

### ► Learning Objectives

1. Describe linear cost functions and three common ways in which they behave
2. Explain the importance of causality in estimating cost functions
3. Understand various methods of cost estimation
4. Outline six steps in estimating a cost function using quantitative analysis
5. Describe three criteria used to evaluate and choose cost drivers
6. Explain nonlinear cost functions, in particular those arising from learning curve effects
7. Be aware of data problems encountered in estimating cost functions

### What is the value of looking at the past?

Perhaps it is to recall fond memories you've had or help you understand historical events. Maybe your return to the past enables you to better understand and predict the future. When an organization looks at the past, it typically does so to analyze its results, so that the best decisions can be made for the company's future. This activity requires gathering information about costs and how they behave so that managers can predict what they will be "down the road." Gaining a deeper understanding of cost behavior can also spur a firm to reorganize its operations in innovative ways and tackle important challenges, as the following article shows.

## Management Accountants at Cisco Embrace Opportunities, Enhance Sustainability<sup>1</sup>

Understanding how costs behave is a valuable technical skill.

Managers look to management accountants to help them identify cost drivers, estimate cost relationships, and determine the fixed and variable components of costs. To be effective, management accountants must have a clear understanding of the business's strategy and operations to identify new opportunities to reduce costs and increase profitability. At Cisco Systems, management accountants' in-depth understanding of the company's costs and operations led to reduced costs, while also helping the environment.

Cisco, makers of computer networking equipment including routers and wireless switches, traditionally regarded the used equipment it received back from its business customers as scrap and recycled it at a cost of about \$8 million a year. As managers looked at the accumulated costs and realized that they may literally be "throwing away money," they decided to reassess their treatment of scrap material. In 2005, managers at Cisco began trying to find uses for the equipment, mainly because 80% of the returns were in working condition. A value recovery team at Cisco identified groups within the company that could use the returned equipment. These included its customer service group, which supports warranty claims and service

<sup>1</sup> Source: Nidumolu, R., C. Prahalad, and M. Rangaswami. 2009. Why sustainability is now the key driver of innovation. *Harvard Business Review*, September 2009; Cisco Systems, Inc. 2009. *2009 corporate social responsibility report*. San Jose, CA: Cisco Systems, Inc.

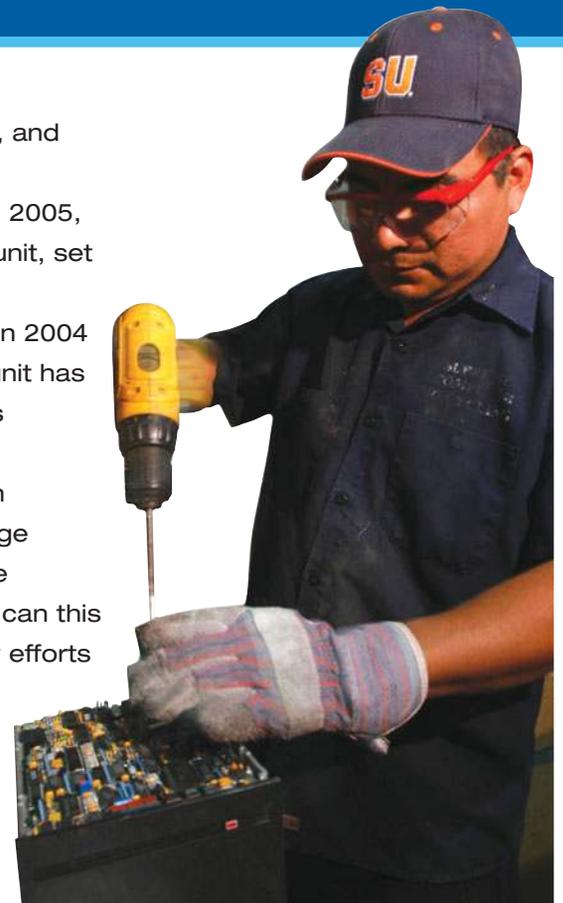
contracts, and the labs that provide technical support, training, and product demonstrations.

Based on the initial success of the value recovery team, in 2005, Cisco designated its recycling group as a company business unit, set clear objectives for it, and assigned the group its own income statement. As a result, the reuse of equipment rose from 5% in 2004 to 45% in 2008, and Cisco's recycling costs fell by 40%. The unit has become a profit center that contributed \$153 million to Cisco's bottom line in 2008.

With product returns reducing corporate profitability by an average of about 4% a year, companies like Cisco can leverage management accountants' insight to reduce the cost of these returns while decreasing its environmental footprint. Not only can this turn a cost center into a profitable business, but sustainability efforts like these signals that the company is concerned about preventing environmental damage by reducing waste.

As the Cisco example illustrates, managers must understand how costs behave to make strategic and operating decisions that have a positive environmental impact. Consider several other examples. Managers at FedEx decided to replace old planes with new Boeing 757s that reduced fuel consumption by 36%, while increasing capacity by 20%. At Clorox, managers decided to create a new line of non-synthetic cleaning products that were better for the environment and helped create a new category of 'green' cleaning products worth about \$200 million annually.

In each situation, knowledge of cost behavior was needed to answer key questions. This chapter will focus on how managers determine cost-behavior patterns—that is, how costs change in relation to changes in activity levels, in the quantity of products produced, and so on.



## Basic Assumptions and Examples of Cost Functions

Managers are able to understand cost behavior through cost functions. A **cost function** is a mathematical description of how a cost changes with changes in the level of an activity relating to that cost. Cost functions can be plotted on a graph by measuring the level of an activity, such as number of batches produced or number of machine-hours used, on the horizontal axis (called the  $x$ -axis) and the amount of total costs corresponding to—or, preferably, dependent on—the levels of that activity on the vertical axis (called the  $y$ -axis).

### Learning Objective 1

Describe linear cost functions

... graph of cost function is a straight line

and three common ways in which they behave

... variable, fixed, and mixed

## Basic Assumptions

Managers often estimate cost functions based on two assumptions:

1. Variations in the level of a single activity (the cost driver) explain the variations in the related total costs.
2. Cost behavior is approximated by a linear cost function within the relevant range. Recall that a relevant range is the range of the activity in which there is a relationship between total cost and the level of activity. For a **linear cost function** represented graphically, total cost versus the level of a single activity related to that cost is a straight line within the relevant range.

We use these two assumptions throughout most, but not all, of this chapter. Not all cost functions are linear and can be explained by a single activity. Later sections will discuss cost functions that do not rely on these assumptions.

## Linear Cost Functions

To understand three basic types of linear cost functions and to see the role of cost functions in business decisions, consider the negotiations between Cannon Services and World Wide Communications (WWC) for exclusive use of a videoconferencing line between New York and Paris.

- **Alternative 1:** \$5 per minute used. Total cost to Cannon changes in proportion to the number of minutes used. The number of minutes used is the only factor whose change causes a change in total cost.

Panel A in Exhibit 10-1 presents this *variable cost* for Cannon Services. Under alternative 1, there is no fixed cost. We write the cost function in Panel A of Exhibit 10-1 as

$$y = \$5X$$

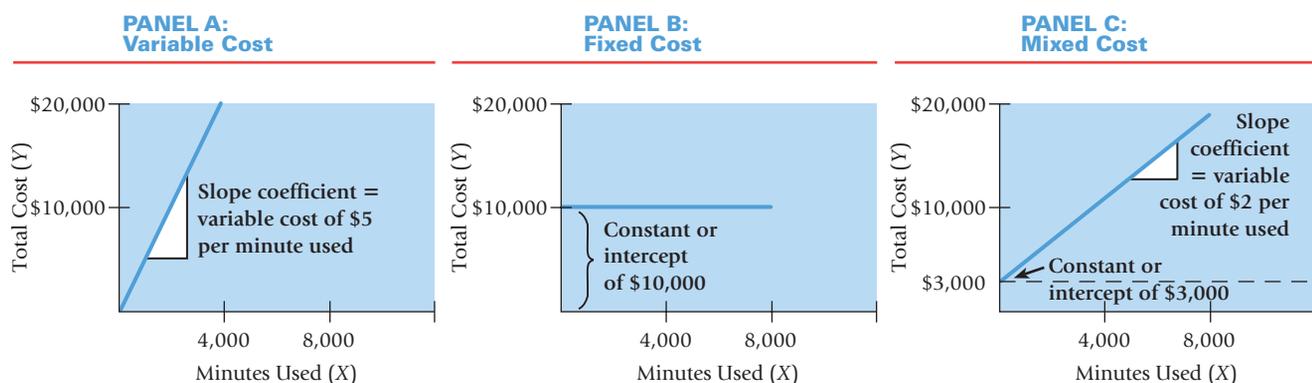
where  $X$  measures the number of minutes used (on the  $x$ -axis), and  $y$  measures the total cost of the minutes used (on the  $y$ -axis) calculated using the cost function. Panel A illustrates the **\$5 slope coefficient**, the amount by which total cost changes when a one-unit change occurs in the level of activity (one minute of usage in the Cannon example). *Throughout the chapter, uppercase letters, such as  $X$ , refer to the actual observations, and lowercase letters, such as  $y$ , represent estimates or calculations made using a cost function.*

- **Alternative 2:** Total cost will be fixed at \$10,000 per month, regardless of the number of minutes used. (We use the same activity measure, number of minutes used, to compare cost-behavior patterns under the three alternatives.)

Panel B in Exhibit 10-1 presents this *fixed cost* for Cannon Services. We write the cost function in Panel B as

$$y = \$10,000$$

**Exhibit 10-1** Examples of Linear Cost Functions



The fixed cost of \$10,000 is called a **constant**; it is the component of total cost that does not vary with changes in the level of the activity. Under alternative 2, the constant accounts for all the cost because there is no variable cost. Graphically, the slope coefficient of the cost function is zero; this cost function intersects the  $y$ -axis at the constant value, and therefore the *constant* is also called the **intercept**.

- **Alternative 3:** \$3,000 per month plus \$2 per minute used. This is an example of a mixed cost. A **mixed cost**—also called a **semivariable cost**—is a cost that has both fixed and variable elements.

Panel C in Exhibit 10-1 presents this *mixed cost* for Cannon Services. We write the cost function in Panel C of Exhibit 10-1 as

$$y = \$3,000 + \$2X$$

Unlike the graphs for alternatives 1 and 2, Panel C has both a constant, or intercept, value of \$3,000 and a slope coefficient of \$2. In the case of a mixed cost, total cost in the relevant range increases as the number of minutes used increases. Note that total cost does not vary strictly in proportion to the number of minutes used within the relevant range. For example, with 4,000 minutes of usage, the total cost equals \$11,000 [ $\$3,000 + (\$2 \text{ per minute} \times 4,000 \text{ minutes})$ ], but when 8,000 minutes are used, total cost equals \$19,000 [ $\$3,000 + (\$2 \text{ per minute} \times 8,000 \text{ minutes})$ ]. Although the usage in terms of minutes has doubled, total cost has increased by only about 73% [ $(\$19,000 - \$11,000) \div \$11,000$ ].

Cannon's managers must understand the cost-behavior patterns in the three alternatives to choose the best deal with WWC. Suppose Cannon expects to do at least 4,000 minutes of videoconferencing per month. Its cost for 4,000 minutes under the three alternatives would be as follows:

- **Alternative 1:** \$20,000 ( $\$5 \text{ per minute} \times 4,000 \text{ minutes}$ )
- **Alternative 2:** \$10,000
- **Alternative 3:** \$11,000 [ $\$3,000 + (\$2 \text{ per minute} \times 4,000 \text{ minutes})$ ]

Alternative 2 is the least costly. Moreover, if Cannon were to use more than 4,000 minutes, as is likely to be the case, alternatives 1 and 3 would be even more costly. Cannon's managers, therefore, should choose alternative 2.

Note that the graphs in Exhibit 10-1 are linear. That is, they appear as straight lines. We simply need to know the constant, or intercept, amount (commonly designated  $a$ ) and the slope coefficient (commonly designated  $b$ ). For any linear cost function based on a single activity (recall our two assumptions discussed at the start of the chapter), knowing  $a$  and  $b$  is sufficient to describe and graphically plot all the values within the relevant range of number of minutes used. We write a general form of this linear cost function as

$$y = a + bX$$

Under alternative 1,  $a = \$0$  and  $b = \$5$  per minute used; under alternative 2,  $a = \$10,000$  and  $b = \$0$  per minute used; and under alternative 3,  $a = \$3,000$  and  $b = \$2$  per minute used. To plot the mixed-cost function in Panel C, we draw a line starting from the point marked \$3,000 on the  $y$ -axis and increasing at a rate of \$2 per minute used, so that at 1,000 minutes, total costs increase by \$2,000 ( $\$2 \text{ per minute} \times 1,000 \text{ minutes}$ ) to \$5,000 ( $\$3,000 + \$2,000$ ) and at 2,000 minutes, total costs increase by \$4,000 ( $\$2 \text{ per minute} \times 2,000 \text{ minutes}$ ) to \$7,000 ( $\$3,000 + \$4,000$ ) and so on.

## Review of Cost Classification

Before we discuss issues related to the estimation of cost functions, we briefly review the three criteria laid out in Chapter 2 for classifying a cost into its variable and fixed components.

### Choice of Cost Object

A particular cost item could be variable with respect to one cost object and fixed with respect to another cost object. Consider Super Shuttle, an airport transportation company. If the fleet of vans it owns is the cost object, then the annual van registration and

license costs would be variable costs with respect to the number of vans owned. But if a particular van is the cost object, then the registration and license costs for that van are fixed costs with respect to the miles driven during a year.

### Time Horizon

Whether a cost is variable or fixed with respect to a particular activity depends on the time horizon being considered in the decision situation. The longer the time horizon, all other things being equal, the more likely that the cost will be variable. For example, inspection costs at Boeing Company are typically fixed in the short run with respect to inspection-hours used because inspectors earn a fixed salary in a given year regardless of the number of inspection-hours of work done. But, in the long run, Boeing's total inspection costs will vary with the inspection-hours required: More inspectors will be hired if more inspection-hours are needed, and some inspectors will be reassigned to other tasks or laid off if fewer inspection-hours are needed.

### Relevant Range

Managers should never forget that variable and fixed cost-behavior patterns are valid for linear cost functions only within a given relevant range. Outside the relevant range, variable and fixed cost-behavior patterns change, causing costs to become nonlinear (non-linear means the plot of the relationship on a graph is not a straight line). For example, Exhibit 10-2 plots the relationship (over several years) between total direct manufacturing labor costs and the number of snowboards produced each year by Ski Authority at its Vermont plant. In this case, the nonlinearities outside the relevant range occur because of labor and other inefficiencies (first because workers are learning to produce snowboards and later because capacity limits are being stretched). Knowing the relevant range is essential to properly classify costs.

#### Decision Point

What is a linear cost function and what types of cost behavior can it represent?

#### Learning Objective 2

Explain the importance of causality in estimating cost functions

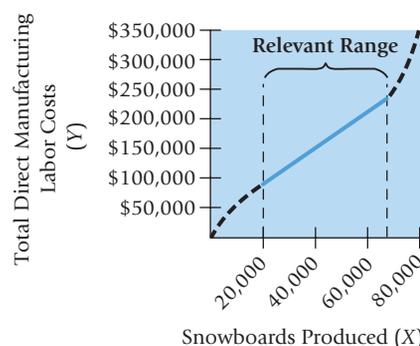
... only a cause-and-effect relationship establishes an economically plausible relationship between an activity and its costs

## Identifying Cost Drivers

The Cannon Services/WWC example illustrates variable-, fixed-, and mixed-cost functions using information about *future* cost structures proposed to Cannon by WWC. Often, however, cost functions are estimated from *past* cost data. Managers use **cost estimation** to measure a relationship based on data from past costs and the related level of an activity. For example, marketing managers at Volkswagen could use cost estimation to understand what causes their marketing costs to change from year to year (for example, the number of new car models introduced or a competitor's sudden recall) and the fixed and variable components of these costs. Managers are interested in estimating past cost-behavior functions primarily because these estimates can help them make more-accurate **cost predictions**, or forecasts, of future costs. Better cost predictions help managers make more-informed planning and control decisions, such as preparing next year's marketing budget. But better management decisions, cost predictions, and estimation of cost functions can be achieved only if managers correctly identify the factors that affect costs.

#### Exhibit 10-2

Linearity Within Relevant Range for Ski Authority, Inc.



## The Cause-and-Effect Criterion

The most important issue in estimating a cost function is determining whether a cause-and-effect relationship exists between the level of an activity and the costs related to that level of activity. Without a cause-and-effect relationship, managers will be less confident about their ability to estimate or predict costs. Recall from Chapter 2 that when a cause-and-effect relationship exists between a change in the level of an activity and a change in the level of total costs, we refer to the activity measure as a *cost driver*. We use the terms *level of activity* and *level of cost driver* interchangeably when estimating cost functions. Understanding the drivers of costs is crucially important for managing costs. The cause-and-effect relationship might arise as a result of the following:

- **A physical relationship between the level of activity and costs.** An example is when units of production are used as the activity that affects direct material costs. Producing more units requires more direct materials, which results in higher total direct material costs.
- **A contractual arrangement.** In alternative 1 of the Cannon Services example described earlier, number of minutes used is specified in the contract as the level of activity that affects the telephone line costs.
- **Knowledge of operations.** An example is when number of parts is used as the activity measure of ordering costs. A product with many parts will incur higher ordering costs than a product with few parts.

*Managers must be careful not to interpret a high correlation, or connection, in the relationship between two variables to mean that either variable causes the other.* Consider direct material costs and labor costs. For a given product mix, producing more units generally results in higher material costs and higher labor costs. Material costs and labor costs are highly correlated, but neither causes the other. Using labor costs to predict material costs is problematic. Some products require more labor costs relative to material costs, while other products require more material costs relative to labor costs. If the product mix changes toward more labor-intensive products, then labor costs will increase while material costs will decrease. Labor costs are a poor predictor of material costs. By contrast, factors that drive material costs such as product mix, product designs, and manufacturing processes, would have more accurately predicted the changes in material costs.

Only a cause-and-effect relationship—not merely correlation—establishes an economically plausible relationship between the level of an activity and its costs. Economic plausibility is critical because it gives analysts and managers confidence that the estimated relationship will appear again and again in other sets of data from the same situation. Identifying cost drivers also gives managers insights into ways to reduce costs and the confidence that reducing the quantity of the cost drivers will lead to a decrease in costs.

To identify cost drivers on the basis of data gathered over time, always use a long time horizon. Why? Because costs may be fixed in the short run (during which time they have no cost driver), but they are usually variable and have a cost driver in the long run.

## Cost Drivers and the Decision-Making Process

Consider Elegant Rugs, which uses state-of-the-art automated weaving machines to produce carpets for homes and offices. Management has made many changes in manufacturing processes and wants to introduce new styles of carpets. It would like to evaluate how these changes have affected costs and what styles of carpets it should introduce. It follows the five-step decision-making process outlined in Chapter 1.

**Step 1: Identify the problem and uncertainties.** The changes in the manufacturing process were specifically targeted at reducing indirect manufacturing labor costs, and management wants to know whether costs such as supervision, maintenance, and quality control did, in fact, decrease. One option is to simply compare indirect manufacturing labor costs before and after the process change. The problem with this approach is that the volume of activity before and after the process change was very different so costs need to be compared after taking into account the change in activity volume.

Managers were fairly confident about the direct material and direct manufacturing labor costs of the new styles of carpets. They were less certain about the impact that the choice of different styles would have on indirect manufacturing costs.

**Step 2: Obtain information.** Managers gathered information about potential cost drivers—factors such as machine-hours or direct manufacturing labor-hours that cause indirect manufacturing labor costs to be incurred. They also began considering different techniques (discussed in the next section) such as the industrial engineering method, the conference method, the account analysis method, the high-low method, and the regression method for estimating the magnitude of the effect of the cost driver on indirect manufacturing labor costs. Their goal was to identify the best possible single cost driver.

**Step 3: Make predictions about the future.** Managers used past data to estimate the relationship between cost drivers and costs and used this relationship to predict future costs.

**Step 4: Make decisions by choosing among alternatives.** As we will describe later (pp. 353–355), Elegant Rugs chose machine-hours as the cost driver of indirect manufacturing labor costs. Using the regression analysis estimate of indirect manufacturing labor cost per machine-hour, managers estimated the costs of alternative styles of carpets and chose to introduce the most profitable styles.

