

3

CRITERIA FOR OPTIMAL INVESTMENT DECISIONS

In this chapter we consider the investment decisions of firms whose shares are traded in perfect capital markets and whose managements are following the market value criterion, that is, maximize the market value of the shares outstanding before the investment decision is made. Strictly speaking, such decisions are technological rather than "financial" problems and so belong to the field of "production." For a variety of reasons, however, the general subject of "capital budgeting" has come to be taught in finance courses, and a considerable part of the literature in finance is focused on this class of problems.

One reason is that the criteria for optimal investment, as we shall see, are closely bound up with the present value apparatus developed in Chapters 1 and 2, and, in fact, capital budgeting is often introduced in finance courses as essentially merely one important application of this apparatus. Perhaps even more decisive in associating capital budgeting with finance is the fact that under uncertainty, one of the key, and most controversial,

questions that immediately arise is whether and to what extent the firm's choice of financing method—bonds versus stock issues versus retained earnings, and so on—does or should influence investment decisions. This class of questions, often referred to as the “cost of capital problem,” has been a major preoccupation in the field of corporation finance. Some aspects of this problem, notably, the issue of internal versus external financing, have already been treated in the previous chapter, but we postpone most of the discussion of this problem to Part II of the book in the context of models allowing for uncertainty.

Don't use the Her ↓
 Because our concern is mainly with the literature on capital budgeting, we restrict our attention here fairly narrowly to the central issue of the appropriate criterion—actually, appropriate equivalent alternative criteria—for investment decisions. We shall have little to say, and certainly little encouraging to say, about computational problems, nor except incidentally and in passing, shall we discuss how the model can be specialized and adapted to deal with different kinds of technological and market settings. Issues of the latter kind, along with the related ones of the construction of macroinvestment functions, are best left to standard economics courses.

The chapter itself is divided into two main sections. The first is concerned mainly with the formal derivation and representation of the criteria for optimality consistent with the market value rule. Much of the discussion here is also concerned with showing how the apparatus for optimal investment decisions, as presented in the theory of finance, is related to the apparatus used to represent optimal output decisions in the standard microeconomic theory of the firm. The second section of the chapter then attempts to show how the criteria for optimal investment decisions can be used as a critical guide in surveying the literature of applied capital budgeting. The two sections of the chapter, however, can be read independently.

I. THE REPRESENTATION OF OPTIMAL INVESTMENT DECISIONS¹

I.A. The Case of a Single Capital Good and Two Time Periods

As usual, we find it helpful to begin, not by considering the problem in all its complexity, but by focusing on a simple special case and then adding

¹ To simplify and shorten the presentation of the optimality conditions in Section I of this chapter, we make extensive use of the fact that any constrained maximization problem of the form $\max F(\bullet)$ subject to $G(\bullet) = 0$ can be restated as an equivalent unconstrained maximization problem of the form $\max F(\bullet) - \lambda G(\bullet)$, where λ is a so-called “lagrangian multiplier.” Readers for whom the lagrangian technique is unfamiliar should simply skim over the equations in which it is used—or if they are more venturesome, translate them back into the more familiar constrained form—and move on to the graphical and verbal discussions.

complications one by one. In particular we start by assuming the existence of a firm that meets the following conditions:

1. The firm produces a single, homogeneous, nonstorable commodity in each time period t and sells it in a perfect market at known prices. In the language of the ordinary theory of the firm, the firm is a pure competitor or price taker or quantity adjuster. Extension to the case of monopoly, in which price as well as quantity must be determined, is relatively straightforward, but for our purposes the additional insights of such an extension are not worth the cost of carrying along the additional decision variable.

2. The production can be accomplished by the services of a stock of a single, homogeneous capital good or machine, owned by the firm, in conjunction with a single, homogeneous type of cooperating labor not owned by the firm. Both machines and man-hours of labor are purchased in perfect markets at known prices and wages, respectively, and there are no internal costs or lags involved in adjusting either factor to its desired level. For both productive factors, the volume of services rendered is taken as strictly proportional to the number of units employed.²

3. The technological production possibilities in any period can be described by a "production function" of the form

$$q_t = F_t(l_t, k_t),$$

where q_t represents the number of units of the commodity to be produced, l_t the number of units of labor services used, and k_t the number of machines, with the function F_t assumed to be everywhere concave. In the language of the theory of the firm it is characterized by "diminishing returns to factor proportions" and "decreasing returns to scale."³ The assumption that returns to scale are decreasing everywhere is much stronger than needed and is introduced solely to simplify the presentation.

4. The firm has a finite planning horizon of two periods. At the end of the second period, production ceases, and all machines held are sold.

5. The shares of the firm are traded in a perfect capital market, and the firm's management follows the market value criterion. Our conventions as to timing are that sales of product produced during t and payments to labor

² This means among other things that the only way the firm can alter the capital intensity of its production process is by varying the number of machines that it holds. In practice, of course, a firm typically has many other ways to vary its effective stock of capital goods—for example, by using equipment that is more durable. For simplicity of presentation, however, and especially with a view to making the optimality conditions stand out as sharply as possible, all consideration of these additional dimensions to the capital stock are deferred until Sec. II.

³ Mathematically, these assumptions imply that (1) $F'_{l,t}, F'_{k,t} > 0$, (2) $F''_{l,t}, F''_{k,t} \leq 0$, and (3) $F_t(\lambda l_t, \lambda k_t) \leq \lambda F_t(l_t, k_t)$.

employed in this production take place at the start of period $t + 1$ but the purchase of machines takes place at the beginning of the production period. The sale of any machines remaining at the end of a period takes place at the price as of the start of the next period.⁴

Given these conditions, the decisions of the management with respect to output, employment, and the purchase, and sale, of machines in each of the two periods can be represented formally as the solution to the problem

$$\begin{aligned} \max V_1 &= K_1 + \frac{K_2}{1 + r_2} = -I_1 + \frac{X_2 + V_2}{1 + r_2} \\ &= -\pi_1(k_1 - k_0) + \frac{p_2q_1 - w_2l_1 + \pi_2k_1}{1 + r_2} \end{aligned} \quad (3.1)$$

subject to

$$q_1 = F_1(l_1, k_1), \quad (3.2)$$

where p_2 is the unit price of output at period 2, w_2 the wage rate per unit of labor, π_1 and π_2 the purchase and sale prices, respectively, of a unit of machinery, and the K 's, V 's, X 's, and I 's have the same meanings as in Chapter 2. Expressions (3.1) and (3.2) can be reformulated as the unconstrained problem

$$\max_{q_1, l_1, k_1, \lambda_1} V_1 = \pi_1(k_1 - k_0) + \frac{p_2q_1 - w_2l_1 + \pi_2k_1}{1 + r_2} - \lambda_1(q_1 - F_1(l_1, k_1)), \quad (3.1')$$

where λ_1 is the lagrangian multiplier for the production function constraint.⁵

I.B. Graphical Representation of the Complete Solution

Because there are so many distinct decision variables, no very simple representation of the optimal decisions is possible in a single, all-inclusive graph. We can, however, regroup the component terms in the maximand so as to present the full, simultaneous solution as if it were actually being

⁴These conventions are the same as those in Chap. 2. It is important to emphasize that they are mere conventions and the reader who prefers different ones can readily adjust all our results to suit his own tastes without affecting anything of substance. For example, if one wants to regard the wages as being paid in advance rather than at the end of the period, one merely redefines the wage rate as $w'_1 = w_2/(1 + r_2)$, and so on.

⁵In principle, there are also nonnegativity constraints for q , l , and k , but as has been our practice throughout, we assume that these are always satisfied. We also continue to omit the second-order or sufficiency conditions for optimality, beyond noting here that the assumptions in Sec. I.A. do guarantee that they will indeed be met.

reached in a sequence of separate steps. In fact, there are several ways in which this can be accomplished, depending on which particular decisions happen to be of most interest.

For example, given a firm that sells its product in a perfect market, the main interest in the ordinary theory of the firm usually attaches to the output decision. If so, and assuming, for simplicity, that $k_0 = 0$, the problem can be restated as

$$\begin{aligned} \max_{q_1, l_1, k_1} V_1 &= \frac{p_2 q_1 - w_2 l_1}{1 + r_2} - k_1 \left(\pi_1 - \frac{\pi_2}{1 + r_2} \right) - \lambda_1 (q_1 - F_1(l_1, k_1)) \\ &= \max_{q_1} \left[\frac{p_2 q_1}{1 + r_2} + \max_{l_1, k_1} \left[- \frac{w_2 l_1}{1 + r_2} - k_1 \left(\pi_1 - \frac{\pi_2}{1 + r_2} \right) - \lambda_1 (q_1 - F_1(l_1, k_1)) \right] \right] \\ &= \max_{q_1} \left[\frac{p_2 q_1}{1 + r_2} - \min_{l_1, k_1} \left[\frac{w_2 l_1}{1 + r_2} + k_1 \left(\pi_1 - \frac{\pi_2}{1 + r_2} \right) + \lambda_1 (q_1 - F_1(l_1, k_1)) \right] \right]. \end{aligned} \tag{3.3}$$

In words, the "scenario" visualized in this regrouping would be as follows: First, pick some arbitrary value for q_1 , and solve the minimum problem in the inner brackets; that is, find the minimum total discounted costs, say, C , of producing this particular level of output. Once the minimum cost has

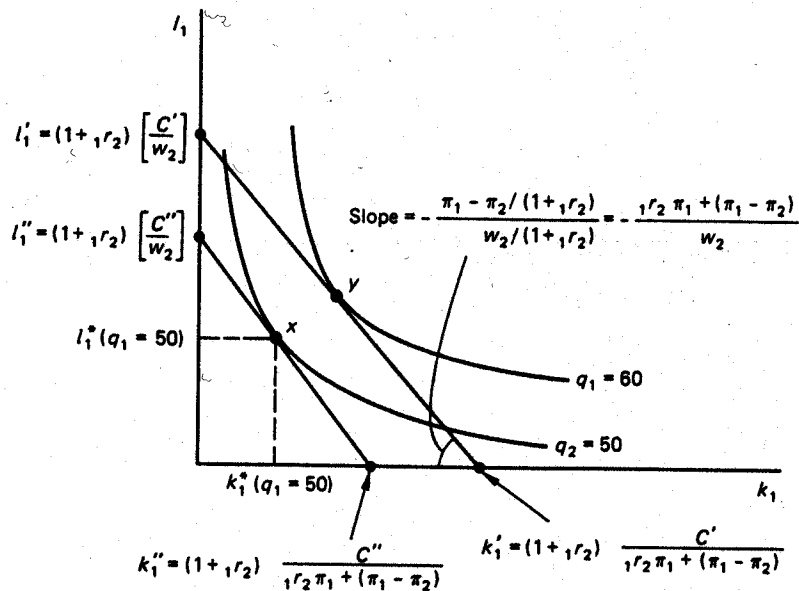


Figure 3.1 Optimal Combinations of Labor and Machines for Specified Levels of Output

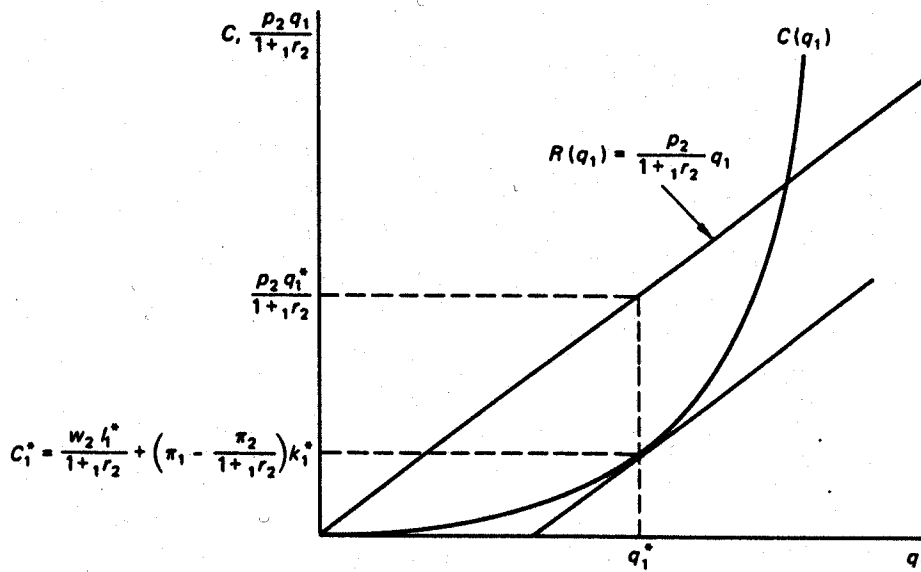


Figure 3.2 Total Revenue, Total Cost, and Optimal Output Decision

been found for every value of q_1 , select the one that maximizes the “profit” for the period, that is, the difference between total (discounted) revenue obtained from the sale of the output and the minimum cost of producing it.

