed with a mean of nine weeks and a standard deviation of six weeks. A European supplier has offered to deliver washing machines with a guaranteed lead time of one week at a cost of 210 euro. Karstadt has a holding cost of 25 percent and targets a cycle service level of 99 percent. Should Karstadt accept the local supplier's offer?

12. Croma is an Indian retail chain for consumer electronics. The company currently has 25 stores located in major metropolitan areas. Weekly demand for smartphones at each store is normally distributed, with a mean of 300 and a standard deviation of 300. The supplier currently takes four weeks to fulfill a replenishment order, which is placed separately by each store. Croma is targeting a CSL of 95 percent and monitors its inventory continuously. How much safety inventory of smartphones should Croma carry at each retail store? Croma is considering moving smartphones to the online channel, where they would be stocked in a single national warehouse. Assume that Croma can move smartphones to the online channel without losing demand (the online demand is a sum of demand at each retail store). How much saving in safety inventory can Croma expect from going online if demand across stores is independent? How much saving in safety inventory can Croma expect from going online if demand across stores has a correlation coefficient of $\rho = 0.5$?

13. Magazine Luiza is a Brazilian retail chain for consumer electronics. The company currently has 100 stores distributed across Brazil. It also operates an online channel. It is considering introduction of a new printer and must decide whether to offer it at retail stores or the online channel. Weekly demand for the printer at each store has been forecast to be normally distributed with a mean of 100 and a standard deviation of 80. The company has also forecast that the demand at the online channel would be the sum of demand across all 100 stores. The supplier will take four weeks to fulfill a replenishment order, whether placed separately by each store or by the online DC. Magazine Luiza is targeting a CSL of 95 percent and monitors its inventory continuously. How much safety inventory will Magazine Luiza carry if the printer is carried at all 100 stores? How much safety inventory will Magazine Luiza carry if the printer is carried online and demand across stores is independent? How much safety inventory will Magazine Luiza carry if the printer is carried online and demand across stores has a correlation coefficient of $\rho = 0.3$?

14. Gap has started selling through its online channel along with its retail stores. Management has to decide which products to carry at the retail stores and which products to carry at a central warehouse to be sold only via the online channel. Gap currently has 900 retail stores in the United States. Weekly demand for size large khaki pants at each store is normally distributed, with a mean of 800 and a standard deviation of 100. Each pair of pants costs \$30. Weekly demand for purple cashmere sweaters at each store is normally distributed, with a mean of 50 and a standard deviation of 50. Each sweater costs \$100. Gap has a holding cost of 25 percent. Gap manages all inventories using a continuous review policy, and the supply lead time for both products is four weeks. The targeted CSL is 95 percent. How much reduction in holding cost per unit sold can Gap expect on moving each of the two products from the stores to the online channel? Which of the two products should Gap carry at the stores, and which should it carry at the

TABLE 12-6 Weekly Demand for Epson Printers in Europe

Country	Mean Demand	Standard Deviation
France	3,000	2,000
Germany	4,000	2,200
Spain	2,000	1,400
Italy	2,500	1,600
Portugal	1,000	800
UK	4,000	2,400

central warehouse for the online channel? Why? Assume demand from one week to the next to be independent.

15. Epson produces printers in its Taiwan factory for sale in Europe. Printers sold in different countries differ in terms of the power outlet as well as the language of the manuals. Currently, Epson assembles and packs printers for sale in individual countries. The distribution of weekly demand in different countries is normally distributed, with means and standard deviations as shown in Table 12-6.

Assume demand in different countries to be independent. Given that the lead time from the Taiwan factory is eight weeks, how much safety inventory does Epson require in Europe if it targets a CSL of 95 percent?

Epson decides to build a central DC in Europe. It will ship base printers (without power supply) to the DC. When an order is received, the DC will assemble power supplies, add manuals, and ship the printers to the appropriate country. The base printers are still to be manufactured in Taiwan with a lead time of eight weeks. How much saving of safety inventory can Epson expect as a result?

16. Return to the Epson data in Exercise 15. Each printer costs Epson $200, and the holding cost is 25 percent. What saving in holding cost can Epson expect as a result of building the European DC? If final assembly in the European DC adds $5 to the production cost of each printer, would you recommend the move? Suppose that Epson is able to cut the production and delivery lead time from its Taiwan factory to four weeks using good information systems. How much savings in holding cost can Epson expect without the European DC? How much savings in holding cost can the firm expect with the European DC?