**Step 5: Implement the decision, evaluate performance, and learn.** After the managers at Elegant Rugs introduced the new carpet styles, they focused on evaluating the results of their decision. Comparing predicted to actual costs helped managers to learn how accurate the estimates were, to set targets for continuous improvement, and to constantly seek ways to improve efficiency and effectiveness.

### Decision Point

What is the most important issue in estimating a cost function?

### Learning Objective 3

Understand various methods of cost estimation

... for example, the regression analysis method determines the line that best fits past data

## Cost Estimation Methods

As we mentioned in Step 2, four methods of cost estimation are the industrial engineering method, the conference method, the account analysis method, and the quantitative analysis method (which takes different forms). These methods differ with respect to how expensive they are to implement, the assumptions they make, and the information they provide about the accuracy of the estimated cost function. They are not mutually exclusive, and many organizations use a combination of these methods.

### Industrial Engineering Method

The **industrial engineering method**, also called the **work-measurement method**, estimates cost functions by analyzing the relationship between inputs and outputs in physical terms. Consider Elegant Rugs. It uses inputs of cotton, wool, dyes, direct manufacturing labor, machine time, and power. Production output is square yards of carpet. Time-and-motion studies analyze the time required to perform the various operations to produce the carpet. For example, a time-and-motion study may conclude that to produce 10 square yards of carpet requires one hour of direct manufacturing labor. Standards and budgets transform these physical input measures into costs. The result is an estimated cost function relating direct manufacturing labor costs to the cost driver, square yards of carpet produced.

The industrial engineering method is a very thorough and detailed way to estimate a cost function when there is a physical relationship between inputs and outputs, but it can be very time consuming. Some government contracts mandate its use. Many organizations, such as Bose and Nokia, use it to estimate direct manufacturing costs but find it too costly or impractical for analyzing their entire cost structure. For example, physical relationships between inputs and outputs are difficult to specify for some items, such as indirect manufacturing costs, R&D costs, and advertising costs.

### Conference Method

The **conference method** estimates cost functions on the basis of analysis and opinions about costs and their drivers gathered from various departments of a company (purchasing, process engineering, manufacturing, employee relations, etc.). The Cooperative Bank

in the United Kingdom has a cost-estimating department that develops cost functions for its retail banking products (checking accounts, VISA cards, mortgages, and so on) based on the consensus of estimates from personnel of the particular departments. Elegant Rugs gathers opinions from supervisors and production engineers about how indirect manufacturing labor costs vary with machine-hours and direct manufacturing labor-hours.

The conference method encourages interdepartmental cooperation. The pooling of expert knowledge from different business functions of the value chain gives the conference method credibility. Because the conference method does not require detailed analysis of data, cost functions and cost estimates can be developed quickly. However, the emphasis on opinions rather than systematic estimation means that the accuracy of the cost estimates depends largely on the care and skill of the people providing the inputs.

## Account Analysis Method

The **account analysis method** estimates cost functions by classifying various cost accounts as variable, fixed, or mixed with respect to the identified level of activity. Typically, managers use qualitative rather than quantitative analysis when making these cost-classification decisions. The account analysis approach is widely used because it is reasonably accurate, cost-effective, and easy to use.

Consider indirect manufacturing labor costs for a small production area (or cell) at Elegant Rugs. Indirect manufacturing labor costs include wages paid for supervision, maintenance, quality control, and setups. During the most recent 12-week period, Elegant Rugs ran the machines in the cell for a total of 862 hours and incurred total indirect manufacturing labor costs of \$12,501. Using qualitative analysis, the manager and the cost analyst determine that over this 12-week period indirect manufacturing labor costs are mixed costs with only one cost driver—machine-hours. As machine-hours vary, one component of the cost (such as supervision cost) is fixed, whereas another component (such as maintenance cost) is variable. The goal is to use account analysis to estimate a linear cost function for indirect manufacturing labor costs with number of machine-hours as the cost driver. The cost analyst uses experience and judgment to separate total indirect manufacturing labor costs (\$12,501) into costs that are fixed (\$2,157, based on 950 hours of machine capacity for the cell over a 12-week period) and costs that are variable (\$10,344) with respect to the number of machine-hours used. Variable cost per machine-hour is  $\$10,344 \div 862 \text{ machine-hours} = \$12 \text{ per machine-hour}$ . The linear cost equation,  $y = a + bX$ , in this example is as follows:

$$\begin{aligned} \text{Indirect manufacturing labor costs} &= \$2,157 + \\ &(\$12 \text{ per machine-hour} \times \text{Number of machine-hours}) \end{aligned}$$

Management at Elegant Rugs can use the cost function to estimate the indirect manufacturing labor costs of using, say, 950 machine-hours to produce carpet in the next 12-week period. Estimated costs equal  $\$2,157 + (950 \text{ machine-hours} \times \$12 \text{ per machine-hour}) = \$13,557$ .

To obtain reliable estimates of the fixed and variable components of cost, organizations must take care to ensure that individuals thoroughly knowledgeable about the operations make the cost-classification decisions. Supplementing the account analysis method with the conference method improves credibility.

## Quantitative Analysis Method

Quantitative analysis uses a formal mathematical method to fit cost functions to past data observations. Excel is a useful tool for performing quantitative analysis. Columns B and C of Exhibit 10-3 show the breakdown of Elegant Rugs' total machine-hours (862) and total indirect manufacturing labor costs (\$12,501) into weekly data for the most recent 12-week period. Note that the data are paired; for each week, there is data for the number of machine-hours and corresponding indirect manufacturing labor costs. For example, week 12 shows 48 machine-hours and indirect manufacturing labor costs of \$963. The next section uses the data in Exhibit 10-3 to illustrate how to estimate a cost

### Decision Point

What are the different methods that can be used to estimate a cost function?

**Exhibit 10-3**

Weekly Indirect  
Manufacturing Labor  
Costs and Machine-  
Hours for Elegant Rugs

	A	B	C
1	Week	Cost Driver: Machine-Hours	Indirect Manufacturing Labor Costs
2		(X)	(Y)
3	1	68	\$ 1,190
4	2	88	1,211
5	3	62	1,004
6	4	72	917
7	5	60	770
8	6	96	1,456
9	7	78	1,180
10	8	46	710
11	9	82	1,316
12	10	94	1,032
13	11	68	752
14	12	<u>48</u>	<u>963</u>
15	Total	<u>862</u>	<u>\$12,501</u>
16			

function using quantitative analysis. We examine two techniques—the relatively simple high-low method as well as the more common quantitative tool used to examine and understand data, regression analysis.

## Steps in Estimating a Cost Function Using Quantitative Analysis

There are six steps in estimating a cost function using quantitative analysis of a past cost relationship. We illustrate the steps as follows using the Elegant Rugs example.

**Step 1: Choose the dependent variable.** Choice of the dependent variable (the cost to be predicted and managed) will depend on the cost function being estimated. In the Elegant Rugs example, the dependent variable is indirect manufacturing labor costs.

**Step 2: Identify the independent variable, or cost driver.** The independent variable (level of activity or cost driver) is the factor used to predict the dependent variable (costs). When the cost is an indirect cost, as it is with Elegant Rugs, the independent variable is also called a cost-allocation base. Although these terms are sometimes used interchangeably, we use the term *cost driver* to describe the independent variable. Frequently, the cost analyst, working with the management team, will cycle through the six steps several times, trying alternative economically plausible cost drivers to identify a cost driver that best fits the data.

A cost driver should be measurable and have an *economically plausible* relationship with the dependent variable. Economic plausibility means that the relationship (describing how changes in the cost driver lead to changes in the costs being considered) is based on a physical relationship, a contract, or knowledge of operations and makes economic sense to the operating manager and the management accountant. As we saw in Chapter 5, all the individual items of costs included in the dependent variable should have the same cost driver, that is, the cost pool should be homogenous. When all items of costs in the dependent variable do not have the same cost driver, the cost analyst should investigate the possibility of creating homogenous cost pools and estimating more than one cost function, one for each cost item/cost driver pair.

### Learning Objective 4

Outline six steps in estimating a cost function using quantitative analysis

... the end result (Step 6) is to evaluate the cost driver of the estimated cost function

As an example, consider several types of fringe benefits paid to employees and the cost drivers of the benefits:

<b>Fringe Benefit</b>	<b>Cost Driver</b>
Health benefits	Number of employees
Cafeteria meals	Number of employees
Pension benefits	Salaries of employees
Life insurance	Salaries of employees

The costs of health benefits and cafeteria meals can be combined into one homogenous cost pool because they have the same cost driver—the number of employees. Pension benefits and life insurance costs have a different cost driver—the salaries of employees—and, therefore, should not be combined with health benefits and cafeteria meals. Instead, pension benefits and life insurance costs should be combined into a separate homogenous cost pool. The cost pool comprising pension benefits and life insurance costs can be estimated using salaries of employees receiving these benefits as the cost driver.

**Step 3: Collect data on the dependent variable and the cost driver.** This is usually the most difficult step in cost analysis. Cost analysts obtain data from company documents, from interviews with managers, and through special studies. These data may be time-series data or cross-sectional data.

*Time-series data* pertain to the same entity (organization, plant, activity, and so on) over successive past periods. Weekly observations of indirect manufacturing labor costs and number of machine-hours at Elegant Rugs are examples of time-series data. The ideal time-series database would contain numerous observations for a company whose operations have not been affected by economic or technological change. A stable economy and technology ensure that data collected during the estimation period represent the same underlying relationship between the cost driver and the dependent variable. Moreover, the periods used to measure the dependent variable and the cost driver should be consistent throughout the observations.

*Cross-sectional data* pertain to different entities during the same period. For example, studies of loans processed and the related personnel costs at 50 individual, yet similar, branches of a bank during March 2012 would produce cross-sectional data for that month. The cross-sectional data should be drawn from entities that, within each entity, have a similar relationship between the cost driver and costs. Later in this chapter, we describe the problems that arise in data collection.

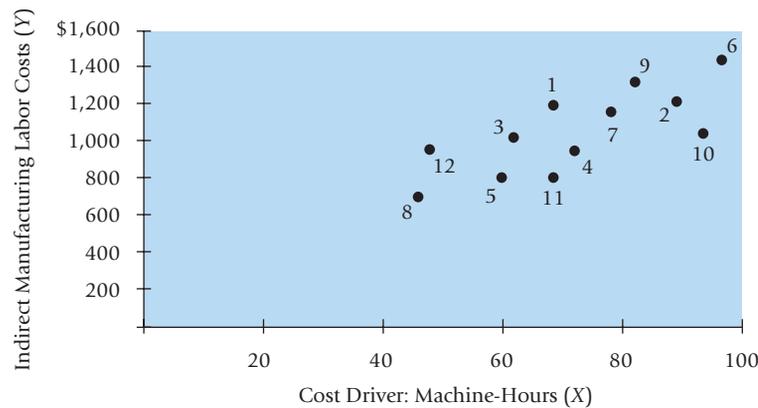
**Step 4: Plot the data.** The general relationship between the cost driver and costs can be readily observed in a graphical representation of the data, which is commonly called a plot of the data. The plot provides insight into the relevant range of the cost function, and reveals whether the relationship between the driver and costs is approximately linear. Moreover, the plot highlights extreme observations (observations outside the general pattern) that analysts should check. Was there an error in recording the data or an unusual event, such as a work stoppage, that makes these observations unrepresentative of the normal relationship between the cost driver and the costs?

Exhibit 10-4 is a plot of the weekly data from columns B and C of the Excel spreadsheet in Exhibit 10-3. This graph provides strong visual evidence of a positive linear relationship between number of machine-hours and indirect manufacturing labor costs (that is, when machine-hours go up, so do indirect manufacturing labor costs). There do not appear to be any extreme observations in Exhibit 10-4. The relevant range is from 46 to 96 machine-hours per week (weeks 8 and 6, respectively).

**Step 5: Estimate the cost function.** We will show two ways to estimate the cost function for our Elegant Rugs data. One uses the high-low method, and the other uses regression analysis, the two most frequently described forms of quantitative analysis. The widespread availability of computer packages such as Excel makes regression analysis much more easy to use. Still, we describe the high-low method to provide some basic intuition for the idea of drawing a line to “fit” a number of data points. We present these methods after Step 6.

**Exhibit 10-4**

Plot of Weekly Indirect Manufacturing Labor Costs and Machine-Hours for Elegant Rugs



**Step 6: Evaluate the cost driver of the estimated cost function.** In this step, we describe criteria for evaluating the cost driver of the estimated cost function. We do this after illustrating the high-low method and regression analysis.

### High-Low Method

The simplest form of quantitative analysis to “fit” a line to data points is the **high-low method**. It uses only the highest and lowest observed values of the cost driver within the relevant range and their respective costs to estimate the slope coefficient and the constant of the cost function. It provides a first cut at understanding the relationship between a cost driver and costs. We illustrate the high-low method using data from Exhibit 10-3.

	Cost Driver: Machine-Hours (X)	Indirect Manufacturing Labor Costs (Y)
Highest observation of cost driver (week 6)	96	\$1,456
Lowest observation of cost driver (week 8)	46	710
Difference	<u>50</u>	<u>\$ 746</u>

The slope coefficient,  $b$ , is calculated as follows:

$$\begin{aligned} \text{Slope coefficient} &= \frac{\text{Difference between costs associated with highest} \\ &\quad \text{and lowest observations of the cost driver}}{\text{Difference between highest and lowest} \\ &\quad \text{observations of the cost driver}} \\ &= \$746 \div 50 \text{ machine-hours} = \$14.92 \text{ per machine-hour} \end{aligned}$$

To compute the constant, we can use either the highest or the lowest observation of the cost driver. Both calculations yield the same answer because the solution technique solves two linear equations with two unknowns, the slope coefficient and the constant. Because

$$\begin{aligned} y &= a + bX \\ a &= y - bX \end{aligned}$$

At the highest observation of the cost driver, the constant,  $a$ , is calculated as follows:

$$\text{Constant} = \$1,456 - (\$14.92 \text{ per machine-hour} \times 96 \text{ machine-hours}) = \$23.68$$

And at the lowest observation of the cost driver,

$$\text{Constant} = \$710 - (\$14.92 \text{ per machine-hour} \times 46 \text{ machine-hours}) = \$23.68$$

Thus, the high-low estimate of the cost function is as follows:

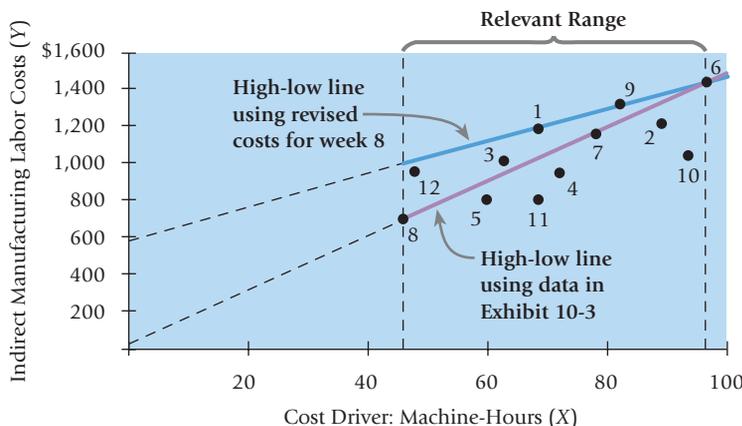
$$\begin{aligned} y &= a + bX \\ y &= \$23.68 + (\$14.92 \text{ per machine-hour} \times \text{Number of machine-hours}) \end{aligned}$$

The purple line in Exhibit 10-5 shows the estimated cost function using the high-low method (based on the data in Exhibit 10-3). The estimated cost function is a straight line joining the observations with the highest and lowest values of the cost driver (number of machine-hours). Note how this simple high-low line falls “in-between” the data points with three observations on the line, four above it and five below it. The intercept ( $a = \$23.68$ ), the point where the dashed extension of the purple line meets the  $y$ -axis, is the constant component of the equation that provides the best linear approximation of how a cost behaves *within the relevant range* of 46 to 96 machine-hours. The intercept should *not* be interpreted as an estimate of the fixed costs of Elegant Rugs if no machines were run. That’s because running no machines and shutting down the plant—that is, using zero machine-hours—is *outside the relevant range*.

Suppose indirect manufacturing labor costs in week 6 were \$1,280, instead of \$1,456, while 96 machine-hours were used. In this case, the highest observation of the cost driver (96 machine-hours in week 6) will not coincide with the newer highest observation of the costs (\$1,316 in week 9). How would this change affect our high-low calculation? Given that the cause-and-effect relationship runs *from* the cost driver *to* the costs in a cost function, we choose the highest and lowest observations of the cost driver (the factor that causes the costs to change). The high-low method would still estimate the new cost function using data from weeks 6 (high) and 8 (low).

There is a danger of relying on only two observations to estimate a cost function. Suppose that because a labor contract guarantees certain minimum payments in week 8, indirect manufacturing labor costs in week 8 were \$1,000, instead of \$710, when only 46 machine-hours were used. The blue line in Exhibit 10-5 shows the cost function that would be estimated by the high-low method using this revised cost. Other than the two points used to draw the line, all other data lie on or below the line! In this case, choosing the highest and lowest observations for machine-hours would result in an estimated cost function that poorly describes the underlying linear cost relationship between number of machine-hours and indirect manufacturing labor costs. In such situations, the high-low method can be modified so that the two observations chosen to estimate the cost function are a *representative high* and a *representative low*. By using this adjustment, managers can avoid having extreme observations, which arise from abnormal events, influence the estimate of the cost function. The modification allows managers to estimate a cost function that is representative of the relationship between the cost driver and costs and, therefore, is more useful for making decisions (such as pricing and performance evaluation).

The advantage of the high-low method is that it is simple to compute and easy to understand; it gives a quick, initial insight into how the cost driver—number of machine-hours—affects indirect manufacturing labor costs. The disadvantage is that it ignores information from all but two observations when estimating the cost function. We next describe the regression analysis method of quantitative analysis that uses all available data to estimate the cost function.



**Exhibit 10-5**  
High-Low Method for Weekly Indirect Manufacturing Labor Costs and Machine-Hours for Elegant Rugs

## Regression Analysis Method

**Regression analysis** is a statistical method that measures the average amount of change in the dependent variable associated with a unit change in one or more independent variables. In the Elegant Rugs example, the dependent variable is total indirect manufacturing labor costs. The independent variable, or cost driver, is number of machine-hours. **Simple regression** analysis estimates the relationship between the dependent variable and *one* independent variable. **Multiple regression** analysis estimates the relationship between the dependent variable and *two or more* independent variables. Multiple regression analysis for Elegant Rugs might use as the independent variables, or cost drivers, number of machine-hours and number of batches. The appendix to this chapter will explore simple regression and multiple regression in more detail.

In later sections, we will illustrate how Excel performs the calculations associated with regression analysis. The following discussion emphasizes how managers interpret and use the output from Excel to make critical strategic decisions. Exhibit 10-6 shows the line developed using regression analysis that best fits the data in columns B and C of Exhibit 10-3. Excel estimates the cost function to be

$$y = \$300.98 + \$10.31X$$

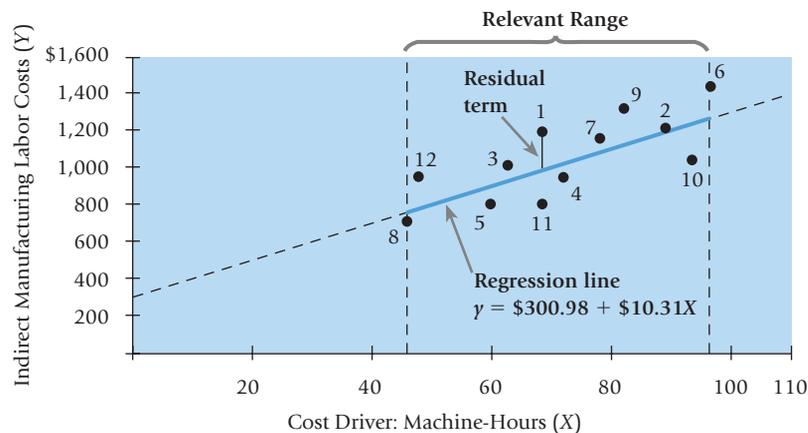
The regression line in Exhibit 10-6 is derived using the least-squares technique. The least-squares technique determines the regression line by minimizing the sum of the squared vertical differences from the data points (the various points in the graph) to the regression line. The vertical difference, called the **residual term**, measures the distance between actual cost and estimated cost for each observation of the cost driver. Exhibit 10-6 shows the residual term for the week 1 data. The line from the observation to the regression line is drawn perpendicular to the horizontal axis, or *x*-axis. The smaller the residual terms, the better the fit between actual cost observations and estimated costs. *Goodness of fit* indicates the strength of the relationship between the cost driver and costs. The regression line in Exhibit 10-6 rises from left to right. The positive slope of this line and small residual terms indicate that, on average, indirect manufacturing labor costs increase as the number of machine-hours increases. The vertical dashed lines in Exhibit 10-6 indicate the relevant range, the range within which the cost function applies.

