APPLIED SYSTEMS ANALYSIS

Engineering Planning and Technology Management

Richard de Neufville
Massachusetts Institute of Technology

McGraw-Hill Publishing Company
New York St. Louis San Francisco Auckland Bogotá Caracas Hamburg
Lisbon London Madrid Mexico Milan Montreal New Delhi Oklahoma City
Paris San Juan São Paulo Singapore Sydney Tokyo Toronto

3.5. *Road Work*
Assume that the number of highway miles that can be graded, H, is a function of both the hours of labor, L, and of machines, M:

$$H = 0.5L^{0.2}M^{0.8}$$

(*a*) Minimize the cost of grading 20 miles, given that the hourly rates for labor and machines are: $C_L = \$20$; $C_M = \$160$.
(*b*) Interpret the significance of the Lagrangean multiplier.

3.6. *Power Plant*
A cooperative operates an oil-fired power plant, selling electricity and steam to its members. The value of a unit of electricity, e, is a constant P_e per unit, which is its price from an alternative supplier. The value of steam, s, is a monotonically decreasing function $P_s(s)$. The cooperative wishes to maximize the net benefits of its members:

$$Z = P_e\,(e) + \int_0^s P_s\,(s)\,ds - P_q\,Q$$

where P_q and Q are the unit price and quantity of oil used in the plant.
The technical constraints on the operation are that the electricity and steam produced are less than their thermal efficiency times the quantity of oil used:

$$e \le E_e\,Q \qquad s \le E_s\,Q$$

and that maximum production of electricity and steam are incompatible:

$$\frac{e}{E_e} + \frac{s}{E_s} \le 1.5Q$$

(*a*) Formulate the problem and state the Kuhn-Tucker conditions.
(*b*) By setting various combinations of the $\lambda_i = 0$, eight potential solutions are logically possible. Examine each to determine which are feasible, and explain their significance.

3.7. *Tim Burr, Jr.*
See Problem 2.10. Junior, happy with his father's identification of significant returns to scale, is all set to maximize production to take advantage of their benefits. Do so for him, with a budget B and input costs for pine and balsam proportional to the $\frac{5}{4}$ power of their quantity—since Tim can only get more lumber by acquiring it further away. Do his economics improve as he builds a larger plant?

3.8. *Heat Exchanger*
160 m of tubes must be installed in a heat exchanger to provide the necessary surface area. The cost of the installation is the sum of the cost of the tubes, the shell of the exchanger, and of the floor space it occupies:

$$\text{Cost} = 700 + 25\,D^{2.5}L + 12.5\,DL$$

where D and L are the diameter and the length of the cylindrical exchanger. No more than 20 tubes can fit into 1 m^2 of cross-section. Find the dimensions that minimize cost.

CHAPTER 4

MARGINAL ANALYSIS

4.1 CONCEPT

Marginal Analysis is a basic form of optimization of design. It is a means of selecting the best choice from among many technically efficient ways to achieve a stated objective or product.

Marginal analysis combines two sorts of models, as all procedures for optimizing design must. One model represents the technical possibilities; the other represents the relative values of the several inputs of the production process. Specifically, marginal analysis combines the production function, which represents only the technically efficient production possibilities; and the *input cost function*, which describes the cost of the inputs used (see Section 4.2).

Marginal analysis is based on three specific assumptions. The first is that the feasible regions are convex for the portions of both models being considered. A key word in this statement is "portion." As Sections 4.2 and 4.3 explain, this qualification permits the application of marginal analysis even for processes that have increasing returns to scale, and for which the entire feasible region is not convex.

The second assumption of marginal analysis is that the only constraint on the system is the amount of money available, the budget. The resources themselves are presumed to be available indefinitely, provided there is enough money to buy them. This situation is quite common in practice, since the materials required for any one system are generally far less than the supply available. A company setting out to build a factory can thus presume that there will be as much steel and concrete available for purchase as it might require, and that it can thus optimize its design for the factory as if the supply were unconstrained.

Exceptions to this pattern are common, however. It may frequently happen that the quantity of some particular resource available is small compared to a large system. A computer manufacturer might have to recognize that his production was limited by the supply of advanced chips; a development program in bioengineering might similarly have to recognize that the number of qualified researchers was limited. A key effect of such constraints is indicated at the end of Section 4.3.

The third assumption of standard marginal analysis is that the models are analytic, that is, that they are continuously differentiable. This premise smooths the mathematics and justifies specific results. In several respects it is not a crucial assumption. The general results of marginal analysis can often still be applied even when this condition is not met. The applications in Section 4.6 illustrate this fact.

Marginal analysis is both a useful practical tool and a good basis for further analysis. Many of the ideas introduced by this method apply to more complicated situations.

4.2 OPTIMALITY CONDITIONS

Any particular objective or Product, Y', can in general be obtained in many ways. Moreover, there are also in general many technically efficient designs that will achieve that objective; these are all represented by the isoquant for Y'.

The optimal design for any Y' will be the one that provides the best value. It is the optimal configuration of inputs $\mathbf{X}^*$ that defines the point on the isoquant which delivers the product at the least cost. This cost can be defined either narrowly in terms of money alone, or more broadly in terms of social values such as environmental effects. For this discussion it does not matter which definition applies, the main result is the same. In practice, one should use the definition appropriate to the designer of the system.

Mathematically, the optimization of the design consists of minimizing the cost of the product $C(Y')$, subject to the possibilities defined by the production function $Y = g(\mathbf{X})$. The cost of the product is itself the cost of the inputs used. This is defined as the *input cost function*, $c(\mathbf{X})$. When it is assumed that there is no practical limit to the availability of any of the resources, the complete problem is

$$\text{Minimize:} \quad C(Y') = c(\mathbf{X})$$

$$\text{Subject to:} \quad g(X) = Y'$$

The optimization is solved by the Lagrangean:

$$L = c(\mathbf{X}) - \lambda\,[g(\mathbf{X}) - Y'].$$

The key result is that, at the optimum,

$$\frac{\partial c(\mathbf{X})}{\partial X_i} = \lambda\left(\frac{\partial g(\mathbf{X})}{\partial X_i}\right) \qquad \text{all } i$$

This will define an overall or global optimum insofar as the feasible region defined by the isoquant for Y' is convex, that is, so long as there are diminishing marginal products.

This result is easily interpreted. The rate of change on the right hand side is the marginal product, $MP_i = \partial g(\mathbf{X})/\partial X_i)$ (see Section 2.4). The quantity on the left hand side is, by definition, known as the *marginal cost*, MC_i. It is the rate of change of $c(\mathbf{X})$ with respect to a single input X_i, as stated. The Lagrangean thus leads to the *Optimality Conditions of Marginal Analysis*:

$$\frac{1}{\lambda} = \frac{MP_i}{MC_i} \qquad \text{all } i$$

The optimality conditions imply balanced design, in which the contribution of each resource X_i is equally effective per unit cost as any other resource X_j. The best design is one in which each input finally provides, at the margin, the same "bang for the buck," the same MP_i/MC_i.

The nature of the input cost function needs to be considered carefully in this context. It can be quite complex. The simplest version of the cost function is one that supposes that each resource X_i has a specific unit price, p_i, which is the same regardless of the quantity purchased. This leads to a simple linear model:

$$c(\mathbf{X}) = \sum p_i X_i$$

This simple model is the one most often used in economic analysis. It frequently does not apply in systems design, however. Attempts to buy large quantities of any resource may easily lead to prices quite different from those that prevail for small quantities. At higher volumes the trend of unit costs may be either higher or lower. One may benefit from advantages such as volume discounts and wholesale prices, or one may encounter higher prices as local supplies are exhausted. The latter is a common phenomenon in the construction of large systems, for example: wage rates for workers rise enormously as labor must be either imported or paid higher rates for extra hours.