17. Return to the Epson data in Exercise 15. Assume that demand in different countries is not independent. Demand in any pair of countries is correlated with a correlation coefficient of ρ. Evaluate the holding cost savings that Epson gains as a result of building a European DC as ρ increases from 0 (independent demand) to 1 (perfectly positively correlated demand) in intervals of 0.2.

18. Motorola obtains cell phones from its contract manufacturer located in China to supply the U.S. market, which is served from a warehouse located in Memphis, Tennessee. Daily demand at the Memphis warehouse is normally distributed, with a mean of 5,000 and a standard deviation of 4,000. The warehouse aims for a CSL of 99 percent. The company is debating whether to use sea or air transportation from China. Sea transportation results in a lead time of 36 days and costs $0.50 per phone. Air transportation results in a lead time of 4 days and costs $1.50 per phone. Each phone costs $100, and Motorola uses a holding cost of 20 percent. Given the minimum lot sizes, Motorola would order 100,000 phones at a time (on average, once every 20 days) if using sea transport and 5,000 phones at a time (on average, daily) if using air transport. To begin with, assume that Motorola takes ownership of the inventory on delivery.

a. Assuming that Motorola follows a continuous review policy, what reorder point and safety inventory should the warehouse aim for when using sea or air transportation? How many days of safety and cycle inventory will Motorola carry under each policy?

b. How many days of cycle inventory does Motorola carry under each policy?

c. Under a continuous review policy, do you recommend sea or air transportation if Motorola does not own the inventory while it is in transit? Does your answer change if Motorola has ownership of the inventory while it is in transit?

19. Weekly demand for gaming consoles at Liverpool, a Mexican department store chain, is normally distributed with a mean of 1,000 and a standard deviation of 400. The replenishment lead time from the supplier is four weeks. Liverpool is targeting a CSL of 95 percent and uses a periodic review policy under which it reorders consoles every eight weeks. What is the average order size? How much safety inventory of consoles should Liverpool carry? What should its order up to level be? How much safety inventory would be required if Liverpool switched to a continuous review policy?

20. Weekly demand for handbags at Liverpool, a Mexican department store chain, is normally distributed with a mean of 3,000 and a standard deviation of 1,000. The replenishment lead time from the supplier is 4 weeks. Liverpool uses a periodic review policy under which it reorders handbags every 12 weeks. It currently uses an order-up-to level of 50,000. What is the average order size? How much safety inventory of handbags does Liverpool carry? What CSL does it achieve? What order up to level should it use if it wants a CSL of 99 percent?

21. Return to the problem data in Exercise 18. Assume that Motorola follows a periodic review policy. Given lot sizes by sea and air, Motorola would have to place an order every 20 days using sea transport but could order daily using air transport.

a. Assume that Motorola follows a periodic review policy. What OUL and safety inventory should the warehouse aim for when using sea or air transportation? How many days of safety inventory will Motorola carry under each policy?

b. How many days of cycle inventory does Motorola carry under each policy?

c. Under a periodic review policy, do you recommend sea or air transportation? Does your answer change if Motorola has ownership of the inventory while it is in transit?

22. DoorRed Pharmacy replenishes one of its best-selling drugs using a continuous review policy. Daily demand for the drug is normally distributed, with a mean of 300 and a standard deviation of 100. The wholesaler can process a replenishment request in two days. The current replenishment policy is to order 1,500 units when there are 750 units on hand.

a. What is the cycle service level that DoorRed achieves with its policy?
b. What is the fill rate that DoorRed achieves with its policy?
c. What change in fill rate would DoorRed achieve if it increased its ROP from 750 to 800?

23. Return to the DoorRed Pharmacy in Exercise 22. For the drug under discussion, DoorRed wants to adjust its ROP from 750 to achieve a fill rate of 99.9 percent. What ROP should it use?

