We begin with the oldest, and simplest, model—the economic order quantity (EOQ), and we work our way up to the more sophisticated reorder point (ROP) models. For each model we give a motivating example, a presentation of its development, and a discussion of its underlying insight.

## 2.2 The Economic Order Quantity Model

One of the earliest applications of mathematics to factory management was the work of Ford W. Harris (1913) on the problem of setting manufacturing lot sizes. Although the original paper was evidently incorrectly cited for many years (see Erlenkotter 1989, 1990), Harris's EOQ model has been widely studied and is a staple of virtually every introductory production and operations management textbook.

### 2.2.1 Motivation

Consider the situation of MedEquip, a small manufacturer of operating-room monitoring and diagnostic equipment, which produces a variety of final products by mounting electronic components in standard metal racks. The racks are purchased from a local metalworking shop, which must set up its equipment (presses, machining stations, and welding stations) each time it produces a "run" of racks. Because of the time wasted setting up the shop, the metalworking shop can produce (and sell) the racks more cheaply if MedEquip purchases them in quantities greater than one. However, because MedEquip does not want to tie up too much of its precious cash in stores of racks, it does not want to buy in excessive quantities.

This dilemma is precisely the one studied by Harris in his paper "How Many Parts to Make at Once." He puts it thus:

> Interest on capital tied up in wages, material and overhead sets a maximum limit to the quantity of parts which can be profitably manufactured at one time; "set-up" costs on the job fix the minimum. Experience has shown one manager a way to determine the economical size of lots. (Harris 1913)

The problem Harris had in mind was that of a factory producing various products and switching between products entails a costly setup. As an example, he described a metalworking shop that produced copper connectors. Each time the shop changed from one type of connector to another, machines had to be adjusted, clerical work had to be done, and material might be wasted (e.g., copper used up as test parts in the adjustment process). Harris defined the sum of the labor and material costs to ready the shop to produce a product to be the **setup cost.** (Notice that if the connectors had been purchased, instead of manufactured, then the problem would remain similar, but setup cost would correspond to the cost of placing a purchase order.)

The basic tradeoff is the same in the MedEquip example and Harris's copper connector case. Large lots reduce setup costs by requiring less frequent changeovers. But small lots reduce inventory by bringing in product closer to the time it is used. The EOQ model was Harris's systematic approach to striking a balance between these two concerns.

### 2.2.2 The Model

Despite his claim in the above quote that the EOQ is based on experience, Harris was consistent with the scientific management emphasis of his day on precise mathematical

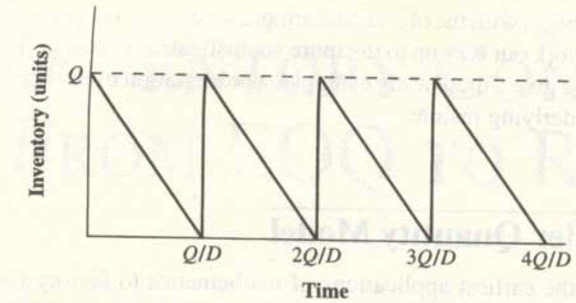



**FIGURE 2.1**
*Inventory versus time in the EOQ model*

approaches to factory management. To derive a lot size formula, he made the following assumptions about the manufacturing system:[1]

1. *Production is instantaneous.* There is no capacity constraint, and the entire lot is produced simultaneously.
2. *Delivery is immediate.* There is no time lag between production and availability to satisfy demand.
3. *Demand is deterministic.* There is no uncertainty about the quantity or timing of demand.
4. *Demand is constant over time.* In fact, it can be represented as a straight line, so that if annual demand is 365 units, this translates to a daily demand of one unit.
5. *A production run incurs a fixed setup cost.* Regardless of the size of the lot or the status of the factory, the setup cost is the same.
6. *Products can be analyzed individually.* Either there is only a single product or there are no interactions (e.g., shared equipment) between products.