This scenario is shown graphically in Figures 3.1 and 3.2. In Figure 3.1, the indifference curves, usually referred to as “isoquants” in this application, show the various combinations of labor units and machine units that can be used to produce a given level of output. Each of the parallel straight lines shows a different level of total discounted costs C , and the points along the line represent the quantities of labor services and machine services that could be purchased with this level of total expenditure at the given prices of $w_2/(1+r_2)$ per unit of labor and $\pi_1 - \pi_2/(1+r_2)$ per machine. Note that the economic “cost” of the capital services provided by a machine is not simply the initial outlay π_1 but the initial outlay minus the resale value after the period’s use (discounted back to the present). Or to approach it another way, because

$$\pi_1 - \frac{\pi_2}{1+r_2} = \frac{r_2\pi_1 + (\pi_1 - \pi_2)}{1+r_2} = \pi_1 \left[r_2 + \left(\frac{\pi_1 - \pi_2}{\pi_1} \right) \right] \frac{1}{1+r_2},$$

the (discounted) cost of capital services can be seen to be the interest foregone on the initial purchase price plus the rate of depreciation (or minus the rate of appreciation if $\pi_2 > \pi_1$).⁶

⁶ The term $[r_2\pi_1 + (\pi_1 - \pi_2)]/(1+r_2)$ is often referred to in the literature as the “rental value” of the machine.

Each of the tangency points, such as x or y , represents the minimum cost combination of productive services for the indicated quantity of output, and the actual value of total cost C that this combination represents is easily determined from the value of either intercept.⁷

Having found the minimum cost combinations, we can plot the minimum value of total cost for each value of q_1 as in the curve $C(q_1)$ in Figure 3.2. The remaining term in the maximand (3.3), $p_2 q_1 / (1 + r_2)$, represents the sales proceeds or total (discounted) revenue as a function of q_1 and is graphed in Figure 3.2 as the straight line with slope $p_2 / (1 + r_2)$. The optimal q_1^* is found at that value of q_1 for which the distance between the $R(q_1)$ and $C(q_1)$ curves is a maximum, which can easily be shown to be the point at which the slopes of the two functions are identical. Alternatively and equivalently, as in Figure 3.3, we could directly graph the slopes of the

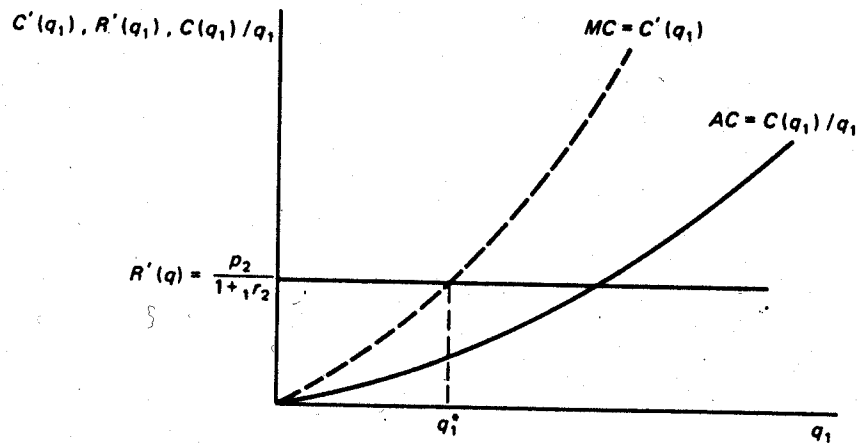


Figure 3.3 Marginal Conditions and Optimal Output

$R(q_1)$ and $C(q_1)$ functions, obtaining the optimal q_1^* at the intersection of the marginal revenue and marginal cost curves in the manner standard in the elementary theory of the firm.

⁷ The first-order conditions for minimum total cost for a given output are

$$\frac{\partial C}{\partial l_1} = 0 = \frac{w_2}{1 + r_2} - \lambda_1 F'_{1l_1}$$

$$\frac{\partial C}{\partial k_1} = 0 = \pi_1 - \frac{\pi_2}{1 + r_2} - \lambda_1 F'_{1k_1}$$

which together imply

$$F'_{1k_1} / F'_{1l_1} = [\pi_1 - \pi_2 / (1 + r_2)] / w_2 / (1 + r_2) = [r_2 \pi_1 + (\pi_1 - \pi_2)] / w_2;$$

that is, the ratio of the marginal physical product of capital to the marginal physical product of labor—the slope of the isoquant at the given value of q_1 —equals the ratio of the cost of a unit of capital services to the cost of a unit of labor services—the slope of the “price line” in Fig. 3.1.

I.C. An Alternative Representation Highlighting the Investment Decision

But in contrast with the elementary theory of the firm the main concern of the theory of finance is the investment decision rather than the output decision. Once again, however, we can obtain a highlighting of the variable of interest k_1 by a regrouping and reinterpretation of the maximand (3.1'). In particular, we can rewrite it as

$$\max_{q_1, l_1, k_1} V_1 = \max_{k_1} \left[\max_{q_1, l_1} \left[\frac{p_2 q_1 - w_2 l_1}{1 + r_2} - \lambda_1 (q_1 - F_1(l_1, k_1)) \right] - k_1 \left(\pi_1 - \frac{\pi_2}{1 + r_2} \right) \right]. \quad (3.4)$$

In words, Equation (3.4) tells us to pick some value for k_1 and then determine values of q_1 and l_1 that yield the maximum possible discounted "quasirent" or discounted "cash flow"

$$\frac{X_2}{1 + r_2} = \max_{q_1, l_1} \left[\frac{p_2 q_1 - w_2 l_1}{1 + r_2} - \lambda_1 (q_1 - F_1(l_1, k_1)) \right]$$

that could be attained with this amount of k_1 . This maximum is, of course, a constrained maximum, the constraint being the production function $q_1 = F_1(l_1, k_1)$. Attainment of the maximum discounted cash flow implies that we keep adding labor to the given stock of machines until the sales value of the additional output so obtained no longer exceeds the cost of the added labor, that is, until the value of the marginal product of labor equals the wage rate.⁸

Having found the maximum return for any given number of machines, call it $X_2(k_1)/(1 + r_2)$, we can plot these maxima for all values of k_1 as in Figure 3.4. The function $C(k_1) = k_1[\pi_1 - \pi_2/(1 + r_2)]$ in this figure is the cost of the capital services of k_1 machines, and the optimum number of machines k_1^* is found at the point of maximum distance between the two curves.⁹

⁸ Because maximizing $X_2/(1 + r_2)$ is equivalent to maximizing X_2 , the first-order conditions for maximum discounted cash flow for a given stock of capital are

$$\frac{\partial X_2}{\partial l_1} = 0 = \frac{-w_2}{1 + r_2} + \lambda_1 F'_{l_1},$$

$$\frac{\partial X_2}{\partial q_1} = 0 = \frac{p_2}{1 + r_2} - \lambda_1,$$

which together imply $w_2 = p_2 F'_{l_1}$.

⁹ The values of q_1 and l_1 corresponding to this solution are, of course, precisely the same as illustrated in Fig. 3.2, because Eqs. (3.3) and (3.4) are mathematically equivalent.

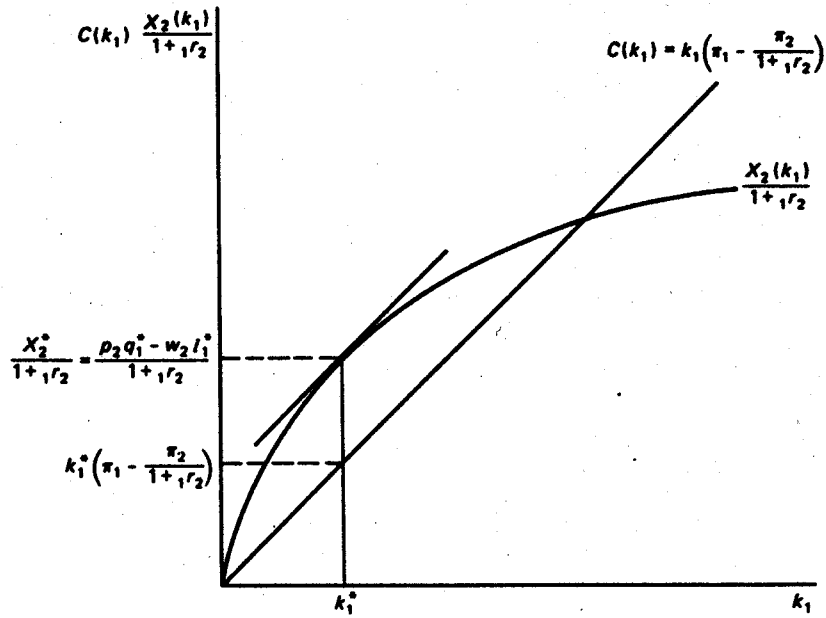


Figure 3.4 Total Cash Flow, Total Cost of Capital Services, and Optimal Stock of Machines

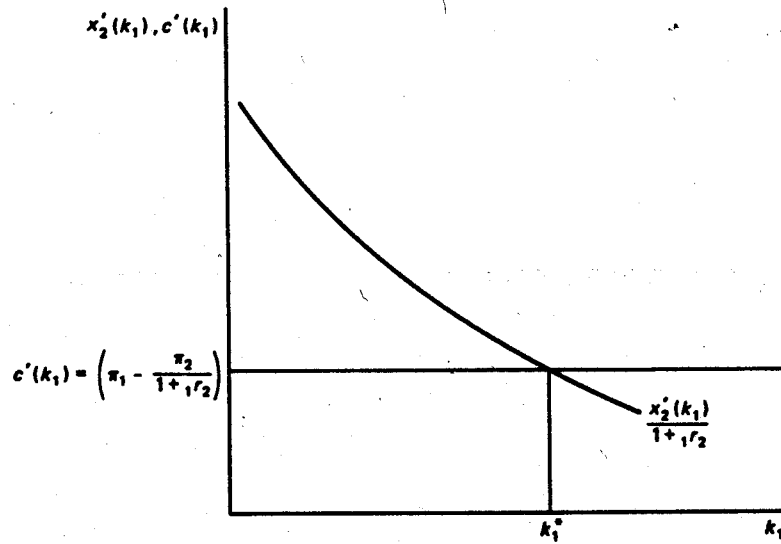


Figure 3.5 Marginal Cash Flow, Marginal Cost of Capital Services, and Optimal Stock of Machines

