Comment:

Optimal Economic Decision-Making

for

Engineers, Accountants and Economists

BERNARD GOLDBERG

University of Houston

Industrial Engineering Department

This is the first revision of a work entitled:

OPTIMAL ECONOMIC DECISION-MAKING

for

ENGINEERS, ACCOUNTANTS and ECONOMISTS

PREFACE

Economic decision-making is designed to achieve the greatest profitability and cost-effectiveness in private as well as public economic organizations. Because profitability and cost-effectiveness are not well-defined, conventional methods of decision-making use multiple criteria, none of which are provably optimal. Consequently, different methods of decision-making are used by various organizations in the same industry, and also by different parts of an organization that compete for the same capital constraint. In contrast, a single profitability criterion is proposed for optimal economic decision-making that is independent of the structure of the organization. For this reason, the definitions of single and multiple profitability criteria are systematically confronted with observations in order to explain their similarities and differences in a more concrete fashion.

The book provides a method of selecting engineering and financial alternatives whose cash flows maximize the net present-value of an organization for a given capital constraint. All parts of an organization would use a single profitability criterion for selecting (1) the best way of doing each project, (2) the best projects to do and (3) which alternatives should be funded. Because forecasting errors are inherent in engineering and financial data, the proof of optimality assumes each alternative has accurate cash flow descriptions that account for engineering risks of input costs, marketing risks of output revenues and external costs of borrowing money after income taxes. The results of the convex-envelope proof of optimal economic decision-making presented here are applicable to industrial firms, financial institutions, government agencies and nonprofit organizations.

Conventional methods of economic decision-making evaluate engineering and financial risks by determining net present-values with internal discount rates such as minimum-attractive-rates-of-return (MARR) and weighted-average-costs-of-capital (WACC). Accounting for engineering and financial risks with internal discount rates can seriously distort the net present-values of cash flow forecasts derived from available engineering and financial information. Moreover, the proof of optimal economic decision-making cannot be verified by observations from future financial statements of an organization unless discount rates reflect the after-tax cost of interest incurred for borrowing money.

Single and multiple-variable optimization problems are treated in the eighth and ninth chapters of the book. Linear programming (LP) translates input and output variables that are subject to inequality and equality constraints into linear relationships from which some objective quantity can be optimized. The simplex algorithm for solving LP problems was developed by George B. Dantzig in 1947 and it is currently applied to many problems in industry and finance. The simplex algorithm is only one of a number of linear and nonlinear techniques used to solve constrained optimization problems.

Engineering production functions (EPF) provide more complete and accurate solutions of small-scale input/output production problems than LP analysis could provide. The EPF analysis uses a single budget constraint in optimizing input and output variables on the basis of their market prices. Consequently, the need for multiple input constraints used in LP analysis is eliminated. The EPF input and output variables may have increasing, decreasing as well as constant returns to scale that LP problems always require.

In writing this book, the author tried to provide a common language for engineers, accountants and economists in the study of optimal economic decision-making. A unified approach is provided for both private and public economic organizations. Because this book is designed for wide audiences of students and practitioners in engineering, accounting and finance, the main text assumes the reader has only a basic background in analytic geometry and calculus. A liberal use of appendices is made where more rigorous mathematical treatments are required. The basic theory of economic decision-making is developed in an intuitively logical fashion with simple computations and graphical displays. Practical examples have been selected from typical engineering and financial alternatives in public as well as private organizations where income taxes are involved.

Present values of future cash flows discounted at constant or varying interest rates are readily determined with spreadsheet computer programs that are also useful for solving many other engineering and financial problems. Therefore, several appendices of this book are designed to help readers utilize popular spreadsheet programs on PC and Macintosh computers. Constant interest rate tables with discrete and continuous compounding of cash flows with uniform, arithmetic and geometric gradients are given using American Society for Engineering Education symbols. Commissioners 1980 Standard Ordinary Mortality Tables with 4% Commutation Columns are given for females and males.

The author wishes to express appreciation to students and colleagues in the Industrial Engineering Department at the University of Houston whose feedback made this book possible. In particular, I would like to thank Drs. Navaratna S. Rajaram and Charles E. Donaghey, and Messrs. Chaitanya Nagappa, Lujun Liu and Joel Jones for many helpful suggestions. My thanks are also due to coworkers at Shell Development Company, especially Drs. Arthur S. Ginzbarg and Harold J. Vinegar, whose varied theoretical and practical backgrounds provided a blend of experiences indispensable to the evolution of my views of optimal economic decision-making. I am deeply grateful to my wife, Charlotte, for her infinite patience, meaningful criticisms and endless encouragement, and to my son, Dr. David A. Goldberg, for his help in publishing the book.

August 2001 Bernard Goldberg

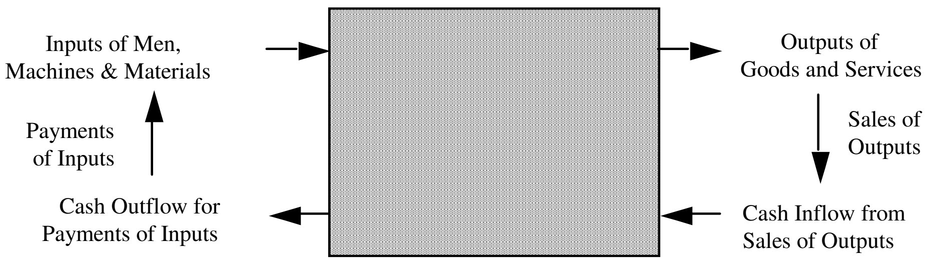
Table of Contents

Chapter One - Introduction

Section 1.1 - Description of Economic Organizations

Every economic organization can be viewed as a black box. Inputs of men, machines and materials enter at the upper left where they are converted into outputs of goods and services that exit at the upper right. Sales of outputs provide cash inflows in the lower right part of the box. The cash outflow at the lower left pays for the inputs of men, machines and materials at the upper left. Thus, in the upper part of the box, inputs are transformed to outputs. In the lower part of the box, output sales provide cash payments for input costs in a countercurrent direction as shown in Figure 1.1.1 below.

Figure 1.1.1 - Black Box of an Economic Organization



Every economic organization can be subdivided into a set of projects whose inputs can be changed into outputs in various ways, called *engineering alternatives*. Sales of the outputs can provide for the costs of the inputs in various ways, called *financial alternatives*. The goal of optimal economic decision-making is to determine the best way of doing each project and the best projects to do in order to maximize the net present-value added to an economic organization under the constraint of its present-value input costs.

Problems of selecting best alternatives in various parts of a company was recognized relatively recently in the literature on economic decision-making. The earliest book on engineering economy was Wellington's "The Economic Theory of the Location of Railways" (Wiley, New York, 1887) in which he observed engineers who chose railway locations with almost a total disregard of how their decisions would affect the costs and revenues of the railway as a whole. The railway engineers' narrow concern for laying track was, as Wellington wrote, "... the neglect of larger questions of where to build and when to build, and whether to build them at all, has in it something at once astounding and discouraging."

Economic decision-making deals with criteria for determining the best engineering and financial alternatives. In major projects such as building a railroad, engineers, accountants and economists often use different criteria for selecting the best alternatives. Consequently, there are conflicting assessments of the objective of an organization and how to achieve those objectives.

This book provides a provably optimal method of selecting engineering and financial alternatives which maximize the net present-value added to an economic organization for a given capital constraint. Because forecasting errors are inseparable from the decision-making process, the proof of optimality is based on deterministic cash flow forecasts of each alternative which takes into account its input engineering risks and output marketing risks.

Net present-values are discounted on the basis of external costs of borrowing money. This approach enables results of economic decision-making to be verified in future financial statements of any organization. Consequently, the proposed method of economic decision-making is applicable to industrial firms, financial institutions, government agencies and nonprofit organizations. This promises to be the best guide for promoting economic progress in the real world where none of the economic decision-making criteria are provably optimal.

Section 1.2 - History of Economic Doctrines

Attempts to optimize economic efficiency has a long history. In every age, goods have been produced, exchanged and distributed by apportioning labor and material resources according to the wants and needs of society. When labor and material resources are joined randomly, inadequate supplies and scarcity result. Man is then forced to organize labor and material resources in ways which are regarded as best for the good of the people.

The main body of economic science began about 200 years ago when problems of organizing the production, exchange and distribution of goods and services became widespread. Medieval communities were headed by feudal lords who protected their communities in war while peasants produced food and shelter. In feudal societies, problems of economic organization were regulated by dictates of the lords and their emissaries.