With these assumptions, we can use Harris's notation, with slight modifications for ease of presentation, to develop the EOQ model for computing optimal production lot sizes. The notation we will require is as follows:

$D$ = demand rate (in units per year)

$c$ = unit production cost, not counting setup or inventory costs (in dollars per unit)

$A$ = fixed setup (ordering) cost to produce (purchase) a lot (in dollars)

$h$ = holding cost (in dollars per unit per year); if the holding cost consists entirely of interest on money tied up in inventory, then $h = ic$, where $i$ is the annual interest rate

$Q$ = lot size (in units); this is the decision variable

For modeling purposes, Harris represented both time and product as continuous quantities. Since he assumed constant, deterministic demand, ordering $Q$ units each time the inventory reaches zero results in an average inventory level of $Q/2$ (see Figure 2.1). The holding cost associated with this inventory is therefore $hQ/2$ per year. The setup cost is $A$ per order, or $AD/Q$ per year, since we must place $D/Q$ orders per year to satisfy demand. The production cost is $c$ per unit, or $cD$ per year. Thus, the total (inventory,

[1]The reader should keep in mind that *all* models are based on simplifying assumptions of some sort. The real world is too complex to analyze directly. Good modeling assumptions are those that facilitate analysis while capturing the essence of the real problem. We will be explicit about the underlying assumptions of the models we discuss in order to allow the reader to personally gauge their reasonableness.

setup, and production) cost per year can be expressed as

$$Y(Q) = \frac{hQ}{2} + \frac{AD}{Q} + cD \tag{2.1}$$

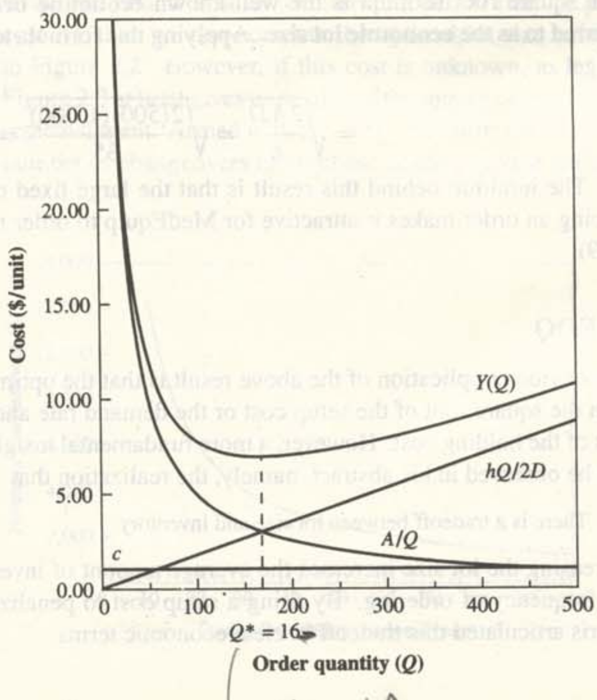
**Example:**

To illustrate the nature of $Y(Q)$, let us return to the MedEquip example. Suppose that its demand for metal racks is fairly steady and predictable at $D = 1{,}000$ units per year. The unit cost of the racks is $c = \$250$, but the metalworking shop also charges a fixed cost of $A = \$500$ per order, to cover the cost of shutting down the shop to set up for a MedEquip run. MedEquip estimates its opportunity cost or hurdle rate for money at 10 percent per year. It also estimates that the floorspace required to store a rack costs roughly \$10 per year in annualized costs. Hence, the annual holding cost per rack is $h = (0.1)(250) + 10 = \$35$. Substituting these values into expression (2.1) yields the plots in Figure 2.2.