As before, we can also express this solution in terms of the marginal rather than the total conditions. In Figure 3.5, for example, the function $x'_2(k_1)/(1 + r_2)$ represents the marginal discounted cash flow from investment in machines, assuming optimal output and manning, and is, of course, simply the slope of the total cash flow function in Figure 3.4. The optimum stock of capital k_1^* is then found at the point at which the marginal cash flow is equal to the marginal cost of capital services per unit, that is, at the point satisfying the condition

$$\frac{x'_2(k_1)}{1 + r_2} = \pi_1 - \frac{\pi_2}{1 + r_2} \equiv c'(k_1) \quad (3.5a)$$

or more compactly

$$x'_2(k_1) = r_2\pi_1 + (\pi_1 - \pi_2).^{10} \quad (3.5b)$$

I.C.1. Alternative forms for the optimizing conditions

There are, of course, still other ways in which this fundamental criterion (3.5b) can be expressed. A popular one, which we make great use of in subsequent sections, is the "present value" form:

$$\pi_1 = \frac{x'_2(k_1)}{1 + r_2} + \frac{\pi_2}{1 + r_2}. \quad (3.6)$$

In words, add to the capital stock until the present value of the marginal cash flow plus the resale value of the last unit exactly equals the initial purchase price per unit. An equally if not more popular alternative is the internal yield or rate of return form. In particular, if we define the marginal (one-period) internal yield as

$$r_2^*(k_1) \equiv \frac{x'_2(k_1)}{\pi_1} + \frac{\pi_2 - \pi_1}{\pi_1}, \quad (3.7)$$

¹⁰ Note that the marginal cash flow $x'_2(k_1)$, which is the total differential $dx_2(k_1)/dk_1$, is not the same as the marginal product of capital F'_{1k} , because the function $x'_2(k_1)$ does not keep the quantity of labor constant as k_1 varies but allows it to be adjusted optimally to the particular value of k_1 . In equilibrium, of course, but only there, the two are the same, as is easily seen by noting that the first-order conditions for a maximum of (3.3) include

$$\frac{\partial V_1}{\partial k_1} = 0 = -\pi_1 + \frac{\pi_2}{1 + r_2} + \lambda_1 F'_{1k},$$

$$\frac{\partial V_1}{\partial q_1} = 0 = \frac{p_2}{1 + r_2} - \lambda_1,$$

which together imply

$$p_2 F'_{1k} = r_2\pi_1 + (\pi_1 - \pi_2).$$

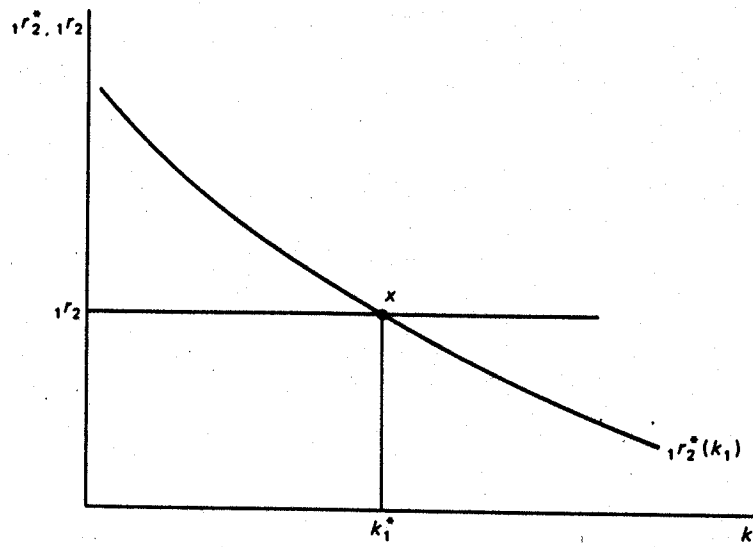


Figure 3.6 Marginal Conditions in Rate of Return Form

the criterion becomes simply

$$r_2^*(k_1) = r_2. \quad (3.8)$$

In words, add to the capital stock until the marginal internal yield on the capital stock, optimally utilized, exactly equals the market rate of interest. The condition is shown graphically in Figure 3.6.¹¹

1.C.2. Optimal capital stock and optimal investment

The optimality conditions have so far been stated in terms of the physical stock variable k_1 , but in the field of finance and in applied capital budgeting the variable more directly of interest is "total net investment," which is the value in money units of the net change in the stock of capital. In the present context, however, going from one variable to the other really involves nothing more than a relabeling of the abscissa in graphs, such as Figure 3.6, to obtain one like Figure 3.7a. Under our assumptions, the initial capital stock k_0 is a known constant—for simplicity, we set its value at zero—so that the value of the investment I_1 at the beginning of the period is simply $I_1 = \pi_1 \Delta k = \pi_1 k_1 - \pi_1 k$, which is just a scale-changing, linear transformation of the variable k_1 , that is, a transformation of the form $y =$

¹¹ Still another way of saying the same thing would be that the market rate of interest is the "cutoff" rate for capital budgeting or the "cost of capital" in the sense of the minimum yield that an addition to the capital stock must offer to be just worth undertaking from the standpoint of the owners.

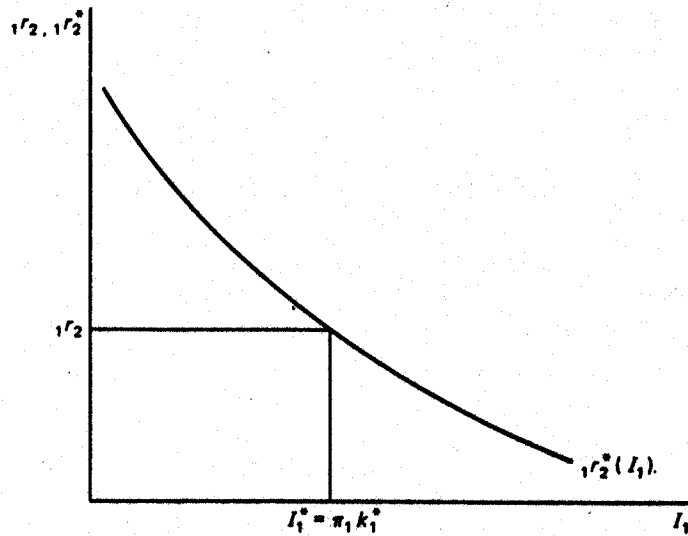


Figure 3.7a Optimal Level of Investment

$a + bz$. In what follows we make use of both variables in discussing the optimality conditions, depending on the context.

I.D. Extension to the Case of Many Different Machines

We can also extend the analysis in fairly straightforward if tedious fashion to allow for the existence of many different types of machines, including of course, the same machine at different ages and hence different efficiencies. Suppose, to be concrete, that there are two machines, type i and type j . Then the full problem corresponding to Equation (3.1) would be, assuming $k_{i0} = k_{j0} = 0$,

$$\max_{q_1, l_1, k_{i1}, k_{j1}} V_1 = -\pi_{i1}k_{i1} - \pi_{j1}k_{j1} + \frac{p_2q_1 - w_2l_1 + \pi_{i2}k_{i1} + \pi_{j2}k_{j1}}{1 + r_2} \quad (3.9)$$

subject to

$$q_1 = F_1(l_1, k_{i1}, k_{j1}). \quad (3.10)$$

Once again, we can highlight the decision for either machine, say, type i , by regrouping and treating type j as merely another cooperating factor along with l_1 . In particular, corresponding to Equation (3.4), we should have

$$\max_{q_1, l_1, k_{i1}, k_{j1}} V_1 = \max_{k_{j1}} \left[\max_{q_1, l_1, k_{i1}} \left[\frac{p_2q_1 - w_2l_1 - k_{j1}[\pi_{j2} + \pi_{j1} - \pi_{j2}]}{1 + r_2} - \lambda_1(q_1 - F_1(l_1, k_{i1}, k_{j1})) - k_{i1} \left[\frac{\pi_{i2} + \pi_{i1} - \pi_{i2}}{1 + r_2} \right] \right] \right]. \quad (3.11)$$

Despite the added complexity, the only substantive change, insofar as graphical representations, such as Figures 3.4 to 3.6, are concerned, is that the net cash flow X_{i2} of the machine i , and the measures derived from it, such as $x'_2(k_{i1})$ or ${}_1r_{i2}^*$, must be taken as net not only of wages but also of the net costs of the services of machine j , adjusted optimally to the specified value of k_{i1} . The last phrase is particularly important; in representing the decision problem for a single machine in graphical form, we must not forget that the optimum investment decision for the machine is embedded in a much larger optimizing problem.

When the concern is not with a single machine in isolation but with total investment in all machines, that is, with the total capital budget and its allocation, the desired representation can be obtained by simple aggregation of relations like those in Figure 3.7a, as shown for a two-machine example in Figure 3.7b. The scenario amounts to first specifying a value for the total investment budget I_1 and then solving the complete production and investment problem (3.9) and (3.10) subject to an additional provisional constraint of the form

$$\pi_{i1}k_{i1} + \pi_{j1}k_{j1} = I_1. \quad (3.12)$$

The optimality conditions require that the total budget of I_1 dollars be allocated between the two machines so that the marginal one-period yields are the same for each machine. Otherwise it would clearly pay to transfer funds from one machine to the other. This common value for the marginal internal yields at any level of total investment ${}_1r_2^*$, which, of course, is not the same as ${}_1r_2$ except at the optimum value of I_1 , gives the required values for the vertical axes for each of the machines. And the combined investment

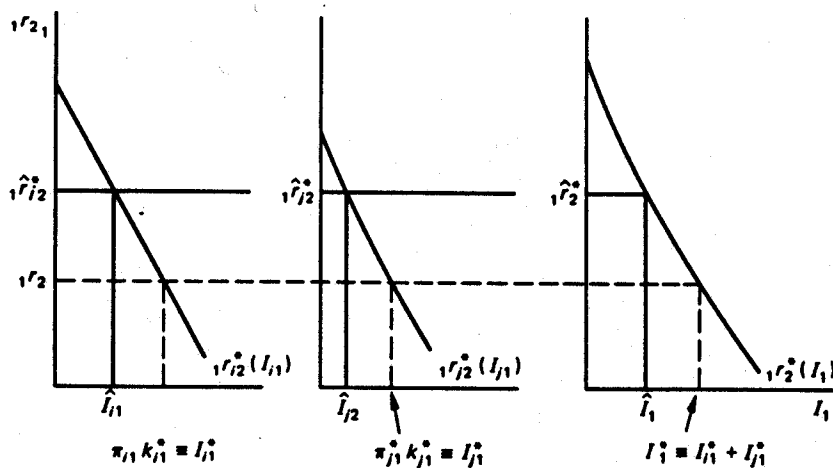


Figure 3.7b Capital Budget in a Two-Machine Case

function for the firm is obtained simply by adding the two curves together horizontally, as shown in the third panel of the figure.

I.E. The Investment Decision and the Transformation Curve

The preceding analysis permits us at long last to justify the use of transformation functions $T(K_1, K_2) = 0$ in contexts other than those of the "seed corn" variety. In Chapter 2, it will be recalled, the transformation function was defined as an implicit function showing the maximum consumption possibilities, or "withdrawals," in period 2, K_2 , that could be obtained for any given level of consumption withdrawals in period 1, K_1 . Thus, to go back to a single-machine context, points on the transformation function are really nothing more than solutions, for different values of K_1 , to a series of problems of the type

$$\max_{q_1, l_1} K_2 = \max_{q_1, l_1} [p_2 q_1 - w_2 l_1] + \pi_2 k_1$$

subject to the constraints

$$K_1 = K'_1 - I_1 = K'_1 - \pi_1 k_1,$$

$$q_1 = F_1(k_1, l_1),$$

where K'_1 represents the resources withdrawable at the start of period 1.