*Instructors and students who want to explore the technical details of estimating the least-squares regression line, can go to the appendix, pages 367–371 and return to this point without any loss of continuity.*

The estimate of the slope coefficient, *b*, indicates that indirect manufacturing labor costs vary at the average amount of \$10.31 for every machine-hour used within the relevant range. Management can use the regression equation when budgeting for future indirect manufacturing labor costs. For instance, if 90 machine-hours are budgeted for the upcoming week, the predicted indirect manufacturing labor costs would be

$$y = \$300.98 + (\$10.31 \text{ per machine-hour} \times 90 \text{ machine-hours}) = \$1,228.88$$

**Exhibit 10-6**  
Regression Model for Weekly Indirect Manufacturing Labor Costs and Machine-Hours for Elegant Rugs



As we have already mentioned, the regression method is more accurate than the high-low method because the regression equation estimates costs using information from all observations, whereas the high-low equation uses information from only two observations. The inaccuracies of the high-low method can mislead managers. Consider the high-low method equation in the preceding section,  $y = \$23.68 + \$14.92 \text{ per machine-hour} \times \text{Number of machine-hours}$ . For 90 machine-hours, the predicted weekly cost based on the high-low method equation is  $\$23.68 + (\$14.92 \text{ per machine-hour} \times 90 \text{ machine-hours}) = \$1,366.48$ . Suppose that for 7 weeks over the next 12-week period, Elegant Rugs runs its machines for 90 hours each week. Assume average indirect manufacturing labor costs for those 7 weeks are \$1,300. Based on the high-low method prediction of \$1,366.48, Elegant Rugs would conclude it has performed well because actual costs are less than predicted costs. But comparing the \$1,300 performance with the more-accurate \$1,228.88 prediction of the regression model tells a much different story and would probably prompt Elegant Rugs to search for ways to improve its cost performance.

Accurate cost estimation helps managers predict future costs and evaluate the success of cost-reduction initiatives. Suppose the manager at Elegant Rugs is interested in evaluating whether recent strategic decisions that led to changes in the production process and resulted in the data in Exhibit 10-3 have reduced indirect manufacturing labor costs, such as supervision, maintenance, and quality control. Using data on number of machine-hours used and indirect manufacturing labor costs of the previous process (not shown here), the manager estimates the regression equation,

$$y = \$546.26 + (\$15.86 \text{ per machine-hour} \times \text{Number of machine-hours})$$

The constant (\$300.98 versus \$545.26) and the slope coefficient (\$10.31 versus \$15.86) are both smaller for the new process relative to the old process. It appears that the new process has decreased indirect manufacturing labor costs.

## Evaluating Cost Drivers of the Estimated Cost Function

How does a company determine the best cost driver when estimating a cost function? In many cases, the choice of a cost driver is aided substantially by understanding both operations and cost accounting.

To see why the understanding of operations is needed, consider the costs to maintain and repair metal-cutting machines at Helix Corporation, a manufacturer of treadmills. Helix schedules repairs and maintenance at a time when production is at a low level to avoid having to take machines out of service when they are needed most. An analysis of the monthly data will then show high repair costs in months of low production and low repair costs in months of high production. Someone unfamiliar with operations might conclude that there is an inverse relationship between production and repair costs. The engineering link between units produced and repair costs, however, is usually clear-cut. Over time, there is a cause-and-effect relationship: the higher the level of production, the higher the repair costs. To estimate the relationship correctly, operating managers and analysts will recognize that repair costs will tend to lag behind periods of high production, and hence, they will use production of prior periods as the cost driver.

In other cases, choosing a cost driver is more subtle and difficult. Consider again indirect manufacturing labor costs at Elegant Rugs. Management believes that both the number of machine-hours and the number of direct manufacturing labor-hours are plausible cost drivers of indirect manufacturing labor costs. However, management is not sure which is the better cost driver. Exhibit 10-7 presents weekly data (in Excel) on indirect manufacturing labor costs and number of machine-hours for the most recent 12-week period from Exhibit 10-3, together with data on the number of direct manufacturing labor-hours for the same period.

### Decision Point

What are the steps to estimate a cost function using quantitative analysis?

### Learning Objective 5

Describe three criteria used to evaluate and choose cost drivers

... economically plausible relationships, goodness of fit, and significant effect of the cost driver on costs

**Exhibit 10-7**

Weekly Indirect Manufacturing Labor Costs, Machine-Hours, and Direct Manufacturing Labor-Hours for Elegant Rugs

	Home	Insert	Page Layout	Formulas	Data	Review
	A	B	C	D		
		Original Cost Driver: Machine-Hours	Alternate Cost Driver: Direct Manufacturing Labor-Hours (X)	Indirect Manufacturing Labor Costs (Y)		
1	Week					
2	1	68	30	\$ 1,190		
3	2	88	35	1,211		
4	3	62	36	1,004		
5	4	72	20	917		
6	5	60	47	770		
7	6	96	45	1,456		
8	7	78	44	1,180		
9	8	46	38	710		
10	9	82	70	1,316		
11	10	94	30	1,032		
12	11	68	29	752		
13	12	<u>48</u>	<u>38</u>	<u>963</u>		
14	Total	<u>862</u>	<u>462</u>	<u>\$12,501</u>		
15						

### Choosing Among Cost Drivers

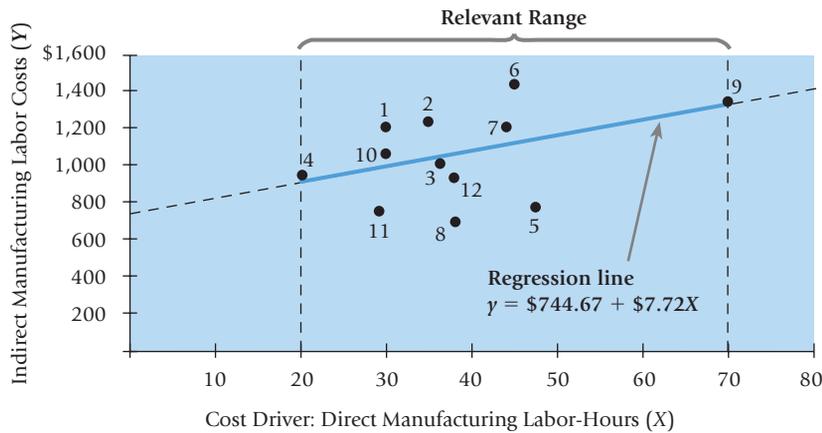
What guidance do the different cost-estimation methods provide for choosing among cost drivers? The industrial engineering method relies on analyzing physical relationships between cost drivers and costs, relationships that are difficult to specify in this case. The conference method and the account analysis method use subjective assessments to choose a cost driver and to estimate the fixed and variable components of the cost function. In these cases, managers must rely on their best judgment. Managers cannot use these methods to test and try alternative cost drivers. The major advantages of quantitative methods are that they are objective—a given data set and estimation method result in a unique estimated cost function—and managers can use them to evaluate different cost drivers. We use the regression analysis approach to illustrate how to evaluate different cost drivers.

First, the cost analyst at Elegant Rugs enters data in columns C and D of Exhibit 10-7 in Excel and estimates the following regression equation of indirect manufacturing labor costs based on number of direct manufacturing labor-hours:

$$y = \$744.67 + \$7.72X$$

Exhibit 10-8 shows the plot of the data points for number of direct manufacturing labor-hours and indirect manufacturing labor costs, and the regression line that best fits the data. Recall that Exhibit 10-6 shows the corresponding graph when number of machine-hours is the cost driver. To decide which of the two cost drivers Elegant Rugs should choose, the analyst compares the machine-hour regression equation and the direct manufacturing labor-hour regression equation. There are three criteria used to make this evaluation.

1. **Economic plausibility.** Both cost drivers are economically plausible. However, in the state-of-the-art, highly automated production environment at Elegant Rugs, managers familiar with the operations believe that costs such as machine maintenance are likely to be more closely related to number of machine-hours used than to number of direct manufacturing labor-hours used.
2. **Goodness of fit.** Compare Exhibits 10-6 and 10-8. The vertical differences between actual costs and predicted costs are much smaller for the machine-hours regression than for the direct manufacturing labor-hours regression. Number of machine-hours used, therefore, has a stronger relationship—or goodness of fit—with indirect manufacturing labor costs.

**Exhibit 10-8**

Regression Model for Weekly Indirect Manufacturing Labor Costs and Direct Manufacturing Labor-Hours for Elegant Rugs

3. **Significance of independent variable.** Again compare Exhibits 10-6 and 10-8 (both of which have been drawn to roughly the same scale). The machine-hours regression line has a steep slope relative to the slope of the direct manufacturing labor-hours regression line. *For the same (or more) scatter of observations about the line (goodness of fit), a flat, or slightly sloped regression line indicates a weak relationship between the cost driver and costs.* In our example, changes in direct manufacturing labor-hours appear to have a small influence or effect on indirect manufacturing labor costs.

Based on this evaluation, managers at Elegant Rugs select number of machine-hours as the cost driver and use the cost function  $y = \$300.98 + (\$10.31 \text{ per machine-hour} \times \text{Number of machine-hours})$  to predict future indirect manufacturing labor costs.

*Instructors and students who want to explore how regression analysis techniques can be used to choose among different cost drivers can go to the appendix, pages 371–374 and return to this point without any loss of continuity.*

Why is choosing the correct cost driver to estimate indirect manufacturing labor costs important? Because identifying the wrong drivers or misestimating cost functions can lead management to incorrect (and costly) decisions along a variety of dimensions. Consider the following strategic decision that management at Elegant Rugs must make. The company is thinking of introducing a new style of carpet that, from a manufacturing standpoint, is similar to the carpets it has manufactured in the past. Prices are set by the market and sales of 650 square yards of this carpet are expected each week. Management estimates 72 machine-hours and 21 direct manufacturing labor-hours would be required per week to produce the 650 square yards of carpet needed. Using the machine-hour regression equation, Elegant Rugs would predict indirect manufacturing labor costs of  $y = \$300.98 + (\$10.31 \text{ per machine-hour} \times 72 \text{ machine-hours}) = \$1,043.30$ . If it used direct manufacturing labor-hours as the cost driver, it would incorrectly predict costs of  $\$744.67 + (\$7.72 \text{ per labor-hour} \times 21 \text{ labor-hours}) = \$906.79$ . If Elegant Rugs chose similarly incorrect cost drivers for other indirect costs as well and systematically underestimated costs, it would conclude that the costs of manufacturing the new style of carpet would be low and basically fixed (fixed because the regression line is nearly flat). But the actual costs driven by number of machine-hours used and other correct cost drivers would be higher. By failing to identify the proper cost drivers, management would be misled into believing the new style of carpet would be more profitable than it actually is. It might decide to introduce the new style of carpet, whereas if Elegant identifies the correct cost driver it might decide not to introduce the new carpet.

Incorrectly estimating the cost function would also have repercussions for cost management and cost control. Suppose number of direct manufacturing labor-hours were used as the cost driver, and actual indirect manufacturing labor costs for the new carpet were \$970. Actual costs would then be higher than the predicted costs of \$906.79. Management would feel compelled to find ways to cut costs. In fact, on the basis of the preferred machine-hour cost driver, the plant would have actual costs lower than the \$1,043.30 predicted costs—a performance that management should seek to replicate, not change!

## Concepts in Action

### Activity-Based Costing: Identifying Cost and Revenue Drivers



Many cost estimation methods presented in this chapter are essential to service, manufacturing, and retail-sector implementations of activity-based costing across the globe. To determine the cost of an activity in the banking industry, ABC systems often rely on expert analyses and opinions gathered from operating personnel (the conference method). For example, the loan department staff at the Co-operative Bank in the United Kingdom subjectively estimate the costs of the loan processing activity and the quantity of the related cost driver—the number of loans processed, a batch-level cost driver, as distinguished from the amount of the loans, an output unit-level cost driver—to derive the cost of processing a loan.

Elsewhere in the United Kingdom, the City of London police force uses input-output relationships (the industrial engineering method) to identify cost drivers and the cost of an activity. Using a surveying methodology, officials can determine the total costs associated with responding to house robberies, dealing with burglaries, and filling out police reports. In the United States, the Boeing Commercial Airplane Group's Wichita Division used detailed analyses of its commercial airplane-manufacturing methods to support make/buy decisions for complex parts required in airplane assembly. The industrial engineering method is also used by U.S. government agencies such as the U.S. Postal Service to determine the cost of each post office transaction and the U.S. Patent and Trademark Office to identify the costs of each patent examination.

Regression analysis is another helpful tool for determining the cost drivers of activities. Consider how fuel service retailers (that is, gas stations with convenience stores) identify the principal cost driver for labor within their operations. Two possible cost drivers are gasoline sales and convenience store sales. Gasoline sales are batch-level activities because payment transactions occur only once for each gasoline purchase, regardless of the volume of gasoline purchased; whereas convenience store sales are output unit-level activities that vary based on the amount of food, drink, and other products sold. Fuel service retailers generally use convenience store sales as the basis for assigning labor costs because multiple regression analyses confirm that convenience store sales, not gasoline sales, are the major cost driver of labor within their operations.

While popular, these are not the only methods used to evaluate cost drivers. If you recall from chapter five, Charles Schwab is one of the growing number of companies using time-driven activity based costing, which uses time as the cost driver. At Citigroup, the company's internal technology infrastructure group uses time to better manage the labor capacity required to provide reliable, secure, and cost effective technology services to about 60 Citigroup business units around the world.

The trend of using activity-based costing to identify cost and revenue drivers also extends into emerging areas. For example, the U.S. government allocated \$19 billion in 2009 to support the adoption of electronic health records. Using the input-output method, many health clinics and doctor's offices are leveraging activity-based costing to identify the cost of adopting this new health information technology tool.

*Sources:* Barton, T., and J. MacArthur. 2003. Activity-based costing and predatory pricing: The case of the retail industry. *Management Accounting Quarterly* (Spring); Carter, T., A. Sedaghat, and T. Williams. 1998. How ABC changed the post office. *Management Accounting*, (February); The Cooperative Bank. Harvard Business School. Case No. N9-195-196; Federowicz, M., M. Grossman, B. Hayes, and J. Riggs. 2010. A tutorial on activity-based costing of electronic health records. *Quality Management in Health Care* (January–March); Kaplan, Robert, and Steven Anderson. 2008. *Time-driven activity-based costing: A simpler and more powerful path to higher profits*. Boston: Harvard Business School Publishing; Leapman, B. 2006. Police spend £500m filling in forms. *The Daily Telegraph*, January 22; Paduano, Rocco, and Joel Cutcher-Gershenfeld. 2001. Boeing Commercial Airplane Group Wichita Division (Boeing Co.). MIT Labor Aerospace Research Agenda Case Study. Cambridge, MA: MIT; Peckenpaugh, J. 2002. Teaching the ABCs. *Government Executive*, April 1; The United Kingdom Home Office. 2007. *The police service national ABC model: Manual of guidance*. London: Her Majesty's Stationary Office.

## Cost Drivers and Activity-Based Costing

Activity-based costing (ABC) systems focus on individual activities—such as product design, machine setup, materials handling, distribution, and customer service—as the fundamental cost objects. To implement ABC systems, managers must identify a cost driver for each activity. For example, using methods described in this chapter, the manager must decide whether the number of loads moved or the weight of loads moved is the cost driver of materials-handling costs.

To choose the cost driver and use it to estimate the cost function in our materials-handling example, the manager collects data on materials-handling costs and the quantities of the two competing cost drivers over a reasonably long period. Why a long period? Because in the short run, materials-handling costs may be fixed and, therefore, will not vary with changes in the level of the cost driver. In the long run, however, there is a clear cause-and-effect relationship between materials-handling costs and the cost driver. Suppose number of loads moved is the cost driver of materials-handling costs. Increases in the number of loads moved will require more materials-handling labor and equipment; decreases will result in equipment being sold and labor being reassigned to other tasks.

ABC systems have a great number and variety of cost drivers and cost pools. That means ABC systems require many cost relationships to be estimated. In estimating the cost function for each cost pool, the manager must pay careful attention to the cost hierarchy. For example, if a cost is a batch-level cost such as setup cost, the manager must only consider batch-level cost drivers like number of setup-hours. In some cases, the costs in a cost pool may have more than one cost driver from different levels of the cost hierarchy. In the Elegant Rugs example, the cost drivers for indirect manufacturing labor costs could be machine-hours and number of production batches of carpet manufactured. Furthermore, it may be difficult to subdivide the indirect manufacturing labor costs into two cost pools and to measure the costs associated with each cost driver. In these cases, companies use multiple regression to estimate costs based on more than one independent variable. The appendix to this chapter discusses multiple regression in more detail.

As the Concepts in Action feature (p. 356) illustrates, managers implementing ABC systems use a variety of methods—industrial engineering, conference, and regression analysis—to estimate slope coefficients. In making these choices, managers trade off level of detail, accuracy, feasibility, and costs of estimating cost functions.

## Nonlinear Cost Functions

In practice, cost functions are not always linear. A **nonlinear cost function** is a cost function for which the graph of total costs (based on the level of a single activity) is not a straight line within the relevant range. To see what a nonlinear cost function looks like, return to Exhibit 10-2 (p. 344). The relevant range is currently set at 20,000 to 65,000 snowboards. But if we extend the relevant range to encompass the region from 0 to 80,000 snowboards produced, it is evident that the cost function over this expanded range is graphically represented by a line that is not straight.

Consider another example. Economies of scale in advertising may enable an advertising agency to produce double the number of advertisements for less than double the costs. Even direct material costs are not always linear variable costs because of quantity discounts on direct material purchases. As shown in Exhibit 10-9 (p. 358), Panel A, total direct material costs rise as the units of direct materials purchased increase. But, because of quantity discounts, these costs rise more slowly (as indicated by the slope coefficient) as the units of direct materials purchased increase. This cost function has  $b = \$25$  per unit for 1–1,000 units purchased,  $b = \$15$  per unit for 1,001–2,000 units purchased, and  $b = \$10$  per unit for 2,001–3,000 units purchased. The direct material cost per unit falls at each price break—that is, the cost per unit decreases with larger purchase orders. If managers are interested in understanding cost behavior over the relevant range from 1 to 3,000 units, the cost function is nonlinear—not a straight line. If, however, managers are only interested in understanding cost behavior over a more narrow relevant range (for example, from 1 to 1,000 units), the cost function is linear.

*Step cost functions* are also examples of nonlinear cost functions. A **step cost function** is a cost function in which the cost remains the same over various ranges of the level of activity, but the cost increases by discrete amounts—that is, increases in steps—as the level of activity increases from one range to the next. Panel B in Exhibit 10-9 shows a *step variable-cost function*, a step cost function in which cost remains the same over narrow ranges of the level of activity in each relevant range. Panel B presents the relationship between units of production and setup costs. The pattern is a step cost function because, as we described in Chapter 5 on activity-based costing, setup costs are

### Decision Point

How should a company evaluate and choose cost drivers?