We can make the following observations about the cost function $Y(Q)$ from Figure 2.2:

1. The holding cost term $hQ/D$ increases linearly in the lot size $Q$ and eventually becomes the dominant component of total annual cost for large $Q$.
2. The setup cost term $AD/Q$ diminishes quickly in $Q$, indicating that while increasing lot size initially generates substantial savings in setup cost, the returns from increased lot sizes decrease rapidly.
3. The unit-cost term $cD$ does not affect the relative cost for different lot sizes, since it does not include a $Q$ term.
4. The total annual cost $Y(Q)$ is minimized by some lot size $Q$. Interestingly, this minimum turns out to occur precisely at the value of $Q$ for which the holding cost and setup cost are exactly balanced (i.e., the $hQ/D$ and $AD/Q$ cost curves cross).

**FIGURE 2.2**
*Costs in the EOQ model*




(Q* = 169

Harris wrote that finding the value of $Q$ that minimizes $Y(Q)$ "involves higher mathematics" and simply gives the solution without further derivation. The mathematics he is referring to (calculus) does not seem quite as high today, so we will fill in some of the details he omitted in the following technical note. Those not interested in such details can skip this and subsequent technical notes without loss of continuity.

---

**Technical Note**

The standard approach for finding the minimum of an unconstrained function, such as $Y(Q)$, is to take its derivative with respect to $Q$, set it equal to zero, and solve the resulting equation for $Q^*$. This will find a point where the slope is zero (i.e., the function is flat). If the function is convex (as we will verify below), then the zero-slope point will be unique and will correspond to the minimum of $Y(Q)$.

Taking the derivative of $Y(Q)$ and setting the result equal to zero yields

$$\frac{dY(Q)}{dQ} = \frac{h}{2} - \frac{AD}{Q^2} = 0 \tag{2.2}$$

This equation represents the *first-order condition* for $Q$ to be a minimum. The *second-order condition* makes sure that this zero-slope point corresponds to a minimum (i.e., as opposed to a maximum or a saddle point) by checking the second derivative of $Y(Q)$:

$$\frac{d^2Y(Q)}{dQ^2} = 2\frac{AD}{Q^3} \tag{2.3}$$

Since this second derivative is positive for any positive $Q$ (that is, $Y(Q)$ is convex), it follows that solving (2.2) for $Q^*$ (as we do in (2.4) below) does indeed minimize $Y(Q)$.

---

The lot size that minimizes $Y(Q)$ in cost function (2.1) is

$$Q^* = \sqrt{\frac{2AD}{h}} \tag{2.4}$$

This square root formula is the well-known **economic order quantity (EOQ),** also referred to as the **economic lot size.** Applying this formula to the example in Figure 2.2, we get

$$Q^* = \sqrt{\frac{2AD}{h}} = \sqrt{\frac{2(500)(1{,}000)}{35}} = 169$$

The intuition behind this result is that the large fixed cost ($500) associated with placing an order makes it attractive for MedEquip to order racks in fairly large batches (169).

### 2.2.3 The Key Insight of EOQ

The obvious implication of the above result is that the optimal order quantity increases with the square root of the setup cost or the demand rate and decreases with the square root of the holding cost. However, a more fundamental insight from Harris's work is the one he observed in his abstract, namely, the realization that

> There is a tradeoff between lot size and inventory.

Increasing the lot size increases the average amount of inventory on hand, but reduces the frequency of ordering. By using a setup cost to penalize frequent replenishments, Harris articulated this tradeoff in clear economic terms.

The basic insight on the previous page is incontrovertible. However, the specific mathematical result (i.e., the EOQ square root formula) depends on the modeling assumptions, some of which we could certainly question (e.g., how realistic is instantaneous production?). Moreover, the usefulness of the EOQ formula for computational purposes depends on the realism of the input data. Although Harris claimed that "The set-up cost proper is generally understood" and "may, in a large factory, exceed *one dollar* per order," estimating setup costs may actually be a difficult task. As we will discuss in detail later in Parts II and III, setups in a manufacturing system have a variety of other impacts (e.g., on capacity, variability, and quality) and are therefore not easily reduced to a single invariant cost. In purchasing systems, however, where some of these other effects are not an issue and the setup cost can be cleanly interpreted as the cost of placing a purchase order, the EOQ model can be very useful.