24. The DoorRed pharmacy has twenty-five retail outlets in the Chicago region. The current policy is to carry every drug in each retail outlet. DoorRed is investigating the possibility of centralizing some of the drugs in one location. Delivery charge would increase by $0.02 per unit if a drug were centralized. The increase in delivery charge comes from the additional cost of operating the shuttle from the central location to each of the other locations. At each retail outlet, DoorRed follows a periodic review policy with weekly replenishment (a replenishment order is placed once every seven days). The replenishment lead time is three days. DoorRed plans to stick to once-a-week ordering even if a drug is centralized. DoorRed uses an inventory holding cost of 20 percent and aims for a cycle service level of 99 percent. Assume that demand across stores is independent.

a. Consider a drug with daily demand at each store that is normally distributed, with a mean of 300 and a standard deviation of 50. The drug costs $10 per unit. What is the annual holding cost of safety inventory across all retail stores? If the drug were centralized in one location, what would the annual cost of holding safety inventory at the central location be? What would the annual increase in delivery charge be? Do you recommend centralization?
b. Now consider a drug with daily demand at each store that is normally distributed, with a mean of 5 and a standard deviation of 4. The drug costs $10 per unit. What is the annual holding cost of safety inventory across all retail stores? If the drug were centralized in one location, what would the annual cost of holding safety inventory at the central location be? What would the annual increase in delivery charge be? Do you recommend centralization?
c. Do your answers to (a) and (b) change if the demand across stores has a correlation coefficient of 0.5?

25. Toyota has decided to set up regional warehouses where some variants of the Scion will be customized and shipped to dealers on demand. Customizing and shipping on demand will raise production and transportation cost per car by $100. Each car costs $20,000, and Toyota has a holding cost of 20 percent. Cars at the dealer are owned by Toyota for the first 90 days. Thus, for all practical purposes, Toyota owns all inventory, whether at the dealers or at the regional warehouse. Consider a region with five large dealers and thirty small dealers. Toyota has partitioned the variants into two groups—popular variants and uncommon variants. Weekly demand for the two types of variants at the two types of dealers is shown in Table 12-7. The goal is to provide a 95 percent cycle service level using a continuous review policy. Replenishment lead times for both dealers and regional warehouses are four weeks. Customization and shipping from a regional warehouse to a dealer can be done in a day, and this time can be ignored. Assume demand to be independent across all dealers.

a. How much safety inventory of a popular variant is required at a large or small dealer?
b. What is the safety inventory required if inventory for the popular variant (for both large and small dealers) is centralized at the regional warehouse by Toyota?
c. What is the safety inventory required if inventory for the popular variant at small dealers is centralized at the regional warehouse but that for large dealers is decentralized?
d. Given the additional customization and transportation cost, which structure do you recommend for the popular variant?
e. Repeat parts (a) to (d) for the uncommon variant.
f. How should Toyota structure inventories given its regional warehouses?

26. Orion is a global company that sells copiers. Orion currently sells ten variants of a copier, with all inventory kept in finished-goods form. The primary component that differentiates the copiers is the printing subassembly. An idea being discussed is to introduce commonality in the printing subassembly so that final assembly can be postponed and inventories kept in component form. Currently, each copier costs $1,000 in terms of components. Introducing commonality in the print subassembly will increase component costs to $1,025. One of the ten variants represents 80 percent of the total demand. Weekly demand for this variant is normally

TABLE 12-7 Weekly Demand at Car Dealers

	Popular Variant		Uncommon Variant	
	Mean	Standard Deviation	Mean	Standard Deviation
Large dealer	50	15	8	5
Small dealer	10	5	2	2

distributed, with a mean of 1,000 and a standard deviation of 200. Each of the remaining nine variants has a weekly demand of 28 with a standard deviation of 20. Orion aims to provide a 95 percent level of service. Replenishment lead time for components is four weeks. Copier assembly can be completed in a matter of hours. Orion manages all inventories using a continuous review policy and uses a holding cost of 20 percent.

a. How much safety inventory of each variant must Orion keep without component commonality? What is the annual holding cost?
b. How much safety inventory must be kept in component form if Orion uses common components for all variants? What is the annual holding cost? What is the increase in component cost using commonality? Is commonality justified across all variants?
c. At what cost of commonality will complete commonality be justified?
d. Now consider the case in which Orion uses component commonality for only the nine low-demand variants. How much reduction in safety inventory does Orion achieve in this case? What are the savings in terms of annual holding cost? Is this more restricted form of commonality justified?
e. At what cost of commonality will commonality across the low-volume variants be justified?

Bibliography

Barnes, Jim. "The Myths and Truths About Inventory Optimization." *Supply Chain Management Review* (March/April 2014): 10–19.