The optimality conditions are easily applied in practice when the functions are analytic (see box). They can also be used when the production function does not have continuous derivatives. This may occur, for example, when the resources are only available in specific sizes, for example, of pipe diameters for a plant, of aircraft, or steel beams. In such cases one can obtain an optimum design by tentatively adjusting the design until the MP_i/MC_i ratios are approximately equal and the optimality conditions are met approximately (see application in Section 4.6).

The optimality conditions can be manipulated to give a useful graphical interpretation. To see this, we must first illustrate the input cost function. We do this by plotting curves of equal cost, much as we plotted isoquants for production functions. These curves of equal costs are called *budget curves*, because they show how much $\mathbf{X}$ can be obtained with any specific budget. Figure 4.1 presents budget lines for both a simple linear input cost function and a more complex one, showing unit costs increasing with quantity. Note that the feasible region for costs

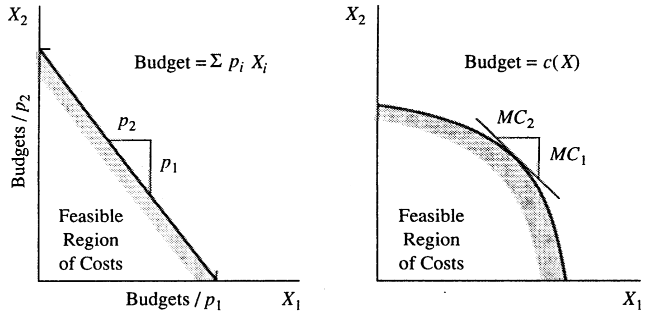

FIGURE 4.1
Budget lines for a simple linear cost function (left), and for a more complex one, showing increasing marginal costs with quantity (right).

are inside of the curves: it is always possible to obtain less, difficult to obtain the most for your money.

The graphical interpretation of the optimality conditions thus follows directly. From

$$\frac{MP_i}{MC_i} = \frac{MP_j}{MC_j} \quad \text{all } i, j$$

we write

$$-\frac{MP_j}{MP_i} = MRS_{ij} = -\frac{MC_j}{MC_i}$$

where *MRS* is the marginal rate of substitution, as defined in Section 2.4. This is to say that, at optimality, the slope of the isoquant should equal the slope of the budget line. Both functions should be tangent to each other, as Figure 4.2 shows.

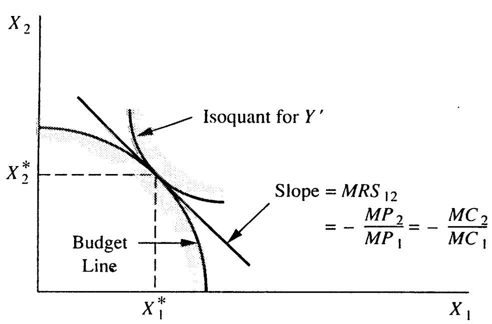

FIGURE 4.2
Graphical interpretation of the optimality conditions.

Application of Optimality Conditions

Assume a simple Cobb-Douglas production function:

$$Y = a_0 X_1^{a_1} X_2^{a_2}$$

Also, assume that the input cost function is a simple function of unit prices p_i:

$$c(\mathbf{X}) = \sum p_i X_i$$

The optimality conditions are then:

$$\frac{MP_1}{MC_1} = \frac{MP_2}{MC_2}$$

Since in this case $MP_i = (a_i/X_i)\,Y$ we can state that at the optimum:

$$\frac{a_1}{p_1 X_1^*} = \frac{a_2}{p_2 X_2^*}$$

For any specific Y', this can be solved explicitly for the optimum amount of each resource, X_i^*, and design $\mathbf{X}^*$.

4.3 EXPANSION PATH

The *expansion path* is the locus of all the optimum designs for every level of output Y. It represents the optimal sequence of designs as the scale of the desired production increases.

The expansion path is an important concept for the designer. Indeed, it should not be supposed that the best way to increase the output is to increase all inputs proportionately. While this may be so, it is in general not the case. The optimal ratio of inputs more usually changes as the size of the design changes. For example, the optimal crew for a 200,000 ton tanker is not 10 times greater than that for a 20,000 ton tanker; it is closer to being the same.

The expansion path is not an inherent property of the technical process. It is defined by the optimality conditions and thus depends on the ratios of the prevailing prices for the resources as well as how they may change with the quantities of resources. The concept is perfectly general, and expansion paths exist even when they cannot be defined by marginal analysis.

In marginal analysis, the expansion path is obtained simply by solving the optimality conditions for any resource X_i in terms of the others. Since there are $(N - 1)$ independent conditions on the N resources, the path is a curve in N-dimensional space. Note that when we construct the expansion path from the optimality conditions of marginal analysis, we do not require that the feasible region of the production function be convex as regards scale. The expansion path, constructed for different sizes of production, is simply a chain of optimal solutions obtained for each level of product. We can thus use marginal analysis to construct an expansion path when there are increasing returns to scale (and the feasible region of the production function is nonconvex) so long as there are diminishing marginal returns and convexity of the feasible region defined by the isoquants. (See box for an example in two dimensions.)

Whenever constraints on the resources make it impossible to achieve the optimal design defined by the optimality conditions of Section 4.2, the expansion path diverts from what it would otherwise be. Figure 4.3*b* illustrates the phenomenon. This shows what happens when a constraint is imposed on the situation described in the box. In this case we suppose that there is a fixed limit on the amount of X_1 available: $X_1 \leq 10$. The expansion path proceeds normally (as

---

Expansion Path

A. Linear Input Cost Function

Using the optimality conditions obtained earlier for the two resource production processes, the expansion path is

$$X_1^* = \left(\frac{1}{p_1}\right)\left(\frac{a_1}{a_2}\right)p_2 X_2^*$$

In this special case, it is a straight line indicating the same ratio of inputs, regardless of size. This would be unusual in practice.

B. Non-Linear Input Cost Function

Consider a Cobb-Douglas production process with increasing returns to scale:

$$Y = 2X_1^{0.48}\, X_2^{0.72}$$

and an input cost function:

$$c(\mathbf{X}) = X_1 + X_2^{1.5}$$

The optimality conditions can be stated as

$$\frac{MP_1}{MP_2} = \frac{MC_1}{MC_2}$$

to give

$$\frac{0.48/X_1}{0.72/X_2} = \frac{1}{1.5X_2^{0.5}}$$

which defines the expansion path as

$$X_2^* = (X_1^*)^{2/3}$$

a situation in which larger designs optimally use proportionately less and less X_2 compared to X_1. Figure 4.3*a* illustrates this situation.

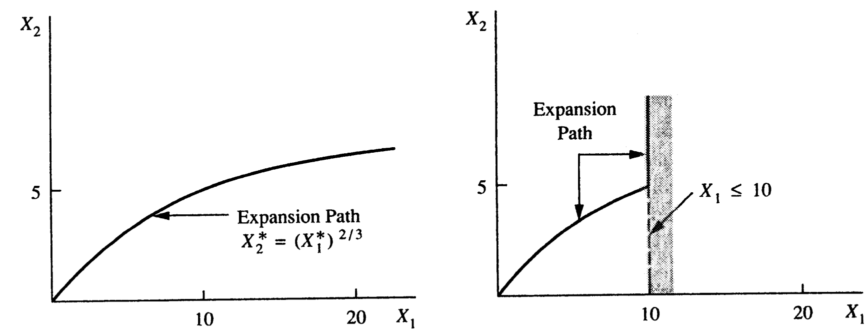

FIGURE 4.3
(*a*) Unconstrained nonlinear expansion path for example problem (left); and (*b*) expansion path constrained by the limit on a resource (right).

in Figure 4.3*a*) until this constraint makes it impossible to use more X_1. At this juncture, the optimal designs, the ones that achieve the highest output with the resources available, are located along the constraint, with ever more use of the other resource.