It is worth noting that we can use the insight that there is a tradeoff between lot size and inventory without even resorting to Harris's square root formula. Since the average number of lots per year $F$ is

$$F = \frac{D}{Q} \tag{2.5}$$

and the total inventory investment is

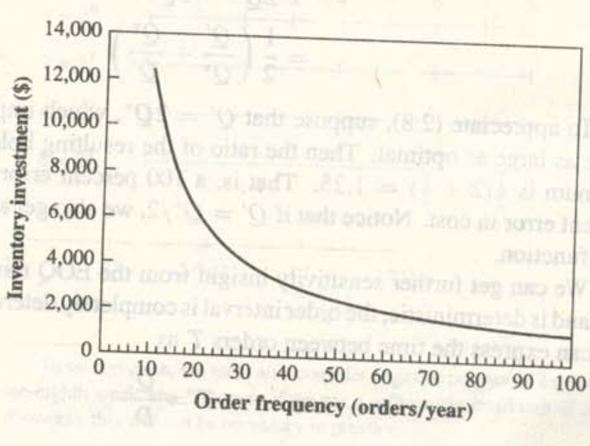
$$I = \frac{cQ}{2} = \frac{cD}{2F} \tag{2.6}$$

we can simply plot inventory investment $I$ as a function of replenishment frequency $F$ in lots per year. We do this for the MedEquip example with $D = 1{,}000$ and $c = \$250$ in Figure 2.3. Notice that this graph shows us that the inventory is cut in half (from \$12,500 to \$6,250) when we produce or order 20 times per year rather than 10 times per year (i.e., change the lot size from 100 to 50). However, if we replenish 30 times per year instead of 20 times per year (i.e., decrease the lot size from 50 to 33), inventory only falls from \$6,250 to \$4,125, a 34 percent decrease.

This analysis shows that there are decreasing returns to additional replenishments. If we can attach a value to these production runs or purchase orders (i.e., the setup cost $A$), then we can compute the optimal lot size using the EOQ formula as we did in Figure 2.2. However, if this cost is unknown, as it may well be, then the curve in Figure 2.3 at least gives us an idea of the impact on total inventory of an additional annual replenishment. Armed with this tradeoff information, a manager can select a reasonable number of changeovers or purchase orders per year and thereby specify a lot size.

**FIGURE 2.3**

*Inventory investment versus lots per year*



### 2.2.4 Sensitivity

A second insight that follows from the EOQ model is that

Holding and setup costs are fairly insensitive to lot size.

We can see this in Figure 2.2, where the total cost only varies between seven and eight for values of $Q$ between 96 and 306. This implies that if, for any reason, we use a lot size that is slightly different than $Q^*$, the increase in the holding plus setup costs will not be large. This feature was qualitatively observed by Harris in his original paper. The earliest quantitative treatment of it of which we are aware is by Brown (1967, 16).

To examine the sensitivity of the cost to lot size, we begin by substituting $Q^*$ for $Q$ into expression (2.1) for $Y$ (but omitting the $c$ term, since this is not affected by lot size), and we find that the minimum holding plus setup cost per unit is given by

$$\begin{aligned} Y^* = Y(Q^*) &= \frac{hQ^*}{2} + \frac{AD}{Q^*} \\ &= \frac{h\sqrt{2AD/h}}{2} + \frac{AD}{\sqrt{2AD/h}} \\ &= \sqrt{2ADh} \end{aligned} \tag{2.7}$$