DeHoratius, Nicole, and Ananth Raman. "Inventory Record Inaccuracy: An Empirical Analysis." *Management Science* 54 (April 2008): 627–641.

Federgruen, Awi, and Yu-Sheng Zheng. "An Efficient Algorithm for Computing an Optimal (*r,Q*) Policy in Continuous Review Stochastic Inventory Systems." *Operations Research* (July–August 1992): 40, 808–813.

Feitzinger, Edward, and Hau L. Lee. "Mass Customization at Hewlett Packard." *Harvard Business Review* (January–February 1997): 116–121.

Gallego, Guillermo. "New Bounds and Heuristics for (Q,r) Policies." *Management Science* 44 (February 1998): 219–233.

Geary, Steve, Paul Childerhouse, and Denis Towill. "Uncertainty and the Seamless Supply Chain." *Supply Chain Management Review* (July–August 2002): 52–61.

Kopczak, Laura, and Hau L. Lee. *Hewlett-Packard Co.: Deskjet Printer Supply Chain (A)*. Stanford University Case GS3A, 2001.

Lee, Hau L. "Design for Supply Chain Management: Concepts and Examples." In R. Sarin, ed. *Perspectives in Operations Management.* Norwell, MA: Kluwer Academic Publishers, 1993, pp. 45–65.

Lee, Hau L., and Corey Billington. "Managing Supply Chain Inventory." *Sloan Management Review* (Spring 1992): 65–73.

Lee, Hau L., Corey Billington, and Brent Carter. "Hewlett-Packard Gains Control of Inventory and Service Through Design for Localization." *Interfaces* (July–August 1993): 1–11.

Nahmias, Steven. *Production and Operations Analysis*. Burr Ridge, IL: Richard P. Irwin, 1997.

Signorelli, Sergio, and James L. Heskett. 1984. "Benetton (A)." Harvard Business School Case 9–685–014, 1984.

Silver, Edward A., David Pyke, and Rein Petersen. *Inventory Management and Production Planning and Scheduling*. New York: Wiley, 1998.

Strang, Randy. "Retail Without Boundaries." *Supply Chain Management Review* (November 2013): 32–39.

Tayur, Sridhar, Ram Ganeshan, and Michael Magazine, eds. *Quantitative Models for Supply Chain Management*. Boston: Kluwer Academic Publishers, 1999.

Trent, Robert J. "Managing Inventory Investment Effectively." *Supply Chain Management Review* (March–April 2002): 28–35.

Zipkin, Paul H. *Foundations of Inventory Management*. Boston: Irwin McGraw-Hill, 2000.

CASE STUDY

Managing Inventories at ALKO Inc.

ALKO began in 1943 in a garage workshop set up by John Williams at his Cleveland home. John had always enjoyed tinkering, and in February 1948 he obtained a patent for one of his designs for lighting fixtures. He decided to produce it in his workshop and tried marketing it in the Cleveland area. The product sold well, and by 1957 ALKO had grown to a $3 million company. Its lighting fixtures were well known for their outstanding quality. By then, it sold five products.

In 1963, John took the company public. Since then, ALKO had been very successful, and the company had started distributing its products nationwide. As competition intensified in the 1980s, ALKO introduced many new lighting fixture designs. The company's profitability,

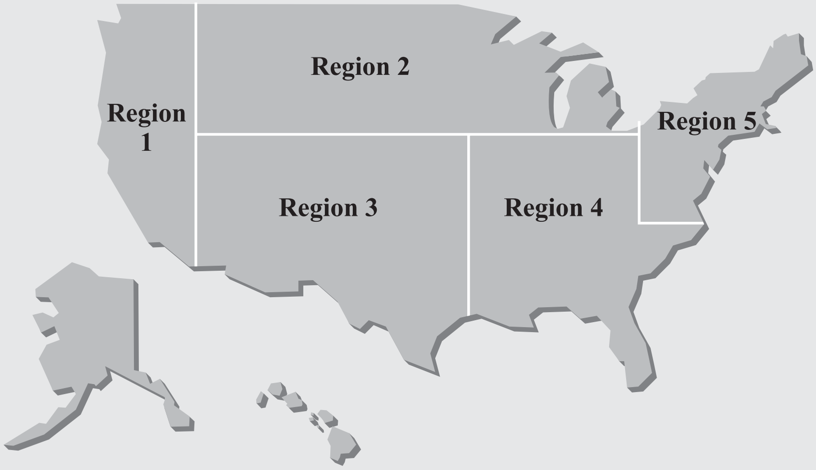


FIGURE 12-8 Sales Regions for ALKO

however, began to worsen despite the fact that ALKO had taken great care to ensure that product quality did not suffer. The problem was that margins had begun to shrink as competition in the market intensified. At this point, the board decided that a complete reorganization was needed, starting at the top. Gary Fisher was hired to reorganize and restructure the company.