When feudal systems became obsolete, new forms of economic organization emerged to produce, exchange and distribute goods and services. Merchants of the 16th and 17th centuries were primarily profit oriented. They sought to increase revenues relative to costs in order to accumulate as much monetary wealth as possible. Similarly, a mercantilist nation sought a favorable balance of trade by exporting more than it imports, levying tariffs and quotas to restrict imports, providing subsides to encourage exports, and extending colonial empires to maintain plentiful supplies of cheap raw materials and labor. The world, at present, is frequently guided by mercantilist doctrines.

The results of mercantilism were both good and bad. It greatly expanded the exchange of foreign goods and services which benefitted trading countries. Local economic systems which didn't trade had to be self sufficient without exchanging exportable surpluses for needed imports. But in pursuit of trade, mercantilism fostered colonial imperialism, now a discredited policy. Also, when money accumulates faster than the needs of circulation, the resulting inflation could destabilize the production and distribution of goods and services.

As new production techniques developed during the Industrial Revolution of the 18th and 19th centuries, critics of mercantilism arose. Most notably, *Adam Smith* (1723-1790), took basic issue with the mercantilist doctrine that wealth is derived from monetary accumulations and export surpluses. Smith argued that wealth increases according to the skill and efficiency of labor. His illuminating example was a pin factory which greatly increased productivity by the division of labor and the specialization of manufacturing tools.

Smith reasoned that productivity which springs forth from the division of labor and industrial specialization is distributed to consumers through a medium of exchange. Unlike animals, humans increase their productivity from a natural tendency to barter and trade. In this vein, Smith wrote "Nobody ever saw a dog make a fair and deliberate exchange of one bone for another bone with another dog". Smith believed the net benefits of industrialization would be distributed equitably among the contributing input factors after taking into account the costs of the specialized tools and work places of labor.

The productivity insights of Adam Smith were overshadowed by his accounting for the distribution of input costs and output revenues Money, he said, is merely a medium of exchange - what circulates is not only money one way but also goods the other way. The wealth of nations increases only as more goods are exchanged. If the amount of exchanged goods was constant, price increases would only increase the quantity of money in circulation.

Smith also argued the converse was true, that price increases would only increase the money in circulation without exchanging more goods. However, exchanging more goods may affect both prices and the quantity of money in circulation. Nevertheless, Smith thought industrialization would occur at constant prices and an "invisible hand" would guide people in a market economy to exchange more goods and achieve greater collective wealth.

As the division of labor and industrial specialization progressed, there was more room to question how much land, labor and capital contributed to the market value of goods. During the period from 1811 to 1815 just after the Napoleonic wars, major debates in England centered on high prices of corn (a general term used for all types of grain). It was commonly thought that high prices of corn were caused by high rents of agricultural lands. Industrialists who had to pay their workers high wages because of high corn prices, exhorted landowners to lower rents of agricultural lands.

Others - notably *David Ricardo* (1772-1823) - claimed high rents were not the cause of high corn prices, but conversely that high corn prices were the cause of high rents. He said corn prices were high because of shortages during the Napoleonic wars. Owing to high corn prices, the profitability of producing corn was high and there was keen competition among farmers for land on which to grow corn. The competition for growing corn drove up the rents of agricultural lands. If tariffs on imported corn were reduced, supplies of domestic corn would increase and bring down prices of corn and lower rents of agricultural lands.

The Ricardian theory of rental income is based on a competitive equilibrium of *marginal profitability and productivity*. Ricardo observed that 'fertile' lands earn higher rents because they are more productive and/or closer to the market. As more land is used keeping all other inputs fixed, the relative increments of output ultimately decrease as implied by the law of diminishing returns. Consequently, decreasing marginal productivity of high-rent agricultural lands would make domestic corn less competitive with corn raised in low-rent foreign lands if protective tariffs of local farmers and landowners were lowered.

Ricardo's theory of rent also applies to wage rates of labor and interest rates of capital. If labor input was increased keeping all other inputs fixed, a marginal increase of productivity would result. The difference between the *value* of the *marginal productivity* of labor and its wage rates is defined as the *marginal profitability* of labor. If the marginal profitability of labor was positive, the quest for profits would tend to increase the demand for labor and its wage rates. But if the marginal profitability of labor was negative, it would tend to decrease the demand for labor and its wage rates. Thus, in a competitive equilibrium, the marginal profitability and productivity of labor would tend to be equal.

Karl Marx (1818-1883) differed from Ricardo's explanation of wage determination. Marx's theory of wages stems from his economic interpretation of history. According to Marx, human labor produces more than its needs for subsistence. Wages are less than the value of labors' output because of exploitation by capitalists who owned the means of production. In order for labor to receive the full fruits of its productivity, Marx believed it was necessary for the state to own all "means of production". Marx thought that production and distribution under state socialism would soon evolve into a communist state guided by his famous motto "From each according to his ability, and to each according to his needs".

Marx claimed that ruthless market competition in capitalist systems forced wages to a subsistence level below the marginal productivity of labor. Since capitalists receive the total value of the output while paying labor only subsistence wages, the unearned surplus (or profit) accumulated by capitalists rightfully belongs to labor. By failing to distribute the full value of labors' output, more is produced than markets can absorb. Until the surplus is consumed or dissipated, unemployment and depression spreads among the working classes. Marx blamed private ownership of capital for the disparity between wages and labor productivity which causes the malfunctioning of capitalist economic systems.

Marx's theory of capital value assumes the quantity of labor input is the sole source of output value. Capital is generally defined as output used to increase future productivity. Marx has three capital categories. *Variable capital* is defined as the output used for the subsistence needs of workers in production. *Constant capital* is defined as the output used to replace machines and materials consumed in production. *Capitalist surplus* (or profit) is defined as the difference between total output value and quantities of labor input in variable and constant capitals as shown in Figure 1.2.1 (see L. L. Pasinetti, "Lectures on the Theory of Production", Columbia University Press, New York 1977).

Figure 1.2.1 - Marxist Theory of Capital Value

Value of Total Output	=	Quantity of Labor in Variable Capital	+	Quantity of Labor in Constant Capital	+	Quantity of Labor in Capitalist Surplus

The Marxist equation is an accounting system in which the value of total output is measured in units of quantities of labor input. However, quantities of labor input are ambiguous units of measurement. The same quantities of skilled and unskilled labor input could have vastly different output values. Busy work may entail large quantities of labor input with little or no output value. Conversely, valuable oil fields have been discovered from exploration efforts using relatively little quantities of labor input. The output value of a hammer bears no relation to the quantity of labor input used in making the hammer. Because of these ambiguities, Marx's theory of capital is not a useful system of accounting.

Marx sought statistical evidence to substantiate his theory of capital value by comparing the earnings of labor and capital. Using long chains of statistical inference, he correlated subsistence wages and capitalist profits which he interpreted as a causal relationship. He claimed wages could exceed subsistence levels of workers only if privately-owned capital was expropriated and capital ownership resided entirely in a state monopoly. The state would then act as custodian for all the people and labor would be remunerated according to the value of its productivity rather than the essential needs for its subsistence. This would clear the market of all surpluses as in a barter economy, and unemployment and depressions of capitalist economic systems would be eliminated.

Marx's theory of capital value came during the industrial revolution in Europe when market competition from the productivity of capital pressured wages to a subsistence level. During that period, capital had returns which probably exceeded its marginal productivity, but the marginal productivity of capital could not be zero as Marx argued (see Section 9.5). State ownership of capital is no guarantee that labor would receive the value of its marginal productivity. Economic decision-making is blind to who owns the capital. As Chinese Communist leader Deng Xiaoping said, "It doesn't matter whether a cat is black or white as long as it catches mice". In practice, wages in both capitalist and socialist systems are based on the productivity of labor (Adam Smith), on wages paid elsewhere (David Ricardo) and on the subsistence needs of labor (Karl Marx).

Section 1.3 - Financial and Managerial Accounting

The ideas of great economists stem not only from general observations, but also from bookkeeping and accounting records. Economics and accounting are inseparable and neither can stand alone. Hence, besides a history of economic doctrines, we need to examine the role of financial and managerial accounting in economic decision-making.

Bookkeeping and accounting practices evolved from the need *to record* exchanges of goods and services without having to remember every transaction. Bookkeeping consists of posting all transactions of an organization in an orderly chronological record. The role of an accountant is *to classify* those transactions into categories that form an integrated system of accounts in order *to summarize* the financial position of an organization in periodic reports.

Single-entry bookkeeping was commonplace in the feudal manors of England. Because of the intricate subdivision of service duties, each party would report by single entry each activity for which it was responsible and from which its responsibilities were removed. The classification of each charge and discharge of responsibilities in a single account provided a simple method of reviewing the performance of all parties in feudal manors.