In Figure 3.7c, for example, we start by picking a trial value for K_1 , such as \hat{K}_1 . As drawn, the firm is assumed to have no withdrawable resources in the first period, so that the "consumption possibility" would actually be negative, that is, $\hat{K}_1 = -\hat{I}_1$. At the ruling price for machines, π_1 , the

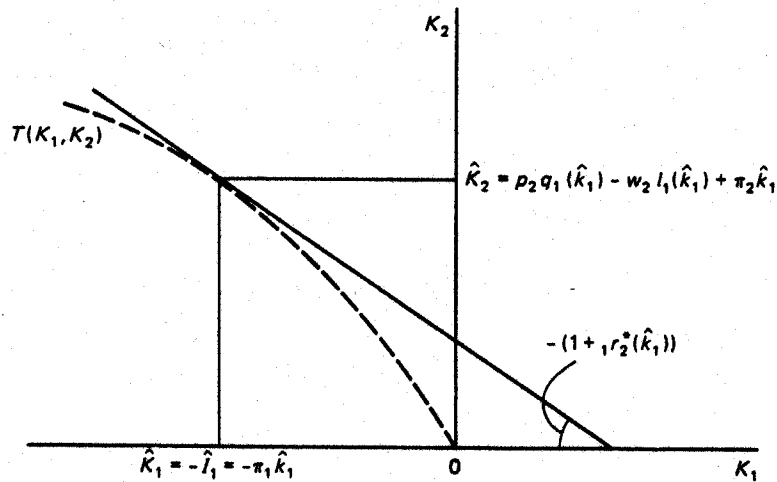


Figure 3.7c Investment Decision and Transformation Curve

investment of I_1 dollars translates into $I_1/\pi_1 = \hat{k}_1$ units of physical capital service to which are then added cooperating labor services $l_1(\hat{k}_1)$ until the value of the marginal product of labor equals the wage rate. The sales proceeds from the output $q_1(\hat{k}_1)$ so determined plus the resale value of the \hat{k}_1 machines minus the wage bill then constitutes the consumption possibility \hat{K}_2 for the next period; that is, $\hat{K}_2 = p_2 q_1(\hat{k}_1) - w_2 l_1(\hat{k}_1) + \pi_2 \hat{k}_1 = X_2(\hat{k}_1) + \pi_2 \hat{k}_1$, and so on, for every other value of K_1 until the whole function is traced out.

It should also be easy to see that the slope of the transformation function at any point such as \hat{K}_1, \hat{K}_2 is indeed really nothing more than $-[1 + r_2^*(\hat{k}_1)]$, where $r_2^*(\hat{k}_1)$ is the one-period internal yield for a capital stock of size \hat{k}_1 . We leave the derivation as an exercise for the reader.

I.F. Extension to More than Two Time Periods

I.F.1. The case of perfect markets for capital goods

The extension of the model to allow for many periods is a simple matter as long as we maintain the assumption that capital goods can be bought or sold in a perfect market; that is, at the beginning of any period t the firm can buy or sell as many units as it likes of the capital good at a known fixed price π_t . In such a case, the selection of the optimal capital stock in each period, in the many-machine as well as the single-machine case, turns out to be merely a sequence of independent, two-period decisions of exactly the kind that we have been discussing; that is, just as the optimal k_1^* in the one-machine case was specified in terms of π_1, π_2, p_2, w_2 , and r_2 , so the optimal k_t^* depends solely on $\pi_t, \pi_{t+1}, p_{t+1}, w_{t+1}$, and r_{t+1} independently of all prices and decisions of all other periods. And similarly for the many-machine version.

This one-period horizon or "myopic" property of the model may seem somewhat paradoxical at first glance, because the machines are durable after all and must in general be considered as producing net returns for the firm over many more than a single period. Remember, however, that as long as the markets for machines are perfect, the firm can adjust its capital at will. Hence, no matter what stock it may happen to have acquired in any period t , it can sell out the whole lot at the start of $t + 1$ and then buy back at the same price whatever quantity that it deems appropriate to carry through $t + 1$. Because no costs are incurred in such a rollover, the firm can clearly never be worse off as a result of proceeding in this one step at a time manner.

Another way of making the same point is to note that, for the purposes of the investment at t , the assumption that the market for capital goods is perfect implies that the value to the firm at $t + 1$ of any unit of capital equipment is precisely π_{t+1} , its resale price at this time. For if the value of

any unit to the firm were less than its resale price, the firm could and would sell it at $t + 1$. On the other hand, if the value to the firm at $t + 1$ of any unit of capital were more than the then ruling market price, precisely this excess of value over cost could be obtained by purchasing the unit at $t + 1$. Thus to justify purchasing the unit at t , the present value at t of (1) the cash flows that it generates at $t + 1$ plus (2) the resale price at $t + 1$ must be at least as great as π_t , the price of the unit at period t .

More formally, let

$$\frac{X_{t+1}(k_t)}{1 + r_{t+1}} = \max_{a_t, l_t} \left[\frac{p_{t+1}q_t - w_{t+1}l_t}{1 + r_{t+1}} - \lambda_t(q_t - F_t(l_t, k_t)) \right];$$

that is, $X_{t+1}(k_t)$ is the maximum of revenues minus labor costs, that is, net cash flow, at $t + 1$ consistent with k_t units of capital goods held at t . And let $x'_{t+1}(k_t)$, as before, be the corresponding marginal net cash flow. Then the preceding argument indicates that the decision criterion for t , expressed in present value form, involves pushing the stock of capital goods to the point at which

$$\pi_t = \frac{x'_{t+1}(k_t) + \pi_{t+1}}{1 + r_{t+1}},$$

which is the exact counterpart of the two-period rule (3.6).

1.F.2. The case of fixed capital

We have postponed until this late point any consideration of cases in which capital goods could not be freely bought and sold at fixed prices in a perfect market. This delay is not to suggest, of course, that such cases are of less importance empirically. Quite the contrary—although situations involving active secondhand markets are much more frequent than the distribution of emphasis in the standard capital budgeting literature would lead one to suspect. Our strategy rather reflects the fact that the representation of the optimality conditions is inevitably a good deal more complex without the perfect market assumption and the exposition can be at least somewhat facilitated by using the results in the simpler case for comparison and contrast.

To see the nature of the difficulties, consider the case of a firm using a single "fixed" capital good that cannot be resold at any time once installed, and for further simplicity suppose that the installation must take place at the start of the first period. Then the firm's decision problem for a three-period case can be stated as

$$\max_{q_1, q_2; l_1, l_2; k_1} V_1 = -\pi_1 k_1 + \frac{p_2 q_1 - w_2 l_1}{1 + r_2} + \frac{p_3 q_2 - w_3 l_2}{(1 + r_2)(1 + r_3)}$$

subject to

$$q_1 = F_1(l_1, k_1),$$

$$q_2 = F_2(l_2, k_1).$$

The major and decisive point of contrast between this problem and the corresponding one in Section I.A. lies in the second of the two production function constraints. The production possibilities in the second period now depend on the capital installation decisions made during the previous period. We can, in other words, no longer break the decision into a series of independent, one-period problems but must consider simultaneously the decisions in all periods.

Fortunately, however, our concern is with the representation rather than the computation of the optimality conditions, and much of the earlier graphical representation can in fact be salvaged with only minor reinterpretations, at least as long as we continue to assume the existence of a finite planning horizon.¹³ Consider first the decision problem as it appears to the firm at the start of the second period. The stock of capital having been determined in the previous period, call it \hat{k}_1 , the firm's problem is then simply that of determining the optimal amounts of cooperating labor to employ and total output to produce during the second period. Formally,

$$V_2(\hat{k}_1) = \max_{q_2, l_2} \left[\frac{p_2 q_2 - w_2 l_2}{1 + r_2} - \lambda (q_2 - F_2(l_2, \hat{k}_1)) \right],$$

where $V_2(\hat{k}_1)$ is the maximum market value of the firm at period 2 consistent with holding k_1 units of the capital good at period 1. This will be recognized, of course, as exactly the kind of one-period problem that we have already considered, and the optimizing condition is the by now familiar equality between the value of the marginal product of labor and the wage rate. By repeating this calculation for every possible value of the initial capital stock, we can obtain the function $V_2(k_1)$, which shows the value as of the start of period 2 of any amount of capital carried over and optimally employed during this period. In terms of the earlier representation, this function is essentially the same as the function $X_2(k_1)/(1 + r_2)$ in Figure 3.4, and the curve marginal to $V_2(k_1)$, to be denoted by $v'_2(k_1)$, corresponds to the function $x'(k_1)/(1 + r_2)$ in Figure 3.5.

To complete the solution, and in particular to obtain the optimal value of k_1 , we take one step backward in time to the start of the first period. The

¹³ An infinite horizon would impose no insuperable mathematical difficulties, but it would unduly complicate the exposition at this stage of the proceedings. We shall consider some (small-scale) problems involving infinite horizons later in the applications discussion in Section II.

problem can then be stated formally as

$$\max_{q_1, l_1, k_1} V_1 = \max_{k_1} \left[\max_{q_1, l_1} \left[\frac{p_2 q_1 - w_2 l_1}{1 + r_2} - \lambda (q_1 - F_1(l_1, k_1)) \right] - k_1 \left(\pi_1 - \frac{V_2(k_1)/k_1}{1 + r_2} \right) \right]. \quad (3.13)$$

Note that Equation (3.13) is a one-period problem exactly the same as that defined earlier in Equation (3.4) except that the terminal value of the stock of capital is given by the function $V_2(k_1)$ rather than $\pi_2 k_1$. Or to put it another way, Equation (3.4) is merely that special case of Equation (3.13) for which, by virtue of the perfect market assumption, $V_2(k_1)/k_1$, the average terminal value per unit, is exactly equal to the market price per unit of capital, π_2 , for all values of k_1 .

Once this relation between the two problems is understood, it is easy to see how the optimality conditions must be restated in the fixed capital case.¹³ In particular, the optimum stock of capital k_1 is found at the point at which the present value of the marginal cash flow, with other factors and total output optimally adjusted, plus the marginal terminal value of the stock, assuming an optimally adjusted cash flow to this stock in period 3, is exactly equal to the initial purchase price per unit, that is, the point for which

$$\pi_1 = \frac{x'_2(k_1) + v'_2(k_1)}{1 + r_2}. \quad (3.14)$$

In equivalent one-period rate of return form the criterion is thus

$$r_2 = \frac{x'_2(k_1)}{\pi_1} + \frac{v'_2(k_1) - \pi_1}{\pi_1} \equiv r_2^*(k_1); \quad (3.15)$$

that is, the capital stock is increased until the optimal marginal cash flow yield plus the optimal marginal rate of appreciation, or minus the marginal rate of depreciation of the stock, exactly equals the first-period rate of interest. And in cost of capital services form, it is

$$x'_2(k_1) = r_2 \pi_1 + (\pi_1 - v'_2(k_1)); \quad (3.16)$$

that is, expand the stock until the marginal cash flow in the coming period exactly equals the interest on the capital invested in the marginal unit plus the marginal appreciation or depreciation.

¹³ It is also easy to see how the model can be generalized to the case of many machines along the lines of Section I.D.

Note finally that we can, if we choose, string together present value conditions, such as Equation (3.14), and obtain an exactly similar criterion running in terms of the cash flows and interest rates over the entire life span of the equipment. In particular, for the example we have

$$v_2'(k_1) = \frac{x_3'(k_1)}{1 + 2r_3}$$

Substituting this expression into Equation (3.14), we obtain as our criterion

$$\pi_1 = \frac{x_2'(k_1)}{1 + 1r_2} + \frac{x_3'(k_1)}{(1 + 1r_2)(1 + 2r_3)} \quad (3.17)$$

Or in words, expand the stock of capital until the present value of the marginal cash flows, with all other cooperating factors and total output optimally adjusted in each period, exactly equals the initial purchase price per unit of capital. And the extension of the criterion to the general n -period case is rather obvious.