When Gary arrived in 2009, he found a company teetering on the edge. He spent his first few months trying to understand the company business and the way it was structured. Gary realized that the key was in the operating performance. Although the company had always been outstanding at developing and producing new products, it had historically ignored its distribution system. The belief within the company was that once one makes a good product, the rest takes care of itself. Gary set up a task force to review the company's current distribution system and come up with recommendations.

Current Distribution System

The task force noted that ALKO had 100 products in its 2009 line. All production occurred at three facilities located in the Cleveland area. For sales purposes, the contiguous United States was divided into five regions, as shown in Figure 12-8. A DC owned by ALKO operated in each of these regions. Customers placed orders with the DCs, which tried to supply them from product in inventory. As the inventory for any product diminished, the DC, in turn, ordered from the plants. The plants scheduled production based on DC orders. Orders were transported from plants to the DCs in TL quantities because order sizes tended to be large. On the other hand, shipments from the DC to the customer were LTL. ALKO used a third-party trucking company for both transportation legs. In 2009, TL costs from the plants to DCs averaged $0.09 per unit. LTL shipping costs from a DC to a customer averaged $0.10 per unit. On average, five days were necessary between the time a DC placed an order with a plant and the time the order was delivered from the plant.

The policy in 2009 was to stock each item in every DC. A detailed study of the product line had shown that there were three basic categories of products in terms of the volume of sales. They were categorized as types High, Medium, and Low. Demand data for a representative product in each category is shown in Table 12-8. Products 1, 3, and 7 are representative of High, Medium, and Low demand products, respectively. Of the 100 products that ALKO sold, 10 were of type High, 20 of type Medium, and 70 of type Low. Each of their demands was identical to those of the representative products 1, 3, and 7, respectively.

The task force identified that plant capacities allowed any reasonable order to be produced and delivered in five days. The replenishment lead time was thus five days. The DCs ordered using a periodic review policy with a reorder interval of six days. The holding cost incurred was $0.15 per unit per day whether the unit was in transit or in storage. All DCs carried safety inventories to ensure a CSL of 95 percent.

TABLE 12-8 Distribution of Daily Demand at ALKO

	Region 1	Region 2	Region 3	Region 4	Region 5
Part 1 Mean	35.48	22.61	17.66	11.81	3.36
Part 1 SD	6.98	6.48	5.26	3.48	4.49
Part 3 Mean	2.48	4.15	6.15	6.16	7.49
Part 3 SD	3.16	6.20	6.39	6.76	3.56
Part 7 Mean	0.48	0.73	0.80	1.94	2.54
Part 7 SD	1.98	1.42	2.39	3.76	3.98

Alternative Distribution Systems

The task force recommended that ALKO build a national distribution center (NDC) outside Chicago. The task force recommended that ALKO close its five DCs and move all inventory to the NDC. Warehouse capacity was measured in terms of the *total number of units handled per year* (i.e., the warehouse capacity was given in terms of the annual demand supplied from the warehouse). The cost of constructing a warehouse is shown in Figure 12-9. However, ALKO expected to recover $50,000 for each warehouse that it closed. The CSL out of the NDC would continue to be 95 percent.

Given that Chicago is close to Cleveland, the inbound transportation cost from the plants to the NDC would fall to $0.05 per unit. The total replenishment lead time for orders from the Chicago NDC would still be five days. Given the increased average distance, however, the outbound transportation cost to customers from the NDC would increase to $0.24 per unit.

Other possibilities the task force considered include building a national distribution center while keeping the regional DCs open. In this case, some products would be stocked at the regional DCs, whereas others would be stocked at the NDC.

Gary Fisher's Decision

Gary Fisher pondered the task force report. It had not detailed any of the numbers supporting the decision. He decided to evaluate the numbers before making his decision.