Double-entry bookkeeping was first described by a Franciscan monk, *Luca Paciolo*, in his book, "Summa de Arithmetica, Geometria, Proportioni et Proportionalita" [1494], which is a landmark in financial accounting history. It was this work which was largely responsible for introducing Italian double-entry bookkeeping to Western Europe. The fact that Italian bookkeeping supplemented many single-entry bookkeeping systems in the 16th century indicates the wide applicability of the methods described by Paciolo. From the 17th century to the present, Italian methods of bookkeeping were modified, but the basic procedures and fundamental ideas of Paciolo were preserved.

Financial accounting is a system of reporting data of past decisions in the form of *income statements* and *balance sheets*. The audiences of financial accounting are largely investors, creditors, banking institutions and government agencies who need timely reports in uniform *accounting periods*. Because financial accounting has a major influence on outsiders assessment of an organization, it is considered an *external* form of accounting.

The income equation depicted in Figure 1.3.1 summarizes revenues and expenses during an accounting period which are reported in a single-entry bookkeeping format. Revenues and expenses during the accounting period of an income statement are summed on an “as is’ basis to provide historical data that are grounded in fact. The certification of financial statements by an independent certified public accountant (CPA) in accordance with “generally accepted accounting principles” is perhaps the best available assurance of the accuracy of the data presented.

Figure 1.3.1 - Income Equation of Single-Entry Income Statements

Revenues - Expenses = Profits or Losses

The difference between revenues and expenses during the accounting period of an income statement is defined as the *profits* or *losses* (i.e., net income) of an organization. Because revenues and expenses in an accounting period may come from different projects, the profits or losses often consist of revenues caused by past expenses and expenses intended for future revenues. A *matching principle* may be used for time shifting expenses to accounting periods that match revenues for which expenses were incurred.

Taxes are major expenses of private enterprises and a basic source of government revenues. Property taxes are included in operating costs of private enterprises along with material, labor and overhead expenses. Sales taxes, if any, are usually included with the costs of purchasing equipment and materials. A primary source of government revenues is income taxation of private enterprises. For income tax purposes, the original cost of capital equipment lasting more than one year cannot be fully expensed when it is acquired. Instead, the original cost is spread over a time schedule, called *depreciation allowances*, which are deductible as expenses against net operating incomes of private enterprises in subsequent accounting periods to determine their *taxable incomes* upon which income taxes are levied.

Income statements express outcomes of past decisions, but they do not explain the alternatives from which past decisions were made. In particular, cost-increasing and cost-decreasing alternatives of each project are not analyzed in in relation to the cost of borrowing money and capital constraints. Overall costs of borrowed money and income taxes are aggregated for all projects in each income statement. This leaves little or no room for analyzing how decisions were made from past project alternatives.

Balance sheets appear in double-entry bookkeeping format at the beginning and end points of each income statement in order to document how the financial position of an organization changed during an accounting period. Each balance sheet has two-way classifications of assets that are tied to entries in adjoining income statements. This synchronizes balance sheets and income statements which is essential for the purposes of checking and auditing the financial accounting system.

The accounting equation of a double-entry balance sheet is shown in Figure 1.3.2. *Assets* are things of monetary value an organization possesses. *Liabilities* and *owners' equity* are asset values claimed by creditors and owners, respectively. The accounting equation always equates assets to liabilities plus owners' equity because the balance comes from a two-way classification of the same assets. More specifically, assets are first classified in descending order of liquidity such as cash, accounts receivable, inventory, equipment, buildings and land. Assets are classified again on the liabilities side of balance sheet according to claims of creditors and owners.

Figure 1.3.2 - Accounting Equation of Double-Entry Balance Sheets

Assets = Liabilities + Owners' Equity

In capitalist economies, claims of creditors and owners are fundamentally different. The investment of creditors is called *debt* for which there is a formal assurance of a return *on* their investment, called *interest*, and a return *of* their investment, called *principal*. If both interest and principal are not paid, a company may be sued or liquidated by its creditors. The fixed investment in a company is called *equity capital* whose return upon the sale of ownership rights is called *capital gains or losses*. Current returns *on* equity capital, called *dividends*, may be distributed to owners after interest and principal payments on the debt.

Both debt and equity are posted on the liability side of the balance sheet which is concerned with who owns the capital, partly for auditing purposes and partly for enabling equity investors to evaluate their share of the assets. Leased equipment or buildings are owned by *lessors* who rent the use of their property to a *lessee* for a period of time stated in a lease agreement. Since the lessee uses but does not own leased capital, it is listed as *contingency liability footnotes* in *off-the-balance-sheet bookkeeping* of the lessee. Rental payments for leased capital are expenses which appear in the lessee's income statements.

Banks and financial institutions invert assets and liabilities on their balance sheets because money is a medium of exchange rather than an object of ownership. The money for bank loans comes primarily from deposits which are treated as liabilities because depositors receive interest when they relinquish control of their money to the bank. Bank loans to borrowers are treated as assets because interest earned from those loans is a major source of bank revenues. Thus, double-entry bookkeeping is applicable to balance sheets of many different types of economic organizations.

In financial accounting, depreciation allowances are systematic expense allocations of the original cost of an asset during its lifetime which may be charged against the income of an organization in order to determine its income tax liability.. In the course of time, the undepreciated portion of the original cost is listed as the asset value in the balance sheet. The *market value* of capital assets at any time may differ appreciably from undepreciated asset values in balance sheets because it is *cost*, not *value*, that is apportioned in depreciation allowances.

Managerial accounting, also called *cost accounting*, deals with optimizing economic decision-making. It is an *internal* language of an organization for improving operations and making new investments. What you do at present are *ongoing alternatives* of past decisions which may be improved by altering input costs, output revenues, lifespans and financing. Cash flows of each alternative must be forecast as time streams of causally connected inputs and outputs over a number of accounting periods. Assuming accurate forecasts, cash flows of optimally selected alternatives could be partitioned into income statements of future accounting periods in order to check the accuracy of the cash flow forecasts.

In order to select the best alternatives, managerial accounting requires cash flow definitions of alternatives that are independent of whether or not the alternative is selected. For this reason, the cash flow forecast of an *engineering alternative* is defined as the **change** of input costs and output revenues of the organization as a whole if the alternative is **accepted** as opposed to the input costs and output revenues if the alternative is **rejected**. A project may be carried out as an *ongoing alternative* or by other alternatives, but at most one project alternative can be funded in the final budget at any time.

Besides engineering alternatives, cash flow forecasts may have *financial alternatives* of debt or equity financing or leasing. If an alternative is **accepted**, the organization as a whole has less money to lend, or more money to borrow at market rates of interest. Conversely, if an alternative is **rejected**, there is more money to lend, or less money to borrow at market rates of interest. In either case, financial requirements should be evaluated against market rates of interest. It is shown in Chapters Two and Seven that the net present value of the cash flow of an alternative discounted at the **cost of borrowing money** is independent of the relative amounts of debt or equity financing.

Income taxes not only reduce the net present value of each alternative, but they also reduce costs of borrowing money by deducting interest expenses from taxable incomes of private enterprises. Consequently, net present values of **after-tax cash flows** should be discounted by **after-tax costs of borrowing money** as shown in Chapter Seven. It is commonly thought discount rates for determining net present values should be based on *investor opportunity costs* and *risks of inflation*. However, such discount rates confound the issues of discounting and cash flow forecasting. For best use of available engineering and financial information, discount rates should only reflect costs of borrowing money at market rates of interest. The cash flow forecasts of each alternative should take into separate account the engineering risks of input costs and marketing risks of output revenues.

Section 1.4 - Definition of Optimal Economic Decision-Making

Even that which is true can be proved - Oscar Wilde

Optimal economic decision-making is defined as selecting engineering and financial alternatives whose net present-value of input costs and output revenues is a maximum for a given capital constraint of present-value input costs of an organization. The proof of optimal economic decision-making presented here requires accurate forecasts of input costs and output revenues of every alternative as well as after-tax costs of borrowing money used for discounting cash flows to their present values. Accurate forecasts are needed for a proof of optimal economic decision-making that can be validated mathematically and verified in financial accounting statements of an organization.

The accuracy of forecasting is a major problem in both the theory and practice of economic decision-making. The failure to make best use of available engineering and financial information in forecasting can often be traced to ambiguities in the process of economic decision-making. Therefore, the development of a proof of optimal economic decision-making which is presented here under conditions of economic certainty promises to make substantial progress in both economic forecasting and decision-making procedures. The proposed method of economic decision-making is not only applicable to different parts of an organization, but also to different types of taxable and tax-exempt organizations such as industrial firms, financial institutions, government agencies and nonprofit organizations.

A major problem of economic decision-making is to measure the contribution of each alternative to the objective of an organization at the margin of its capital constraint. For this purpose, an organization is subdivided into non-overlapping and collectively exhaustive projects, each of which may have a number of mutually exclusive alternatives. We now ask:

1. What is the best way of doing each project?
2. Which are the best projects to do?
3. Which alternatives should be funded?