4.4 COST FUNCTION

The *cost function* describes the optimal, the least cost of producing any level of product Y. It follows directly from the optimality conditions and the expansion path. Each optimal design $\mathbf{X}^*$ has a cost, $c(\mathbf{X}^*)$ associated with the product Y'. The locus of all these points is the cost function, $C(Y)$. By means of the optimality conditions it can be defined solely in terms of costs and product, without reference to any of the inputs.

> **Semantic caution:** The cost function as defined above is to be carefully distinguished from the input cost function that characterizes the cost of any set of resources. These two functions refer to different quantities. The former defines an optimal cost; whereas the latter defines any cost; and their functional forms will generally be different (see box on following page).

The cost function is very useful and important in practice. It helps the system designer focus on and answer the vital question of how large a system to create. It is the means of deciding what size of system we want to have—a most crucial piece of information before we engage in any specific configuration of a system.

The information conveyed by the cost function is particularly valuable because it fills a relative void in systems engineering. Indeed, while designers typically have considerable experience in designing systems of various sizes, and may often be able to outline an optimal configuration for any size, they typically

Cost Function

Consider as before the Cobb-Douglas production process with increasing returns to scale:

$$Y = 2X_1^{0.48}X_2^{0.72}$$

and the input cost function:

$$c(\mathbf{X}) = X_1 + X_2^{1.5}$$

for which the expansion path is as shown in the previous box:

$$X_1^* = (X_2^*)^{1.5}$$

Stating Y in terms of X_2^* we obtain

$$Y = 2(X_2^*)^{1.44}$$

This leads to

$$X_2^* = \left(\frac{Y}{2}\right) \exp\left(\frac{1}{1.44}\right) = \left(\frac{Y}{2}\right)^{0.69}$$

Stating the input cost function in terms of X_2^* we obtain

$$c(\mathbf{X}^*) = 2(X_2^*)^{1.5}$$

Expressing X_2^* as a function of Y we get the cost function as

$$C(Y) = c(\mathbf{X}^*) = (2^{-0.04})Y^{1.04}$$

do not have either much experience or basis for specifying the optimal size of the system. The cost function is a means to explore this question.

The cost function has the advantage of being a particularly easy tool to use in practice. This is because, in any specific situation, it is expressed in just two dimensions: cost and product. It thus can be easily displayed and discussed with all kinds of decisionmakers, even those who are not technically trained. Being in just two dimensions, it also permits anyone to grasp the salient points.

A typical cost function appears in Figure 4.4. This presentation, with costs on the vertical axis, is the form most often used by economists. This is the way it is generally displayed in the discussion of economics issues. Notice that the feasible region for the cost function is above the curve in this case; it is always easier to spend more than required. (See also the discussion of convex feasible regions in Section 2.4.)

In engineering and other applied discussions, the cost function is generally presented as a *cost-effectiveness function*, as in Figure 4.5. Here the emphasis is on what can be obtained for any specific cost or budget. The name is different, also. One refers to effectiveness rather than product because the output of a system is often some measure of performance rather than a product to be sold. The output

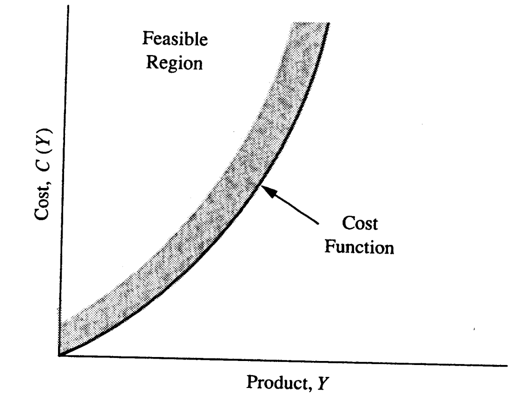

FIGURE 4.4
A cost function.

of a process for desulfurizing smokestack gasses might be "reductions in sulfur in parts per million," for example. The concept of cost-effectiveness is discussed further in Section 13.3, which covers criteria of evaluation.

One way to use the cost or cost-effectiveness function is to focus on the amount of money available. One can ask: How much can we obtain with a given budget? This information is often desired by many decisionmakers and planners. It also leads to more interesting questions of: How much more effectiveness could we obtain with a greater budget? How much more effectiveness would we have to give up if the budget were smaller?

The answers to these questions define a ratio of change in total cost, $\Delta C(Y)$, to change in effectiveness, ΔY; this is the

$$\textit{Cost-Effectiveness Ratio}: \frac{\Delta C(Y)}{\Delta Y}$$

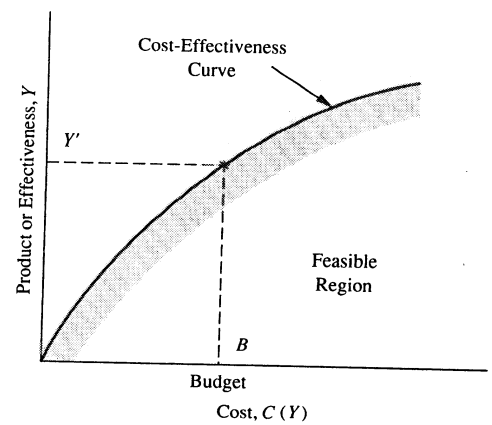

FIGURE 4.5
The cost-effectiveness curve, an alternate view of the cost function.

This gives the designer the unit cost of extra units of effectiveness. This actual cost can be compared to the value of effectiveness to determine whether one is paying too much or too little. The size of the design can then be adjusted to the right amount. Section 4.6 gives an example of an application of this analysis.

The cost function can be derived analytically whenever the production function and the input cost function are both known and analytic. To do this, one simply uses the information contained in the expansion path to state both production function and the input cost function in terms of a single X_i^* above. Thus:

$$c(\mathbf{X}^*) = c(X_i^*)$$
$$Y = g(X_i^*)$$

The production function thus stated can be inverted to obtain an expression for X_i^* in terms of output:

$$X_i^* = g^{-1}(Y)$$

This result, substituted in the input cost function gives the cost function:

$$C(Y) = c(X^*)$$

The box on page 48 gives an example of the procedure for obtaining the cost function. The cost function for a Cobb-Douglas production process and a simple linear cost function has a particularly convenient form. Its derivation is as in the box, but as it requires a messy notation, it is omitted. The result is simply

$$C(Y) = \left(A \prod p_i^{(a_i/r)}\right) Y^{1/r}$$

where $r = \sum a_i$ and indicates whether the returns to scale are increasing or not (see Section 2.4), and A is a constant. The above equation is an important result both theoretically and practically.

> **Semantic caution:** The parameter r is defined in agreement with the traditional form used in the economic literature. It should not be confused with the discount rate, also usually referred to as r in the relevant literature, as well as in Chapters 11 through 14.

Economists use the cost function to determine the nature of the production function. They do this statistically from observation on the prices, p_i, prevailing in different locations and times for each resource X_i, the amount of product and the total cost. The statistical analysis then generates estimates of the exponents of the prices, a_i/r, and of the product, $1/r$. The estimates then permit one to derive estimates of each a_i and thus to characterize the Cobb-Douglas production function that has been presumed to underlie the situation. This process is in fact how most analytic production functions are deduced.