Questions

1. What is the annual inventory and distribution cost of the current distribution system?
2. What are the savings that would result from following the task force recommendation and setting up an NDC? Evaluate the savings as the correlation coefficient of demand in any pair of regions varies from 0 to 0.5 to 1.0. Do you recommend setting up an NDC?
3. Suggest other options that Gary Fisher should consider. Evaluate each option and recommend a distribution system for ALKO that would be most profitable. How dependent is your recommendation on the correlation coefficient of demand across different regions?

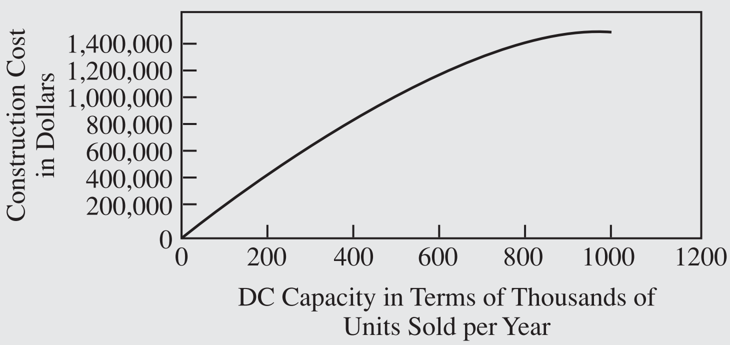


FIGURE 12-9 Construction Costs for NDC

CASE STUDY

Shall It Be Postponed?

Foxcompany (Fox) Co. Ltd. was founded in 1974 in Taiwan as a manufacturer of electrical components for computers. With strong research and development efforts, by 2011, it had accumulated more than 25,000 patents granted worldwide. Fox is now one of the world's 500 biggest companies, according to *Fortune* magazine. Its biggest production operation is located in Shenzhen Longlong Science & Technology Park, which covers more than 3 square km with 15 factories. Not only does Fox have dormitories, a hospital, and a fire brigade, but it also broadcasts its own TV channel within the park.

Fox is highly specialized in producing computer components and produces and packages private-label components for various famous brand names, including Acer, Apple, Dell, and Hewlett-Packard. The components manufactured are basically identical but are labeled and packaged differently for the various customers. The Shenzhen manufacturing facility replenishes a distribution center (DC) in Taiwan where the lead time is nine weeks. Fox adopts a continuous review policy to manage the inventory at its DC and wants to maintain a cycle service level of 95 percent for all orders.

The previous month had been challenging: Apple asked for 5,000 extra units than were available at the DC, whereas Acer and Dell ordered 3,500 units and 4,000 units fewer, respectively. Although there was sufficient inventory available at the DC in the form of basic product, Fox was not able to meet Apple's demand because the excess inventory available was labeled and packaged for Acer and Dell. As a result, Fox lost the extra business opportunity and surplus inventory because of the wrong labels and packaging.

Labeling and Packaging at the DC

To allow more flexibility for Fox production to accept such additional orders from customers by simply switching the inventory, the senior logistics supply chain manager proposes to postpone the labeling and packaging work to the DC, where the lead time of manufacturing and transportation remains unchanged. As a consequence, Fox would be able to meet Apple's sudden additional order more readily if other customers (e.g., Acer) placed a smaller order.

However, the management at the DC worried about the additional labeling and packaging work. Moreover, a detailed study revealed that the postponement would cost $1 more per unit. In particular, the DC managers believed that those $1 increases in cost per unit would be held against them once the process was changed and they would be under pressure to lower costs. They also thought the added workload would affect the overall service level of the DC.

Evaluating the Two Options

A task force was set up to look into this matter. It would focus its study mainly on three major components—motherboards, graphics cards, and chassis—and the four key customers—Acer, Apple, Dell, and HP. Weekly demand is shown in Table 12-9. In each case, the mean denotes the average demand per week, and SD denotes the standard deviation of the demand per week. Furthermore, all demands follow the normal distribution pattern. Fox incurred a total cost of $100 per motherboard, $50 per graphics card, and $30 per chassis. As per the rule of thumb of the industry, Fox used a holding cost of 30 percent when making all inventory decisions. The task force studied the impact of postponement on safety inventories before providing its final recommendation.