1. Only differences between mutually exclusive alternatives can decide the best one. *Inclusive 'or'* cash flows are common to all alternatives and cannot determine the best one. *Exclusive 'or'* cash flows that differ between alternatives are deciding factors. If alternatives have the same present-value output revenues or input costs, the best one minimizes present-value input costs or maximizes present-value output revenues. If alternatives differ in present-value output revenues and input costs, other projects may affect which is best.

More specifically, let mutually exclusive alternatives of each project be represented by a bundle of two-dimensional vectors whose components are *present-value input costs* [ΔC] and *present-value output revenues* [ΔR]. All vectors in the bundle of project alternatives have a common initial point placed at a relocatable origin of the [ΔC,ΔR] diagram. The terminal points of the vectors would then lie in the first quadrant of the [ΔC,ΔR] diagram. Thus, each vector bundle in the sample diagrams of Figure 1.4.1 below has distinct terminal points and a common initial point at a relocatable origin of the [ΔC,ΔR] diagram.

Difference $\Delta R - \Delta C \equiv \Delta NPV$ measures the *N̲et P̲resent-V̲alue* or *Absolute Profitability* which a vector adds to the organization as a whole. Projecting the terminal point of a vector at 45° on the ΔR-axis also measures the ΔNPV of the vector. Ratio $\Delta R/\Delta C \equiv Ø$ is the vector's slope which measures the *Capital Efficiency* or *Relative Profitability* of a vector in converting present-value input costs, ΔC, into present-value output revenues, ΔR.

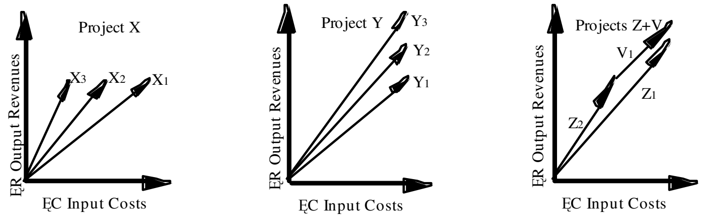


Figure 1.4.1 - Vector representation of mutually exclusive project alternatives.

Mutually exclusive alternatives $X_1[8,9]$, $X_2[7,9]$ and $X_3[6,9]$ of Figure 1.4.1 concern minimizing input costs, ΔC, for given output revenues, $\Delta R=9$. Projections of the X_1, X_2 and X_3 terminal points at 45° on the ΔR-axis show $\Delta NPV\{X_3\}=3$ is the greatest absolute profitability of the three vectors. The slopes of X_1, X_2 and X_3 show $Ø\{X_3\} = 1.5$ is the greatest relative profitability of the three vectors. Thus, X_3 is best for doing project X.

Mutually exclusive alternatives $Y_1[5,7]$, $Y_2[5,8]$ and $Y_3[5,9]$ of Figure 1.4.1 concern maximizing output revenues, ΔR, for given input costs, $\Delta C=5$. Projections of Y_1, Y_2 and Y_3 terminal points at 45° on the ΔR-axis show $\Delta NPV\{Y_3\}=4$ is the greatest absolute profitability of the three vectors. The slopes of Y_1, Y_2 and Y_3 show $Ø\{Y_3\} = 1.8$ is the greatest relative profitability of the three vectors. Thus, Y_3 is the best way of doing project Y.

It is not clear if mutually exclusive alternative $Z_1[8,14]$ or $Z_2[5,10]$ of Figure 1.4.1 is better because neither has both greater absolute and relative profitabilities than the other. Thus, $\Delta NPV\{Z_1\}=6 > \Delta NPV\{Z_2\}=5$, but $Ø\{Z_1\}=1.75 < Ø\{Z_2\}=2$. Alternatives Z_2 could have both greater absolute and relative profitabilities than Z_1 depending on opportunities for investing the difference $\Delta C\{Z_1\text{-}Z_2\} = 8\text{-}5 = 3$ in another project to yield more than $\Delta R\{Z_1\text{-}Z_2\} = 14\text{-}10 = 4$. Vector difference $\{Z_1\text{-}Z_2\}[8\text{-}5,14\text{-}10] = \{Z_1\text{-}Z_2\}[3,4]$ is not an alternative because Z_1 and Z_2 are indivisible. But $\{Z_1\text{-}Z_2\}[3,4]$ is an important criterion for selecting vectors of other projects.

2. Consider project V whose alternative $V_1[3,5]$ has $\Delta NPV\{V_1\}=2$ and $Ø\{V_1\}=1.67$, both of which are less than those of Z_1 and Z_2. But resultant vector $\{Z_2+V_1\}[5+3,10+5] = \{Z_2+V_1\}[8,15]$ has $\Delta NPV\{Z_2+V_1\}=7$ and $Ø\{Z_2+V_1\}=1.875$, both of which are greater than those of Z_1. Therefore, $\{Z_2+V_1\}[8,15]$ is better than $Z_1[8,14]$ because net present-value increases from $\Delta NPV\{Z_1\}=6$ to $\Delta NPV\{Z_2+V_2\}=7$ for the same input costs $\Delta C\{Z_1\}= \Delta C\{Z_2+V_2\}=8$.

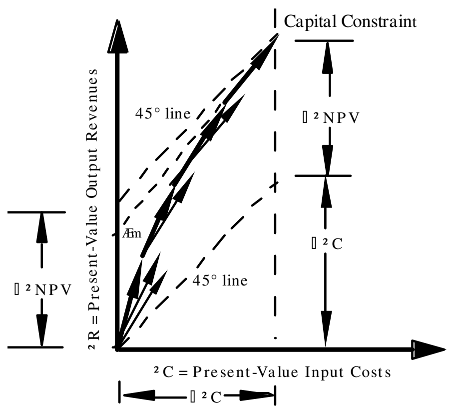
One method of determining the optimal mix of project alternatives for a given capital constraint is to compute the ΔNPV of all possible combinations. But this may be impractical due to the exponential increase of possible combinations that must be examined. Thus, if 100 non-overlapping projects have two alternatives each, $3^{100} \approx 5 \cdot 10^{46}$ budget choices are possible since each project could be done in either one of two ways, or not at all. Such astronomical numbers of combinations could not be evaluated in a reasonable amount of time. This discourages the consideration of project alternatives which could greatly benefit the net present-value objective of economic organizations subject to their capital constraints.

3. The convex-envelope proof in Section 1.5 satisfies the necessary and sufficient conditions for optimal economic decision-making based on Ricardo's marginal principle. This requires that *any* accepted alternative must have greater present-value output revenues for its present-value input cost than *any* alternative with the same present-value input cost that could not be funded because of the capital constraint. A single profitability criterion then enables optimal decisions to be made in a timely manner for all projects in an economic organization whenever their alternatives are under consideration.

Section 1.5 - Convex-envelope Proof of Optimal Capital Budgeting

The convex-envelope proof of optimal capital budgeting provides a rapid method of scanning the best way of doing each project and the best projects to do for a given capital constraint without exhaustively evaluating every possible combination. The vector bundles of each project are first ranked in descending order of their steepest-slope vectors which are added geometrically to form an initial convex envelope as shown in Figure 1.5.1 below.

Figure 1.5.1 - Initial steepest-slope convex envelope of project alternatives.



Vectors of the steepest-slope convex envelope from the origin to the capital-constraint line are the most capital-efficient alternatives for each project. But they could be replaced by vectors of the same or other bundles to maximize ∑ΔNPV for the capital constraint ∑ΔC. The resultant of convex-envelope vectors is independent of the order of adding the vectors. The only reason for ordering vectors in a convex envelope is to prove that a single profitability criterion can maximize both the *absolute profitability*, ∑ΔNPV, and the *average capital efficiency*, $\sum\Delta R/\sum\Delta C \equiv Ø_{avg}$, of the resultant for the capital constraint ∑ΔC.

The slope of the last vector of the convex envelope is $\Delta R_m/\Delta C_m \equiv Ø_m < Ø_{avg}$ which is the smallest slope of the convex envelope and its relative profitability is called the *marginal capital efficiency slope* of the convex envelope. The slope of the last vector of the convex envelope must have an angle that is greater than 45°. Otherwise, the last vector would add nothing to ∑ΔNPV while consuming ΔC_m of capital constraint ∑ΔC.