In practice, engineering handbooks and similar references for a particular discipline frequently state the cost function for a process as

$$C(Y) = A_0 Y^a$$

where the (A_0, a) parameters have been chosen to match experience in the field. Such formulas can be good rules of thumb for a specific time and place. They are not valid in general, however, since they do not reflect prices specifically. Relative prices of resources such as labor and materials do indeed vary substantially from country to country and over time. These changes would change the parameter A_0. The exponent of Y in the cost function should remain valid, so long as the production function itself does not change, that is, so long as there is no substantial technological advance.

4.5 CHARACTERISTICS OF COST FUNCTIONS

Cost functions can have special characteristics important for design. The most significant feature concerns "economies of scale." As this concept does not apply conveniently to production processes with multiple outputs, two new concepts are also being developed and applied; those of "economies of scope" and "advantages of scale." The following paragraphs define all three, plus the complementary concept of the learning curve.

Economies of scale. Economies of scale exist for a production process when it is cheaper to produce in quantity. Formally, they exist when the average optimal cost per unit of output decreases as the level of production increases.

Mathematically, a cost function with economies of scale is generally expressed in the form:

$$C(Y) = A_0 Y^a \qquad a < 1.0$$

The average cost per unit of output, AC_Y, is then

$$AC_Y = A_0 Y^{a-1}$$

and decreases with greater Y. It should also be noted that the incremental cost for each additional unit of production, that is the marginal cost MC_Y, is both decreasing and less than the average cost:

$$MC_Y = aA_0 Y^{a-1} < AC_Y$$

Figure 4.6 illustrates these relationships.

> **Semantic caution:** Economies of scale are similar to increasing returns to scale, but these two concepts are definitely not identical. The one does not imply the other. See discussion in text.

The concept of economies of scale is closely related to—but different from—that of increasing returns to scale (see Section 2.4). Both refer to the idea of somehow getting proportionately more as the scale of production increases. Their principal difference arises from the fact that the notion of economies of scale incorporates information about the input cost function, $c(\mathbf{X})$. The form of this function can thus affect the increasing returns to scale.

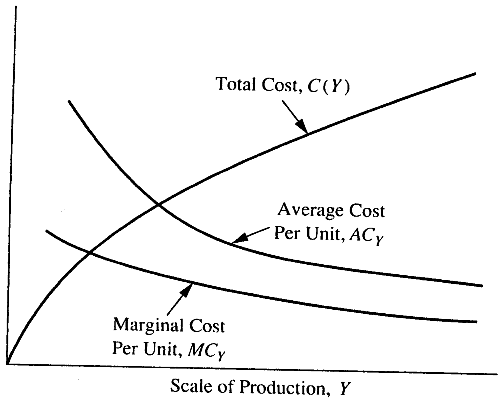

FIGURE 4.6 General relationship of total, average, and marginal costs of production for a process with economies of scale.

Economies of scale imply increasing returns to scale only in one class of situations: when the production function is Cobb-Douglas and the input cost function is linear. In this special case, the cost function is

$$C(Y) = \left(A \prod p_i^{(a_i/r)}\right) Y^{1/r}$$

and increasing returns to scale on the production process imply $r = \sum a_i > 1.0$ and thus decreasing average costs, since $1/r < 1.0$.

For the more general case in which the input cost function is nonlinear, increasing returns to scale do not imply economies of scale, or vice-versa. The box on cost functions in Section 4.4 provides an example of this possibility. In that case a production process with increasing returns to scale:

$$Y = 2X_1^{0.48} X_2^{0.72}$$

combines with a nonlinear input cost function:

$$c(\mathbf{X}) = X_1 + X_2^{1.5}$$

to define the cost function:

$$C(Y) = (2^{-0.04})Y^{1.04}$$

which does not have economies of scale. In fact, this situation illustrates *dis-economies of scale*, where increasing the size of the production process leads to higher average costs.

The concepts of economies of scale and increasing returns to scale differ in yet another respect. Returns to scale reflect the proportionate increase in output when all inputs to the design change proportionately. Economies of scale describe what happens when the scale of the output changes, whatever the optimal design and its ratio of inputs may be. Indeed, the optimal design for a larger version of the production process is often, possibly generally, not simply a scaled up version of a smaller process. For example, the optimum design for a small 20,000 ton tanker might involve a crew of 12 whereas the optimum design for a much larger 200,000 ton tanker might have a crew of only 30 rather than 120; only one captain, first mate, radio operator, et cetera may be necessary regardless of size. The Cobb-Douglas production function combined with a linear input cost function do not reflect this reality; they imply that the expansion path is a straight line with constant proportions of inputs.

Economies of scale occur in many industries. In engineering practice these are usually expressed by empirical equations of the form:

$$C(\text{Capacity}) = (\text{Constant})(\text{Capacity})^a = A_0 Y^a$$

In different fields the exponent a has been observed to vary from about 0.4 to 0.9. For example:

- In chemical engineering it often seems that $a = 0.6$; this relationship has been known as the "six-tenths rule."
- A common rule of thumb for computers has been that Cost = (Constant)(Capacity)$^{1/2}$, which is known as "Grosch's Law."
- In analyzing large water tunnels over 10 ft in diameter for New York City, we found that Cost = (Constant)(Cross-Sectional Area)$^{0.6}$.

Economies of scale may occur in all industries for which there are increasing returns to scale, but—as previously indicated—they do not necessarily. The physical increasing returns may be counterbalanced by the economic disadvantage of input costs that increase exponentially. As indicated in Section 6.6, economies of scale are frequently also obtained by shifting to a different technology.

The existence of economies of scale in a situation or industry has significant implications both for design and policy. For design, it means that there is an important reason to concentrate production in big units: concentration drives the average cost of production down. This is the principal reason why electric generating facilities, refineries, and chemical plants tend to be so large.

For national or international policy on the organization of industry, economies of scale imply that the company that can build the largest facilities will be able to produce products cheaper than its competitors, and therefore will have a definite financial advantage over them. The largest companies can therefore drive the others out of business and will tend toward monopolies. Industries characterized by economies of scale are thus said to have *natural monopolies*, the situation in which a monopoly might be in the consumer's best interests, because it could produce at the lowest average cost. But a company with a monopoly might also be able to raise prices excessively, because it would have no competition.

Economies of scale in an industry thus present a difficult policy issue: how does society ensure the most efficient solution? Should it grant franchises to a single producer, and then regulate the industry to insure that the consumers are charged a fair price? This solution is typically implemented for electric power and communication utilities. Or should the government prohibit monopolies and

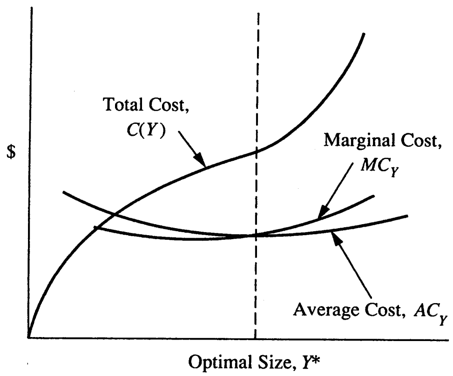

FIGURE 4.7
General relationship of total, average, and marginal costs of production for a process with an optimal size.

cartels, counting on competition to force prices down to some reasonable if not minimum level? This is the solution embedded in the "antitrust" legislation in the United States and applied to virtually every major American industry. There is no clear answer to this dilemma; each solution has its advantages. The problem must be faced whenever economies of scale exist.