### Learning Objective 6

Explain nonlinear cost functions

... graph of cost function is not a straight line, for example, because of quantity discounts or costs changing in steps

in particular those arising from learning curve effects

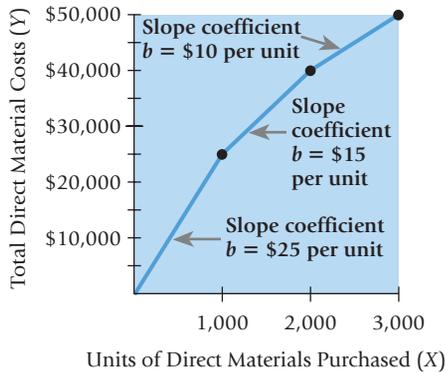
... either cumulative average-time learning, where cumulative average time per unit declines by a constant percentage, as units produced double

... or incremental unit-time learning, in which incremental time to produce last unit declines by constant percentage, as units produced double

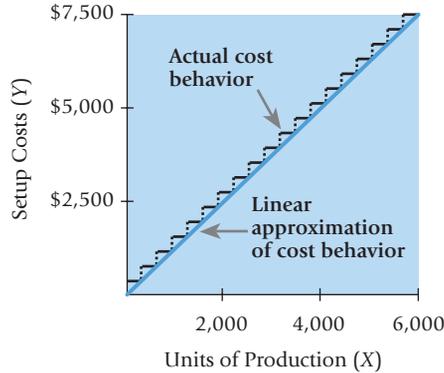
## Exhibit 10-9

## Examples of Nonlinear Cost Functions

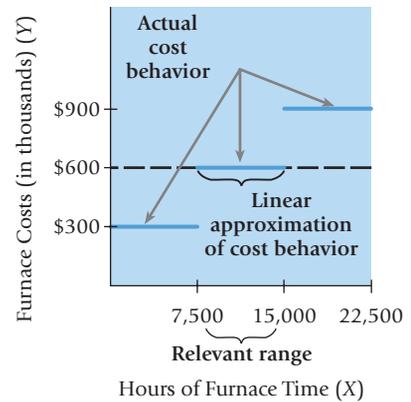
**PANEL A:**  
Effects of Quantity  
Discounts on Slope  
Coefficient of Direct  
Material Cost Function



**PANEL B:**  
Step Variable-Cost  
Function



**PANEL C:**  
Step Fixed-Cost  
Function



related to each production batch started. If the relevant range is considered to be from 0 to 6,000 production units, the cost function is nonlinear. However, as shown by the blue line in Panel B, managers often approximate step variable costs with a continuously-variable cost function. This type of step cost pattern also occurs when production inputs such as materials-handling labor, supervision, and process engineering labor are acquired in discrete quantities but used in fractional quantities.

Panel C in Exhibit 10-9 shows a *step fixed-cost function* for Crofton Steel, a company that operates large heat-treatment furnaces to harden steel parts. Looking at Panel C and Panel B, you can see that the main difference between a step variable-cost function and a step fixed-cost function is that the cost in a step fixed-cost function remains the same over *wide* ranges of the activity in each relevant range. The ranges indicate the number of furnaces being used (each furnace costs \$300,000). The cost increases from one range to the next higher range when the hours of furnace time needed require the use of another furnace. The relevant range of 7,500 to 15,000 hours of furnace time indicates that the company expects to operate with two furnaces at a cost of \$600,000. Management considers the cost of operating furnaces as a fixed cost within this relevant range of operation. However, if the relevant range is considered to be from 0 to 22,500 hours, the cost function is nonlinear: The graph in Panel C is not a single straight line; it is three broken lines.

## Learning Curves

Nonlinear cost functions also result from learning curves. A **learning curve** is a function that measures how labor-hours per unit decline as units of production increase because workers are learning and becoming better at their jobs. Managers use learning curves to predict how labor-hours, or labor costs, will increase as more units are produced.

The aircraft-assembly industry first documented the effect that learning has on efficiency. In general, as workers become more familiar with their tasks, their efficiency improves. Managers learn how to improve the scheduling of work shifts and how to operate the plant more efficiently. As a result of improved efficiency, unit costs decrease as productivity increases, and the unit-cost function behaves nonlinearly. These nonlinearities must be considered when estimating and predicting unit costs.

Managers have extended the learning-curve notion to other business functions in the value chain, such as marketing, distribution, and customer service, and to costs other than labor costs. The term *experience curve* describes this broader application of the learning curve. An **experience curve** is a function that measures the decline in cost per unit in various

business functions of the value chain—marketing, distribution, and so on—as the amount of these activities increases. For companies such as Dell Computer, Wal-Mart, and McDonald’s, learning curves and experience curves are key elements of their strategies. These companies use learning curves and experience curves to reduce costs and increase customer satisfaction, market share, and profitability.

We now describe two learning-curve models: the cumulative average-time learning model and the incremental unit-time learning model.

### Cumulative Average-Time Learning Model

In the **cumulative average-time learning model**, cumulative average time per unit declines by a constant percentage each time the cumulative quantity of units produced doubles. Consider Rayburn Corporation, a radar systems manufacturer. Rayburn has an 80% learning curve. The 80% means that when the quantity of units produced is doubled from  $X$  to  $2X$ , cumulative average time *per unit* for  $2X$  units is 80% of cumulative average time *per unit* for  $X$  units. Average time per unit has dropped by 20% ( $100\% - 80\%$ ). Exhibit 10-10 is an Excel spreadsheet showing the calculations for the cumulative average-time learning model for Rayburn Corporation. Note that as the number of units produced doubles from 1 to 2 in column A, cumulative average time per unit declines from 100 hours to 80% of 100 hours ( $0.80 \times 100$  hours = 80 hours) in column B. As the number of units doubles from 2 to 4, cumulative average time per unit declines to 80% of 80 hours = 64 hours, and so on. To obtain the cumulative total time in column D, multiply cumulative average time per unit by the cumulative number of units produced. For example, to produce 4 cumulative units would require 256 labor-hours ( $4 \text{ units} \times 64 \text{ cumulative average labor-hours per unit}$ ).

**Exhibit 10-10** Cumulative Average-Time Learning Model for Rayburn Corporation

Cumulative Average-Time Learning Model for Rayburn Corporation									
80% Learning Curve									
Cumulative Number of Units (X)	Cumulative Average Time per Unit (y)*: Labor-Hours	Cumulative Total Time: Labor-Hours	Individual Unit Time for Xth Unit: Labor-Hours						
				D = Col A × Col B					
1	100.00	100.00	100.00	E13 = D13 - D12 = 210.63 - 160.00					
2	80.00 = (100 × 0.8)	160.00	60.00						
3	70.21	210.63	50.63						
4	64.00 = (80 × 0.8)	256.00	45.37						
5	59.56	297.82	41.82						
6	56.17	337.01	39.19						
7	53.45	374.14	37.13						
8	51.20 = (64 × 0.8)	409.60	35.46						
9	49.29	443.65	34.05						
10	47.65	476.51	32.86						
11	46.21	508.32	31.81						
12	44.93	539.22	30.89						
13	43.79	569.29	30.07						
14	42.76	598.63	29.34						
15	41.82	627.30	28.67						
16	40.96 = (51.2 × 0.8)	655.36	28.06						

\*The mathematical relationship underlying the cumulative average-time learning model is as follows:

$$y = aX^b$$

where  $y$  = Cumulative average time (labor-hours) per unit  
 $X$  = Cumulative number of units produced  
 $a$  = Time (labor-hours) required to produce the first unit  
 $b$  = Factor used to calculate cumulative average time to produce units

The value of  $b$  is calculated as

$$\frac{\ln(\text{learning-curve \% in decimal form})}{\ln 2}$$

For an 80% learning curve,  $b = \ln 0.8 / \ln 2 = -0.2231 / 0.6931 = -0.3219$   
 For example, when  $X = 3$ ,  $a = 100$ ,  $b = -0.3219$ ,  
 $y = 100 \times 3^{-0.3219} = 70.21$  labor-hours

The cumulative total time when  $X = 3$  is  $70.21 \times 3 = 210.63$  labor-hours.  
 Numbers in table may not be exact because of rounding.

### Incremental Unit-Time Learning Model

In the **incremental unit-time learning model**, incremental time needed to produce the last unit declines by a constant percentage each time the cumulative quantity of units produced doubles. Again, consider Rayburn Corporation and an 80% learning curve. The 80% here means that when the quantity of units produced is doubled from  $X$  to  $2X$ , the time needed to produce the last unit when  $2X$  total units are produced is 80% of the time needed to produce the last unit when  $X$  total units are produced. Exhibit 10-11 is an Excel spreadsheet showing the calculations for the incremental unit-time learning model for Rayburn Corporation based on an 80% learning curve. Note how when units produced double from 2 to 4 in column A, the time to produce unit 4 (the last unit when 4 units are produced) is 64 hours in column B, which is 80% of the 80 hours needed to produce unit 2 (the last unit when 2 units are produced). We obtain the cumulative total time in column D by summing individual unit times in column B. For example, to produce 4 cumulative units would require 314.21 labor-hours (100.00 + 80.00 + 70.21 + 64.00).

Exhibit 10-12 presents graphs using Excel for the cumulative average-time learning model (using data from Exhibit 10-10) and the incremental unit-time learning model (using data from Exhibit 10-11). Panel A graphically illustrates cumulative average time per unit as a function of cumulative units produced for each model (column A in Exhibit 10-10 or 10-11). The curve for the cumulative average-time learning model is plotted using the data from Exhibit 10-10, column B, while the curve for the incremental unit-time learning model is plotted using the data from Exhibit 10-11, column E. Panel B graphically illustrates cumulative total labor-hours, again as a function of cumulative units produced for each model. The curve for the cumulative average-time learning model is plotted using the data from Exhibit 10-10, column D, while that for the incremental unit-time learning model is plotted using the data from Exhibit 10-11, column D.

**Exhibit 10-11** Incremental Unit-Time Learning Model for Rayburn Corporation

Incremental Unit-Time Learning Model for Rayburn Corporation										
80% Learning Curve										
Cumulative Number of Units (X)	Individual Unit Time for Xth Unit (y)*: Labor-Hours	Cumulative Total Time: Labor-Hours	Cumulative Average Time per Unit: Labor-Hours							
				E = Col D ÷ Col A						
1	100.00	100.00	100.00							
2	80.00 = (100 × 0.8)	180.00	90.00	D14 = D13 + B14 = 180.00 + 70.21						
3	70.21	250.21	83.40							
4	64.00 = (80 × 0.8)	314.21	78.55							
5	59.56	373.77	74.75							
6	56.17	429.94	71.66							
7	53.45	483.39	69.06							
8	51.20 = (64 × 0.8)	534.59	66.82							
9	49.29	583.89	64.88							
10	47.65	631.54	63.15							
11	46.21	677.75	61.61							
12	44.93	722.68	60.22							
13	43.79	766.47	58.96							
14	42.76	809.23	57.80							
15	41.82	851.05	56.74							
16	40.96 = (51.2 × 0.8)	892.01	55.75							

\*The mathematical relationship underlying the incremental unit-time learning model is as follows:

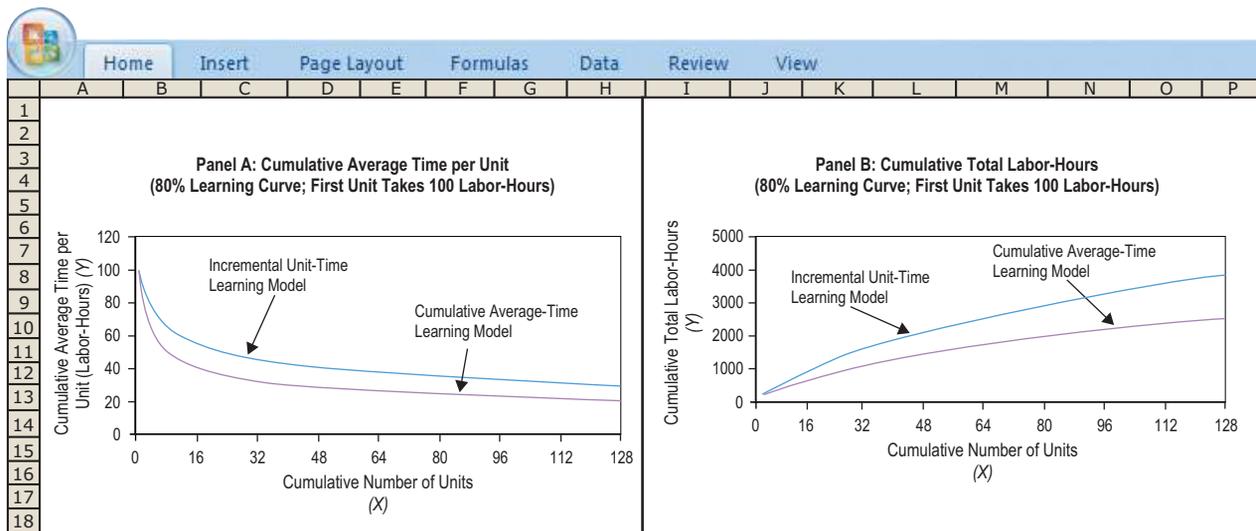
$$y = aX^b$$

where  $y$  = Time (labor-hours) taken to produce the last single unit  
 $X$  = Cumulative number of units produced  
 $a$  = Time (labor-hours) required to produce the first unit  
 $b$  = Factor used to calculate incremental unit time to produce units  
 $= \frac{\ln(\text{learning-curve \% in decimal form})}{\ln 2}$

For an 80% learning curve,  $b = \ln 0.8 \div \ln 2 = -0.2231 \div 0.6931 = -0.3219$   
 For example, when  $X = 3$ ,  $a = 100$ ,  $b = -0.3219$ ,  
 $y = 100 \times 3^{-0.3219} = 70.21$  labor-hours  
 The cumulative total time when  $X = 3$  is  $100 + 80 + 70.21 = 250.21$  labor-hours.  
 Numbers in the table may not be exact because of rounding.

## Exhibit 10-12

Plots for Cumulative Average-Time Learning Model and Incremental Unit-Time Learning Model for Rayburn Corporation



The incremental unit-time learning model predicts a higher cumulative total time to produce 2 or more units than the cumulative average-time learning model, assuming the same learning rate for both models. That is, in Exhibit 10-12, Panel B, the graph for the 80% incremental unit-time learning model lies above the graph for the 80% cumulative average-time learning model. If we compare the results in Exhibit 10-10 (column D) with the results in Exhibit 10-11 (column D), to produce 4 cumulative units, the 80% incremental unit-time learning model predicts 314.21 labor-hours versus 256.00 labor-hours predicted by the 80% cumulative average-time learning model. That's because under the cumulative average-time learning model *average labor-hours needed to produce all 4 units* is 64 hours; the labor-hour amount needed to produce unit 4 is much less than 64 hours—it is 45.37 hours (see Exhibit 10-10). Under the incremental unit-time learning model, the labor-hour amount needed to produce unit 4 is 64 hours, and the labor-hours needed to produce the first 3 units are more than 64 hours, so average time needed to produce all 4 units is more than 64 hours.

How do managers choose which model and what percent learning curve to use? It is important to recognize that managers make their choices on a case-by-case basis. For example, if the behavior of manufacturing labor-hour usage as production levels increase follows a pattern like the one predicted by the 80% learning curve cumulative average-time learning model, then the 80% learning curve cumulative average-time learning model should be used. Engineers, plant managers, and workers are good sources of information on the amount and type of learning actually occurring as production increases. Plotting this information and estimating the model that best fits the data is helpful in selecting the appropriate model.<sup>2</sup>

## Incorporating Learning-Curve Effects into Prices and Standards

How do companies use learning curves? Consider the data in Exhibit 10-10 for the cumulative average-time learning model at Rayburn Corporation. Suppose variable costs subject to learning effects consist of direct manufacturing labor, at \$20 per hour, and related overhead, at \$30 per direct manufacturing labor-hour. Managers should predict the costs shown in Exhibit 10-13.

These data show that the effects of the learning curve could have a major influence on decisions. For example, managers at Rayburn Corporation might set an extremely low selling price on its radar systems to generate high demand. As its production increases to meet this growing demand, cost per unit drops. Rayburn “rides the product down the

<sup>2</sup> For details, see C. Bailey, “Learning Curve Estimation of Production Costs and Labor-Hours Using a Free Excel Add-In,” *Management Accounting Quarterly*, (Summer 2000: 25–31). Free software for estimating learning curves is available at Dr. Bailey’s Web site, [www.profbaily.com](http://www.profbaily.com).

**Exhibit 10-13**Predicting Costs Using  
Learning Curves at  
Rayburn Corporation

	A	B	C	D	E	F
1		<b>Cumulative</b>				
2	<b>Cumulative</b>	<b>Average Time</b>	<b>Cumulative</b>	<b>Cumulative Costs</b>		<b>Additions to</b>
3	<b>Number of</b>	<b>per Unit:</b>	<b>Total Time:</b>	<b>at \$50 per</b>		<b>Cumulative</b>
4	<b>Units</b>	<b>Labor-Hours<sup>a</sup></b>	<b>Labor-Hours<sup>a</sup></b>	<b>Labor-Hour</b>		<b>Costs</b>
5	1	100.00	100.00	\$ 5,000	(100.00 × \$50)	\$ 5,000
6	2	80.00	160.00	8,000	(160.00 × \$50)	3,000
7	4	64.00	256.00	12,800	(256.00 × \$50)	4,800
8	8	51.20	409.60	20,480	(409.60 × \$50)	7,680
9	16	40.96	655.36	32,768	(655.36 × \$50)	12,288
10						
11	<sup>a</sup> Based on the cumulative average-time learning model. See Exhibit 10-10 for the computations					
12	of these amounts.					

learning curve” as it establishes a larger market share. Although it may have earned little operating income on its first unit sold—it may actually have lost money on that unit—Rayburn earns more operating income per unit as output increases.

Alternatively, subject to legal and other considerations, Rayburn’s managers might set a low price on just the final 8 units. After all, the total labor and related overhead costs per unit for these final 8 units are predicted to be only \$12,288 (\$32,768 – \$20,480). On these final 8 units, the \$1,536 cost per unit (\$12,288 ÷ 8 units) is much lower than the \$5,000 cost per unit of the first unit produced.

Many companies, such as Pizza Hut and Home Depot, incorporate learning-curve effects when evaluating performance. The Nissan Motor Company expects its workers to learn and improve on the job and evaluates performance accordingly. It sets assembly-labor efficiency standards for new models of cars after taking into account the learning that will occur as more units are produced.

The learning-curve models examined in Exhibits 10-10 to 10-13 assume that learning is driven by a single variable (production output). Other models of learning have been developed (by companies such as Analog Devices and Hewlett-Packard) that focus on how quality—rather than manufacturing labor-hours—will change over time, regardless of whether more units are produced. Studies indicate that factors other than production output, such as job rotation and organizing workers into teams, contribute to learning that improves quality.

**Decision Point**

What is a nonlinear cost function and in what ways do learning curves give rise to nonlinearities?

**Learning Objective 7**

Be aware of data problems encountered in estimating cost functions

... for example, unreliable data and poor record keeping, extreme observations, treating fixed costs as if they are variable, and a changing relationship between a cost driver and cost

**Data Collection and Adjustment Issues**

The ideal database for estimating cost functions quantitatively has two characteristics:

1. The database should contain numerous reliably measured observations of the cost driver (the independent variable) and the related costs (the dependent variable). Errors in measuring the costs and the cost driver are serious. They result in inaccurate estimates of the effect of the cost driver on costs.
2. The database should consider many values spanning a wide range for the cost driver. Using only a few values of the cost driver that are grouped closely considers too small a segment of the relevant range and reduces the confidence in the estimates obtained.

Unfortunately, cost analysts typically do not have the advantage of working with a database having both characteristics. This section outlines some frequently encountered data problems and steps the cost analyst can take to overcome these problems.

1. The time period for measuring the dependent variable (for example, machine-lubricant costs) does not properly match the period for measuring the cost driver. This problem often arises when accounting records are not kept on the accrual basis. Consider a cost function with machine-lubricant costs as the dependent variable and number of machine-hours as the cost driver. Assume that the lubricant is purchased sporadically

and stored for later use. Records maintained on the basis of lubricants purchased will indicate little lubricant costs in many months and large lubricant costs in other months. These records present an obviously inaccurate picture of what is actually taking place. The analyst should use accrual accounting to measure cost of lubricants consumed to better match costs with the machine-hours cost driver in this example.