Possible replacements of initial convex-envelope vectors are derived by projecting the terminal points of all vectors in each bundle on the ΔR-axis at the marginal capital efficiency slope, $Ø_m$, of the convex envelope. By definition, the *marginal profitability* of vector V[ΔC,ΔR] is $\Delta NPV_m\{V\} \equiv \Delta R\{V\} - Ø_m\Delta C\{V\}$ where $Ø_m > 1$. When we determine the marginal profitabilities of every vector in a bundle, we may find that the vector with the greatest marginal profitability is either

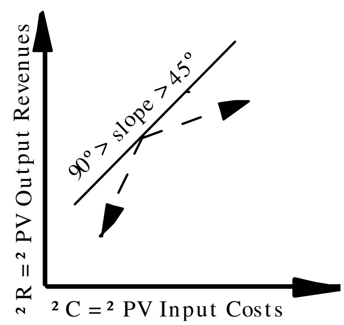
a) the steepest-slope vector in a bundle that was already in the convex envelope, or
b) a smaller capital-efficiency vector with a larger input cost, or
c) two or more vectors with smaller capital efficiencies and larger input costs.

In cases a) and b), the vectors with the greatest marginal profitabilities in each bundle are ranked again in descending order of their slopes and then added geometrically to form a new convex envelope of vectors. The vectors from case c) need to be merged with vectors from cases a) and b) to obtain the greatest $\sum\Delta NPV$ and $\sum\Delta R/\sum\Delta C \equiv Ø_{avg}$.

Hence, if the marginal capital efficiency, $Ø_{m}$, is increased continuously from 45^{o}, convex envelopes could be altered to maximize $\sum\Delta NPV$ subject to a capital constraint $\sum\Delta C$. Each vector in the final convex envelope represents the best way of doing each project as well as the best projects to do. Since the range of capital constraints has finite upper and lower bounds, only a finite number of marginal capital efficiency slopes are needed to determine the final convex envelope (see Appendix 1B - Programming Algorithm for Capital Budgeting).

The convex-envelope proof of optimal economic decision-making is validated by proving that any change of the proposed solution gives suboptimal results. For this purpose, the marginal capital efficiency slope, $Ø_{m}$, is drawn in Figure 1.5.2 below.

Figure 1.5.2 - Marginal Capital Efficiency Slope, $Ø_{m}$, of Optimal Economic Decision-Making



Three types of changes in the solution are possible:

1) The removal of the final convex-envelope vectors which appear in the solution.
2) The vector differences between bundle vectors that are and are not in the solution.
3) The introduction of vectors from bundles which are not in the solution.

All of these possible changes in the solution would be represented by dashed vectors pointed in the direction of the plane below the line passing through the marginal capital efficiency slope. The resultant of the possible changes cannot, therefore, be in the portion of the plane above the marginal capital efficiency slope. Hence, no change in the proposed solution could increase the net present-value added to the organization for the given capital constraint.

The proof indicates that the marginal capital efficiency slope of the organization for a given capital constraint is the key to optimal economic decision-making. Each project may have many engineering and financial alternatives for carrying out its production objective just like other projects that compete for the capital constraint of an organization. Therefore, it is not sufficient to ask "What is the best way of doing each project?" and "Which are the best projects to do?" because the marginal capital efficiency slope must be determined for the capital constraint of the organization before those questions could be answered. The marginal capital efficiency slope, $Ø_{m}$, measures the trade off between increasing the absolute profitability and average capital efficiency of an economic organization and increasing its capital constraint of present-value input costs.

Marginal capital efficiency is a single profitability criterion which applies to all project alternatives. If projects use different profitability criteria for economic optimization, the resulting under-investment, over-investment and missed opportunities degrade the productivity and profitability of the organization. Projects should compete for same capital constraint with the same marginal capital efficiency criterion in order to optimize results for the organization as a whole.

If an alternative has a sufficiently large ΔNPV and $\Delta R/\Delta C > Ø_m$ but *cannot* be undertaken because it would exceed the capital constraint, then it is worth seeking sources of debt, equity, leasing and joint venture financing to increase the amount of available capital. If surplus funds are available after all profitable alternatives have been undertaken, then those funds should be used to reduce external indebtedness or kept in short-term financial instruments while waiting for better project opportunities.

The effectiveness of capital constraints is difficult to determine in practical situations. An alternative should be sufficiently more profitable than the cost of borrowing money to make it worthwhile to borrow funds for undertaking that alternative. Otherwise, project opportunities generated by an organization will be used up largely for the benefit of external sources of capital. However, many enterprises cannot get nearly as much debt and equity financing as they would like to undertake profitable project opportunities. The capital constraint then depends on the delicate balance between the availability of financing and the ability of the organization to find profitable project opportunities.

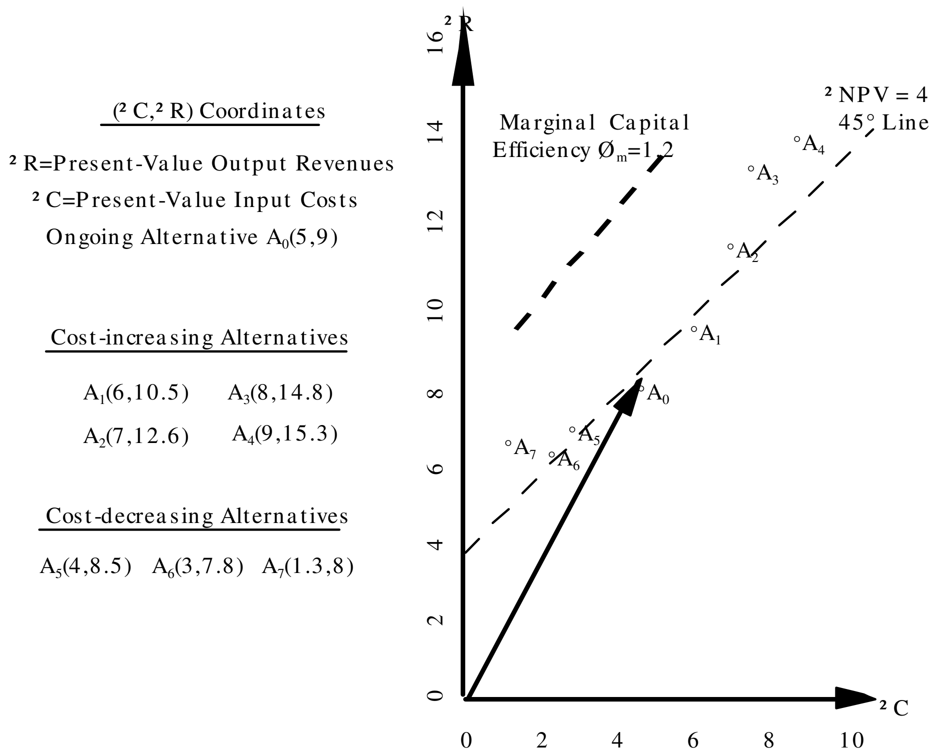
Section 1.6 - Optimal Decision-making for Project Alternatives

To fix ideas on determining the best way of doing each project, let us consider an organization with an $Ø_m = 1.2$ marginal capital efficiency ratio. Project A has eight mutually exclusive alternatives for carrying out its production objective. *Ongoing alternative* of project A is the continuation of past decisions, and is represented by a vector whose terminal point is denoted by $A_0(5,9)$ in the $(\Delta C,\Delta R)$ coordinates of Figure 1.6.1. Relative to ongoing alternative $A_0(5,9)$, there are four cost-increasing and three cost-decreasing alternatives.

The best way of doing project A can be found by plotting the terminal points of its alternatives in the $(\Delta C,\Delta R)$ plane as shown in Figure 1.6.1. The marginal capital efficiency slope, $\tan^{-1}\{Ø_m=1.2\} = 50.2^o$, is drawn as a dashed line in the upper left region of the $(\Delta C,\Delta R)$ plane and it is translated parallel to itself towards the terminal points of the alternatives. The first encounter of the dashed line with a terminal point identifies the alternative which is the best way of doing the project. If two or more terminal points are encountered simultaneously, the alternative with the greatest ΔR coordinate is best. The plotting procedure described above can be determined more conveniently by determining the alternative with the greatest marginal profitability $\Delta NPV_m\{V\} \equiv \Delta R\{V\}-Ø_m\Delta C\{V\}$.

The geometric solution can also be carried out by a method of binary comparisons in a sequential elimination process called *dynamic programming*. Sequential elimination enables one to find the best of N alternatives from N-1 binary comparisons. At each binary comparison, the worse alternative is eliminated because it is <u>known</u> not to be the best alternative. After N-1 worse alternatives are eliminated, only the best one is left. It is important to note that the sequence of binary comparisons does not affect the solution for the best alternative. To fix ideas, let us first consider replacing the ongoing alternative by cost-increasing alternatives which increase present-value input costs in order to increase absolute profitability while possibly decreasing capital efficiency.