TABLE 12-9 Distribution of Weekly Demand by Product and Customer

	Motherboards		Graphics cards		Chassis	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
Acer	1,000	700	2,000	1,000	4,000	1,000
Apple	700	600	1,500	800	4,500	900
Dell	800	600	1,200	600	2,000	700
HP	500	400	900	500	1,400	500

Questions

1. What is the annual inventory cost before postponement?
2. How would the inventory cost change if postponement were implemented? Evaluate the change in inventory costs as the correlation coefficient of demand between any pair of customer varies from 0 to 0.5 to 1.0.
3. Should Fox postpone its labeling and packaging process to the DC? Would the answer change if the additional cost of labeling and packaging at the DC were reduced to $0.5 (from the current $1)?

APPENDIX 12A

The Normal Distribution

A continuous random variable X has a *normal distribution* with mean μ and standard deviation $\sigma > 0$ if the probability density function $f(x, \mu, \sigma)$ of the random variable is given by

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad \textbf{(12.21)}$$

The normal density function is as shown in Figure 12-10.

The *cumulative normal distribution function* is denoted by $F(x, \mu, \sigma)$ and is the probability that a normally distributed random variable with mean μ and standard deviation σ takes on a value less than or equal to x. The cumulative normal distribution function and the density function are related as follows:

$$F(x, \mu, \sigma) = \int_{X=-\infty}^{x} f(X, \mu, \sigma)\,dX$$

A normal distribution with a mean $\mu = 0$ and standard deviation $\sigma = 1$ is referred to as the *standard normal distribution.* The standard normal density function is denoted by $f_S(x)$ and the cumulative standard normal distribution function is denoted by $F_S(x)$. Thus,

$$f_S(x) = f(x, 0, 1) \text{ and } F_S(x) = F(x, 0, 1)$$

Given a probability p, the inverse normal $F^{-1}(p, \mu, \sigma)$ is the value x such that p is the probability that the normal random variable takes on a value x or less. Thus, if $F(x, \mu, \sigma) = p$ then $x = F^{-1}(p, \mu, \sigma)$. The inverse of the standard normal distribution is denoted by $F_S^{-1}(p)$. Thus,

$$F_S^{-1}(p) = F^{-1}(p, 0, 1).$$

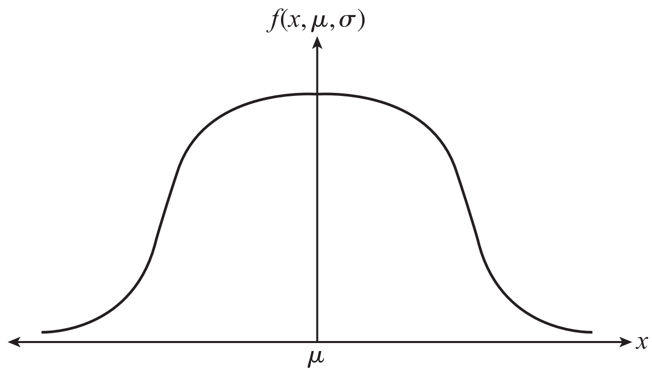


FIGURE 12-10 Normal Density Function

APPENDIX 12B

The Normal Distribution in Excel

The following Excel functions can be used to evaluate various normal distribution functions:

$$F(x, \mu, \sigma) = NORMDIST(x, \mu, \sigma, 1) \tag{12.22}$$

$$f(x, \mu, \sigma) = NORMDIST(x, \mu, \sigma, 0) \tag{12.23}$$

$$F^{-1}(p, \mu, \sigma) = NORMINV(p, \mu, \sigma) \tag{12.24}$$

The Excel functions to evaluate various standard normal distribution functions are listed next.

$$F_S(x) = NORMDIST(x, 0, 1, 1) \text{ or } NORMSDIST(x) \tag{12.25}$$

$$f_S(x) = NORMDIST(x, 0, 1, 0) \tag{12.26}$$

$$F_S^{-1}(p) = NORMSINV(p) \tag{12.27}$$

APPENDIX 12C

Expected Shortage per Replenishment Cycle

Objective

Establish an alternative formula for expected shortage per replenishment cycle (ESC) to be evaluated using Excel.