Economies of scale may exist in an industry up to a certain size. Beyond this point, average costs may rise. The industry is then said to have an optimum size of plant or process. Figure 4.7 illustrates how the costs would vary in such a situation. Notice that the curve of marginal cost crosses that of average cost at the optimal size, that is, the point of lowest costs. This is because average costs decrease so long as marginal costs are less. In practice, the optimal size is nearly impossible to determine with any certainty. The existence of an optimum size is, however, important in principle: it indicates that a natural monopoly does not exist in this industry, and therefore that monopolies cannot produce and sell to the consumer more cheaply. The manufacture of automobiles is almost certainly an industry with an optimum size of production: large economies of scale to a point, and then diseconomies of scale.

Learning curve. The concept of the learning curve reflects another way in which the marginal costs of a product decrease as we manufacture more. It is, however, quite different from economies of scale in a subtle way. This must be carefully noted as it has substantially different implications for design.

> **Semantic caution:** The learning curve is also known as the "experience curve," particularly in management consulting.

The learning curve embodies the idea that any person or any plant becomes more efficient as it gains experience. The producers learn how to organize their work better, discover acceptable short cuts, and generally become more skilled. Consequently, their cost of production decreases as they produce more and more.

The strength of the learning curve effect is commonly stated in terms of the ratio, expressed as a percentage, of the marginal costs as the total number of units produced doubles between time $t = 0$ and $t = 1$. Thus:

$$\left(\frac{MC_1}{MC_0}\right) = \text{Percent Learning Curve}$$

In practice, we may hope to observe 70% to 90% learning curves.

To compare the effect of the learning curve to that of economies of scale, it is useful to express the total cost of production as

$$C(Y) = (\text{Constant})\ (\text{Number Produced})^u$$

where u is the learning curve exponent. The marginal cost of any unit at time t is then

$$MC_t = (\text{Constant})\ (\text{Number Produced})_t^{u-1}$$

This defines the percent learning curve as

$$\left(\frac{MC_1}{MC_0}\right) = (2)^{u-1}$$

For a median 85% learning curve, this implies: $(2)^{u-1} = 0.85$ and then $u = 0.78$. The effect of learning is thus comparable to that of economies of scale.

To determine the learning curve experimentally, one plots the marginal costs against the total number produced over time. This is usefully done on a logarithmic scale, in which a constant exponent for the learning curve would imply a straight line (see Figure 4.8).

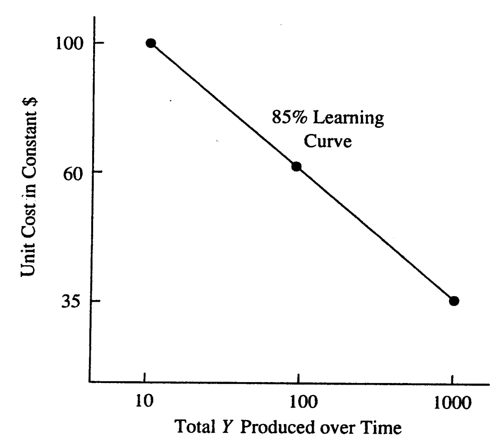

FIGURE 4.8
Typical learning curve, plotted on logarithmic scales.

When comparing the learning curve to the decreasing costs associated with economies of scale, we should carefully note the different causes of the economies, as indicated by the horizontal axes in Figures 4.6 and 4.7. With economies of scale, the costs are reduced because larger plants or facilities increase productivity; with the learning curve, the costs depend on the total number produced over time. Thus, with "learning," a smaller plant may produce more cheaply over time as it gains experience.

The learning curve effect is obviously most important in new industries, such as electronics and other producers of high technology. Correspondingly, it is less likely to be apparent in established, traditional industries, such as brick manufacture.

The implications of the learning curve for the design of systems are parallel to those of economies of scale. For new industries where the learning curve effects are strong, it is advantageous to be the first and largest producer: it enables one to produce more cheaply than competitors, thus to undersell them and gain a larger share of the market in both profits and volume. Some experts argue that the Japanese have been successful in many industries by following the policy of exploiting the learning curve effect.

Since the two effects, economies of scale and "learning," stem from independent causes, note that they can exist separately and jointly. An industry may face the possibility of both economies of scale and learning.

Economies of scope. Economies of scope is one of two ways of referring to advantages of larger size for systems producing multiple outputs. This concept is quite similar to that of economies of scale, the difference being in the nature of what causes decreasing average costs.

Economies of scope exist when the average cost of a unit output decreases with the "scope" rather than with the scale or size of output. Scope is a rather loose term embodying notions of the spread or the distribution of the production process, as distinguished from scale, a precise term relating strictly to the number of units produced. To illustrate the distinction between scale and scope, consider a transportation system for freight. Its costs in fact do depend both on the amount carried (the size or scale of the operation) and on the geographic distribution served (the scope). Both elements are in fact outputs of the transport system. The cost function would thus be

$$C(Y) = c(\text{ton-miles; mileage})$$

Economies of scale would refer to $\partial C(Y)/\partial(\text{ton-miles})$ whereas economies of scope would refer to $\partial C(Y)/\partial(\text{mileage})$.

Economies of scope exist in a number of industries where economies of scale do not. Air transport is a prime example. In air transport, the basic unit of production is the aircraft; to get more product one has to replicate this unit, so that the average cost for the output is essentially constant: there are no economies of scale. On the other hand, there may be economies of scope. If the average length of an airline's flights increases, for example, it becomes cheaper to carry a ton-mile of cargo because the processes of take-off and landing use a lot of fuel, and average costs decrease when these portions of the flight take a smaller fraction of the time. There are thus economies of scope in airlines with respect to stage length. Current research indicates that economies of scope may generally be quite significant, even when economies of scale do not exist.

Advantages of scale. Advantages of scale is the other way of referring to the benefits that may be associated with a larger system producing multiple outputs. This concept differs from those of economies of scale and of scope in that it does not describe changes in average cost.

Advantages of scale exist when the production of more units concurrently improves the quality of the service. The customers may then see that for a given price of output they get better value for money as scale increases. Strictly speaking, the cost per unit may be constant (as the price does not change), but to the customer getting better quality, advantages of scale are quite comparable to economies of scope.

To illustrate this difficult concept, consider a mass transit system operating buses in a fixed route. Because the production process simply consists of replicating identical units, by putting more buses on the line, one may presume that this system has no economies of scale, that average costs per seat-mile are constant. Yet as the size of the operation increases, the quality of the service increases: more buses on the line mean a greater frequency of service and less wait. The bigger the operation, the better it seems to the customer.

Advantages of scale occur most frequently in service industries, in which the simple fact of producing more insures a more accessible, reliable service. Examples are transportation, as just discussed; communications; banks with many branches; and so on. Formally, they are likely to occur if the multiple outputs of the production process, $\mathbf{Y}$ = (Quantity, Quality), are closely connected, if greater quantity implies greater quality.

The policy implications of advantages of scale are similar to those of economies of scale. This is natural, since both concepts reflect better value for money with increasing size of the operation. Advantages of scale provide a strong rationale for concentrating production in an industry.

4.6 APPLICATIONS

This section provides three applications to illustrate the use of marginal and cost-effectiveness analyses in practice. They deal with river transportation, as a continuation of the application for Chapter 2; the supply of water for a city; and the design of a safety system. Yet the principles used are applicable to many situations.

River transportation. In designing inland water transportation for the U.S. Agency for International Development and Colombia particularly, we built on the production function described in Section 2.6. To obtain the optimal, most cost-effective design, we had to incorporate the costs of the two main inputs, the horsepower and deck area of a system of barges and a towboat.