Figure 1.6.1 - Cost-Increasing, Cost-Decreasing & Ongoing Alternatives of Project A



Let us consider replacing ongoing alternative $A_0(5,9)$, called the *defender*, by the first cost-increasing alternative $A_1(6,10.5)$, called the *challenger*. Alternatives A_0 and A_1 are compared by projecting their terminal points with marginal capital efficiency slope $Ø_m = 1.2$ on the ΔR-axis to determine their marginal profitabilities $\Delta NPV_m \equiv \Delta R\text{-}Ø_m\Delta C$ as follows:

Defender A_0: $\Delta NPV_m\{A_0\}[5,9] = 9\text{-}1.2\cdot 5 = 3.0$
Challenger A_1: $\Delta NPV_m\{A_1\}[6,10.5] = 10.5\text{-}1.2\cdot 6 = 3.3$

Replacing $A_0(5,9)$ by $A_1(6,10.5)$ at the margin of the capital constraint would increase the absolute profitability $\sum\Delta NPV$ of the organization by $\Delta NPV_m\{A_1\text{-}A_0\} = 3.3\text{-}3.0 = 0.3$ units. Therefore, we should replace $A_0(5,9)$ and retain challenger $A_1(6,10.5)$ as a defender to be compared to the next challenger $A_2(7,12.6)$.

Defender A_1: $\Delta NPV_m\{A_1\}[6,10.5] = 10.5\text{-}1.2\cdot 6 = 3.3$
Challenger A_2: $\Delta NPV_m\{A_2\}[7,12.6] = 12.6\text{-}1.2\cdot 7 = 4.2$

Replacing $A_1(6,10.5)$ by $A_2(7,12.6)$ at the margin of the capital constraint would increase the absolute profitability $\sum\Delta NPV$ of the organization by $\Delta NPV_m\{A_2\text{-}A_1\} = 4.2\text{-}3.3 = 0.9$ units. Therefore, we should replace $A_1(6,10.5)$ and retain challenger $A_2(7,12.6)$ as a defender to be compared to the next challenger $A_3(8,14.8)$.

Defender A_2: $\Delta NPV_m\{A_2\}[7,12.6] = 12.6\text{-}1.2\cdot 7 = 4.2$
Challenger A_3: $\Delta NPV_m\{A_3\}[8,14.8] = 14.8\text{-}1.2\cdot 8 = 5.2$

Replacing $A_2(7,12.6)$ by $A_3(8,14.8)$ at the margin of the capital constraint would increase the absolute profitability $\sum\Delta NPV$ of the organization by $\Delta NPV_m\{A_3\text{-}A_2\} = 5.2\text{-}4.2 = 1.0$ units. Therefore, we should replace $A_2(7,12.6)$ and retain challenger $A_3(8,14.8)$ as a defender to be compared to the last cost-increasing alternative, challenger $A_4(9,15.3)$.

Defender A_3: $\Delta NPV_m\{A_3\}[8,14.8] = 14.8\text{-}1.2 \cdot 8 = 5.2$
Challenger A_4: $\Delta NPV_m\{A_4\}[9,15.3] = 15.3\text{-}1.2 \cdot 9 = 4.5$

Replacing $A_3(8,14.8)$ by $A_4(9,15.3)$ at the margin of the capital constraint would *decrease* the absolute profitability $\sum\Delta NPV$ of the organization by $\Delta NPV_m\{A_4\text{-}A_3\} = 4.5\text{-}5.2 = \text{-}0.7$ units. Therefore, $A_4(9,15.3)$ should not replace $A_3(8,14.8)$ which would be retained as a defender to be compared to the first cost-decreasing alternative, challenger $A_5(4,8.5)$.

Defender A_3: $\Delta NPV_m\{A_3\}[8,14.8] = 14.8\text{-}1.2 \cdot 8 = 5.2$
Challenger A_5: $\Delta NPV_m\{A_5\}[4,8.5] = 8.5\text{-}1.2 \cdot 4 = 3.7$

Replacing $A_3(8,14.8)$ by $A_5(4,8.5)$ at the margin of the capital constraint would *decrease* the absolute profitability $\sum\Delta NPV$ of the organization by $\Delta NPV_m\{A_5\text{-}A_3\} = 3.7\text{-}5.2 = \text{-}1.5$ units. Therefore, $A_5(4,8.5)$ should not replace $A_3(8,14.8)$ which would be retained as a defender to be compared to the next challenger $A_6(3,7.8)$.

Defender A_3: $\Delta NPV_m\{A_3\}[8,14.8] = 14.8\text{-}1.2 \cdot 8 = 5.2$
Challenger A_6: $\Delta NPV_m\{A_6\}[3,7.8] = 7.8\text{-}1.2 \cdot 3 = 4.2$

Replacing $A_3(8,14.8)$ by $A_6(3,7.8)$ at the margin of the capital constraint would *decrease* the absolute profitability $\sum\Delta NPV$ of the organization by $\Delta NPV_m\{A_6\text{-}A_3\} = 4.2\text{-}5.2 = \text{-}1.0$ units. Therefore, $A_6(3,7.8)$ should not replace $A_3(8,14.8)$ which would be retained as a defender to be compared to the last cost-decreasing alternative, challenger $A_7(1.3,8)$.

Defender A_3: $\Delta NPV_m\{A_3\}[8,14.8] = 14.8\text{-}1.2 \cdot 8 = 5.2$
Challenger A_7: $\Delta NPV_m\{A_7\}[1.3,8] = 8\text{-}1.2 \cdot 1.3 = 6.44$

Replacing $A_3(8,14.8)$ by $A_7(1.3,8)$ at the margin of the capital constraint would increase the absolute profitability $\sum\Delta NPV$ of the organization by $\Delta NPV_m\{A_7\text{-}A_3\} = 6.44\text{-}5.2 = 1.24$ units. Therefore, $A_7(1.3.8)$ should replace $A_3(8,14.8)$. In the absence of other challengers, $A_7(1.3,8)$ is the best way for doing project A.

Although $A_7(1.3,8)$ is the best alternative, its $\Delta NPV\{A_7\}[1.3,8] = 6.7$ is smaller than $\Delta NPV\{A_3\}[8,14.8] = 6.8$. This raises the question, Should the best project alternative have the greatest ΔNPV? The answer is *not always* because of the *capital constraint*. If $A_7(1.3,8)$ *did not* replace $A_3(8,14.8)$, $\Delta C\{A_3\text{-}A_7\} = 6.7$ units of input costs would be used to get only $\Delta NPV\{A_3\text{-}A_7\} = 6.8\text{-}6.7 = 0.1$ units more of net present value. At the margin of the capital constraint, 6.7 units of input costs could have been used either to increase net present value by $\text{Ø}_m\Delta C\{A_3\text{-}A_7\} = 1.2 \cdot 6.7 = 8.04$ units, or to reduce debt by 6.7 units.

Replacements by cost-increasing and cost-decreasing alternatives affect both the capital constraint and marginal capital efficiency slope. The range of capital constraints is limited by costs of borrowing money. Cost-increasing and cost-decreasing alternatives affect capital efficiency by having either increasing, decreasing or constant returns to scale. Marginal capital efficiencies are discussed in Appendix 1A - Algebra of Marginal Capital Efficiency Ratios and in Chapter Nine, Section 9.4 - Engineering Production Functions.

Section 1.7 - Objectives of Economic Decision-making Models

A myriad of details clouds economic decisions to be made from complex alternatives subject to divergent objectives. In order to rationalize the decision-making process, economic models must be represented in simpler terms in order to simulate reality without too many details. However, methods of economic modeling are not universal, and there is room for questioning how close economic models are to reality. This brings up a need to discuss the objectives of economic decision-making models and how their results could be verified.

A decision cannot be optimized without a well-defined objective. For example, let us ask "Which is more, one table or four chairs?". This question is meaningless unless 'more' is defined as an additive property of table and chairs. If the objective is to have more firewood, the table and chairs could be weighed to see which alternative provides more firewood. Weight is a *criterion* of deciding which alternative is better, and the availability of firewood is a *constraint* on satisfying the objective. Conditions may arise which change constraints, but objectives and criteria cannot be changed without changing the underlying question.

It is essential to point out that there is universal agreement on the objectives of maximizing output revenues for given input costs and minimizing input costs for given output revenues. These objectives apply to industrial firms, financial institutions, government agencies and nonprofit organizations. However, when alternatives do not have either the same input costs or output revenues, the objectives are not well-defined. As a consequence, economic decision-making often gives conflicting choices of project alternatives, none of which is clearly optimal. Moreover, the ambiguity of objectives give rise to economic criteria which differ from one organization to the next and within the same organization.

The objective of economic decision-making is defined here as the selection of engineering and financial alternatives which maximize the net present-value added to an organization as a whole for a given capital constraint of present-value input costs. However, present values are not defined until the discount rate is specified. If discount rates are too low, large net present-values may come from output revenues to be received in the distant future. The present value of distant revenues can be reduced by using higher discount rates. How high the discount rates should be is an open question in both theory and practice.