2. Fixed costs are allocated as if they are variable. For example, costs such as depreciation, insurance, or rent may be allocated to products to calculate cost per unit of output. *The danger is to regard these costs as variable rather than as fixed. They seem to be variable because of the allocation methods used.* To avoid this problem, the analyst should carefully distinguish fixed costs from variable costs and not treat allocated fixed cost per unit as a variable cost.
3. Data are either not available for all observations or are not uniformly reliable. Missing cost observations often arise from a failure to record a cost or from classifying a cost incorrectly. For example, marketing costs may be understated because costs of sales visits to customers may be incorrectly recorded as customer-service costs. Recording data manually rather than electronically tends to result in a higher percentage of missing observations and erroneously entered observations. Errors also arise when data on cost drivers originate outside the internal accounting system. For example, the accounting department may obtain data on testing-hours for medical instruments from the company's manufacturing department and data on number of items shipped to customers from the distribution department. One or both of these departments might not keep accurate records. To minimize these problems, the cost analyst should design data collection reports that regularly and routinely obtain the required data and should follow up immediately whenever data are missing.
4. Extreme values of observations occur from errors in recording costs (for example, a misplaced decimal point), from nonrepresentative periods (for example, from a period in which a major machine breakdown occurred or from a period in which a delay in delivery of materials from an international supplier curtailed production), or from observations outside the relevant range. Analysts should adjust or eliminate unusual observations before estimating a cost relationship.
5. There is no homogeneous relationship between the cost driver and the individual cost items in the dependent variable-cost pool. A homogeneous relationship exists when each activity whose costs are included in the dependent variable has the same cost driver. In this case, a single cost function can be estimated. As discussed in Step 2 for estimating a cost function using quantitative analysis (p. 348), when the cost driver for each activity is different, separate cost functions (each with its own cost driver) should be estimated for each activity. Alternatively, as discussed on pages 372–374, the cost function should be estimated with more than one independent variable using multiple regression.
6. The relationship between the cost driver and the cost is not stationary. That is, the underlying process that generated the observations has not remained stable over time. For example, the relationship between number of machine-hours and manufacturing overhead costs is unlikely to be stationary when the data cover a period in which new technology was introduced. One way to see if the relationship is stationary is to split the sample into two parts and estimate separate cost relationships—one for the period before the technology was introduced and one for the period after the technology was introduced. Then, if the estimated coefficients for the two periods are similar, the analyst can pool the data to estimate a single cost relationship. When feasible, pooling data provides a larger data set for the estimation, which increases confidence in the cost predictions being made.
7. Inflation has affected costs, the cost driver, or both. For example, inflation may cause costs to change even when there is no change in the level of the cost driver. To study the underlying cause-and-effect relationship between the level of the cost driver and costs, the analyst should remove purely inflationary price effects from the data by dividing each cost by the price index on the date the cost was incurred.

In many cases, a cost analyst must expend considerable effort to reduce the effect of these problems before estimating a cost function on the basis of past data.

**Decision Point**

What are the common data problems a company must watch for when estimating costs?

## Problem for Self-Study

The Helicopter Division of GLD, Inc., is examining helicopter assembly costs at its Indiana plant. It has received an initial order for eight of its new land-surveying helicopters. GLD can adopt one of two methods of assembling the helicopters:

Home Insert Page Layout Formulas Data Review View					
	A	B	C	D	E
1		<b>Labor-Intensive Assembly Method</b>		<b>Machine-Intensive Assembly Method</b>	
2	Direct material cost per helicopter	\$ 40,000		\$36,000	
3	Direct-assembly labor time for first helicopter	2,000	labor-hours	800	labor-hours
4	Learning curve for assembly labor time per helicopter	85%	cumulative average time*	90%	incremental unit time**
5	Direct-assembly labor cost	\$ 30	per hour	\$ 30	per hour
6	Equipment-related indirect manufacturing cost	\$ 12	per direct-assembly labor-hour	\$ 45	per direct-assembly labor-hour
7	Material-handling-related indirect manufacturing cost	50%	of direct material cost	50%	of direct material cost
8					
9					
10	*Using the formula (p. 359), for an 85% learning curve, $b = \frac{\ln 0.85}{\ln 2} = \frac{-0.162519}{0.693147} = -0.234465$				
11					
12					
13					
14	**Using the formula (p. 360), for a 90% learning curve, $b = \frac{\ln 0.90}{\ln 2} = \frac{-0.105361}{0.693147} = -0.152004$				
15					
16					
17					

**Required**

- How many direct-assembly labor-hours are required to assemble the first eight helicopters under (a) the labor-intensive method and (b) the machine-intensive method?
- What is the total cost of assembling the first eight helicopters under (a) the labor-intensive method and (b) the machine-intensive method?

**Solution**

- a. The following calculations show the labor-intensive assembly method based on an 85% cumulative average-time learning model (using Excel):

Home Insert Page Layout Formulas Data Review View					
	G	H	I	J	K
1	<b>Cumulative</b>	<b>Cumulative</b>		<b>Cumulative</b>	<b>Individual</b>
2	<b>Number</b>	<b>Average Time</b>		<b>Total Time:</b>	<b>time for</b>
3	<b>of Units</b>	<b>per Unit (y):</b>		<b>Labor-Hours</b>	<b>Xth unit:</b>
4		<b>Labor-Hours</b>			<b>Labor-Hours</b>
5				<b>Col J = Col G × Col H</b>	
6	1	2,000		2,000	2,000
7	2	1,700	(2,000 × 0.85)	3,400	1,400
8	3	1,546		4,637	1,237
9	4	1,445	(1,700 × 0.85)	5,780	1,143
10	5	1,371		6,857	1,077
11	6	1,314		7,884	1,027
12	7	1,267		8,871	987
13	8	1,228.25	(1,445 × 0.85)	9,826	955
14					

Cumulative average-time per unit for the Xth unit in column H is calculated as  $y = aX^b$ ; see Exhibit 10-10 (p. 359). For example, when  $X = 3$ ,  $y = 2,000 \times 3^{-0.234465} = 1,546$  labor-hours.

- b. The following calculations show the machine-intensive assembly method based on a 90% incremental unit-time learning model:

	G	H	I	J	K
1	<b>Cumulative</b>	<b>Individual</b>		<b>Cumulative</b>	<b>Cumulative</b>
2	<b>Number</b>	<b>Unit Time</b>		<b>Total Time:</b>	<b>Average Time</b>
3	<b>of Units</b>	<b>for Xth Unit (y):</b>		<b>Labor-Hours</b>	<b>Per Unit:</b>
4		<b>Labor-Hours</b>			<b>Labor-Hours</b>
5					<b>Col K = Col J ÷ Col G</b>
6	1	800		800	800
7	2	720	(800 × 0.9)	1,520	760
8	3	677		2,197	732
9	4	648	(720 × 0.9)	2,845	711
10	5	626		3,471	694
11	6	609		4,081	680
12	7	595		4,676	668
13	8	583	(648 × 0.9)	5,258	657

Individual unit time for the Xth unit in column H is calculated as  $y = aX^b$ ; see Exhibit 10-11 (p. 360). For example, when  $X = 3$ ,  $y = 800 \times 3^{-0.152004} = 677$  labor-hours.

- 2. Total costs of assembling the first eight helicopters are as follows:

	O	P	Q
1		<b>Labor-Intensive</b>	<b>Machine-Intensive</b>
2		<b>Assembly Method</b>	<b>Assembly Method</b>
3		<b>(using data from part 1a)</b>	<b>(using data from part 1b)</b>
4	Direct materials:		
5	8 helicopters × \$40,000; \$36,000 per helicopter	\$320,000	\$288,000
6	Direct-assembly labor:		
7	9,826 hrs.; 5,258 hrs. × \$30/hr.	294,780	157,740
8	Indirect manufacturing costs		
9	Equipment related		
10	9,826 hrs. × \$12/hr.; 5,258 hrs. × \$45/hr.	117,912	236,610
11	Materials-handling related		
12	0.50 × \$320,000; \$288,000	<u>160,000</u>	<u>144,000</u>
13	Total assembly costs	<u>\$892,692</u>	<u>\$826,350</u>

The machine-intensive method’s assembly costs are \$66,342 lower than the labor-intensive method (\$892,692 – \$826,350).

## Decision Points

The following question-and-answer format summarizes the chapter's learning objectives. Each decision presents a key question related to a learning objective. The guidelines are the answer to that question.

### Decision

1. What is a linear cost function and what types of cost behavior can it represent?

2. What is the most important issue in estimating a cost function?

3. What are the different methods that can be used to estimate a cost function?

4. What are the steps to estimate a cost function using quantitative analysis?

5. How should a company evaluate and choose cost drivers?

6. What is a nonlinear cost function and in what ways do learning curves give rise to nonlinearities?

7. What are the common data problems a company must watch for when estimating costs?

### Guidelines

A linear cost function is a cost function in which, within the relevant range, the graph of total costs based on the level of a single activity is a straight line. Linear cost functions can be described by a constant,  $a$ , which represents the estimate of the total cost component that, within the relevant range, does not vary with changes in the level of the activity; and a slope coefficient,  $b$ , which represents the estimate of the amount by which total costs change for each unit change in the level of the activity within the relevant range. Three types of linear cost functions are variable, fixed, and mixed (or semivariable).

The most important issue in estimating a cost function is determining whether a cause-and-effect relationship exists between the level of an activity and the costs related to that level of activity. Only a cause-and-effect relationship—not merely correlation—establishes an economically plausible relationship between the level of an activity and its costs.

Four methods for estimating cost functions are the industrial engineering method, the conference method, the account analysis method, and the quantitative analysis method (which includes the high-low method and the regression analysis method). If possible, the cost analyst should apply more than one method. Each method is a check on the others.

There are six steps to estimate a cost function using quantitative analysis: (a) Choose the dependent variable; (b) identify the cost driver; (c) collect data on the dependent variable and the cost driver; (d) plot the data; (e) estimate the cost function; and (f) evaluate the cost driver of the estimated cost function. In most situations, working closely with operations managers, the cost analyst will cycle through these steps several times before identifying an acceptable cost function.

Three criteria for evaluating and choosing cost drivers are (a) economic plausibility, (b) goodness of fit, and (c) significance of independent variable.

A nonlinear cost function is one in which the graph of total costs based on the level of a single activity is not a straight line within the relevant range. Nonlinear costs can arise because of quantity discounts, step cost functions, and learning-curve effects. With learning curves, labor-hours per unit decline as units of production increase. In the cumulative average-time learning model, cumulative average-time per unit declines by a constant percentage each time the cumulative quantity of units produced doubles. In the incremental unit-time learning model, the time needed to produce the last unit declines by a constant percentage each time the cumulative quantity of units produced doubles.

The most difficult task in cost estimation is collecting high-quality, reliably measured data on the costs and the cost driver. Common problems include missing data, extreme values of observations, changes in technology, and distortions resulting from inflation.

## Appendix

### Regression Analysis

This appendix describes estimation of the regression equation, several commonly used regression statistics, and how to choose among cost functions that have been estimated by regression analysis. We use the data for Elegant Rugs presented in Exhibit 10-3 (p. 348) and displayed here again for easy reference.

Week	Cost Driver: Machine-Hours (X)	Indirect Manufacturing Labor Costs (Y)
1	68	\$ 1,190
2	88	1,211
3	62	1,004
4	72	917
5	60	770
6	96	1,456
7	78	1,180
8	46	710
9	82	1,316
10	94	1,032
11	68	752
12	48	963
Total	<u>862</u>	<u>\$12,501</u>

### Estimating the Regression Line

The least-squares technique for estimating the regression line minimizes the sum of the squares of the vertical deviations from the data points to the estimated regression line (also called *residual term* in Exhibit 10-6, p. 352). The objective is to find the values of  $a$  and  $b$  in the linear cost function  $y = a + bX$ , where  $y$  is the *predicted* cost value as distinguished from the *observed* cost value, which we denote by  $Y$ . We wish to find the numerical values of  $a$  and  $b$  that minimize  $\Sigma(Y - y)^2$ , the sum of the squares of the vertical deviations between  $Y$  and  $y$ . Generally, these computations are done using software packages such as Excel. For the data in our example,<sup>3</sup>  $a = \$300.98$  and  $b = \$10.31$ , so that the equation of the regression line is  $y = \$300.98 + \$10.31X$ .

### Goodness of Fit

Goodness of fit measures how well the predicted values,  $y$ , based on the cost driver,  $X$ , match actual cost observations,  $Y$ . The regression analysis method computes a measure of goodness of fit, called the **coefficient of determination**. The coefficient of determination ( $r^2$ ) measures the percentage of variation in  $Y$  explained by  $X$  (the independent variable).

<sup>3</sup> The formulae for  $a$  and  $b$  are as follows:

$$a = \frac{(\Sigma Y)(\Sigma X^2) - (\Sigma X)(\Sigma XY)}{n(\Sigma X^2) - (\Sigma X)(\Sigma X)} \text{ and } b = \frac{n(\Sigma XY) - (\Sigma X)(\Sigma Y)}{n(\Sigma X^2) - (\Sigma X)(\Sigma X)}$$

where for the Elegant Rugs data in Exhibit 10-3,

$$\begin{aligned} n &= \text{number of data points} = 12 \\ \Sigma X &= \text{sum of the given } X \text{ values} = 68 + 88 + \dots + 48 = 862 \\ \Sigma X^2 &= \text{sum of squares of the } X \text{ values} = (68)^2 + (88)^2 + \dots + (48)^2 + 4,624 + 7,744 + \dots + 2,304 = 64,900 \\ \Sigma Y &= \text{sum of given } Y \text{ values} = 1,190 + 1,211 + \dots + 963 = 12,501 \\ \Sigma XY &= \text{sum of the amounts obtained by multiplying each of the given } X \text{ values by the associated observed} \\ &\quad Y \text{ value} = (68)(1,190) + (88)(1,211) + \dots + (48)(963) \\ &\quad = 80,920 + 106,568 + \dots + 46,224 = 928,716 \\ a &= \frac{(12,501)(64,900) - (862)(928,716)}{12(64,900) - (862)(862)} = \$300.98 \\ b &= \frac{12(928,716) - (862)(12,501)}{12(64,900) - (862)(862)} = \$10.31 \end{aligned}$$

It is more convenient to express the coefficient of determination as 1 minus the proportion of total variance that is *not* explained by the independent variable—that is, 1 minus the ratio of unexplained variation to total variation. The unexplained variance arises because of differences between the actual values,  $Y$ , and the predicted values,  $y$ , which in the Elegant Rugs example is given by<sup>4</sup>

$$r^2 = 1 - \frac{\text{Unexplained variation}}{\text{Total variation}} = 1 - \frac{\sum(Y - y)^2}{\sum(Y - \bar{Y})^2} = 1 - \frac{290,824}{607,699} = 0.52$$

The calculations indicate that  $r^2$  increases as the predicted values,  $y$ , more closely approximate the actual observations,  $Y$ . The range of  $r^2$  is from 0 (implying no explanatory power) to 1 (implying perfect explanatory power). Generally, an  $r^2$  of 0.30 or higher passes the goodness-of-fit test. However, do not rely exclusively on goodness of fit. It can lead to the indiscriminate inclusion of independent variables that increase  $r^2$  but have no economic plausibility as cost drivers. *Goodness of fit has meaning only if the relationship between the cost drivers and costs is economically plausible.*

An alternative and related way to evaluate goodness of fit is to calculate the *standard error of the regression*. The **standard error of the regression** is the variance of the residuals. It is equal to

$$S = \sqrt{\frac{\sum(Y - y)^2}{\text{Degrees of freedom}}} = \sqrt{\frac{\sum(Y - y)^2}{n - 2}} = \sqrt{\frac{290,824}{12 - 2}} = \$170.54$$

Degrees of freedom equal the number of observations, 12, *minus* the number of coefficients estimated in the regression (in this case two,  $a$  and  $b$ ). On average, actual  $Y$  and the predicted value,  $y$ , differ by \$170.54. For comparison,  $\bar{Y}$ , the average value of  $Y$ , is \$1,041.75. The smaller the standard error of the regression, the better the fit and the better the predictions for different values of  $X$ .

## Significance of Independent Variables

Do changes in the economically plausible independent variable result in significant changes in the dependent variable? Or alternatively stated, is the slope coefficient,  $b = \$10.31$ , of the regression line statistically significant (that is, different from \$0)? Recall, for example, that in the regression of number of machine-hours and indirect manufacturing labor costs in the Elegant Rugs illustration,  $b$  is estimated from a sample of 12 weekly observations. The estimate,  $b$ , is subject to random factors, as are all sample statistics. That is, a different sample of 12 data points would undoubtedly give a different estimate of  $b$ . The **standard error of the estimated coefficient** indicates how much the estimated value,  $b$ , is likely to be affected by random factors. The  $t$ -value of the  $b$  coefficient measures how large the value of the estimated coefficient is relative to its standard error.

The cutoff  $t$ -value for making inferences about the  $b$  coefficient is a function of the number of degrees of freedom, the significance level, and whether it is a one-sided or two-sided test. A 5% level of significance indicates that there is less than a 5% probability that random factors could have affected the coefficient  $b$ . A two-sided test assumes that random factors could have caused the coefficient to be either greater than \$10.31 or less than \$10.31 with equal probability. At a 5% level of significance, this means that there is less than a 2.5% ( $5\% \div 2$ ) probability that random factors could have caused the coefficient to be greater than \$10.31 and less than 2.5% probability that random factors could have caused the coefficient to be less than \$10.31. Under the expectation that the coefficient of  $b$  is positive, a one-sided test at the 5% level of significance assumes that there is less than 5% probability that random factors would have caused the coefficient to be less than \$10.31. The cutoff  $t$ -value at the 5% significance level and 10 degrees of freedom for a two-sided test is 2.228. If there were more observations and 60 degrees of freedom, the cutoff  $t$ -value would be 2.00 at a 5% significance level for a two-sided test.

The  $t$ -value (called  $t$  Stat in the Excel output) for the slope coefficient  $b$  is the value of the estimated coefficient, \$10.31  $\div$  the standard error of the estimated coefficient \$3.12 = 3.30, which exceeds the cutoff  $t$ -value of 2.228. In other words, a relationship exists between the independent variable, machine-hours, and the dependent variable that cannot be attributed to random chance alone. Exhibit 10-14 shows a convenient format (in Excel) for summarizing the regression results for number of machine-hours and indirect manufacturing labor costs.

<sup>4</sup> From footnote 3,  $\sum Y = 12,501$  and  $\bar{Y} = 12,501 \div 12 = 1,041.75$

$$\sum(Y - \bar{Y})^2 = (1,190 - 1,041.75)^2 + (1,211 - 1,041.75)^2 + \dots + (963 - 1,041.75)^2 = 607,699$$

Each value of  $X$  generates a predicted value of  $y$ . For example, in week 1,  $y = \$300.98 + (\$10.31 \times 68) = \$1002.06$ ; in week 2,  $y = \$300.98 + (\$10.31 \times 88) = \$1,208.26$ ; and in week 12,  $y = \$300.98 + (\$10.31 \times 48) = \$795.86$ . Comparing the predicted and actual values,

$$\sum(Y - y)^2 = (1,190 - 1,002.06)^2 + (1,211 - 1,208.26)^2 + \dots + (963 - 795.86)^2 = 290,824$$

## Exhibit 10-14

Simple Regression Results with Indirect Manufacturing Labor Costs as Dependent Variable and Machine-Hours as Independent Variable (Cost Driver) for Elegant Rugs

	A	B	C	D	E	F
1		<b>Coefficients</b>	<b>Standard Error</b>	<b>t Stat</b>		= Coefficient/Standard Error
2		(1)	(2)	(3) = (1) ÷ (2)		= B3/C3
3	Intercept	\$300.98	\$229.75	1.31		= 300.98/229.75
4	Independent Variable: Machine-Hours (X)	\$ 10.31	\$ 3.12	3.30		
5						
6		<b>Regression Statistics</b>				
7	R Square	0.52				
8	Durbin-Watson Statistic	2.05				

An alternative way to test that the coefficient  $b$  is significantly different from zero is in terms of a *confidence interval*: There is less than a 5% chance that the true value of the machine-hours coefficient lies outside the range  $\$10.31 \pm (2.228 \times \$3.12)$ , or  $\$10.31 \pm \$6.95$ , or from  $\$3.36$  to  $\$17.26$ . Because 0 does not appear in the confidence interval, we can conclude that changes in the number of machine-hours do affect indirect manufacturing labor costs. Similarly, using data from Exhibit 10-14, the  $t$ -value for the constant term  $a$  is  $\$300.98 \div \$229.75 = 1.31$ , which is less than 2.228. This  $t$ -value indicates that, within the relevant range, the constant term is *not* significantly different from zero. The Durbin-Watson statistic in Exhibit 10-14 will be discussed in the following section.

## Specification Analysis of Estimation Assumptions

**Specification analysis** is the testing of the assumptions of regression analysis. If the assumptions of (1) linearity within the relevant range, (2) constant variance of residuals, (3) independence of residuals, and (4) normality of residuals all hold, then the simple regression procedures give reliable estimates of coefficient values. This section provides a brief overview of specification analysis. When these assumptions are not satisfied, more-complex regression procedures are necessary to obtain the best estimates.<sup>5</sup>

1. **Linearity within the relevant range.** A common assumption—and one that appears to be reasonable in many business applications—is that a linear relationship exists between the independent variable  $X$  and the dependent variable  $Y$  within the relevant range. If a linear regression model is used to estimate a nonlinear relationship, however, the coefficient estimates obtained will be inaccurate.

When there is only one independent variable, the easiest way to check for linearity is to study the data plotted in a scatter diagram, a step that often is unwisely skipped. Exhibit 10-6 (p. 352) presents a scatter diagram for the indirect manufacturing labor costs and machine-hours variables of Elegant Rugs shown in Exhibit 10-3 (p. 348). The scatter diagram reveals that linearity appears to be a reasonable assumption for these data.