With the development of this generalized present value rule, we have completed the derivation of the optimality conditions for investment decisions by the firm. We may turn now to consider some of the problems involved in the application of such rules to capital budgeting.

II. INVESTMENT DECISIONS AND CAPITAL BUDGETING

II.A. Problems in the Application of the Present Value Criterion

The generalized present or market value rule with which we concluded the previous section may seem so simple and familiar that no further discussion is required. Actually, however, it is by no means always immediately clear how it is to be implemented in specific concrete applications even quite apart from the (serious) problems of determining the values of the optimally adjusted marginal cash flows or allowing for the inevitable uncertainties.¹⁴

II.A.1. Marginal versus average present values

To see the nature of one of the difficulties, consider the following typical problem in capital budgeting: A firm is considering building a large cen-

¹⁴ We remind readers that we are not here attempting to provide a comprehensive or well-rounded survey of the field but merely focusing on a limited number of topics illustrating some of the important concepts and results that we have worked with up to this point. Students not previously exposed to the literature on capital budgeting will find a good introductory survey in the first part of the text by Harold Bierman, Jr., and Seymour Smidt, *The Capital Budgeting Decision*, 2d ed. New York: The Macmillan Co., 1966. An extensive bibliography is also provided.

tralized warehouse to facilitate the distribution of its finished products. Research studies of the savings permitted by the new facility have produced reliable or at least accepted figures for the projected cash flow in each year of the life of the facility. The present value of these flows at the accepted interest rate, or rates if they differ over time, is, say, \$4 million. The construction cost of the facility is \$3 million. Should the firm undertake the project?

At first glance, the answer would seem to be yes, obviously. But a closer look at our criterion shows that on the basis only of the evidence presented, the most we can really say is maybe, for our decision rule was not stated in terms of "projects" but of units of capital, or, equivalently, units of investment in the capital good.¹⁵ There may, perhaps, be cases in which a choice really is of the all or nothing, single-project variety. But such would surely not be the case in general for a warehouse that could be built in many different sizes and with varying degrees of durability. Clearly, then, we cannot make our final decision until all these opportunities have been taken into account.

A more appropriate way of structuring the decision in such cases is given in Table 1. (For simplicity, we assume that the rate of interest is a constant 5 percent per period, and we bypass the question of durability for the moment by assuming that the facilities have infinite life). Column 1 is some physical measure of warehouse size, such as millions of cubic feet of usable storage space. Column 2 is an estimate of the total cost of constructing a facility of this size, and column 3 is the estimate of the corresponding total annual cash flow. Column 4 is the present value of these cash flows. As can readily be seen, the present value exceeds the construction cost for all the listed warehouse sizes.

TABLE 1

(Columns 2 to 7 are in units of millions of dollars)

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Size, Millions of Cu Ft	Cost of Construction	Total Annual Cash Flow	Present Value of Cash Flow at 5 Percent	Marginal Construction Cost	Marginal Present Value of Cash Flow	Net Present Value, Column 4 - Column 2
1	1	0.100	2.0	1	2.0	1.0
2	2.1	0.185	3.7	1.1	1.7	1.6
3	3.3	0.240	4.8	1.2	1.1	1.5
4	4.6	0.290	5.8	1.3	1	1.2
5	6.1	0.315	6.3	1.5	0.5	0.2

¹⁵ See footnote 2, page 110.

Given the data in Table 1, we have two entirely equivalent ways of determining the optimal size. On the one hand, we can compute the series marginal to those in columns 2 and 4, as in the representation of the optimality conditions shown in Figure 3.5. In the present case, for example, we can then see immediately from columns 5 and 6 that a warehouse of a capacity of 1 million cu ft would not be an optimal decision, for if we increased the design by 1 million cu ft, we should increase the value of the facility by \$1.7 million at an added cost of only \$1.1 million. By the same token, sizes of 4 and 5 million cu ft are also not optimal, because the cost of added capacity exceeds its present value at both levels. As for the size that should be chosen, we can say that the optimal size under the circumstances is greater than 1 million and less than 3 million cu ft, and if necessary, we could narrow this range further by using a less coarse grid. (Note that without looking at warehouse sizes between 1 and 2 million cu ft, we cannot conclude that the optimal size is 2 million cu ft or greater, although the change from 1 to 2 million increases net present value.)

We could also have arrived at the identical conclusion, not by looking for the marginal equality, but simply by searching directly for the maximum difference between the value of the cash flow and the cost of acquiring it (in the spirit of the representation in Figure 3.4). These differences, essentially what we have earlier called the "net present value" or "goodwill" of an investment, are listed in column 7.¹⁶

II.A.2. Comparing investments with different lives

The previous example focused on differences in capacity or scale at different levels of investment. There are, however, other ways in which the capital intensity of a project can be varied, such as by changing its dura-

¹⁶ Note that the criterion of choice is that of the maximum absolute difference between columns 4 and 2 and not the maximum relative difference or the maximum ratio of value to cost. This ratio, which has come to be called the "present value index," would clearly give the wrong decision in the present case, because its maximum value is for a warehouse of size 1 (or perhaps even smaller).

It may strike some readers as paradoxical that a warehouse of size 2, which has a cost of \$2.1 million, a present value of \$3.7 million, and a net present value index of 1.8 (or equivalently, a net "rate of return" of about 80 percent) could ever be preferable to a warehouse of size 1, which not only has a higher present value index (and rate of return) but requires a much smaller initial outlay on the part of the firm. Remember, however, that under the assumption of perfect capital markets the firm faces no limitations on obtaining funds to finance profitable projects, so that the absolute size of investments, whether mutually exclusive or not, is completely irrelevant. It is true, of course, that not all firms are in so fortunate a position and may well face, or feel that they face, limits on their total investment budgets. But if so, there would still be no reason to believe that the present value index would be a reliable guide to decision making or even that such an index has any particular economic meaning. We return to a variant of this "rationing" problem in another context below.

bility. Suppose, for example, that we can provide 2 million cu ft of storage space either by a very solidly built structure with a life of 40 years, a gross present value over this period of \$2 million, and costing \$1.5 million or by a much lighter structure with a life of only 30 years, a gross present value of \$1.5 million, and costing \$1.2 million. Despite the fact that the more durable structure has the larger net present value, we cannot conclude that it is the preferable choice in this case. We first must make some allowance for the cash flows that would be earned during the years between 30 and 40 if the firm adopted the less durable alternative.

Precisely how to make this allowance is a problem to which our theory, as such, can make no contribution. In practice and in standard textbook discussions the assumption most typically made about the cash flow in the overlap intervals is that a facility, at the end of its economic life, is replaced by one of exactly the same kind. In the present case, the 30-year warehouse would thus be presumed to be replaced by another of 30 years' life, which would in turn open a gap of 20 years as compared with the single, 40-year warehouse. But this too is presumed to be replaced by a similar unit, and by repeating this process, we eventually match exactly the lengths of the two chains at 120 years, the lowest common multiple of 30 and 40. Applying the net present value rule over this interval, entering the cost of the future replacements as negative cash flows in the appropriate years, then in principle yields the correct choice as to durability.

Fortunately, however, there is a much simpler way to perform the calculation, and this, curiously enough, is to assume that the chain of like replacements actually extends all the way to infinity. Although this greatly increases the presumed time span, it also permits the use of the simple and compact perpetuity results noted earlier in Chapter 1. In particular, and assuming, of course, a given constant rate of interest in all periods, we first compute the net present value of any one link in either of the two chains, call them $V(30)$ and $V(40)$, respectively. The net present value to infinity of a warehouse replaced every 30 years is then¹⁷

$$\begin{aligned} V(30, \infty) &= V(30) \left(1 + \frac{1}{(1+r)^{30}} + \frac{1}{((1+r)^{30})^2} + \dots \right) \\ &= V(30) \frac{1}{1 - 1/(1+r)^{30}} = V(30) \left[\frac{(1+r)^{30}}{(1+r)^{30} - 1} \right] \quad (3.18) \end{aligned}$$

and similarly for the other chain.

The same result can also be expressed in another way, particularly popular among engineers. Instead of working with the stock values $V(30, \infty)$

¹⁷ Note again that for $0 < x < 1$

$$(1 + x + x^2 + x^3 + \dots + \dots) = \frac{1}{1-x}$$

and $V(40, \infty)$, we can, if we choose, work with their flow equivalents, that is, with the flows, in perpetuity, that have the same present values as the stocks. These flows, which we denote by $\bar{X}(30)$ and $\bar{X}(40)$ and which are typically referred to in the engineering literature as "time-adjusted average cash flows," can be obtained simply by multiplying the present value of each infinite chain by the rate of interest, so that we have

$$\bar{X}(30) = rV(30, \infty) = V(30) \left[\frac{r(1+r)^{30}}{(1+r)^{30} - 1} \right] \quad (3.19)$$

and similarly for $\bar{X}(40)$. The basic decision rule restated in terms of these (time-adjusted) average cash flows is to pick the higher of the two, which will always and necessarily be the one with the higher net present value.

Note that the term in brackets in the expression for the time-adjusted average cash flow is the so-called "capital recovery factor," which was derived and discussed in a somewhat different connection in Chapter 1. Thus, for example, $\bar{X}(30)$ is the uniform cash flow per year for 30 years that has a present value of $V(30)$. When it is assumed that the 30-year warehouse is always replaced by one of exactly the same type, $\bar{X}(30)$ is also the cash flow per period in perpetuity that has a present value of $V(30, \infty)$.¹⁸

II.B. Replacement Policies and the Optimal Economic Life of Equipment

In the previous illustrations, we took the length of life of our various alternative warehouses as fixed. But the life of any given capital good is rarely, if ever, determined solely by purely physical or technological considerations. The decision to terminate the life of a machine, by sale or scrapping, and to replace it with a different and presumably younger one is an economic choice. In principle, therefore, our criterion should apply in this case as well. But, once again, a certain amount of care has to be exercised to establish a meaningful comparison of the alternatives.

In particular, consider the following hypothetical problem. A firm now has an old warehouse that it proposes to demolish and replace with a new one of identical capacity whose economic life, for the moment, is assumed to be 30 years. (We later consider how this economic life itself is determined.) For simplicity, it is assumed that the gross cash flow before expenses of the two warehouses would be exactly the same. But of course the operating expenses of a new building—repairs, maintenance, heating, and so on—would be less than those of an old one. The decision would then seem to hinge on whether the present value of the "savings" from the new ware-

¹⁸ Like most of the other present value concepts the time-adjusted cash flow has its counterpart in rate of return form. In particular if I is the amount of initial investment, then $rV(n, \infty)/I$ is the time-adjusted perpetual rate of return on the investment. The concept has found its main usefulness in valuation theory (see the $r^*(t)$ in Chap. 2, Sec. III.E.3).

house, that is, the difference between the operating costs of the old and the new buildings, is larger than the cost of construction of the new building plus the net cost of demolition.

The question remains, however, over precisely what time period these savings are to be computed. The previous discussion of the matching of streams for present value calculations might seem to suggest that the appropriate time span was 30 years, that is, the given full life of the new building. But it would be madness to project operating costs for the old warehouse over the next 30 years, for the firm would surely not keep the old building for so long. Hence, most of the savings being discounted would be purely "phantom savings" that would never in fact be realized.¹⁹

To obtain a proper comparison, we must allow somehow for the fact that the old warehouse will eventually be replaced, which may seem to involve us in a circularity, because this is the decision that we are currently trying to make. The paradox is resolved, however, as soon as we rephrase our initial question so that the decision problem is stated, not as whether to replace the old warehouse, but when to do so.