We propose that managerial accounting should discount future input costs and output revenues of alternatives at the *costs of borrowing money* just as past *costs of borrowed money* are posted in financial accounting. For this reason, discount rates should be forecast from expected costs of borrowing money at market rates of interest. Risk, inflation and investor expectations are often coupled with market rates of interest to form higher discount rates. The problem with artificially high discount rates is that they obscure the future by making long-term investments look unprofitable. In contrast, discount rates based on costs of borrowing money enable one to make best use of available engineering and financial information in forecasting the input costs and output revenues of each alternative.

The assumption of accurate forecasts is needed to develop a provable cause-and-effect approach to economic decision-making. It enables one to focus on the logic of the decision-making process regardless of who, where or when decisions are made. The optimality of decisions cannot depend on taking place only in industrial firms, financial institutions, government agencies or nonprofit organizations. Although economic decision-making is the prerogative of managers, it is implemented by engineers, accountants and economists in various levels of an organization. It is both students and practitioners of these disciplines and activities that this book is addressed.

Current methods of economic decision-making lack a well-defined objective. This has given rise to multiple criteria in economic decision-making such as maximizing net present-values discounted at weighted-average-costs-of-capital (WACC) or minimum-attractive-rates-of-return (MARR), maximizing internal rate-of-return (IRR), maximizing benefit/cost ratio and minimizing payback period. The problem with multiple criteria is that they often lead to arbitrary constraints and conflicting decisions which are resolved by subjective judgements. Because multiple criteria are commonly used in many economic organizations, the similarities and differences between the proposed method of economic decision-making and multiple criteria decision-making are analyzed in greater detail in subsequent chapters.

Section 1.8 - Summary of Chapter One

Every economic organization can be viewed as a black box in which physical inputs and outputs flow one way and cash flows the other way. The study of economic decision-making is aimed at a better understanding of the cause-and-effect relationships between choices of physical inputs and outputs (called *engineering alternatives*) and choices of cash flow inputs and outputs (called *financial alternatives*). The book presents a provably optimal method of selecting engineering and financial alternatives that maximize the net present-value added to an economic organization for a given capital constraint.

The history of economic doctrines (Section 1.2) is traced from medieval times to the works of classical economists. Adam Smith (1723-1790) claimed that wealth increased with exchanges of goods brought about by the division of labor and industrial specialization. David Ricardo (1772-1823) based his theory of land rent on a competitive equilibrium of profitability and productivity. Karl Marx (1818-1883) advocated state ownership of all land and capital in order for labor to receive the full benefit of its productivity. These economists contributed greatly to the study of economics, but much room was left for questioning.

The ideas of great economists largely stem from bookkeeping records. Single-entry bookkeeping in the 13th and 14th century evolved from the need to record transactions without having to remember every transaction. Single-entry bookkeeping was supplemented by the two-way classification system of double-entry bookkeeping which was first described by Luca Paciolo in 1494. It was Paciolo's work which was largely responsible for introducing Italian double-entry bookkeeping to Western Europe from the 16th century to the present.

Financial accounting deals with results of past decisions. *Income statements* record input costs and output revenues occurring within short *accounting periods* of time in single-entry bookkeeping format. The net revenues of income statements are called *profits* or *losses*. However, profits or losses in an accounting period may be the effects of previous accounting periods or the causes of profits or losses in future accounting periods. Although income statements are grounded in fact, it lacks records of past alternatives and it does not allocate the cost of borrowed money and income taxes to individual projects. *Balance sheets* equate the assets and liabilities of an organization in double-entry bookkeeping format at the start and finish of each accounting period. For auditing purposes, entries in single-entry income statements must synchronize with the two adjoining double-entry balance sheets.

Managerial accounting, also called *cost accounting*, is the domain of economic decision-making. Because forecasting errors are inseparable from the economic decision-making process, the proof of optimality is based on deterministic cash flow forecasts of each alternative. These forecasts take into account engineering risks of input costs, marketing risks of output revenues and interest-rate risks of money transactions.

The cash flow of each alternative is defined as cash flows of the organization as a whole if the alternative is accepted as opposed to the cash flow if the alternative is not accepted. The differences in cash flow between accepting and not accepting alternatives affect how much is loaned or borrowed at market rates of interest. To verify the results of economic decision-making, cash flow forecasts of alternatives should be discounted at the *cost of borrowing money* just as the *cost of borrowed money* is posted in financial accounting. Managerial accounting makes no distinction between debt or equity financing in cash flow descriptions of alternatives because borrowed and out-of-pocket money is indistinguishable.

Economic decision-making needs to measure the contribution of each alternative to the objective of an economic organization. In this regard, an organization is subdivided into nonoverlapping and collectively exhaustive projects. We now ask these three questions:

1. What is the best way of doing each project?
2. Which are the best projects to do?
3. Which projects should be funded?

The answers to these questions are interrelated because all projects compete for the same capital constraint. The necessary and sufficient conditions for optimal economic decision-making is that any accepted alternative must have a greater present-value output revenues for its present-value input cost than any alternative with the same present-value input cost that could not be funded because of the capital constraint (i.e., Ricardo's marginal principle).

Mutually exclusive project alternatives are represented by two-dimensional vectors whose components are *present-value input costs* [ΔC] and *present-value output revenues* [ΔR]. The vector bundle has a common initial point and individual terminal points. The difference $\Delta R - \Delta C \equiv \Delta NPV$ measures the *Net Present-Value* or *Absolute Profitability* which a vector adds to the organization's objective. Projecting the terminal point of a vector at 45º on the ΔR-axis also measures the ΔNPV of the vector. Ratio $\Delta R/\Delta C \equiv Ø$ is the vector's slope which measures its *Capital Efficiency* or *Relative Profitability* in converting input ΔC into output ΔR.

The vector bundles are ranked in descending order of their steepest-slope vectors which are added geometrically to form a convex envelope. The last vector of the convex envelope to reach the capital constraint has a *marginal capital efficiency slope*, $\Delta R_m/\Delta C_m \equiv Ø_m$, which helps measure the *marginal profitability*, $\Delta NPV_m\{V\} \equiv \Delta R\{V\} - Ø_m \Delta C\{V\}$, of every vector at the margin of the capital constraint. Convex-envelope vectors are replaceable by vectors in their bundles which have the greatest marginal profitability. Hence, if slope $Ø_m$ is rotated from 45º to 90º, the convex envelopes could be altered to maximize $\sum\Delta NPV$ for a given $\sum\Delta C$. The convex-envelope proof of optimal decision-making indicates that the marginal capital efficiency slope, $Ø_m > 1$, is a single criterion which applies to every set of project alternatives.

Conventional methods of economic decision-making lack a well-defined objective. This has given rise to multiple criteria in economic decision-making such as maximizing net present-values discounted at weighted-average-costs-of-capital (WACC) or minimum-attractive-rates-of-return (MARR), maximizing internal rate-of-return (IRR), maximizing benefit/cost ratio and minimizing payback period. However, multiple criteria often lead to arbitrary constraints and conflicting decisions which are resolved by subjective judgements. Because multiple criteria are commonly used in economic organizations, the similarities and differences between the single criterion of optimal economic decision-making and multiple criteria of conventional decision-making are analyzed in greater detail in later chapters.

Appendix 1A - Algebra of Marginal Capital Efficiency Ratios

The marginal capital efficiency slope, $\Delta R_m/\Delta C_m$, is the ratio of present-value output revenues, ΔR_m, to present-value input costs, ΔC_m, at the margin of the capital constraint. Ratios are dimensionless quantities because they are relationships of two quanities of the same kind. The properties of ratios which are summarized here may be found in greater detail in "Higher Algebra" by H. R. Hall and S. R. Knight, Macmillan and Co., London 1950.

The value of a ratio is unchanged if the numerator and denominator are multiplied by the same quantity. Two or more ratios may be compared by reducing their equivalent fractions to a common denominator. Thus, suppose a/b and x/y are two ratios. Multiply a/b by y, and multiply x/y by b, to get common denominator by. Consequently, a/b is greater than, equal to or smaller than x/y, just as ay is greater than, equal to or smaller than xb.