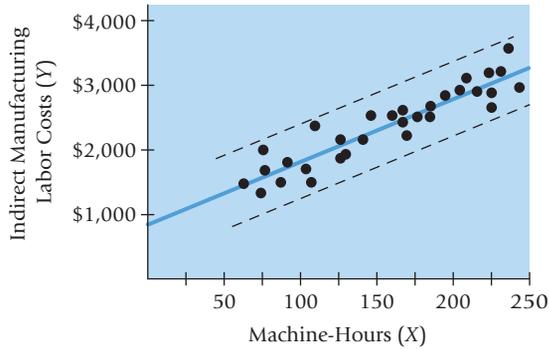
The learning-curve models discussed in this chapter (pp. 358–361) are examples of nonlinear cost functions. Costs increase when the level of production increases, but by lesser amounts than would occur with a linear cost function. In this case, the analyst should estimate a nonlinear cost function that incorporates learning effects.

2. **Constant variance of residuals.** The vertical deviation of the observed value  $Y$  from the regression line estimate  $y$  is called the *residual term*, *disturbance term*, or *error term*,  $u = Y - y$ . The assumption of constant variance implies that the residual terms are unaffected by the level of the cost driver. The assumption also implies that there is a uniform scatter, or dispersion, of the data points about the regression line as in Exhibit 10-15, Panel A. This assumption is likely to be violated, for example, in cross-sectional estimation of costs in operations of different sizes. For example, suppose Elegant Rugs has production areas of varying sizes. The company collects data from these different production areas to estimate the relationship between machine-hours and indirect manufacturing labor costs. It is very possible that the residual terms in this regression will be larger for the larger production

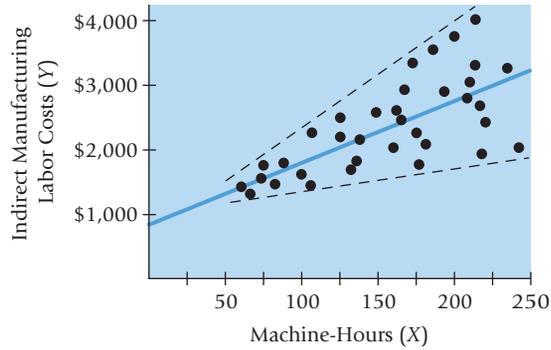
<sup>5</sup> For details see, for example, W. H. Greene, *Econometric Analysis*, 6th ed. (Upper Saddle River, NJ: Prentice Hall, 2007).

**Exhibit 10-15** Constant Variance of Residuals Assumption

**PANEL A:**  
Constant Variance  
(Uniform Scatter of Data  
Points Around Regression Line)



**PANEL B:**  
Nonconstant Variance  
(Higher Outputs Have  
Larger Residuals)



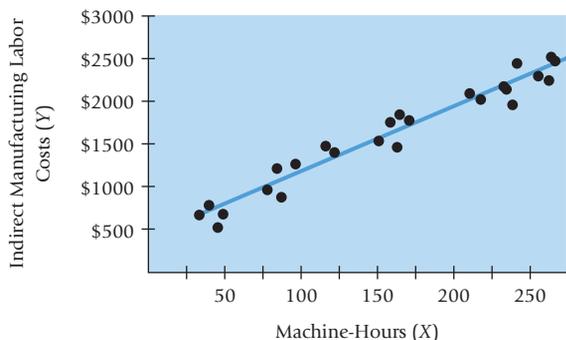
areas that have higher machine-hours and higher indirect manufacturing labor costs. There would not be a uniform scatter of data points about the regression line (see Exhibit 10-15, Panel B). Constant variance is also known as *homoscedasticity*. Violation of this assumption is called *heteroscedasticity*.

*Heteroscedasticity* does not affect the accuracy of the regression estimates  $a$  and  $b$ . It does, however, reduce the reliability of the estimates of the standard errors and thus affects the precision with which inferences about the population parameters can be drawn from the regression estimates.

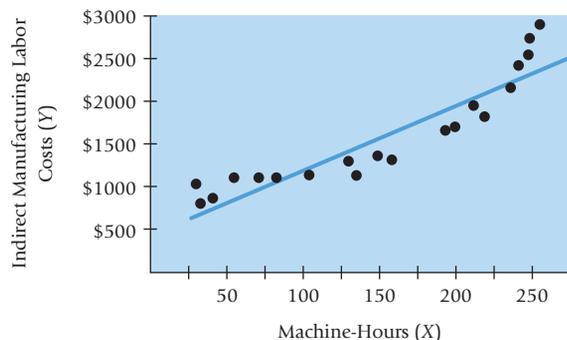
- Independence of residuals.** The assumption of independence of residuals is that the residual term for any one observation is not related to the residual term for any other observation. The problem of *serial correlation* (also called *autocorrelation*) in the residuals arises when there is a systematic pattern in the sequence of residuals such that the residual in observation  $n$  conveys information about the residuals in observations  $n + 1$ ,  $n + 2$ , and so on. Consider another production cell at Elegant Rugs that has, over a 20-week period, seen an increase in production and hence machine-hours. Exhibit 10-16 Panel B is a scatter diagram of machine-hours and indirect manufacturing labor costs. Observe the systematic pattern of the residuals in Panel B—positive residuals for extreme (high and low) quantities of machine-hours and negative residuals for moderate quantities of machine-hours. One reason for this observed pattern at low values of the cost driver is the “stickiness” of costs. When machine-hours are below 50 hours, indirect manufacturing labor costs do not decline. When machine-hours increase over time as production is ramped up, indirect manufacturing labor costs increase more as managers at Elegant Rugs struggle

**Exhibit 10-16** Independence of Residuals Assumption

**PANEL A:**  
Independence of Residuals  
(No Pattern in Residuals)



**PANEL B:**  
Serial Correlation in Residuals  
(A Pattern of Positive Residuals for  
Extreme Machine-Hours Used;  
Negative Residuals for Moderate  
Machine-Hours Used)



to manage the higher volume. How would the plot of residuals look if there were no auto-correlation? Like the plot in Exhibit 10-16, Panel A that shows no pattern in the residuals.

Like nonconstant variance of residuals, serial correlation does not affect the accuracy of the regression estimates  $a$  and  $b$ . It does, however, affect the standard errors of the coefficients, which in turn affect the precision with which inferences about the population parameters can be drawn from the regression estimates.

The Durbin-Watson statistic is one measure of serial correlation in the estimated residuals. For samples of 10 to 20 observations, a Durbin-Watson statistic in the 1.10–2.90 range indicates that the residuals are independent. The Durbin-Watson statistic for the regression results of Elegant Rugs in Exhibit 10-14 is 2.05. Therefore, an assumption of independence in the estimated residuals is reasonable for this regression model.

4. **Normality of residuals.** The normality of residuals assumption means that the residuals are distributed normally around the regression line. The normality of residuals assumption is frequently satisfied when using regression analysis on real cost data. Even when the assumption does not hold, accountants can still generate accurate estimates based on the regression equation, but the resulting confidence interval around these estimates is likely to be inaccurate.

## Using Regression Output to Choose Cost Drivers of Cost Functions

Consider the two choices of cost drivers we described earlier in this chapter for indirect manufacturing labor costs ( $y$ ):

$$y = a + (b \times \text{Number of machine-hours})$$

$$y = a + (b \times \text{Number of direct manufacturing labor-hours})$$

Exhibits 10-6 and 10-8 show plots of the data for the two regressions. Exhibit 10-14 reports regression results for the cost function using number of machine-hours as the independent variable. Exhibit 10-17 presents comparable regression results (in Excel) for the cost function using number of direct manufacturing labor-hours as the independent variable.

On the basis of the material presented in this appendix, which regression is better? Exhibit 10-18 compares these two cost functions in a systematic way. For several criteria, the cost function based on machine-hours is preferable to the cost function based on direct manufacturing labor-hours. The economic plausibility criterion is especially important.

Do not always assume that any one cost function will perfectly satisfy all the criteria in Exhibit 10-18. A cost analyst must often make a choice among “imperfect” cost functions, in the sense that the data of any particular cost function will not perfectly meet one or more of the assumptions underlying regression analysis. For example, both of the cost functions in Exhibit 10-18 are imperfect because, as stated in the section on specification analysis of estimation assumptions, inferences drawn from only 12 observations are not reliable.

### Exhibit 10-17

Simple Regression Results with Indirect Manufacturing Labor Costs as Dependent Variable and Direct Manufacturing Labor-Hours as Independent Variable (Cost Driver) for Elegant Rugs

	A	B	C	D	E	F	G	H
1		<b>Coefficients</b>	<b>Standard Error</b>	<b>t Stat</b>				
2		(1)	(2)	(3) = (1) ÷ (2)				
3	Intercept	\$744.67	\$217.61	3.42				
4	Independent Variable: Direct Manufacturing Labor-Hours (X)	\$ 7.72	\$ 5.40	1.43		= Coefficient/Standard Error = B4/C4 = 7.72/5.40		
5								
6	<b>Regression Statistics</b>							
7	R Square	0.17						
8	Durbin-Watson Statistic	2.26						

**Exhibit 10-18**

Comparison of Alternative Cost Functions for Indirect Manufacturing Labor Costs Estimated with Simple Regression for Elegant Rugs

Criterion	Cost Function 1: Machine-Hours as Independent Variable	Cost Function 2: Direct Manufacturing Labor-Hours as Independent Variable
Economic plausibility	A positive relationship between indirect manufacturing labor costs (technical support labor) and machine-hours is economically plausible in Elegant Rugs' highly automated plant	A positive relationship between indirect manufacturing labor costs and direct manufacturing labor-hours is economically plausible, but less so than machine-hours in Elegant Rugs' highly automated plant on a week-to-week basis.
Goodness of fit <sup>a</sup>	$r^2 = 0.52$ ; standard error of regression = \$170.50. Excellent goodness of fit.	$r^2 = 0.17$ ; standard error of regression = \$224.60. Poor goodness of fit.
Significance of independent variable(s)	The $t$ -value of 3.30 is significant at the 0.05 level.	The $t$ -value of 1.43 is not significant at the 0.05 level.
Specification analysis of estimation assumptions	Plot of the data indicates that assumptions of linearity, constant variance, independence of residuals (Durbin-Watson statistic = 2.05), and normality of residuals hold, but inferences drawn from only 12 observations are not reliable.	Plot of the data indicates that assumptions of linearity, constant variance, independence of residuals (Durbin-Watson statistic = 2.26), and normality of residuals hold, but inferences drawn from only 12 observations are not reliable.

<sup>a</sup>If the number of observations available to estimate the machine-hours regression differs from the number of observations available to estimate the direct manufacturing labor-hours regression, an *adjusted*  $r^2$  can be calculated to take this difference (in degrees of freedom) into account. Programs such as Excel calculate and present *adjusted*  $r^2$ .

## Multiple Regression and Cost Hierarchies

In some cases, a satisfactory estimation of a cost function may be based on only one independent variable, such as number of machine-hours. In many cases, however, basing the estimation on more than one independent variable (that is, *multiple regression*) is more economically plausible and improves accuracy. The most widely used equations to express relationships between two or more independent variables and a dependent variable are linear in the form

$$y = a + b_1X_1 + b_2X_2 + \dots + u$$

where,

$y$  = Cost to be predicted

$X_1, X_2, \dots$  = Independent variables on which the prediction is to be based

$a, b_1, b_2, \dots$  = Estimated coefficients of the regression model

$u$  = Residual term that includes the net effect of other factors not in the model as well as measurement errors in the dependent and independent variables

**Example:** Consider the Elegant Rugs data in Exhibit 10-19. The company's ABC analysis indicates that indirect manufacturing labor costs include large amounts incurred for setup and changeover costs when a new batch of carpets is started. Management believes that in addition to number of machine-hours (an output unit-level cost driver), indirect manufacturing labor costs are also affected by the number of batches of carpet produced during each week (a batch-level driver). Elegant Rugs estimates the relationship between two independent variables, number of machine-hours and number of production batches of carpet manufactured during the week, and indirect manufacturing labor costs.

## Exhibit 10-19

Weekly Indirect Manufacturing Labor Costs, Machine-Hours, Direct Manufacturing Labor-Hours, and Number of Production Batches for Elegant Rugs

	A	B	C	D	E
1	Week	Machine-Hours ( $X_1$ )	Number of Production Batches ( $X_2$ )	Direct Manufacturing Labor-Hours	Indirect Manufacturing Labor Costs ( $Y$ )
2	1	68	12	30	\$ 1,190
3	2	88	15	35	1,211
4	3	62	13	36	1,004
5	4	72	11	20	917
6	5	60	10	47	770
7	6	96	12	45	1,456
8	7	78	17	44	1,180
9	8	46	7	38	710
10	9	82	14	70	1,316
11	10	94	12	30	1,032
12	11	68	7	29	752
13	12	48	14	38	963
14	Total	862	144	462	\$12,501
15					

Exhibit 10-20 presents results (in Excel) for the following multiple regression model, using data in columns B, C, and E of Exhibit 10-19:

$$y = \$42.58 + \$7.60X_1 + \$37.77X_2$$

where  $X_1$  is the number of machine-hours and  $X_2$  is the number of production batches. It is economically plausible that both number of machine-hours and number of production batches would help explain variations in indirect manufacturing labor costs at Elegant Rugs. The  $r^2$  of 0.52 for the simple regression using number of machine-hours (Exhibit 10-14) increases to 0.72 with the multiple regression in Exhibit 10-20. The  $t$ -values suggest that the independent variable coefficients of both number of machine-hours (\$7.60) and number of production batches (\$37.77) are significantly different from zero ( $t = 2.74$  is the  $t$ -value for number of machine-hours, and  $t = 2.48$  is the  $t$ -value for number of production batches compared to the cut-off  $t$ -value of 2.26). The multiple regression model in Exhibit 10-20 satisfies both economic plausibility and statistical criteria, and it explains much greater variation (that

## Exhibit 10-20

Multiple Regression Results with Indirect Manufacturing Labor Costs and Two Independent Variables of Cost Drivers (Machine-Hours and Production Batches) for Elegant Rugs

	A	B	C	D	E	F
1		Coefficients	Standard Error	$t$ Stat		
2		(1)	(2)	(3) = (1) ÷ (2)		
3	Intercept	\$42.58	\$213.91	0.20		
4	Independent Variable 1: Machine-Hours ( $X_1$ )	\$ 7.60	\$ 2.77	2.74	→	= Coefficient/Standard Error = B4/C4 = 7.60/2.77
5	Independent Variable 2: Number of Production Batches ( $X_2$ )	\$37.77	\$ 15.25	2.48		
6						
7	<b>Regression Statistics</b>					
8	R Square	0.72				
9	Durbin-Watson Statistic	2.49				

is,  $r^2$  of 0.72 versus  $r^2$  of 0.52) in indirect manufacturing labor costs than the simple regression model using only number of machine-hours as the independent variable.<sup>6</sup> The standard error of the regression equation that includes number of batches as an independent variable is

$$\sqrt{\frac{\sum(Y - y)^2}{n - 3}} = \sqrt{\frac{170,156}{9}} = \$137.50$$

which is lower than the standard error of the regression with only machine-hours as the independent variable, \$170.50. That is, even though adding a variable reduces the degrees of freedom in the denominator, it substantially improves fit so that the numerator,  $\sum(Y - y)^2$ , decreases even more. Number of machine-hours and number of production batches are both important cost drivers of indirect manufacturing labor costs at Elegant Rugs.

In Exhibit 10-20, the slope coefficients—\$7.60 for number of machine-hours and \$37.77 for number of production batches—measure the change in indirect manufacturing labor costs associated with a unit change in an independent variable (assuming that the other independent variable is held constant). For example, indirect manufacturing labor costs increase by \$37.77 when one more production batch is added, assuming that the number of machine-hours is held constant.

An alternative approach would create two separate cost pools for indirect manufacturing labor costs: one for costs related to number of machine-hours and another for costs related to number of production batches. Elegant Rugs would then estimate the relationship between the cost driver and the costs in each cost pool. The difficult task under this approach is to properly subdivide the indirect manufacturing labor costs into the two cost pools.

## Multicollinearity

A major concern that arises with multiple regression is multicollinearity. **Multicollinearity** exists when two or more independent variables are highly correlated with each other. Generally, users of regression analysis believe that a *coefficient of correlation* between independent variables greater than 0.70 indicates multicollinearity. Multicollinearity increases the standard errors of the coefficients of the individual variables. That is, variables that are economically and statistically significant will appear not to be significantly different from zero.

The matrix of correlation coefficients of the different variables described in Exhibit 10-19 are as follows:

	Indirect Manufacturing Labor Costs	Machine-Hours	Number of Production Batches	Direct Manufacturing Labor-Hours
Indirect manufacturing labor costs	1			
Machine-hours	0.72	1		
Number of production batches	0.69	0.4	1	
Direct manufacturing labor-hours	0.41	0.12	0.31	1

These results indicate that multiple regressions using any pair of the independent variables in Exhibit 10-19 are not likely to encounter multicollinearity problems.

When multicollinearity exists, try to obtain new data that do not suffer from multicollinearity problems. Do not drop an independent variable (cost driver) that should be included in a model because it is correlated with another independent variable. Omitting such a variable will cause the estimated coefficient of the independent variable included in the model to be biased away from its true value.

<sup>6</sup> Adding another variable always increases  $r^2$ . The question is whether adding another variable increases  $r^2$  sufficiently. One way to get insight into this question is to calculate an adjusted  $r^2$  as follows:

$$\text{Adjusted } r^2 = 1 - (1 - r^2) \frac{n - 1}{n - p - 1}, \text{ where } n \text{ is the number of observations and } p \text{ is the number of coefficients estimated.}$$

In the model with only machine-hours as the independent variable, adjusted  $r^2 = 1 - (1 - 0.52) \frac{12 - 1}{12 - 2 - 1} = 0.41$ .

In the model with both machine-hours and number of batches as independent variables, adjusted  $r^2 = 1 - (1 - 0.72) \frac{12 - 1}{12 - 3 - 1} = 0.62$ . Adjusted  $r^2$  does not have the same interpretation as  $r^2$  but the increase in adjusted  $r^2$  when number of batches is added as an independent variable suggests that adding this variable significantly improves the fit of the model in a way that more than compensates for the degree of freedom lost by estimating another coefficient.

## Terms to Learn

This chapter and the Glossary at the end of this book contain definitions of the following important terms:

account analysis method (p. 347)	incremental unit-time learning model (p. 360)	residual term (p. 352)
coefficient of determination ( $r^2$ ) (p. 367)	independent variable (p. 348)	semivariable cost (p. 343)
conference method (p. 346)	industrial engineering method (p. 346)	simple regression (p. 352)
constant (p. 343)	intercept (p. 343)	slope coefficient (p. 342)
cost estimation (p. 344)	learning curve (p. 358)	specification analysis (p. 369)
cost function (p. 341)	linear cost function (p. 342)	standard error of the estimated coefficient (p. 368)
cost predictions (p. 344)	mixed cost (p. 343)	standard error of the regression (p. 368)
cumulative average-time learning model (p. 359)	multicollinearity (p. 374)	step cost function (p. 357)
dependent variable (p. 348)	multiple regression (p. 352)	work-measurement method (p. 346)
experience curve (p. 358)	nonlinear cost function (p. 357)	
high-low method (p. 350)	regression analysis (p. 352)	

## Assignment Material

### Questions



- 10-1** What two assumptions are frequently made when estimating a cost function?
- 10-2** Describe three alternative linear cost functions.
- 10-3** What is the difference between a linear and a nonlinear cost function? Give an example of each type of cost function.
- 10-4** “High correlation between two variables means that one is the cause and the other is the effect.” Do you agree? Explain.
- 10-5** Name four approaches to estimating a cost function.
- 10-6** Describe the conference method for estimating a cost function. What are two advantages of this method?
- 10-7** Describe the account analysis method for estimating a cost function.
- 10-8** List the six steps in estimating a cost function on the basis of an analysis of a past cost relationship. Which step is typically the most difficult for the cost analyst?
- 10-9** When using the high-low method, should you base the high and low observations on the dependent variable or on the cost driver?
- 10-10** Describe three criteria for evaluating cost functions and choosing cost drivers.
- 10-11** Define learning curve. Outline two models that can be used when incorporating learning into the estimation of cost functions.
- 10-12** Discuss four frequently encountered problems when collecting cost data on variables included in a cost function.
- 10-13** What are the four key assumptions examined in specification analysis in the case of simple regression?
- 10-14** “All the independent variables in a cost function estimated with regression analysis are cost drivers.” Do you agree? Explain.
- 10-15** “Multicollinearity exists when the dependent variable and the independent variable are highly correlated.” Do you agree? Explain.