In particular, proceeding systematically, let us first compare replacing the warehouse at the beginning of this year (say, year $t - 1$ in the life of the warehouse) versus next year (year t in its life). If we replace it immediately, the present value of all future costs of warehousing—operating costs, construction costs and demolition costs—can be expressed as $V(30, \infty)$, assuming, as in the previous section, that the new warehouse, in its turn, will be replaced by an identical one with identical costs after 30 years and similarly thereafter. If, on the other hand, we delay the replacement until next year, the present value of our costs to infinity will be $W(t)/(1 + r) + V(30, \infty)/(1 + r)$, where $W(t)$ represents the costs of operating the old warehouse during the current year, assumed to be paid at the beginning of year t . Our decision rule can then be stated simply as follows: Replace now in preference to next year if

$$\frac{W(t)}{1 + r} + \frac{V(30, \infty)}{1 + r} \geq V(30, \infty).^{20} \quad (3.20)$$

¹⁹ We owe the expressive term "phantom savings" as well as many others that have become standard in replacement theory to George Terborgh whose book *Dynamic Equipment Policy*, New York: McGraw-Hill Book Company, 1949, still remains among the best available treatments of the practical as well as theoretical sides of replacement decisions.

²⁰ Stating the criterion in terms of the conditions for minimum cost rather than maximum net present value is permissible mathematically under our assumption that the gross revenue of the warehouses is the same; that is, if R stands for gross revenue, W for operating cost, and Z for the set of decision variables, then

$$\max_z (R - W) = \max_z R + \max_z (-W) = \max_z R - \min_z W = k - \min_z W$$

Note that our calculations could end at this point, although we have explicitly compared only two of the many possible replacement strategies. For example, suppose that Equation (3.20) holds, so that costs of an immediate replacement are less than those of a delay until next year. Then as long as the costs of operating the old warehouse increase with time, we know that it is actually optimal to replace now; that is, if Equation (3.20) holds and $W(t+1) > W(t)$, it is easy to show that

$$\frac{W(t+1)}{(1+r)^2} + \frac{V(30, \infty)}{(1+r)^2} + \frac{W(t)}{1+r} > \frac{W(t)}{1+r} + \frac{V(30, \infty)}{(1+r)^2} \geq V(30, \infty),$$

so that replacement now is more profitable than replacement next year, which is in turn more profitable than replacement in 2 years. Applying this reasoning period by period leads to the conclusion that as long as Equation (3.20) holds, the optimal decision is to replace now.

If, on the other hand, the costs of immediate replacement are greater than the cost of another year's operation, our one-period calculation would certainly not tell us when the replacement will occur. But, in this problem, this is something that we do not really need to know. The calculation has solved the immediate action question, and we can safely postpone any further calculation until the future, when we shall in due course face the problem again.

Note, finally, that as in so many cases before, we can express our decision criterion in other entirely equivalent forms that may provide additional insights into the nature of the problem. In particular, a simple rearrangement of the inequality in Equation (3.20) yields the following rule: Replace now if

$$W(t) \geq rV(30, \infty). \quad (3.21)$$

In words, and recalling our discussion in the previous section, replace if the marginal cost of extending the life of the "defender" for one more year is greater than the (time-adjusted) average cost of the "challenger."²¹

in any case in which R is a given constant k . From the economic point of view, however, we must regard this as, strictly speaking, a suboptimal solution, because the difference in cost should in general lead to differences in output, prices, and their product, gross revenue. In practical applications, however, including also most inventory and related operations research models, the gain in simplicity by way of the minimum cost formulation more than compensates for the (normally) minor error of approximation involved.²¹ The rule in this form can easily be extended to allow for any salvage values connected with the defender. If S_1 is the current salvage value and S_2 the salvage value next period, the marginal cost of extending the life of the defender is $W_1 + rS_1 + (S_1 - S_2)$, that is, the sum of the operating cost, the interest that could otherwise have been earned on the salvage value plus the change, presumably decline, in salvage value. The criterion in this form should have a thoroughly familiar look to it.

Stating the criterion in this form helps to show, among other things, how we can complete the analysis by dropping our provisional assumption that we already know in advance the life span of the best challenger, for the replacement time for any challenger itself must be determined by the criterion (3.21). The details of the calculation for a given challenger are shown in Figure 3.8. The broken line $W(t)$ represents the actual operating costs that would be incurred each period, costs that are assumed to rise steadily as the warehouse ages. The curve $\bar{W}(t)$, lying below it, shows the time-adjusted average annual operating costs. As in Equation (3.19) the time-adjusted average annual operating cost is the present value of the operating costs per period up to period t converted to a uniform flow equivalent by way of the appropriate capital recovery factor. These costs too rise with time but less rapidly than $W(t)$ because of both the "discounting" built into the time adjustment and the fact that $\bar{W}(t)$ is a composite of the operating costs for all periods up to t . The curve $\bar{K}(t)$ represents the time-adjusted average annual "capital cost," that is, the construction cost for the warehouse converted to a flow equivalent by the capital recovery factor. This component of total cost falls steadily with age, as in any other case in which a "fixed" cost is spread over a larger number of time units. The curve $\bar{TC}(t)$, which is just the sum of $\bar{K}(t)$ and $\bar{W}(t)$, is the time-adjusted average combined cost, or average total cost, for short, and it will clearly be U-shaped, given the assumed behavior of $\bar{W}(t)$ and

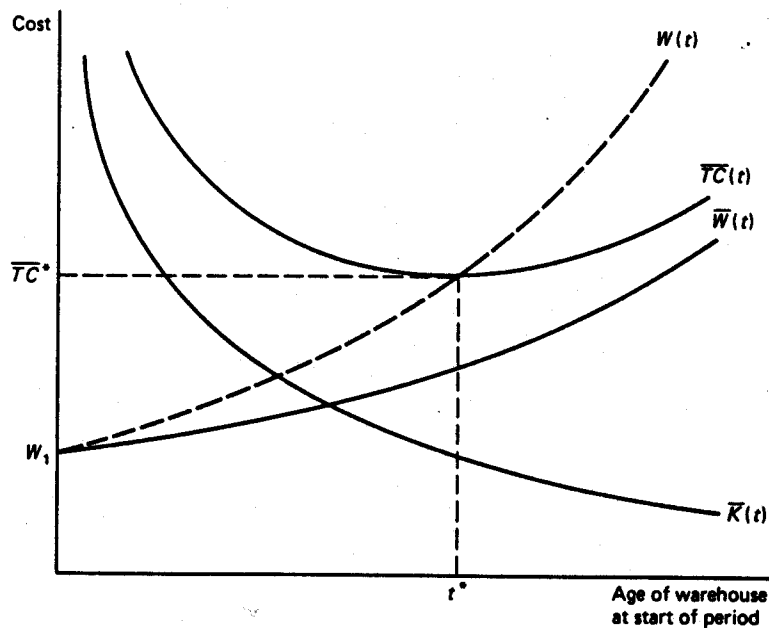


Figure 3.8 Optimal Length of Life of Warehouse

$\bar{K}(t)$. The optimum economic length of life is then found at that age t^* for which the average total cost is a minimum. And at this age the $W(t)$ curve intersects $\bar{TC}(t)$, because, from Equation (3.21), the fact of replacement at t^* must imply that at this point in time the marginal cost of extending the life is equal to the average cost of replacement with an identical warehouse, and this average cost is just

$$\bar{TC}(t^*) = rV(t^*, \infty).^{22}$$

To determine the optimal type of warehouse, that is, the best challenger, an analysis like that in Figure 3.8 would have to be carried out for each type to determine its optimal economic life and minimum average time-adjusted annual cost. The best challenger would then be the one with the overall minimum average annual cost.

II.C. Maximizing Present Value Subject to Constraints

Still another area in which the application of the present value criterion raises some questions and difficulties is the problem of capital budgeting subject to financial constraints. Several reasonably sophisticated formulations of this problem can now be found in the literature, all bringing to bear on the question some of the tools and concepts of mathematical programming.²³ One such formulation, which we can regard as typical for our purposes, visualizes a set of projects x_j , $j = 1, 2, \dots, n$, that the firm may undertake in any period from now (period 1) up to some finite horizon period T . With each such project, there is associated in each period a cash flow c_{jt} that may be negative, as, for example, in the period during which it is purchased; or positive, as during its normal earning life; or zero, implying either that it has not started yet or that its life is already over. Given these cash flows, we can compute a net present value b_j , $j = 1, 2, \dots, n$, for each project by applying the given market rates of interest to the flows in exactly the manner considered in previous sections.

²² From Fig. 3.8 we can also see how the analysis can readily be extended to allow for at least some fairly simple and regular kinds of technological improvement, and hence, by implication, of technological obsolescence. The curve $W(t)$ can be considered the absolute value of the operating cost per year, or we could shift our origin to the point W_1 , the operating costs of the first year, and interpret $W(t)$ as the excess cost above that incurred with a new machine. To the extent that each vintage of new machines incorporates improvements that result in lower operating costs in the first year, and all subsequent years, we can treat this improvement as an additional "obsolescence cost" to be added to $W(t)$ to obtain the complete excess cost of operating with an old rather than the best new machine then available. The rest of the calculation then proceeds as before. For an extensive discussion and illustration see Terborgh, *op. cit.*

²³ One of the best examples, and the one that is serving as the model for our discussion here, is that of H. M. Weingartner, *Mathematical Programming and the Analysis of Capital Budgeting Decisions*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963.

If there were no financial constraints, our problem would then be simply

$$\begin{aligned} & \max \sum_{j=1}^n b_j x_j \\ \text{subject to} & \quad 0 \leq x_j \leq 1, \quad j = 1, 2, \dots, n, \\ & \quad x_j \text{ integer} \end{aligned}$$

which is just another minor variant of our standard "maximize present value" format, the special wrinkle here being the added integer constraint whose function, of course, is to ensure that any project is either rejected completely ($x_j = 0$) or accepted in toto ($x_j = 1$).²⁴

Suppose, however, that we imposed the further restriction that the amount invested in projects in any period could not exceed the available cash throw-off during the period from previously undertaken projects plus a specified amount of outside borrowing B_t , which might be zero. This would constitute an additional set of constraints of the form

$$\sum_{j=1}^n c_{jt} x_j \geq -B_t, \quad t = 1, 2, \dots, T.$$

The optimal x_j for this more tightly constrained problem are clearly different from those of the original problem whenever one or more of the financial constraints is binding. For when such constraints are present, every project now contributes to the total present value in two ways: directly, by way of its b_j term in the maximand, and indirectly, by way of its cash flow terms c_{jt} , which loosen or tighten the financial constraint in a bottleneck period and thus permit or rule out other profitable projects.

But although this procedure certainly yields a solution to the problem of capital budgeting subject to such financial constraints, there are certain reservations that must also be entered about its use in any practical setting, quite apart from such matters as where the numbers are supposed to come from or how to deal with the severe computational difficulties involved in large integer programming problems. For the problem, as formulated, really involves a conceptual inconsistency between the maximand and the constraints. The maximum present value criterion is invoked in investment and other corporate decisions, not as an end itself, but because it can serve

²⁴ The integer constraint also makes it possible to allow for many kinds of dependencies between projects without overly complicating the statement of the problem. For example, to indicate that projects m and n are mutually exclusive, we need add only a constraint of the form $x_m + x_n \leq 1$. To make project r conditional on the undertaking of project s , we impose the restriction $x_r \leq x_s$.

as a surrogate for the best interests of the owners of the firm in certain circumstances. These circumstances, as we have seen, include the existence of a perfect capital market in which firms and individuals can borrow and lend indefinite quantities at the going rate of interest.