The key description of costs for any situation is the ratio of prices of the inputs. In this case we have

$$\frac{MC_{HP}}{MC_{DA}} = \frac{\text{(unit price of hp)}}{\text{(unit price of deck area)}}$$

This can be expected to vary widely from situation to situation, from country to country. In eastern Colombia, for example, the cost of constructing barges was low but the cost of engines that had to be transported over mountains was high, so the ratio was quite high.

We solved for the expansion path for any situation by using our computerized production function. Specifically, we identified the point on several isoquants at which the ratio of marginal costs equalled the marginal rate of substitution or ratio of marginal products of horsepower and deck area:

$$\frac{MC_{HP}}{MC_{DA}} = -MRS = -\frac{MP_{HP}}{MP_{DA}}$$

The results of the analysis defined different expansion paths—and thus quite different designs—for different conditions. Figure 4.9 shows typical results. In this case, as it happened, the expansion paths for water with no current (lakes, for example) were approximately straight lines through the origin. This means that there is an optimal design ratio valid for the entire range of design. The situation can be different for other cases, though. Where the current is strong, much more horsepower is required to obtain any specific output and the production function is changed. So then are the expansion paths. These are shifted to the right, as compared to Figure 4.9, when there is current and the expansion path no longer is a straight line through the origin. The optimal ratio of the inputs then changes drastically as greater output is needed.

Water supply. A cost-effectiveness analysis led to major improvements, worth hundreds of millions of dollars, in the design of the water supply system for New York City. This case provides a good example of the strengths and possibilities of marginal analysis. (See de Neufville, 1970 for full details).

The water supply system for New York City consists of several distinct elements: a series of dams in the mountains 100 or more miles (150 to 200 km) away from the city, aqueducts that bring this water to reservoirs in the hills at the edge of the city, a primary distribution network consisting of a few enormous

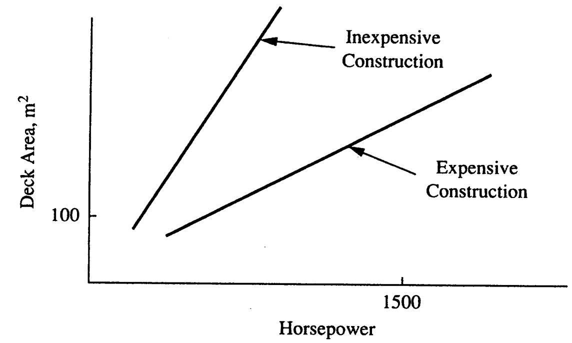

FIGURE 4.9
Expansion paths and associated optimal designs for River Transport differ according to the ratio of costs of inputs. (Results are for specific water conditions: stream width = 18 m; depth = 2.4 m; and current = 0.)

tunnels, a much more extensive secondary distribution network of pipes up to a few feet (about 1 m) in diameter, and finally the detailed distribution to each individual user.

The focus of the analysis was on the primary distribution network. As of 1975 this system consisted of City Tunnels Number 1 and 2. Each is about 20 mi (30 km) long, stretching from the Croton reservoir throughout the city, and around 15 ft (5 m) in diameter. Because of their size, the water flows through them quite slowly at about 3 ft/s (1 m/s) or approximately 2 mi/hr (3.2 km/hr). The tunnels can be visualized as large, slow-moving, underground rivers, almost as lakes.

The reason the tunnels are so large is to minimize the energy required to push the water through the system. This energy equals the frictional resistance of water in the tunnel, and is proportional to the square of the velocity of the water. For any given amount of flow, the velocity is inversely proportional to the cross-section of the stream, and thus to the square of the diameter. The net result is that the energy required to distribute the water is inversely proportional to the fourth power of the diameter:

$$\text{Energy Lost} \sim \text{Order of} \left(\frac{1}{\text{diameter}}\right)^4$$

Large tunnels thus conserve energy. For New York City, where water is distributed by gravity and all the energy comes from the elevation of the Croton reservoir, the large tunnels ensure that the pressure throughout the city will be sufficient.

City Tunnels 1 and 2 are becoming obsolete. Since they were built around 1910 and 1925, there is reason to believe that they might soon fail in some way through old age: the concrete linings of the tunnels might collapse somewhere, or valves might jam. Because the distribution of industry and residences have altered substantially in the last half century, it is also certainly true that the existing network is not optimal for the current pattern of loads on the system.

To ensure an adequate water supply for the future, New York City's Board or Water Supply thus proposed to construct a new tunnel. This new tunnel would both provide an adequate margin of safety in case of difficulty with the older tunnels, and would distribute the water more efficiently to where it is now required. Specifically, the Board of Water Supply proposed to construct the Third City Tunnel; it would be up to 28 ft (8.4 m) in diameter and cost about 1 billion dollars.

The mayor and central government of New York City were concerned about this enormous expense and wondered if it were really a good idea. They thus called in a team from the Massachusetts Institute of Technology. The questions put to us were: is the proposal from the Board of Water Supply optimal? Does it represent the best use of our money? This is a classic basis for a cost-effectiveness analysis.

Our first step was to construct the relevant production function. In doing so, the initial issue was to define its terms. The output of the system was easy to specify in principle—it was the pressure delivered by the tunnel system when supplying the demand for water. The practical difficulty is that the pressure varies throughout the system, specifically between each node at which water is taken from the tunnel system when supplying the demand for water. This is because the water is flowing and there are pressure drops. The problem was resolved by defining the output as an average of the pressure delivered at each node, weighted by the volume of water used at that node:

$$Y = \text{Average Pressure}$$

$$= \frac{\sum \text{(Pressure at node)(Volume at node)}}{\sum \text{(Volume at all nodes)}}$$

The inputs to the system were the several sections of tunnel that could be constructed:

$$\mathbf{X} = \text{(Sections of Tunnel)}$$

Each section could be varied from a diameter of zero, meaning that it would not be used, to as large as liked.

The production function itself then described the average pressure based on the design, consisting of some specification of tunnel sections. This relationship is highly complex. It is nonlinear first of all, being in fourth powers. Secondly, the equilibrium solution for the system results from the solution of differential equations describing the conservation of flow at each node. There is no direct solution to the governing set of nonlinear differential systems; an iterative solution must be obtained. The situation was thus represented by an inductive engineering production function. Specifically, we used a computer program that generated the technologically efficient output, in terms of average pressure, for any specified set of **X**. At the time this program was created, about two decades ago, this production function was at the forefront of research. Production functions of this sort are now relatively common.

The input cost function for tunnel sections of different sizes was derived from detailed engineering data available from the Board of Water Supply and their technical consultants. This has the form:

$$c(\mathbf{X}) = A\text{(Length)(Cross-Sectional Area)}^{0.6}$$

indicating economies of scale. The result is typical for tunnels, pipelines and similar items (see Section 2.4).

The optimal design for any level of output was defined iteratively. The optimality conditions could not be defined directly, because the production function was not analytic. They were therefore defined indirectly. To do so we had to calculate the ratio of marginal product, in terms of pressure, to marginal cost for each section of tunnel. We then developed a balanced design by adjusting a trial design until the (*MP*/*MC*) ratios were approximately equal for all sections. The exact procedure consisted of the following steps:

1. Pick a trial design.
2. Estimate (*MP*/*MC*) ratios.
3. Identify which sections have high (*MP*/*MC*) ratios and are thus undersized and require a larger diameter; conversely, identify the sections with low (*MP*/*MC*) ratios, which are oversized and should have smaller diameters.
4. Define and test new designs.
5. Iterate until a balanced—and therefore optimal—design is achieved.

The expansion path was likewise defined iteratively. Because the production and input cost function were highly nonlinear, the optimal design for a larger tunnel system was not a simple scaling up of a smaller design.