Analysis:

Given a reorder point of $ROP = D_L + ss$, the ESC is given as

$$ESC = \int_{x=ROP}^{\infty} (x - ROP) f(x) dx$$

$$= \int_{x=D_L+ss}^{\infty} (x - D_L - ss) f(x) dx$$

Given that the demand during lead time is normally distributed with a mean D_L and a standard deviation σ_L, we have (using Equation 12.21)

$$ESC = \int_{x=D_L+ss}^{\infty} (x - D_L - ss) \frac{1}{\sqrt{2\pi}\sigma_L} e^{-(x-D_L)^2/2\sigma_L^2} dx$$

Substitute the following:

$$z = \frac{(x - D_L)}{\sigma_L}$$

This implies that

$$dx = \sigma_L dz$$

Thus, we have

$$ESC = \int_{z=ss/\sigma_L}^{\infty} (z\sigma_L - ss)\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz$$
$$= -ss\int_{z=ss/\sigma_L}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz$$
$$+ \sigma_L\int_{z=ss/\sigma_L}^{\infty} z\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz$$

Recall that $F_S(.)$ is the cumulative distribution function and $f_S(.)$ is the probability density function for the standard normal distribution with mean 0 and standard deviation 1. Using Equation 12.21 and the definition of the standard normal distribution, we have

$$1 - F_S(y) = \int_{z=y}^{\infty} f_s(z)dz = \int_{z=y}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz$$

Substitute $w = z^2/2$ into the expression for ESC. This implies that

$$ESC = -ss[1 - F_S(ss/\sigma_L)] + \sigma_L\int_{w=ss^2/2\sigma_L^2}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-w}dw$$

or

$$ESC = -ss[1 - F_S(ss/\sigma_L)] + \sigma_L f_s(ss/\sigma_L)$$

Using Equations 12.25 and 12.26, ESC may be evaluated using Excel as follows:

$$ESC = -ss[1 - NORMDIST(ss/\sigma_L, 0, 1, 1)]$$
$$+ \sigma_L NORMDIST\ (ss/\sigma_L, 0, 1, 0)$$

APPENDIX 12D

Evaluating Safety Inventory for Slow-Moving Items

Objective

Devise a procedure for evaluating safety inventory for slow-moving items whose demand can be approximated using a Poisson distribution.

Analysis:

For slow-moving items, the normal distribution is not a good estimation for the demand distribution. A better approach is to use the Poisson distribution with demand arriving at a rate D. In such a setting, (Q, r) policies are known to be optimal. Under a (Q, r) policy, an order is placed whenever the inventory position drops to or below the reorder point r, and the order size is nQ, where n is the number of batches of size Q required to raise the inventory position to be in the interval $(r, r + Q)$.

For the Poisson distribution, given a constant lead time L, the average demand over the lead time is given by LD, and the variance of demand over the lead time is given by $\sigma^2 = \sqrt{LD}$. Efficient algorithms to obtain the Q and r are given by Federgruen and Zheng (1992). The results we present are based on Gallego (1998), who has given effective heuristics to solve the problem.

If H is the holding cost per unit per unit time, p the fixed shortage cost per unit per unit time, and S the fixed order cost per batch, Gallego suggests a batch size of Q^*, where

$$Q^* = Min\left(\sqrt{2}, \sqrt[4]{1 + \left(\frac{(H + p)L}{2S}\right)^2}\right)\sqrt{\frac{2DS}{H}} \qquad \textbf{(12.28)}$$

He shows that the use of batch size Q^* results in a cost that is no more than 7 percent from the optimal batch size. The reorder point r^* can be obtained using a procedure discussed by Federgruen and Zheng (1992). The long-run average cost $C(r, Q)$ of an (r, Q) policy when demand is Poisson is given by

$$C(r, Q) = \frac{DS}{Q} + \frac{1}{Q}\sum_{y=r+1}^{r+Q}\left\{H\sum_{i=0}^{y}(y - i)P_i + p\sum_{i=y+1}^{\infty}(i - y)P_i\right\}, \qquad \textbf{(12.29)}$$

where

$$P_i = \frac{e^{DL}(DL)^i}{i!}, i = 0, 1, \ldots$$

The reorder point r^* is obtained by inserting the batch size Q^* from Equation 12.28 into Equation 12.29 and searching for the value r^* that minimizes the cost $C(r, Q^*)$. Given that $C(r, Q^*)$ is unimodal [as shown by Federgruen and Zheng (1992)], r^* can be obtained using a binary search over the integers.