Given this rationale for the maximand, what sense can we make of the constraints? If the firm really does face approximately perfect markets, the financial constraints are arbitrary impositions of the management contrary to the best interests of the owners. The solution from their point of view would not be optimal, although it had been formally derived by a "maximization" process. On the other hand, if the constraints were genuine and the firm really faced limitations on outside funds, it is the maximand that would be purely arbitrary. For what point is there in discounting a stream with market interest rates that do not represent actual opportunities for the firm in question? The firm might just as well use the rates of a foreign country or any other set of numbers plucked out of the air.²⁵

To say that programming models of this kind suffer from logical difficulties is not to suggest, of course, that such mathematical programming approaches have no value in capital budgeting problems. Setting aside the obvious problems of data collection, allowance for uncertainty, and so on, the programming models may well have an important role to play when the relevant constraints are not financial. For rapidly expanding firms the key limitations are often those of certain specialized kinds of manpower, and the programming approach, which would represent, in effect, an attempt to provide a computationally feasible simplification of one of the fixed capital models considered in Section I, would aim at an optimal allocation of this resource over projects and time periods.

Moreover, even models with financial constraints might have valid uses, provided that these constraints were treated as provisional planning estimates rather than as inviolable policies or even desirable targets; that is, an initial trial run-through of the model with an internal funds restriction could be used to highlight which future periods would be tight and thus provide a basis for an appropriate program of external financing to overcome

²⁵ From an esthetic point of view, "rationing" models, in which the objective is to maximize the terminal value of the firm as of some horizon date rather than its supposed present value, offer some advantages. This formulation does at least maintain consistency between the interest rates in the functional and in the constraints, which can be set up to take account of whatever borrowing and lending opportunities are actually available. Moreover, it is at least possible to imagine special utility functions for the owners for which the resultant policies would be optimal. Examples of this type of formulation can be found in Weingartner, *op. cit.*, and in Charnes, Cooper, and Miller, "An Application of Linear Programming to Financial Budgeting and the Costing of Funds," *Journal of Business*, vol. 32, no. 1 (January 1959).

the binds. In sum, there are more uses for mathematical models in the applied areas of finance than just that of literal, direct, on-line, real-time decision making, and, in the present state of the art, these planning uses, in which a simplified formal model serves mainly to organize and explore some of the grosser implications of policy decisions, are likely to be more important.

II.D. The Rate of Return Criterion: Uses and Abuses

As noted at several points in Section I of this chapter and in Chapter 2, the criterion for optimal investment decisions may be stated either in present value form or in rate of return form. Once again, however, a certain amount of care has to be taken to avoid meaningless comparisons, and this is particularly true when the rate of return computed happens to be that variant known as the "discounted cash flow internal rate of return"—the variant most widely used at least in the standard popular treatments of capital budgeting.

II.D.1. The discounted cash flow rate of return

The discounted cash flow internal rate of return on an investment is defined as that rate of discount for which the net present value of the investment would be exactly zero. Algebraically, this means finding a root ρ^* of the n th-order polynomial

$$v(\rho^*) = \frac{x_1}{(1 + \rho^*)} + \frac{x_2}{(1 + \rho^*)^2} + \dots + \frac{x_n}{(1 + \rho^*)^n} - I_0 = 0, \quad (3.22)$$

where the x_i are the net cash flows per period and I_0 is the amount of any initial outlay. The logic of the calculation is shown in Figure 3.9. First set the discount rate equal to zero and solve Equation (3.22) for $v(0)$, in this case, merely by summing the x_i and subtracting I_0 . Next set the discount rate equal to some small positive amount and again solve Equation (3.22). Assuming the x_i all positive, the net present value so obtained must be lower than $v(0)$, because each x_i element in the summation is multiplied by a number $1/(1 + \rho)^t < 1$. Repeat the process for successively higher values of the discount rate, tracing out the entire curve $v(\rho)$, as in Figure 3.9. The value of the discount rate at which this curve cuts the horizontal axis is the DCF internal rate of return ρ^* .

A glance at Figure 3.9 also helps to make clear the sense in which this rate of return can provide a criterion equivalent to the present value rule. For whenever the actual market rate of interest is less than ρ^* , the net present value of the investment is necessarily positive under our assumptions, and the investment is worth considering. Conversely, if $r > \rho^*$, the

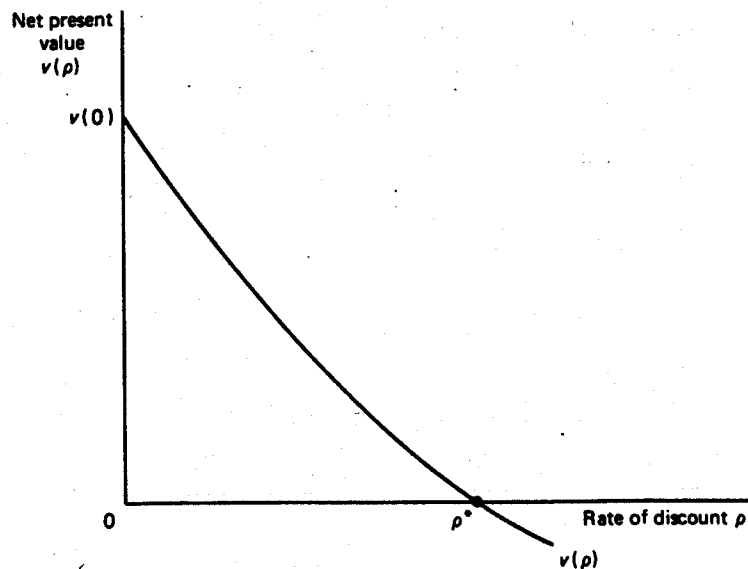


Figure 3.9 Discounted Cash Flow Rate of Return

net present value is negative, and the investment would not meet the present value test.²⁶

In using the DCF rate of return as an alternative to a present value calculation, it should be kept in mind that the equivalence of the two procedures holds only if the rate of interest is assumed to be the same in all periods. If not, there is no single rate of interest to be compared with ρ^* , and the DCF rate of return cannot be used as the basis of an investment decision criterion.²⁷

²⁶ Note that there is no need to refine the calculation further by allowing for the fact that the cash throw-off from the project may have to be reinvested at rates lower than ρ^* . Because the lowest such reinvestment opportunity is always at least r , taking it into account could never reverse the direction of the inequality.

²⁷ This limitation on the use of the DCF rate of return is perhaps not of great practical consequence, because it is rare in actual capital budgeting to use anything but an assumed constant rate of interest or cost of capital. Nevertheless as a reminder of the shortcomings of the DCF rate of return as an investment criterion, consider the following example, involving two infinite streams of cash flows:

$t =$	1	2	3	≥ 3
$t-1r^t$	0.01	0.05	0.10	0.20
Stream 1	0	2.00	4.00	16.00
Stream 2	10.00	10.39	10.54	10.54

When discounted at the market interest rates, both streams have present (=market)

II.D.2. Mutually exclusive investments and multiple rates of return

Although the DCF rate of return thus has its perfectly legitimate and consistent uses as an alternative form of the present value criterion, at least when interest rates are constant in time, there are some types of decisions even then in which its application runs into difficulties. One of the most important of these arises when a choice must be made between mutually exclusive investment opportunities. The nature of the difficulty is shown in Figure 3.10, which shows the $v(\rho)$ curves for two different machines, A and B, being considered as alternatives for the same task.²³ Both machines require an initial outlay of \$25. Machine A has a net revenue stream of \$5 per year at the beginning of each of the following 10 years, but B produces

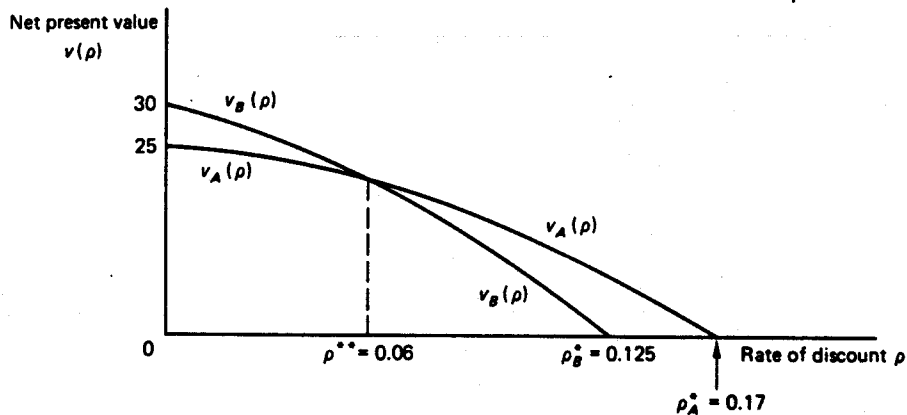


Figure 3.10 Rates of Return for Mutually Exclusive Investment Opportunities

\$1, \$2, \$3, and so on, up to \$10. The DCF internal rate of return on A, ρ_A^* , is about 17 percent, but ρ_B^* is about 12.5 percent.

As for the criterion of choice in such a comparison, it might seem that the rule should be to select the one with the higher rate of return, just as, under the present value criterion, we select the alternative with the higher net present value. In the case pictured, the higher rate of return rule would

value of \$73.89. An investor who bought stream 1 at this price, however, would obtain a DCF rate of return of 15.1 percent; stream 2 would yield only 14.1 percent. The difference in rates of return has absolutely no economic significance, however, because in a perfect capital market a stream of cash flows with a given market value can always be exchanged for any other stream with the same market value.

²³ In an important sense, of course, the mutually exclusive case is the standard one, because, as we have emphasized earlier, any project is really a whole set of investments of different capital intensity in many different dimensions.

The specific example in the text is taken from A. Alchian, "The Rate of Interest, Fisher's Rate of Return over Cost, and Keynes' Internal Rate of Return," *American Economic Review* (December 1955).

lead to the selection of machine A. But, as can readily be seen from the graph, this choice might conflict with that signaled by the present value criterion. For whenever the actual rate of interest is less than ρ^{**} , the net present value of machine B is higher than that of machine A. Which rule should be followed?

The answer is the present value criterion, of course. The reader for whom this conclusion is not obvious by now may perhaps be able to convince himself by asking why the crossover of the two $v(\rho)$ functions occurs at all. He will soon realize that the key to the different rates of decline of the two functions lies in differences in the timing of cash flows for the two projects. Project B has increasing cash flows across its lifetime. Thus at low market interest rates the higher cash flows at the end of its life contribute heavily to its present (\equiv market) value. Project A, on the other hand, has a constant cash inflow throughout its life. It produces larger cash inflows than B in the early years but smaller flows in the later years. When the market interest rate is high, later cash flows have low current market value, so that machine A then has a higher present value than B because of the larger early cash inflows of A.²⁹

II.D.3. A modified rate of return rule for the mutually exclusive case

For the example given, it is relatively easy to define a somewhat more elaborate rate of return criterion that would overcome the difficulties discussed above and guarantee reaching the correct decision with a rate of return rule. In particular, first select the machine with the larger rate of return, in this case, machine A, and compare this rate with the market rate of interest r . If the rate of return is less than r , reject both alternatives; if it is greater than r , accept the machine provisionally as the defender. Then compute the rate of return of the challenger, in this case, machine B, over the defender, that is, compute the rate of return on the differences in the cash flow and outlay streams period by period. The reader should convince himself that, in Figure 3.10, this would be the rate ρ^{**} . The last step in the rule would then be to compare ρ^{**} with the rate of interest; if $\rho^{**} > r$, accept the challenger, and if $\rho^{**} < r$, accept the defender.³⁰

²⁹ Any reader still unconvinced that the present value criterion is always the correct one should turn back to Chap. 2, Sec. II, in which the maximize present value rule was shown to be an implication of maximum utility for the firm's owners under the perfect capital market assumption.

³⁰ It is important not to neglect the first step, because ρ^{**} , which is often called the "relative" rate of return, may be positive and greater than r , although the values of ρ^* , or "absolute" rate of return, may be negative. In such a case, ρ^{**} would merely serve to indicate which of the two alternatives would produce the smaller loss.