The cost-effectiveness function resulting naturally from the above information was as shown in Figure 4.10. This had several key features. First, it would be necessary to spend a great deal, about a quarter of a billion dollars, to obtain any significant result, as shown by point *A* in the figure. Second, the function eventually indicated strong diseconomies of scale. This is because, when the tunnel is very large, the flow is very slow and the pressure loss minuscule, so that an even larger tunnel then does very little to decrease pressure loss. Thirdly, the original design of the Board of Water Supply was at the far end of the cost-effectiveness curve.

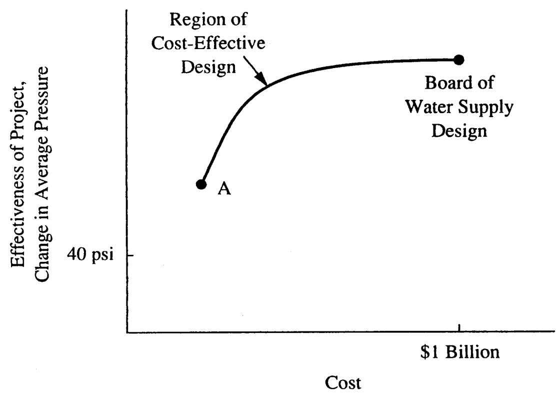

FIGURE 4.10
Cost-effectiveness function for Third City Tunnel in New York City.

The cost-effectiveness function became the focus of the determination of the optimal design. The object was to determine the right size to build and the right amount to spend. The essential criterion for determining this result was the cost-effectiveness ratio. We could use it to determine whether a larger size produced results appropriate to their costs. As Figure 4.10 demonstrates, a really small design, such as A, was not as cost-effective as a larger one; by doubling the costs one would really get a significant improvement. Conversely, a really large design was also cost-ineffective, only small improvements would finally be achieved at tremendous extra cost. The optimal, most cost-effective design was somewhere in between, around the "knee" of the cost-effectiveness function as indicated in Figure 4.10. It could not be pinpointed more precisely, as nobody had any reasonable basis for determining a precise value for the safety and convenience represented by a higher average pressure in the system. Finally, the City and the Board of Water Supply agreed on a design in which the diameter of the Third City Tunnel was to be around 21 ft (6.3 m) instead of as large as 28 ft (8.4 m). This design was then committed to construction.

This cost-effectiveness analysis led to significant savings. The revised design, featuring a tunnel up to 25% narrower, and thus over 40% smaller in cross-section, could be estimated at 20 to 25% less expensive to build. As the final cost of Third City Tunnel is well over a billion dollars, our cost-effectiveness analysis implies savings in the hundreds of millions of dollars.

Railroad crossings. A cost-effectiveness analysis was used to define the optimal allocation of safety devices to railroad crossings for the state of California. The issue was that the state wanted to reduce accidents, to lower the number of deaths at unprotected crossings. They faced several issues:

How many lives can we save with our existing budget?
How much better could we do with a bigger budget? Is this worthwhile?
How much would we lose if the budget were cut?
What is the best way to allocate the safety devices for any budget?

Based on a statistical analysis, the hundreds of different crossings could be classed into a few groups (such as "rural, medium traffic") and all locations in a group were presumed (as a reasonable approximation) to behave similarly. Several devices were available to reduce accidents; to illustrate the analysis, consider two in particular: gates and grade separations (tunnels or bridges). Gates are cheaper but not accident-proof, grade separations offer complete security at great cost. Again, as a first-order approximation the state estimated that gates and grade separations would not vary in cost by location.

This problem can be analyzed by marginal analysis even though the production process (saving lives by spending resources) is neither analytic nor continuous. It is suitable because the resources have diminishing marginal products: intelligent, technically efficient design uses each input in its most productive situations first, then in its next best, and so on. This defines diminishing marginal products by construction.

To conduct the cost-effectiveness analysis we construct a table that gives the ratio of marginal product to marginal cost for each possible type of investment at each location. In our case this ratio is simply

$$\left(\frac{MP}{MC}\right) = \left(\frac{\text{Lives Saved}}{\text{Unit Cost}}\right)_i \qquad \text{for each device } i$$

At the start of the analysis we then have something like Table 4.1.

To conduct the analysis we use the simple principle of using resources, of placing investments where they are most effective. The optimal design explores all the best opportunities first, then the next best, and so on. This procedure generates a cost-effectiveness curve that constantly flattens as in Figure 4.5. The

TABLE 4.1
Initial Matrix for Cost-Effectiveness of Railroad Crossings: each A_{ij} represents the *MP/MC* ratio for an investment at a location.

Device Type	Location Class 1	2	. . .
Gates	A_{11}	A_{12}	. . .
Grade separations	A_{21}	A_{22}	. . .

TABLE 4.2
Second Matrix for Cost-Effectiveness of Railroad Crossings: The A_{ij} have been renewed to take into account prior designs.

Device Type	Location Class 1	2	...
Gates	0	A_{12}	...
Grade Separations	A'_{21}	A_{22}	...

expansion path is then simply the sequence of placements that constitute the optimal strategy.

In our case, suppose that the most cost-effective investment, the one with the greatest *MP/MC* ratio, is gates at locations of class 1 (i.e., $A_{11} > A_{12}, A_{21}, A_{22} \ldots$). This means that we can reduce deaths by $(A_{11})N$ for N locations of class 1 at a cost of N gates. Having fully exploited the possibility of gates at this location we proceed to look for the next best possibility. Before we do so, we must revise our table to take into account the use of gates at location of class 1; specifically, we must recognize that the use of another device at these locations will no longer be as effective as when they had no protection: the A_{ij} for location of class 1 must be reduced as in Table 4.2.

This kind of analysis can be applied quite generally in many practical situations. The Railroad Crossing problem (4.17) attached to this chapter is a good exercise to fix the procedure in mind.

REFERENCES

Baumol, W. J., (1977). *Economic Theory and Operations Analysis*, 4th ed. Prentice-Hall, Englewood Cliffs, NJ.

de Neufville, R., (1970). "Cost-Effectiveness Analysis of Civil Engineering Systems—New York City's Primary Water Supply System," *Operations Research*, Vol.13, September–October, pp. 758–804.

Hax, A., and Majluf, N. S., (1984). "Competitive Cost Dynamics: The Experience Curve," Chapter 6, *Strategic Management, An Integrative Approach*, Prentice-Hall, Englewood Cliffs, NJ.

PROBLEMS

4.1. *Marginal Analysis I*
Using the input cost function: $C = 2X + 3Y^2$; and the results from Problem 2.4:
(*a*) Write the equation for the expansion path and sketch it.
(*b*) What is the significance of the expansion path?

(*c*) Find and sketch the cost-effectiveness function.
(*d*) Given a budget of $C \leq \$80$, what is the maximum output you can achieve? What combination of inputs achieves this result?
(*e*) What is the maximum output that can be achieved if, with no budget constraint:

$$X \leq 2 \quad \text{and} \quad Y \leq 8?$$

4.2. *Marginal Analysis II*
Do (*a*) to (*c*) as above for Problem 2.5, using the input cost function

$$C = 6X + 4Y^3$$

4.3. *Marginal Analysis III*
Given the production function: $Z = 10X^{0.3}Y^{0.4}$
and the input cost function: $C = 3X + 2Y^2$
(*a*) Find the marginal products and costs.
(*b*) State the optimality criteria for this situation.
(*c*) Write an expression in X and Y for the marginal rate of substitution of X for Y.
(*d*) Find the expansion path.
(*e*) Find the cost-effectiveness function.