Now, suppose that instead of using $Q^*$, we use some other arbitrary lot size $Q'$, which might be larger or smaller than $Q^*$. From expression (2.1) for $Y(Q)$, we see that the annual holding plus setup cost under $Q'$ can be written

$$Y(Q') = \frac{hQ'}{2} + \frac{AD}{Q'}$$

Hence, the ratio of the annual cost using lot size $Q'$ to the optimal annual cost (using $Q^*$) is given by

$$\begin{aligned} \frac{Y(Q')}{Y^*} &= \frac{hQ'/2 + AD/Q'}{\sqrt{2ADh}} \\ &= \frac{Q'}{2}\sqrt{\frac{h^2}{2ADh}} + \frac{1}{Q'}\sqrt{\frac{A^2D^2}{2ADh}} \\ &= \frac{Q'}{2}\sqrt{\frac{h}{2AD}} + \frac{1}{2Q'}\sqrt{\frac{2AD}{h}} \\ &= \frac{Q'}{2Q^*} + \frac{Q^*}{2Q'} \\ &= \frac{1}{2}\left(\frac{Q'}{Q^*} + \frac{Q^*}{Q'}\right) \end{aligned} \tag{2.8}$$

To appreciate (2.8), suppose that $Q' = 2Q^*$, which implies that we use a lot size twice as large as optimal. Then the ratio of the resulting holding plus setup cost to the optimum is $\frac{1}{2}(2 + \frac{1}{2}) = 1.25$. That is, a 100 percent error in lot size results in a 25 percent error in cost. Notice that if $Q' = Q^*/2$, we also get an error of 25 percent in the cost function.

We can get further sensitivity insight from the EOQ model by noting that because demand is deterministic, the order interval is completely determined by the order quantity. We can express the time between orders $T$ as

$$T = \frac{Q}{D} \tag{2.9}$$

Hence, dividing (2.4) by $D$, we get the following expression for the optimal order interval

$$T^* = \sqrt{\frac{2A}{hD}} \tag{2.10}$$

and by substituting (2.9) into (2.8), we get the following expression for the ratio of the cost resulting from an arbitrary order interval $T'$ and the optimum cost:

$$\frac{\text{Annual cost under } T'}{\text{Annual cost under } T^*} = \frac{1}{2}\left(\frac{T'}{T^*} + \frac{T^*}{T'}\right) \tag{2.11}$$

Expression (2.11) is useful in multiproduct settings, where it is desirable to order such that different products are frequently replenished at the same time (e.g., to facilitate sharing of delivery trucks). A method for facilitating this that has been widely proposed in the operations research literature is to order items at intervals given by *powers of 2*. That is, make the order interval one week, two weeks, four weeks, eight weeks, etc.[2] The result is that items ordered at $2^n$ week intervals will be placed at the same time as orders for items with $2^k$ intervals for all $k$ smaller than $n$ (see Figure 2.4). This will facilitate sharing of trucks, consolidation of ordering effort, simplification of shipping schedules, etc.

Moreover, the sensitivity results we derived above for the EOQ model imply that the error introduced by restricting order intervals to powers of 2 will not be excessive. To see this, suppose that the optimal order interval for an item $T^*$ lies between $2^m$ and $2^{m+1}$ for some $m$ (see Figure 2.5). Then $T^*$ lies either in the interval $[2^m, 2^m\sqrt{2}]$ or in the interval $[2^m\sqrt{2}, 2^{m+1}]$. All points in $[2^m, 2^m\sqrt{2}]$ are no more than $\sqrt{2}$ times as large as $2^m$. Likewise, all points in the interval $[2^m\sqrt{2}, 2^{m+1}]$ are no less than $2^{m+1}$ divided by $\sqrt{2}$. For instance, in Figure 2.5, $2^m$ is within a multiplicative factor of $\sqrt{2}$ of $T_1^*$, and $2^{m+1}$ is within a multiplicative factor of $1/\sqrt{2}$ of $T_2^*$. Hence, the power-of-2 order interval $T'$ must lie in the interval $[T^*/\sqrt{2}, \sqrt{2}T^*]$ around the optimal order interval $T^*$. Thus, the maximum error in cost will occur when $T' = \sqrt{2}T^*$, or $T' = T^*/\sqrt{2}$. From (2.11), the error from using $T' = \sqrt{2}T^*$ is

$$\frac{1}{2}\left(\sqrt{2} + \frac{1}{\sqrt{2}}\right) = 1.06$$

and is the same when $T' = T^*/\sqrt{2}$. Hence, the error in the holding plus setup cost resulting from using the optimal power-of-2 order interval instead of the optimal order interval is guaranteed to be no more than six percent. Jackson, Maxwell, and Muckstadt (1985); Roundy (1985, 1986); and Federgruen and Zheng (1992) give algorithms for