To see that ρ^{**} is indeed the rate of return of the challenger over the defender, note

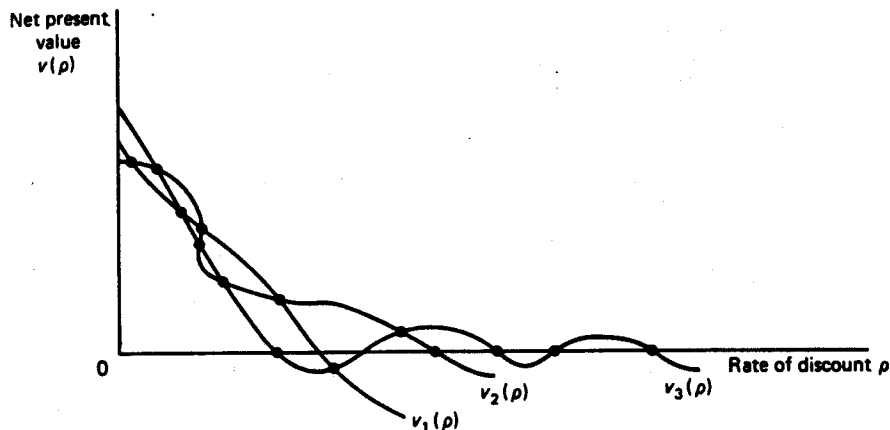


Figure 3.11 Multiple Rates of Return in the Case of Many Alternative Opportunities

Although a simple rule could easily be stated for this special case, the difficulties of statement mount rapidly when we consider more alternatives and alternatives whose $v(\rho)$ curves are less nicely behaved. Consider, for example, the mess shown in Figure 3.11. Note first that machine 3, the one with the highest absolute rate of return, actually has, not one, but five distinct rates of return or crossings of the $v_3(\rho)$ curve with the horizontal axis. There is nothing really remarkable in this, because the basic defining equation (3.21) is, after all, an n th-order polynomial, and from Descartes's rule of signs we can say only that there can be no more such distinct positive roots, in this case, DCF rates of return, than there are changes of sign of the coefficients, that is, the x_i . As a practical matter, moreover, net negative cash flows in some periods are perfectly sensible, the classical illustration being that of periodic pumpings of water or gas into oil wells to restore the pressure and increase the rate of oil flow.²¹ A less exotic case would be that of periodic major overhauls of equipment or major replacements, for remember, in making the comparison between alternatives, we must match the length of the streams.

that the present value at the discount rate ρ of the differences between the cash flow and outlay streams for the two machines is just the difference between their total net present values at the rate ρ , that is, the difference between the two present value curves in Fig. 3.10. The rate of return of the challenger over the defender is the rate of discount for which the present value of the differences between the cash flow and initial outlay stream is 0. But this is just the rate of discount for which $v_B(\rho) - v_A(\rho) = 0$, that is, the rate ρ^{**} .

²¹ A simple example is provided by Ezra Solomon, "The Arithmetic of Capital Budgeting Decisions," *Journal of Business* (April 1956). Suppose that a project has an immediate cash outflow of \$1600, a net inflow next period of \$10,000, and a final net outflow two periods from now of \$10,000. As the reader can check, the project has two DCF rates of return, 25 and 400 percent.

As for the crossings between the curves, we have no less than nine of which one has been drawn to occur below the horizontal axis. Clearly, however, with a little patience, one could devise a routine to search through the maze and present a conditional "decision tree" running solely in terms of rates of return for matching challengers against defenders so as to determine the true champion. In this sense, then, it would still be true that the rate of return is equivalent to the present value criterion. But the rule would be so long and complex as to destroy the rationale for having used a rate of return calculation in the first place.

The popularity of the rate of return in applied capital budgeting is, after all, at least in part a reflection of its simplicity. Many people, apparently, find it more intuitively appealing to compare two rates than two present values. Much more important in its popularity, however, has been the fact that the rate of return approach seems to permit a convenient administrative separation of the decision process. The technicians or engineers are supposed to estimate the cash flows and compute rates of return on the various investment projects that they send upstairs for approval. The finance staff in the treasurer's office then independently computes the appropriate rate of interest or cost of capital for the firm. And top management maintains control over capital spending by putting the two sets of estimates together and selecting the specific projects to be implemented.²² But the previous example makes clear that this neat separation runs into difficulties when many mutually exclusive alternative possibilities for accomplishing a given task are available, as they normally are.

II.E. The Rate of Return, Rankings of Projects, and Financial Constraints

The seeming ability to define a DCF rate of return independently of the rate of interest or cost of capital has given the DCF ρ^* tremendous appeal for those concerned with investment planning for firms, particularly small ones, who face or feel they face limits on the funds available for investment.²³ Some treatises recommend, for example, that the firm compute the DCF ρ^* for each proposed investment project and rank all the projects in decreasing order of rate of return. The optimal "cutoff point" is then supposed to be found simply by marching down this schedule, accepting lower and lower rated projects until the allotted total investment budget has been exhausted.

²² This description, of course, is a statement of the ideal and not of how the system actually works in practice.

²³ For simplicity, we speak here only of a presumed absolute limitation on the funds to be employed. The same strictures apply with equal force, however, to analyses in which the firm is pictured as facing a cost of capital function that rises with the amount invested in any period.

Despite the fact that this procedure may bear some superficial resemblance to the process shown in Figure 3.7, it should be amply clear by now that it is really only a caricature of the underlying theory of investment. For one thing, we have seen that the calculations of investment productivity cannot proceed in any meaningful way without some prior assumptions as to the relevant market rate of interest. Before one could even assemble the list of projects to be ranked, the rate of interest would have had to be invoked in choosing between mutually exclusive alternatives. Moreover, even if it were somehow possible to find a list of independent, all or nothing projects whose DCF rates could be computed without reference to an interest rate, the very notion of a ranking of such projects by rate of return runs counter to the basic rationale of the procedure. The DCF rate of return is an interesting number for such projects only because a rate higher than the market rate of interest signals that the project is worth doing. It provides, as it were, a simple go versus no-go criterion and says nothing about the relative desirability of projects to the firm.³⁴ If the objective is to maximize the net wealth of the owners, all such projects that flash the go signal must be undertaken regardless of how strong the flash.³⁵

These objections to rate of return ranking are perhaps obvious enough in the case of perfect capital markets, in which the simple accept or reject decision is all that we need; but the idea of ranking by rate of return becomes doubly nonsensical in the presence of financial constraints. It is more than just a matter of an inconsistency between the maximand and the constraints of the kind discussed earlier in connection with the constrained version of the present value rule. In this case, it was at least clear what was being maximized, if not why, and the indirect contributions of the projects by way of their effects on the constraints were also taken into account. If, however, we merely march down a rate of return ranking until we run out of money without taking into account the indirect effects on later constraints, we have something like the disembodied smile of the Cheshire cat—a maximizing condition, but one not derived from or even related to any known maximand.

II.F. Conclusion

In this chapter we began by deriving the criteria for optimal investment decisions by the firm under certainty and relating these criteria both to the standard theory of the firm and to our earlier treatment of wealth allocation

³⁴ For example, the reader should find it easy to construct examples where the rankings of projects according to their DCF rates of return is not the same as the rankings provided by their net present values. (See Fig. 3.10 when $r < 0.06$.)

³⁵ Recall the similar remarks with respect to the "present value index" in Sec. II.A.1.

and security valuation in perfect capital markets. We then went on not so much to consider how these criteria could or should be applied in practice—real-world decision problems normally involving uncertainty in crucial ways that our simple apparatus cannot encompass—but rather to call attention to some of the pitfalls and inconsistencies to which mechanical application of standard capital budgeting procedures can all too easily lead. With this, the task begun in Chapter 1—the analysis of the role of capital markets in the allocation of wealth with respect to time—has been completed, and we now turn to the task of extending the analysis to allow for uncertainty.

REFERENCES

The great burst of interest in recent years in both the theoretical and applied aspects of capital budgeting was triggered off by a number of books that appeared in the late 1940s and early 1950s, of which the most influential have been

1. Dean, Joel, *Capital Budgeting*. New York: Columbia University Press, 1951.
2. Lutz, Frederick, and Vera Lutz, *The Theory of Investment of the Firm*. Princeton, N.J.: Princeton University Press, 1951.
3. Terborgh, George, *Dynamic Equipment Policy*. New York: McGraw-Hill Book Company, 1949.

Also extremely important in making capital budgeting one of the main themes in the field of finance was the collection of readings that appeared a decade or so later under the editorship of Ezra Solomon:

4. Solomon, Ezra (ed.), *The Management of Corporate Capital*. New York: The Free Press, 1959.

Most of the articles cited in Section II can be found in this collection.

A good introductory survey of capital budgeting along standard lines with many drill problems can be found in the first section of

5. Bierman, Harold, Jr., and Seymour Smidt, *The Capital Budgeting Decision*, 2d ed. New York: The Macmillan Company, 1966.

One of the best treatments of the capital rationing problem remains that of

6. Weingartner, H. Martin, *Mathematical Programming and the Analysis of Capital Budgeting Problems*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963.



UNCERTAINTY MODELS

The second part of the book is concerned with the same general topics as the first; that is, we discuss models for decision making by individuals and firms and the way these decisions at the micro level interact to determine the nature of equilibrium in the capital market. Moreover, as in the first part of the book, almost all our work is carried out in the context of a perfect capital market. Simply stated, our goal now is to study how the analyses and conclusions of the preceding chapters can be adjusted to allow for the effects of uncertainty.

Thus the major results with respect to the nature of optimal production-financing decisions by firms obtained for a certainty world are the two separation principles in Chapter 2. Specifically, (1) given its production decisions, a firm's financing decisions are a matter of indifference to its security holders, so that production and financing decisions are separable; and (2) optimal production decisions for a firm simply involve adherence to the market value rule, so that such decisions are independent of the details of owner tastes. In Chapter 4, we see that these separation or independence principles hold also in a world of uncertainty, and in fact their validity requires only the assumed existence of a perfect capital market.

In a certainty world, and given a perfect capital market, implementation of the market value rule is a simple matter, because market values are always determined by applying known interest rates to cash flows whose values are also known for certain. In a world of uncertainty, however, a perfect capital market is not in itself a sufficient basis for a model of price determination. To derive meaningful, that is, testable, statements about how the current price of a probability distribution of future payoff is determined in the market, we first need a more detailed specification of investor tastes. Thus, in Chapter 5, we present such a theory of choice under uncertainty: the expected utility model.

But we find that the expected utility model does not carry us all the way to the goal. The model simply says that investors rank probability distributions on the basis of the expected or average value of utility for each distribution. To develop first a testable

theory of investor decision making and then one of market equilibrium, the dimensionality of the investor's decision problem must somehow be reduced; that is, we must somehow determine either additional restrictions on investor tastes—for example, some assumption about the form of utility functions—or assumptions about common properties of probability distributions of returns—for example, all are normal—that allow us to describe the different alternatives available to an investor in terms of a finite number of parameters—for example, means and variances of return distributions—that are common to all alternatives.

Thus Chapter 6 is concerned with two-parameter models of investor decision making under uncertainty. In brief, these models assume that investors are risk-averse, a term whose precise meaning is obtained from the expected utility model, and that they find it possible to summarize any probability distribution of return in terms of two parameters, the mean of the distribution, and some measure of the dispersion of possible return values like the variance or standard deviation.

Chapter 7 is then concerned with the implications of the two-parameter model for the characteristics of capital market equilibrium; that is, given a market in which investors make decisions according to a two-parameter model, what can we say about the determination of the market values of securities and firms? Specifically, what is the appropriate way to measure the risk of an asset, and what kind of relationships might be expected between risk and return?

Finally, Chapters 6 and 7 concentrate entirely on two-period models. In Chapter 8 the analysis is extended to the multiperiod case. As we shall see, this is in keeping with our practice of proceeding from the simple to the more difficult.