4.4. *Marginal Analysis IV*
Given the production function: $Z = 2X^{0.5}Y^{2.5}$
(*a*) Are the marginal products increasing or decreasing? Over what ranges?
(*b*) What type of returns to scale does this function exhibit?
Given the input cost function:

$$C = (64/3)X^3 + (5/3)Y^3 + 17 \quad \text{for } X, Y > 0$$

$$C = 0 \quad \text{for } X, Y = 0$$

(*c*) What is the equation of the expansion path?
(*d*) Write an equation for the cost-effectiveness function.

4.5. *Marginal Analysis V*
Given the production function: $Z = X^{0.7}Y^{1.4}$
(*a*) Are the marginal products increasing or decreasing?
(*b*) What types of returns to scale does this function exhibit?
Given the input cost function: $C = X^3/3 + 16/3Y^3, \quad X, Y \geq 0.$
(*c*) What is the equation of the expansion path?
(*d*) Write an equation for the cost-effectiveness function.

4.6. *Marginal Analysis VI*
The technology of a process is represented by: $Q = 2XY$
(*a*) Does the production function display increasing returns to scale?
(*b*) Find the marginal productivity with respect to X, to Y.
For a particular island with limited resources: $C = 6X^2 + 3Y^4 + 5$
(*c*) Write an equation that defines the expansion path.
(*d*) Write an equation describing the cost-effectiveness function.

4.7. *Marginal Analysis VII*
For Problem 2.6 and the input cost function $C = 2X + 4Y$, determine the expansion path and the cost-effectiveness function.

4.8. *Marginal Analysis VIII*
For Problem 2.7 and the input cost function $C = 2X^2 + Y$:
(*a*) Determine the expansion path and the cost-effectiveness function.
(*b*) Reconcile the result with the returns to scale defined by Problem 2.7.

4.9. *Vi-Tall Again*
Plot of the expansion path for Problem 2.8.

4.10. *Economies of Scale?*
Given: $Z = X^{0.3}Y^{0.8}$
(*a*) What are the marginal products?
(*b*) Are the returns to scale increasing? Why or why not?
For the input cost function $C = X^3 + 4Y^2$
(*c*) Write an equation defining the expansion path.
(*d*) Write an equation defining the cost-effectiveness function.
(*e*) Does the cost-effectiveness function show economies of scale?

4.11. *More Road Work*
(*a*) Use the optimality criteria to solve Problem 3.5.
(*b*) Explain the relation between the optimality criteria and the Lagrangean multipliers obtained in Problem 3.5.

4.12. *Timothy Burr, III*
Bothered by the lack of economies of scale (see Problem 3.7), TB the Third decides to investigate the advantages of experience. He looks over the production figures for the last eight years:

	Year							
	1	2	3	4	5	6	7	8
Product	1700	2900	2500	1400	1800	2400	2600	2700
Av. cost	3.00	2.65	2.55	2.40	2.35	2.30	2.15	2.10

Estimate the percent learning curve for the logging operation.

4.13. *Efficient Design*
Given the production function: $Z = 4^{0.7}X^{0.3}Y^{0.5}$
and the input cost function: $C = 38 + 3X + 10Y^2$
(*a*) Does the production function define a convex feasible region? Explain.
(*b*) What type of returns to scale does the production function exhibit?
(*c*) Define and calculate the marginal rate of substitution.
(*d*) Determine the expansion path.

(*e*) Given a budget of $C = \$60$, what is the maximum output that can be achieved? What combination of inputs is required to achieve this output at $C = \$60$?

4.14. *Chemical Plant*
There are economies of scale if chemical reactor tanks are placed in series. A liquid waste passes through this system of tanks at a constant flow rate and chemicals are added to each tank to reduce the concentration from S_0 to S_n. The control variables for tank reactors in series are:

n = number of tanks

v = volume of each tank, where all tanks have the same volume

The cost of this setup may be expressed as: $C = 0.1\ n^{0.4}v^{0.6}$

The quantity produced may be approximated as: $Q = \dfrac{10v}{(10 - n)}$

(*a*) Compute the values of the control variables that minimize cost, as functions of Q^*, the target production volume.
(*b*) Compute the minimum cost as a function of Q^*.

4.15. *Potamia*
The production function for unpaved roads in Potamia is approximately:

$$M = 1.8L^{0.2}B^{0.8}$$

where: M is the length of road constructed
L is labor used (in 1000 person-hours); and
B is the machine-hours of bulldozers

(*a*) When labor costs \$500/1000 hr and bulldozers \$500/hr, in what ratio should labor and bulldozers be used?
(*b*) If L is constrained (≤ 100), how would this change the optimal ratio of L to B? (You may find it useful to graph the expansion path).
(*c*) At the invitation of the Prime Minister, you attend a meeting of the Potamian cabinet. The Minister for Industrial Development observes that, in developed nations, the B/L ratio is higher than it is in Potamia, and that progress is synonymous with increasing that ratio. The Labor Minister disagrees—the M/L ratio is the one that Potamia must increase, at almost any price, if it is to join the technically advanced world. Comment.

4.16. *Electric Company*
The Electric Company's oil-fired generating plant is currently operating near design capacity. More power can be generated by the plant but not very efficiently.

$Z = 10F^{0.4}L^{0.1}$ — Z = extra megawatts; C = extra cost in dollars/hr

$C = 24F + 6L$ — F = extra barrels of oil; L = extra operating and maintenance employees

(*a*) Find the marginal costs and marginal products.
(*b*) Determine the expansion path (optimal combinations of inputs).
(*c*) Write an equation describing the cost-effectiveness function for extra electricity generation from the current plant.
(*d*) Over the next five years, the company can supply additional power (above the present level) to its customers from three sources:

1. Extra capacity can be wrung from the current plant, as previously described.
2. Power can be bought in any quantity from the regional power grid at \$6/ MW-hr.
3. An antiquated hydro-electric facility can be put back in service, to produce up to 20 MW. The cost of having the facility in service is \$60/hr, regardless of the amount of electricity generated.

Construct the company's cost-effectiveness function for supplying additional power.

4.17. *Railroad Crossing*
You have been appointed project manager to work out the region's railroad crossing protection plan for the new high-speed rail line. At present, all of the 100 intersections along this line are protected by flashing lights. You must decide which crossings should have the additional protection of automatic gates and which should have grade separations.

Your staff has given you the following data describing the options and likely effects for each of the 100 crossings along this line. These are divided into three groups according to such factors as amount of traffic.

Crossing group	Number of crossings	Estimated annual number of accidents avoided per crossing by installing	
		Gates	Grade separations
A	20	0.5	0.8
B	50	0.2	0.6
C	30	0.12	0.3

The annual cost of each gate is \$100K and \$400K for each grade separation.
(*a*) What is the implied product of this system?
(*b*) What would be the logical resources for this problem?
(*c*) Given parts (a) and (b), what are the axes of the production function?
(*d*) Draw the isoquant passing through the point that represents building a gate at eight of the crossings in group A; draw the isoquant associated with an annual reduction of 22 accidents; and draw the line representing a budget of \$4000K.
(*e*) What points on the above isoquants are on the expansion path?
(*f*) If \$2M or less is to be spent annually, the most effective way to spend it is by installing gates at the crossings in group A, where 0.5 lives are saved for every \$100K commitment. How many lives are saved per \$100K for the five other options?

If \$2.3M is to be spent, either one of the group A crossings can be redesigned as a grade separation or three of the group B crossings can be supplied with gates. Which is a more effective use of the marginal \$300K expenditure? What is the marginal benefit per \$100K for each alternative?
(*g*) Examine the marginal benefit of competing uses of succeeding increments of money, and construct the cost-effectiveness function.
(*h*) Describe a few ways to decide what crossing protections should be built. Which method is optimal? Which method would you expect to see used in practice? Are the results the same?