**FIGURE 2.4**
*Powers-of-2 order intervals*

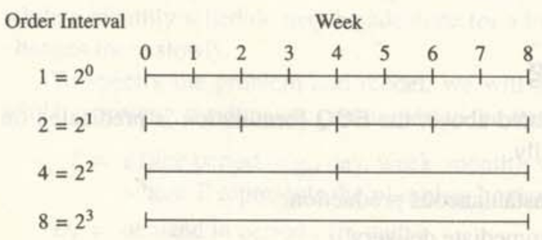
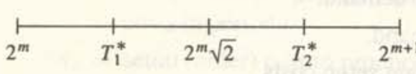



**FIGURE 2.5**
*The "root-2" interval*




---

[2] To be complete, we must also consider negative powers of 2 or one-half week, one-fourth week, one-eighth week, etc. However, if we use a sufficiently small unit of time as our baseline (e.g., days instead of weeks), this will not be necessary in practice.

computing the optimal power-of-2 policy and extend the above results to more general multipart settings.

As a concrete illustration of these concepts, consider once again the MedEquip problem. We computed the optimal order quantity for racks to be $Q^* = 169$. Hence, the optimal order interval is $T^* = Q^*/D = 169/1,000 = 0.169$ year, or $0.169 \times 52 = 8.78$ weeks. Suppose further that MedEquip orders a variety of other parts from the same supplier. The unit price of \$250 for racks is a delivered price, assuming an average shipping cost. However, if MedEquip combines orders for different parts, total shipping costs can be reduced. If the minimum order interval for any of the products under consideration is one week, then the order interval for racks can be rounded to the nearest power of 2 of $T = 8$ weeks or $8/52 = 0.154$ year. This implies an order quantity of $Q = TD = 0.154(1,000) = 154$. The holding plus order cost of this modified order quantity is

$$Y(Q) = \frac{hQ}{2} + \frac{AD}{Q} = \frac{35(154)}{2} + \frac{500(1{,}000)}{154} = \$5{,}942$$

The optimal annual cost (i.e., from using $Q^* = 169$) is given by

$$Y^* = \sqrt{2ADh} = \sqrt{2(500)(1{,}000)(35)} = \$5{,}916$$

So the modified order quantity results in less than a one percent increase in cost. The other parts ordered from the same supplier will have similar increases in holding plus order cost—but none of more than six percent. If these increases are offset by the reduced transportation cost, then the power-of-2 order schedule is worthwhile.

### 2.2.5 EOQ Extensions

Harris's original formula has been extended in a variety of ways over the years. One of the earliest extensions (Taft 1918) was to the case in which replenishment is not instantaneous; instead, there is a finite, but constant and deterministic, production rate. This model is sometimes called the **economic production lot (EPL)** model and results in a similar square root formula to the regular EOQ. Other variations of the basic EOQ include backorders (i.e., orders that are not filled immediately, but have to wait until stock is available), major and minor setups, and quantity discounts among others (see Johnson and Montgomery 1974; McClain and Thomas 1985; Plossl 1985; Silver, Pyke, and Peterson 1998).

## 2.3 Dynamic Lot Sizing

As we noted above, the EOQ formulation is predicated on a number of assumptions, specifically,

1. Instantaneous production.
2. Immediate delivery.
3. Deterministic demand.
4. Constant demand.
5. Known constant setup costs.
6. Single product or separable products.

We have already noted that Taft relaxed the assumption of instantaneous production. Introducing delivery delays is straightforward if delivery times are known and fixed (i.e.,