### Exercises



**10-16 Estimating a cost function.** The controller of the Ijiri Company wants you to estimate a cost function from the following two observations in a general ledger account called Maintenance:

Month	Machine-Hours	Maintenance Costs Incurred
January	6,000	\$4,000
February	10,000	5,400

**Required**

1. Estimate the cost function for maintenance.
2. Can the constant in the cost function be used as an estimate of fixed maintenance cost per month? Explain.

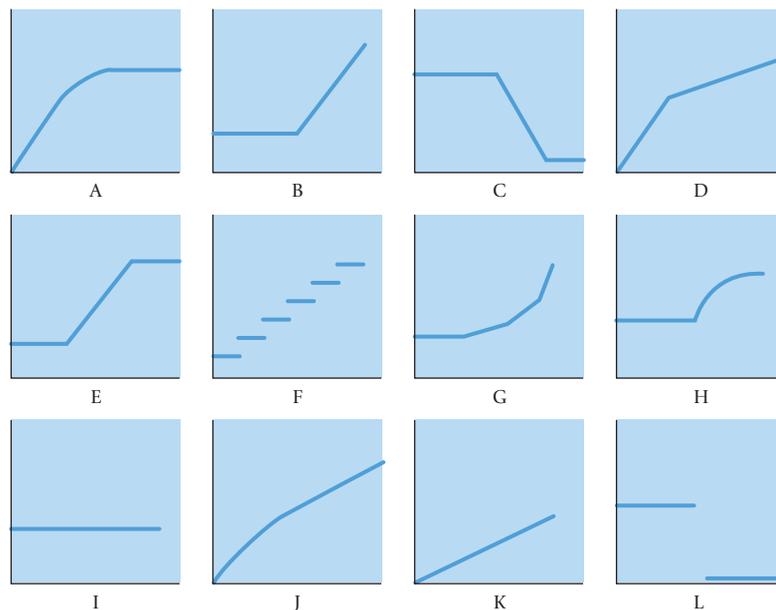
**10-17 Identifying variable-, fixed-, and mixed-cost functions.** The Pacific Corporation operates car rental agencies at more than 20 airports. Customers can choose from one of three contracts for car rentals of one day or less:

- Contract 1: \$50 for the day
- Contract 2: \$30 for the day plus \$0.20 per mile traveled
- Contract 3: \$1 per mile traveled

**Required**

1. Plot separate graphs for each of the three contracts, with costs on the vertical axis and miles traveled on the horizontal axis.
2. Express each contract as a linear cost function of the form  $y = a + bX$ .
3. Identify each contract as a variable-, fixed-, or mixed-cost function.

**10-18 Various cost-behavior patterns.** (CPA, adapted) Select the graph that matches the numbered manufacturing cost data (requirements 1–9). Indicate by letter which graph best fits the situation or item described.



The vertical axes of the graphs represent total cost, and the horizontal axes represent units produced during a calendar year. In each case, the zero point of dollars and production is at the intersection of the two axes. The graphs may be used more than once.

**Required**

1. Annual depreciation of equipment, where the amount of depreciation charged is computed by the machine-hours method.
2. Electricity bill—a flat fixed charge, plus a variable cost after a certain number of kilowatt-hours are used, in which the quantity of kilowatt-hours used varies proportionately with quantity of units produced.
3. City water bill, which is computed as follows:

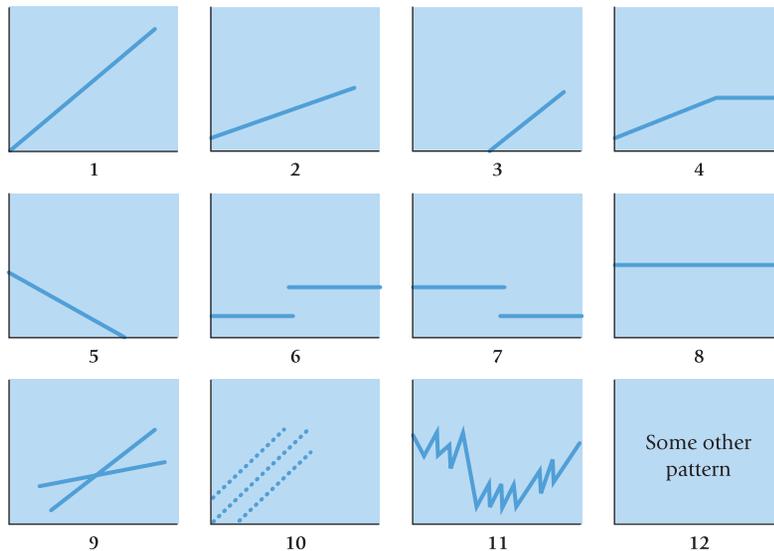
First 1,000,000 gallons or less	\$1,000 flat fee
Next 10,000 gallons	\$0.003 per gallon used
Next 10,000 gallons	\$0.006 per gallon used
Next 10,000 gallons	\$0.009 per gallon used
and so on	and so on

The gallons of water used vary proportionately with the quantity of production output.

4. Cost of direct materials, where direct material cost per unit produced decreases with each pound of material used (for example, if 1 pound is used, the cost is \$10; if 2 pounds are used, the cost is \$19.98; if 3 pounds are used, the cost is \$29.94), with a minimum cost per unit of \$9.20.

5. Annual depreciation of equipment, where the amount is computed by the straight-line method. When the depreciation schedule was prepared, it was anticipated that the obsolescence factor would be greater than the wear-and-tear factor.
6. Rent on a manufacturing plant donated by the city, where the agreement calls for a fixed-fee payment unless 200,000 labor-hours are worked, in which case no rent is paid.
7. Salaries of repair personnel, where one person is needed for every 1,000 machine-hours or less (that is, 0 to 1,000 hours requires one person, 1,001 to 2,000 hours requires two people, and so on).
8. Cost of direct materials used (assume no quantity discounts).
9. Rent on a manufacturing plant donated by the county, where the agreement calls for rent of \$100,000 to be reduced by \$1 for each direct manufacturing labor-hour worked in excess of 200,000 hours, but a minimum rental fee of \$20,000 must be paid.

**10-19 Matching graphs with descriptions of cost and revenue behavior.** (D. Green, adapted) Given here are a number of graphs.



The horizontal axis represents the units produced over the year and the vertical axis represents total cost or revenues. Indicate by number which graph best fits the situation or item described (a–h). Some graphs may be used more than once; some may not apply to any of the situations.

- a. Direct material costs
- b. Supervisors’ salaries for one shift and two shifts
- c. A cost-volume-profit graph
- d. Mixed costs—for example, car rental fixed charge plus a rate per mile driven
- e. Depreciation of plant, computed on a straight-line basis
- f. Data supporting the use of a variable-cost rate, such as manufacturing labor cost of \$14 per unit produced
- g. Incentive bonus plan that pays managers \$0.10 for every unit produced above some level of production
- h. Interest expense on \$2 million borrowed at a fixed rate of interest

**10-20 Account analysis method.** Lorenzo operates a car wash. Incoming cars are put on an automatic conveyor belt. Cars are washed as the conveyor belt carries them from the start station to the finish station. After a car moves off the conveyor belt, it is dried manually. Workers then clean and vacuum the inside of the car. Lorenzo serviced 80,000 cars in 2012. Lorenzo reports the following costs for 2012:

Account Description	Costs
Car wash labor	\$260,000
Soap, cloth, and supplies	42,000
Water	38,000
Electric power to move conveyor belt	72,000
Depreciation	64,000
Salaries	46,000

**Required**

1. Classify each account as variable or fixed with respect to the number of cars washed. Explain.
2. Suppose Lorenzo washed 90,000 cars in 2012. Use the cost classification you developed in requirement 1 to estimate Lorenzo's total costs in 2012. Depreciation is computed on a straight-line basis.

**10-21 Account analysis, high-low.** Java Joe Coffees wants to find an equation to estimate monthly utility costs. Java Joe's has been in business for one year and has collected the following cost data for utilities:

Month	Electricity Bill	Kilowatt Hours Used	Telephone Bill	Telephone Minutes Used	Water Bill	Gallons of Water Used
January	\$360	1,200	\$92.00	1,100	\$60	30,560
February	\$420	1,400	\$91.20	1,060	\$60	26,800
March	\$549	1,830	\$94.80	1,240	\$60	31,450
April	\$405	1,350	\$89.60	980	\$60	29,965
May	\$588	1,960	\$98.00	1,400	\$60	30,568
June	\$624	2,080	\$98.80	1,440	\$60	25,540
July	\$522	1,740	\$93.40	1,170	\$60	32,690
August	\$597	1,990	\$96.20	1,310	\$60	31,222
September	\$630	2,100	\$95.60	1,280	\$60	33,540
October	\$615	2,050	\$93.80	1,190	\$60	31,970
November	\$594	1,980	\$91.00	1,050	\$60	28,600
December	\$633	2,110	\$97.00	1,350	\$60	34,100

**Required**

1. Which of the preceding costs is variable? Fixed? Mixed? Explain.
2. Using the high-low method, determine the cost function for each cost.
3. Combine the preceding information to get a monthly utility cost function for Java Joe's.
4. Next month, Java Joe's expects to use 2,200 kilowatt hours of electricity, make 1,500 minutes of telephone calls, and use 32,000 gallons of water. Estimate total cost of utilities for the month.

**10-22 Account analysis method.** Gower, Inc., a manufacturer of plastic products, reports the following manufacturing costs and account analysis classification for the year ended December 31, 2012.

Account	Classification	Amount
Direct materials	All variable	\$300,000
Direct manufacturing labor	All variable	225,000
Power	All variable	37,500
Supervision labor	20% variable	56,250
Materials-handling labor	50% variable	60,000
Maintenance labor	40% variable	75,000
Depreciation	0% variable	95,000
Rent, property taxes, and administration	0% variable	100,000

Gower, Inc., produced 75,000 units of product in 2012. Gower's management is estimating costs for 2013 on the basis of 2012 numbers. The following additional information is available for 2013.

- a. Direct materials prices in 2013 are expected to increase by 5% compared with 2012.
- b. Under the terms of the labor contract, direct manufacturing labor wage rates are expected to increase by 10% in 2013 compared with 2012.
- c. Power rates and wage rates for supervision, materials handling, and maintenance are not expected to change from 2012 to 2013.
- d. Depreciation costs are expected to increase by 5%, and rent, property taxes, and administration costs are expected to increase by 7%.
- e. Gower expects to manufacture and sell 80,000 units in 2013.

**Required**

1. Prepare a schedule of variable, fixed, and total manufacturing costs for each account category in 2013. Estimate total manufacturing costs for 2013.
2. Calculate Gower's total manufacturing cost per unit in 2012, and estimate total manufacturing cost per unit in 2013.
3. How can you obtain better estimates of fixed and variable costs? Why would these better estimates be useful to Gower?

**10-23 Estimating a cost function, high-low method.** Reisen Travel offers helicopter service from suburban towns to John F. Kennedy International Airport in New York City. Each of its 10 helicopters makes between 1,000 and 2,000 round-trips per year. The records indicate that a helicopter that has made 1,000 round-trips in the year incurs an average operating cost of \$350 per round-trip, and one that has made 2,000 round-trips in the year incurs an average operating cost of \$300 per round-trip.

1. Using the high-low method, estimate the linear relationship  $y = a + bX$ , where  $y$  is the total annual operating cost of a helicopter and  $X$  is the number of round-trips it makes to JFK airport during the year.
2. Give examples of costs that would be included in  $a$  and in  $b$ .
3. If Reisen Travel expects each helicopter to make, on average, 1,200 round-trips in the coming year, what should its estimated operating budget for the helicopter fleet be?

**Required**

**10-24 Estimating a cost function, high-low method.** Laurie Daley is examining customer-service costs in the southern region of Capitol Products. Capitol Products has more than 200 separate electrical products that are sold with a six-month guarantee of full repair or replacement with a new product. When a product is returned by a customer, a service report is prepared. This service report includes details of the problem and the time and cost of resolving the problem. Weekly data for the most recent 8-week period are as follows:

Week	Customer-Service Department Costs	Number of Service Reports
1	\$13,700	190
2	20,900	275
3	13,000	115
4	18,800	395
5	14,000	265
6	21,500	455
7	16,900	340
8	21,000	305

1. Plot the relationship between customer-service costs and number of service reports. Is the relationship economically plausible?
2. Use the high-low method to compute the cost function, relating customer-service costs to the number of service reports.
3. What variables, in addition to number of service reports, might be cost drivers of weekly customer-service costs of Capitol Products?

**Required**

**10-25 Linear cost approximation.** Terry Lawler, managing director of the Chicago Reviewers Group, is examining how overhead costs behave with changes in monthly professional labor-hours billed to clients. Assume the following historical data:

Total Overhead Costs	Professional Labor-Hours Billed to Clients
\$335,000	2,000
400,000	3,000
430,000	4,000
472,000	5,000
533,000	6,500
582,000	7,500

1. Compute the linear cost function, relating total overhead costs to professional labor-hours, using the representative observations of 3,000 and 6,500 hours. Plot the linear cost function. Does the constant component of the cost function represent the fixed overhead costs of the Chicago Reviewers Group? Why?
2. What would be the predicted total overhead costs for (a) 4,000 hours and (b) 7,500 hours using the cost function estimated in requirement 1? Plot the predicted costs and actual costs for 4,000 and 7,500 hours.
3. Lawler had a chance to accept a special job that would have boosted professional labor-hours from 3,000 to 4,000 hours. Suppose Lawler, guided by the linear cost function, rejected this job because it would have brought a total increase in contribution margin of \$35,000, before deducting the predicted increase in total overhead cost, \$38,000. What is the total contribution margin actually forgone?

**Required**

**10-26 Cost-volume-profit and regression analysis.** Goldstein Corporation manufactures a children's bicycle, model CT8. Goldstein currently manufactures the bicycle frame. During 2012, Goldstein made 32,000 frames at a total cost of \$1,056,000. Ryan Corporation has offered to supply as many frames as Goldstein wants at a cost of \$32.50 per frame. Goldstein anticipates needing 35,000 frames each year for the next few years.

**Required**

1.
  - a. What is the average cost of manufacturing a bicycle frame in 2012? How does it compare to Ryan's offer?
  - b. Can Goldstein use the answer in requirement 1a to determine the cost of manufacturing 35,000 bicycle frames? Explain.
2. Goldstein's cost analyst uses annual data from past years to estimate the following regression equation with total manufacturing costs of the bicycle frame as the dependent variable and bicycle frames produced as the independent variable:

$$y = \$435,000 + \$19X$$

During the years used to estimate the regression equation, the production of bicycle frames varied from 31,000 to 35,000. Using this equation, estimate how much it would cost Goldstein to manufacture 35,000 bicycle frames. How much more or less costly is it to manufacture the frames rather than to acquire them from Ryan?

3. What other information would you need to be confident that the equation in requirement 2 accurately predicts the cost of manufacturing bicycle frames?

**10-27 Regression analysis, service company.** (CMA, adapted) Bob Jones owns a catering company that prepares food and beverages for banquets and parties. For a standard party the cost on a per-person basis is as follows:

Food and beverages	\$15
Labor (0.5 hour × \$10 per hour)	5
Overhead (0.5 hour × \$14 per hour)	<u>7</u>
Total cost per person	<u>\$27</u>

Jones is quite certain about his estimates of the food, beverages, and labor costs but is not as comfortable with the overhead estimate. The overhead estimate was based on the actual data for the past 12 months, which are presented here. These data indicate that overhead costs vary with the direct labor-hours used. The \$14 estimate was determined by dividing total overhead costs for the 12 months by total labor-hours.

Month	Labor-Hours	Overhead Costs
January	2,500	\$ 55,000
February	2,700	59,000
March	3,000	60,000
April	4,200	64,000
May	7,500	77,000
June	5,500	71,000
July	6,500	74,000
August	4,500	67,000
September	7,000	75,000
October	4,500	68,000
November	3,100	62,000
December	<u>6,500</u>	<u>73,000</u>
Total	<u>57,500</u>	<u>\$805,000</u>

Jones has recently become aware of regression analysis. He estimated the following regression equation with overhead costs as the dependent variable and labor-hours as the independent variable:

$$y = \$48,271 + \$3.93X$$

**Required**

1. Plot the relationship between overhead costs and labor-hours. Draw the regression line and evaluate it using the criteria of economic plausibility, goodness of fit, and slope of the regression line.
2. Using data from the regression analysis, what is the variable cost per person for a standard party?
3. Bob Jones has been asked to prepare a bid for a 200-person standard party to be given next month. Determine the minimum bid price that Jones would be willing to submit to recoup variable costs.

**10-28 High-low, regression.** Melissa Crupp is the new manager of the materials storeroom for Canton Manufacturing. Melissa has been asked to estimate future monthly purchase costs for part #4599, used in two of Canton's products. Melissa has purchase cost and quantity data for the past nine months as follows:

Month	Cost of Purchase	Quantity Purchased
January	\$10,390	2,250 parts
February	10,550	2,350
March	14,400	3,390
April	13,180	3,120
May	10,970	2,490
June	11,580	2,680
July	12,690	3,030
August	8,560	1,930
September	12,450	2,960

Estimated monthly purchases for this part based on expected demand of the two products for the rest of the year are as follows:

Month	Purchase Quantity Expected
October	2,800 parts
November	3,100
December	2,500

- The computer in Melissa's office is down and Melissa has been asked to immediately provide an equation to estimate the future purchase cost for part # 4599. Melissa grabs a calculator and uses the high-low method to estimate a cost equation. What equation does she get?
- Using the equation from requirement 1, calculate the future expected purchase costs for each of the last three months of the year.
- After a few hours Melissa's computer is fixed. Melissa uses the first nine months of data and regression analysis to estimate the relationship between the quantity purchased and purchase costs of part #4599. The regression line Melissa obtains is as follows:

**Required**

$$y = \$1,779.6 + 3.67X$$

Evaluate the regression line using the criteria of economic plausibility, goodness of fit, and significance of the independent variable. Compare the regression equation to the equation based on the high-low method. Which is a better fit? Why?

- Use the regression results to calculate the expected purchase costs for October, November, and December. Compare the expected purchase costs to the expected purchase costs calculated using the high-low method in requirement 2. Comment on your results.

**10-29 Learning curve, cumulative average-time learning model.** Global Defense manufactures radar systems. It has just completed the manufacture of its first newly designed system, RS-32. Manufacturing data for the RS-32 follow:

			
	A	B	C
1	Direct material cost	\$160,000	per unit of RS-32
2	Direct manufacturing labor time for first unit	6,000	direct manufacturing labor-hours
3	Learning curve for manufacturing labor time per radar system	85%	cumulative average time <sup>a</sup>
4	Direct manufacturing labor cost	\$ 30	per direct manufacturing labor-hour
5	Variable manufacturing overhead cost	\$ 20	per direct manufacturing labor-hour
6			
7	<sup>a</sup> Using the formula (p. 359), for a 85% learning curve, $b = \frac{\ln 0.85}{\ln 2} = \frac{-0.162519}{0.693147} = -0.234465$		
8			

**Required** Calculate the total variable costs of producing 2, 4, and 8 units.

**10-30 Learning curve, incremental unit-time learning model.** Assume the same information for Global Defense as in Exercise 10-29, except that Global Defense uses an 85% incremental unit-time learning model as a basis for predicting direct manufacturing labor-hours. (An 85% learning curve means  $b = -0.234465$ .)

**Required**

1. Calculate the total variable costs of producing 2, 3, and 4 units.
2. If you solved Exercise 10-29, compare your cost predictions in the two exercises for 2 and 4 units. Why are the predictions different? How should Global Defense decide which model it should use?



## Problems

**10-31 High-low method.** Ken Howard, financial analyst at KMW Corporation, is examining the behavior of quarterly maintenance costs for budgeting purposes. Howard collects the following data on machine-hours worked and maintenance costs for the past 12 quarters:

Quarter	Machine-Hours	Maintenance Costs
1	100,000	\$205,000
2	120,000	240,000
3	110,000	220,000
4	130,000	260,000
5	95,000	190,000
6	115,000	235,000
7	105,000	215,000
8	125,000	255,000
9	105,000	210,000
10	125,000	245,000
11	115,000	200,000
12	140,000	280,000

**Required**

1. Estimate the cost function for the quarterly data using the high-low method.
2. Plot and comment on the estimated cost function.
3. Howard anticipates that KMW will operate machines for 100,000 hours in quarter 13. Calculate the predicted maintenance costs in quarter 13 using the cost function estimated in requirement 1.

**10-32 High-low method and regression analysis.** Local Harvest, a cooperative of organic family-owned farms outside of Columbus, Ohio, has recently started a fresh produce club to provide support to the group's member farms, and to promote the benefits of eating organic, locally-produced food to the nearby suburban community. Families pay a seasonal membership fee of \$50, and place their orders a week in advance for a price of \$40 per week. In turn, Local Harvest delivers fresh-picked seasonal local produce to several neighborhood distribution points. Eight hundred families joined the club for the first season, but the number of orders varied from week to week.

Harvey Hendricks has run the produce club for the first 10-week season. Before becoming a farmer, Harvey had been a business major in college, and he remembers a few things about cost analysis. In planning for next year, he wants to know how many orders will be needed each week for the club to break even, but first he must estimate the club's fixed and variable costs. He has collected the following data over the club's first 10 weeks of operation:

Week	Number of Orders per Week	Weekly Total Costs
1	351	\$18,795
2	385	21,597
3	410	22,800
4	453	22,600
5	425	21,900
6	486	24,600
7	455	23,900
8	467	22,900
9	525	25,305
10	510	24,500

1. Plot the relationship between number of orders per week and weekly total costs.
2. Estimate the cost equation using the high-low method, and draw this line on your graph.
3. Harvey uses his computer to calculate the following regression formula:

**Required**

$$\text{Total weekly costs} = \$8,631 + (\$31.92 \times \text{Number of weekly orders})$$

Draw the regression line on your graph. Use your graph to evaluate the regression line using the criteria of economic plausibility, goodness of fit, and significance of the independent variable. Is the cost function estimated using the high-low method a close approximation of the cost function estimated using the regression method? Explain briefly.

4. Did Fresh Harvest break even this season? Remember that each of the families paid a seasonal membership fee of \$50.
5. Assume that 900 families join the club next year, and that prices and costs do not change. How many orders, on average, must Fresh Harvest receive each week to break even?

**10-33 High-low method; regression analysis.** (CIMA, adapted) Anna Martinez, the financial manager at the Casa Real restaurant, is checking to see if there is any relationship between newspaper advertising and sales revenues at the restaurant. She obtains the following data for the past 10 months:

Month	Revenues	Advertising Costs
March	\$50,000	\$2,000
April	70,000	3,000
May	55,000	1,500
June	65,000	3,500
July	55,000	1,000
August	65,000	2,000
September	45,000	1,500
October	80,000	4,000
November	55,000	2,500
December	60,000	2,500

She estimates the following regression equation:

$$\text{Monthly revenues} = \$39,502 + (\$8.723 \times \text{Advertising costs})$$

1. Plot the relationship between advertising costs and revenues.
2. Draw the regression line and evaluate it using the criteria of economic plausibility, goodness of fit, and slope of the regression line.
3. Use the high-low method to compute the function, relating advertising costs and revenues.
4. Using (a) the regression equation and (b) the high-low equation, what is the increase in revenues for each \$1,000 spent on advertising within the relevant range? Which method should Martinez use to predict the effect of advertising costs on revenues? Explain briefly.

**Required**

**10-34 Regression, activity-based costing, choosing cost drivers.** Fitzgerald Manufacturing has been using activity-based costing to determine the cost of product X-678. One of the activities, "Inspection," occurs just before the product is finished. Fitzgerald inspects every 10th unit, and has been using "number of units inspected" as the cost driver for inspection costs. A significant component of inspection costs is the cost of the test-kit used in each inspection.

Neela McFeen, the line manager, is wondering if inspection labor-hours might be a better cost driver for inspection costs. Neela gathers information for weekly inspection costs, units inspected, and inspection labor-hours as follows:

Week	Units Inspected	Inspection Labor-Hours	Inspection Costs
1	1,400	190	\$3,700
2	400	70	1,800
3	1,700	230	4,500
4	2,400	240	5,900
5	2,100	210	5,300
6	700	90	2,400
7	900	110	2,900

Neela runs regressions on each of the possible cost drivers and estimates these cost functions:

$$\text{Inspection Costs} = \$977 + (\$2.05 \times \text{Number of units inspected})$$

$$\text{Inspection Costs} = \$478 + (\$20.31 \times \text{Inspection labor-hours})$$

**Required**

1. Explain why number of units inspected and inspection labor-hours are plausible cost drivers of inspection costs.
2. Plot the data and regression line for units inspected and inspection costs. Plot the data and regression line for inspection labor-hours and inspection costs. Which cost driver of inspection costs would you choose? Explain.
3. Neela expects inspectors to work 140 hours next period and to inspect 1,100 units. Using the cost driver you chose in requirement 2, what amount of inspection costs should Neela budget? Explain any implications of Neela choosing the cost driver you did not choose in requirement 2 to budget inspection costs.

**10-35 Interpreting regression results, matching time periods.** Brickman Apparel produces equipment for the extreme-sports market. It has four peak periods, each lasting two months, for manufacturing the merchandise suited for spring, summer, fall, and winter. In the off-peak periods, Brickman schedules equipment maintenance. Brickman's controller, Sascha Green, wants to understand the drivers of equipment maintenance costs.

The data collected is shown in the table as follows:

Month	Machine-Hours	Maintenance Costs
January	5,000	\$ 1,300
February	5,600	2,200
March	1,500	12,850
April	6,500	1,665
May	5,820	2,770
June	1,730	15,250
July	7,230	1,880
August	5,990	2,740
September	2,040	15,350
October	6,170	1,620
November	5,900	2,770
December	1,500	14,700

A regression analysis of one year of monthly data yields the following relationships:

$$\text{Maintenance costs} = \$18,552 - (\$2.683 \times \text{Number of machine-hours})$$

Upon examining the results, Green comments, "So, all I have to do to reduce maintenance costs is run my machines longer?! This is hard to believe, but numbers don't lie! I would have guessed just the opposite."

**Required**

1. Explain why Green made this comment. What is wrong with her analysis?
2. Upon further reflection, Sascha Green reanalyzes the data, this time comparing quarterly machine-hours with quarterly maintenance expenditures. This time, the results are very different. The regression yields the following formula:

$$\text{Maintenance costs} = \$2,622.80 + (\$1.175 \times \text{Number of machine-hours})$$

What caused the formula to change, in light of the fact that the data was the same?

**10-36 Cost estimation, cumulative average-time learning curve.** The Nautilus Company, which is under contract to the U.S. Navy, assembles troop deployment boats. As part of its research program, it completes the assembly of the first of a new model (PT109) of deployment boats. The Navy is impressed with the PT109. It requests that Nautilus submit a proposal on the cost of producing another six PT109s.

Nautilus reports the following cost information for the first PT109 assembled and uses a 90% cumulative average-time learning model as a basis for forecasting direct manufacturing labor-hours for the next six PT109s. (A 90% learning curve means  $b = -0.152004$ .)

			
	A	B	C
1	Direct material	\$200,000	
2	Direct manufacturing labor time for first boat	15,000	labor-hours
3	Direct manufacturing labor rate	\$ 40	per direct manufacturing labor-hour
4	Variable manufacturing overhead cost	\$ 25	per direct manufacturing labor-hour
5	Other manufacturing overhead	20%	of direct manufacturing labor costs
6	Tooling costs <sup>a</sup>	\$280,000	
7	Learning curve for manufacturing labor time per boat	90%	cumulative average time <sup>b</sup>
8			
9	<sup>a</sup> Tooling can be reused at no extra cost because all of its cost has been assigned to the first deployment boat.		
10			
11	<sup>b</sup> Using the formula (p. 359), for a 90% learning curve, $b = \frac{\ln 0.9}{\ln 2} = \frac{-0.105361}{0.693147} = -0.152004$		
12			

1. Calculate predicted total costs of producing the six PT109s for the Navy. (Nautilus will keep the first deployment boat assembled, costed at \$1,575,000, as a demonstration model for potential customers.)
2. What is the dollar amount of the difference between (a) the predicted total costs for producing the six PT109s in requirement 1, and (b) the predicted total costs for producing the six PT109s, assuming that there is no learning curve for direct manufacturing labor? That is, for (b) assume a linear function for units produced and direct manufacturing labor-hours.

**Required**

**10-37 Cost estimation, incremental unit-time learning model.** Assume the same information for the Nautilus Company as in Problem 10-36 with one exception. This exception is that Nautilus uses a 90% incremental unit-time learning model as a basis for predicting direct manufacturing labor-hours in its assembling operations. (A 90% learning curve means  $b = -0.152004$ .)

1. Prepare a prediction of the total costs for producing the six PT109s for the Navy.
2. If you solved requirement 1 of Problem 10-36, compare your cost prediction there with the one you made here. Why are the predictions different? How should Nautilus decide which model it should use?

**Required**

**10-38 Regression; choosing among models.** Tilbert Toys (TT) makes the popular Floppin' Freddy Frog and Jumpin' Jill Junebug dolls in batches. TT has recently adopted activity-based costing. TT incurs setup costs for each batch of dolls that it produces. TT uses "number of setups" as the cost driver for setup costs.

TT has just hired Bebe Williams, an accountant. Bebe thinks that "number of setup-hours" might be a better cost driver because the setup time for each product is different. Bebe collects the following data.

				
	A	B	C	D
1	Month	Number of Setups	Number of Setup Hours	Setup Costs
2	1	300	1,840	\$104,600
3	2	410	2,680	126,700
4	3	150	1,160	57,480
5	4	480	3,800	236,840
6	5	310	3,680	178,880
7	6	460	3,900	213,760
8	7	420	2,980	209,620
9	8	300	1,200	90,080
10	9	270	3,280	221,040

**Required**

1. Estimate the regression equation for (a) setup costs and number of setups and (b) setup costs and number of setup-hours. You should obtain the following results:

*Regression 1: Setup costs =  $a + (b \times \text{Number of setups})$*

Variable	Coefficient	Standard Error	t-Value
Constant	\$12,890	\$61,365	0.21
Independent variable 1: No. of setups	\$ 426.77	\$ 171	2.49

$r^2 = 0.47$ ; Durbin-Watson statistic = 1.65

*Regression 2: Setup costs =  $a + (b \times \text{Number of setup-hours})$*

Variable	Coefficient	Standard Error	t-Value
Constant	\$6,573	\$25,908	0.25
Independent variable 1: No. of setup-hours	\$ 56.27	\$ 8.90	6.32

$r^2 = 0.85$ ; Durbin-Watson statistic = 1.50

2. On two different graphs plot the data and the regression lines for each of the following cost functions:
  - a. Setup costs =  $a + (b \times \text{Number of setups})$
  - b. Setup costs =  $a + (b \times \text{Number of setup-hours})$
3. Evaluate the regression models for "Number of setups" and "Number of setup-hours" as the cost driver according to the format of Exhibit 10-18 (p. 372).
4. Based on your analysis, which cost driver should Tilbert Toys use for setup costs, and why?

**10-39 Multiple regression (continuation of 10-38).** Bebe Williams wonders if she should run a multiple regression with both number of setups and number of setup-hours, as cost drivers.

**Required**

1. Run a multiple regression to estimate the regression equation for setup costs using both number of setups and number of setup-hours as independent variables. You should obtain the following result:

*Regression 3: Setup costs =  $a (b_1 \times \text{No. of setups}) + (b_2 \times \text{No. of setup-hours})$*

Variable	Coefficient	Standard Error	t-Value
Constant	-\$2,807	\$34,850	-0.08
Independent variable 1: No. of setups	\$ 58.62	\$ 133.42	0.44
Independent variable 2: No. of setup-hours	\$ 52.31	\$ 13.08	4.00

$r^2 = 0.86$ ; Durbin-Watson statistic = 1.38

2. Evaluate the multiple regression output using the criteria of economic plausibility goodness of fit, significance of independent variables, and specification of estimation assumptions. (Assume linearity, constant variance, and normality of residuals.)
3. What difficulties do not arise in simple regression analysis that may arise in multiple regression analysis? Is there evidence of such difficulties in the multiple regression presented in this problem? Explain.
4. Which of the regression models from Problems 10-38 and 10-39 would you recommend Bebe Williams use? Explain.

**10-40 Purchasing department cost drivers, activity-based costing, simple regression analysis.** Fashion Bling operates a chain of 10 retail department stores. Each department store makes its own purchasing decisions. Barry Lee, assistant to the president of Fashion Bling, is interested in better understanding the drivers of purchasing department costs. For many years, Fashion Bling has allocated purchasing department costs to products on the basis of the dollar value of merchandise purchased. A \$100 item is allocated 10 times as many overhead costs associated with the purchasing department as a \$10 item.

Lee recently attended a seminar titled "Cost Drivers in the Retail Industry." In a presentation at the seminar, Couture Fabrics, a leading competitor that has implemented activity-based costing, reported number of purchase orders and number of suppliers to be the two most important cost drivers of purchasing department costs. The dollar value of merchandise purchased in each purchase order was not found to be a significant cost driver. Lee interviewed several members of the purchasing department at the Fashion Bling store in Miami. They believed that Couture Fabrics' conclusions also applied to their purchasing department.

Lee collects the following data for the most recent year for Fashion Bling's 10 retail department stores:

	A	B	C	D	E
1	<b>Department Store</b>	<b>Purchasing Department Costs (PDC)</b>	<b>Dollar Value of Merchandise Purchased (MP\$)</b>	<b>Number of Purchase Orders (No. of POs)</b>	<b>Number of Suppliers (No. of Ss)</b>
2	Baltimore	\$1,522,000	\$ 68,307,000	4,345	125
3	Chicago	1,095,000	33,463,000	2,548	230
4	Los Angeles	542,000	121,800,000	1,420	8
5	Miami	2,053,000	119,450,000	5,935	188
6	New York	1,068,000	33,575,000	2,786	21
7	Phoenix	517,000	29,836,000	1,334	29
8	Seattle	1,544,000	102,840,000	7,581	101
9	St. Louis	1,761,000	38,725,000	3,623	127
10	Toronto	1,605,000	139,300,000	1,712	202
11	Vancouver	1,263,000	130,110,000	4,736	196

Lee decides to use simple regression analysis to examine whether one or more of three variables (the last three columns in the table) are cost drivers of purchasing department costs. Summary results for these regressions are as follows:

*Regression 1:*  $PDC = a + (b \times MP\$)$

Variable	Coefficient	Standard Error	t-Value
Constant	\$1,041,421	\$346,709	3.00
Independent variable 1: MP\$	0.0031	0.0038	0.83

$r^2 = 0.08$ ; Durbin-Watson statistic = 2.41

*Regression 2:*  $PDC = a (b \times \text{No. of POs})$

Variable	Coefficient	Standard Error	t-Value
Constant	\$722,538	\$265,835	2.72
Independent variable 1: No. of POs	\$ 159.48	\$ 64.84	2.46

$r^2 = 0.43$ ; Durbin-Watson statistic = 1.97

*Regression 3:*  $PDC = a + (b \times \text{No. of Ss})$

Variable	Coefficient	Standard Error	t-Value
Constant	\$828,814	\$246,570	3.36
Independent variable 1: No. of Ss	\$ 3,816	\$ 1,698	2.25

$r^2 = 0.39$ ; Durbin-Watson statistic = 2.01

1. Compare and evaluate the three simple regression models estimated by Lee. Graph each one. Also, use the format employed in Exhibit 10-18 (p. 372) to evaluate the information.
2. Do the regression results support the Couture Fabrics' presentation about the purchasing department's cost drivers? Which of these cost drivers would you recommend in designing an ABC system?
3. How might Lee gain additional evidence on drivers of purchasing department costs at each of Fashion Bling's stores?

**Required**

**10-41 Purchasing department cost drivers, multiple regression analysis (continuation of 10-40).** Barry Lee decides that the simple regression analysis used in Problem 10-40 could be extended to a multiple regression analysis. He finds the following results for two multiple regression analyses:

*Regression 4:*  $PDC = a + (b_1 \times \text{No. of POs}) + (b_2 \times \text{No. of Ss})$

Variable	Coefficient	Standard Error	t-Value
Constant	\$484,522	\$256,684	1.89
Independent variable 1: No. of POs	\$ 126.66	\$ 57.80	2.19
Independent variable 2: No. of Ss	\$ 2,903	\$ 1,459	1.99

$r^2 = 0.64$ ; Durbin-Watson statistic = 1.91

*Regression 5:*  $PDC = a + (b_1 \times \text{No. of POs}) + (b_2 \times \text{No. of Ss}) + (b_3 \times \text{MP\$})$

Variable	Coefficient	Standard Error	t-Value
Constant	\$483,560	\$312,554	1.55
Independent variable 1: No. of POs	\$ 126.58	\$ 63.75	1.99
Independent variable 2: No. of Ss	\$ 2,901	\$ 1,622	1.79
Independent variable 3: MP\$	0.00002	0.0029	0.01

$r^2 = 0.64$ ; Durbin-Watson statistic = 1.91

The coefficients of correlation between combinations of pairs of the variables are as follows:

	PDC	MP\$	No. of POs
MP\$	0.28		
No. of POs	0.66	0.27	
No. of Ss	0.62	0.30	0.29

### Required

- Evaluate regression 4 using the criteria of economic plausibility, goodness of fit, significance of independent variables and specification analysis. Compare regression 4 with regressions 2 and 3 in Problem 10-40. Which one of these models would you recommend that Lee use? Why?
- Compare regression 5 with regression 4. Which one of these models would you recommend that Lee use? Why?
- Lee estimates the following data for the Baltimore store for next year: dollar value of merchandise purchased, \$78,000,000; number of purchase orders, 4,000; number of suppliers, 95. How much should Lee budget for purchasing department costs for the Baltimore store for next year?
- What difficulties do not arise in simple regression analysis that may arise in multiple regression analysis? Is there evidence of such difficulties in either of the multiple regressions presented in this problem? Explain.
- Give two examples of decisions in which the regression results reported here (and in Problem 10-40) could be informative.

### Collaborative Learning Problem

**10-42 Interpreting regression results, matching time periods, ethics.** Jayne Barbour is working as a summer intern at Mode, a trendy store specializing in clothing for twenty-somethings. Jayne has been working closely with her cousin, Gail Hubbard, who plans promotions for Mode. The store has only been in business for 10 months, and Valerie Parker, the store's owner, has been unsure of the effectiveness of the store's advertising. Wanting to impress Valerie with the regression analysis skills she acquired in a cost accounting course the previous semester, Jayne decides to prepare an analysis of the effect of advertising on revenues. She collects the following data:

	A	B	C
1	Month	Advertising Expense	Revenue
2	October	4,560	\$35,400
3	November	3,285	44,255
4	December	1,200	56,300
5	January	4,099	28,764
6	February	3,452	49,532
7	March	1,075	43,200
8	April	4,768	30,600
9	May	4,775	52,137
10	June	1,845	49,640
11	July	1,430	29,542

Jayne performs a regression analysis, comparing each month’s advertising expense with that month’s revenue, and obtains the following formula:

$$\text{Revenue} = \$47,801 - (1.92 \times \text{Advertising expense})$$

Variable	Coefficient	Standard Error	t-Value
Constant	\$47,801.72	7,628.39	6.27
Independent variable: Advertising expense	-1.92	2.26	-0.85

$r^2 = 0.43$ ; Standard error = 10,340.18

**Required**

- Plot the preceding data on a graph and draw the regression line. What does the cost formula indicate about the relationship between monthly advertising expense and monthly revenues? Is the relationship economically plausible?
- Jayne worries that if she makes her presentation to the owner as planned, it will reflect poorly on her cousin Gail’s performance. Is she ethically obligated to make the presentation?
- Jayne thinks further about her analysis, and discovers a significant flaw in her approach. She realizes that advertising done in a given month should be expected to influence the following month’s sales, not necessarily the current month’s. She modifies her analysis by comparing, for example, October advertising expense with November sales revenue. The modified regression yields the following:

$$\text{Revenue} = \$23,538 + (5.92 \times \text{Advertising expense})$$

Variable	Coefficient	Standard Error	t-Value
Constant	\$23,538.45	4,996.60	4.71
Independent variable: Previous month’s advertising expense	5.92	1.42	4.18

$r^2 = 0.71$ ; Standard error = 6,015.67

What does the revised cost formula indicate? Plot the revised data on a graph. (You will need to discard October revenue and July advertising expense from the data set.) Is this relationship economically plausible?

- Can Jayne conclude that there is a cause and effect relationship between advertising expense and sales revenue? Why or why not?