A ratio is said to be a ratio of *greater inequality*, of *lesser inequality*, or of *equality*, according as the numerator is *greater than*, *less than,* or *equal to* the denominator. If the same quantity was added to the numerator and denominator, a ratio of greater inequality would be diminished and a ratio of lesser inequality would be increased. The results of such additions may be proved as follows:

Let a/b be the ratio to which x is added to form the new ratio $(a+x)/(b+x)$. Multiply a/b by $(b+x)$ and multiply $(a+x)/(b+x)$ by b to get

$$[a/b] - [(a+x)/(b+x)] = x(a-b)/b(b+x)$$

Consequently, if a is greater than b, then $x(a-b)/b(b+x)$ is positive and $[a/b] > [(a+x)/(b+x)]$. And if a is smaller than b, then $x(a-b)/b(b+x)$ is negative and $[a/b] < [(a+x)/(b+x)]$. Similarly, if the same quantity was subtracted from the numerator and denominator, a ratio of greater inequality would be increased and a ratio of lesser inequality would be diminished.

When two or more ratios are equal, many useful propositions may be proved by equating them to a single symbol k to denote each of the equal ratios. This procedure will be illustrated by proving the following important theorem. If $[a/b] = [c/d] = \ldots = k$, then each ratio $[(pa^n+qc^n)/(pb^n+qd^n)]^{1/n}$ also equals k, where p, q and n are *any quantities whatever*.

Proof: Since $[a/b] = [c/d] = k$, it follows that a = bk and c = dk. Also $pa^n = pb^nk^n$ and $qc^n = qd^nk^n$. Therefore,

$$[(pa^n+qc^n)/(pb^n+qd^n)] = [(pb^nk^n+ qd^nk^n)/(pb^n+qd^n)] = k^n$$

$$[(pa^n+qc^n)/(pb^n+qd^n)]^{1/n} = k = [a/b] = [c/d]$$

By giving different values to p, q and n, many particular cases of this general proposition may be deduced, or they may be proved independently by the same method. A case of frequent utility is that when a series of fractions are equal, then each of them is equal to the sum of all the numerators divided by the sum of all the denominators.

It is also true that when a series of n fractions are equal, then each of them is equal to the *nth*-root of the product of all the numerators divided by the product of all the denominators. Variations of this general principle are illustrated in the following example.

If $[a/b] = [c/d] = k$, then it follows that $[ac/bd] = k^2$ and $[ac/bd]^{1/2} = k$ as stated in the general principal. Let us suppose again that $[a/b] = [c/d] = k$, and we want to show that the ratio $[(2a^2b^3 - 3ac^2d^2)/(2b^5 - 3ad^4)]$ is also equal to $[ac/bd] = k^2$ and $[ac/bd]^{1/2} = k$. To prove this, we have $a = bk$ and $2a^2b^3 = 2b^5k^2$ as well as $c = dk$ and $-3ac^2d^2 = -3ad^4k^2$. Therefore,

$$[(2a^2b^3 - 3ac^2d^2)/(2b^5 - 3ad^4)] = [(2b^5k^2 - 3ad^4k^2)/(2b^5 - 3ad^4)] = k^2 = [ac/bd]$$

$$[ac/bd]^{1/2} = k = [a/b] = [c/d]$$

Empirical samples of marginal capital efficiency ratios tend to be unequal. In this connection, the following theorem is important. If $[a_1/b_1]$, $[a_2/b_2]$, ... , $[a_n/b_n]$ are unequal fractions with positive denominators, then the magnitude of the fraction

$$\frac{a_1 + a_2 + ... + a_n}{b_1 + b_2 + ... + b_n}$$

lies between the smallest fraction $[a_r/b_r]$ and the largest fraction $[a_s/b_s]$.

Proof: Let $[a_r/b_r] = k$. Then $a_1 > kb_1$, $a_2 > kb_2$, ... , $a_n > kb_n$. By addition, we get

$$a_1 + a_2 + ... + a_n > k(b_1 + b_2 + ... + b_n)$$

$$\frac{a_1 + a_2 + ... + a_n}{b_1 + b_2 + ... + b_n} > k = [a_r/b_r]$$

Similarly, we may prove that

$$\frac{a_1 + a_2 + ... + a_n}{b_1 + b_2 + ... + b_n} < [a_s/b_s]$$

Appendix 1B - Programming Algorithm for Capital Budgeting

A capital-budgeting program is presented here which permits a rapid scanning of more profitable combinations of projects and alternatives within a planned range of capital constraints, without exhaustively examining the profitabilities of every possible combination. Let us designate the nonoverlapping and indivisible projects which comprise the planned set of production opportunities by subscript i which may have the values 1, 2, ... , m. Let us number the mutually exclusive alternatives of each project by subscript j which may assume the values 1, 2, ... , n. Consequently, subscript ij is the j-th mutually exclusive alternative of the i-th project.

For any project alternative with the general subscript ij, we can calculate the changes of present-value input costs ΔC_{ij} and present-value output revenues ΔR_{ij} based on the cash flows of the organization as a whole if the ij-th project alternative was accepted as opposed to those cash flows if the ij-th project alternative was replaced or rejected. The difference $\Delta R_{ij} - \Delta C_{ij} = \Delta NPV_{ij}$ would represent, therefore, the net present-value added by the ij-th project alternative. A rectangular array of these alternatives can be made, with each element of a column representing one of the m projects, and with each element of a row representing one of the n mutually exclusive alternatives of the i-th project, as is shown in the matrix notation of (1B.1).

$$\text{Given:} \qquad ||\Delta C_{ij}|| \text{ and } ||\Delta R_{ij}|| \qquad \text{....(1B.1)}$$

Since the number of mutually exclusive alternatives for each project is rarely the same, it would be necessary to fill many positions in these matrices with zeros in order for them to have a rectangular shape.

The capital budgeting problem we seek to solve is to find the projects, with at most one alternative way of fulfilling each project, that give a maximum net present-value added for a given capital cost. Let us denote the given capital cost by the discrete variable C. Then the statement of the generalized capital-budgeting problem is as follows:

$$\text{Find: } X = ||\Delta X_{ij}|| \text{ such that} \qquad \text{....(1B.2)}$$

$$\text{1) } X_{ij} = (X_{ij})^2 \quad \text{(true only if } X_{ij} \text{ equals zero or one)} \qquad \text{....(1B.3)}$$

$$\text{2) } \sum_j X_{ij} \le 1 \quad \text{(at most one alternative in a row)} \qquad \text{....(1B.4)}$$

$$\text{3) } \sum_i \sum_j \Delta C_{ij} X_{ij} = C \qquad \text{....(1B.5)}$$

$$\text{4) } \sum_i \sum_j \Delta NPV_{ij} X_{ij} \text{ is a maximum} \qquad \text{....(1B.6)}$$

The capital-budgeting schedule which will be obtained in the proposed solution to this problem will only pertain to those discrete costs of capital C that correspond to budget choices lying on the convex polygon.

A graphical approach will be used for the solution of this generalized capital-budgeting problem. The solution consists of the following:

a) Calculate the relative profitability ratio $Ø_{ij} = \Delta R_{ij}/\Delta C_{ij}$ for all project alternatives and rearrange the elements of each row of the $Ø = || Ø_{ij} ||$ matrix according to the scheme $Ø_{i1} > Ø_{i2} > \ldots > Ø_{in}$.

b) Rearrange the rows of the $Ø = || Ø_{ij} ||$ matrix in descending order of the first element (i.e., $Ø_{11} > Ø_{21} > \ldots > Ø_{m1}$).

c) Draw in the position vectors of the first column of the $Ø = || Ø_{ij} ||$ matrix in the manner shown in Figure 1.6.1, namely, by drawing the position vector from the origin to the point $(\Delta C_{11},\Delta R_{11})$, and then adding vector $(\Delta C_{21},\Delta R_{21})$ to the point $(\Delta C_{11},\Delta R_{11})$, and then adding vector $(\Delta C_{31},\Delta R_{31})$ to the point $(\Delta C_{11}+\Delta C_{21},\Delta R_{11}+\Delta R_{21})$, and so on. These vectors will form a pattern convex to the ΔC-axis, because the direction angle $\tan^{-1}Ø_{11}$ is greater than $\tan^{-1}Ø_{21}$ which, in turn, is greater than $\tan^{-1}Ø_{m1}$ as a consequence of step b).

d) Draw in all the remaining alternatives in each row, if any, with the same initial point as the first element in the row. The first element in every row will have a greater slope than the succeeding elements in its row because $\tan^{-1}Ø_{i1} > \tan^{-1}Ø_{i2} > \ldots > \tan^{-1}Ø_{in}$ as a consequence of step a).

e) Let $Ø_{m1} = \Delta R_{m1}/\Delta C_{m1}$ be the first estimate of the marginal capital efficiency which is translated parallel to itself until it passes through the extreme terminal point of the bundle of vectors in the first row. We shall earmark with an asterisk (*) that vector of the first row whose terminal point was the most extreme to pass through the marginal comparison slope. In the event several vectors of the first row have terminal points which are colinear with the terminus of the earmarked vector along the line of direction of the marginal comparison slope, then only the vector with the greatest direction angle should be earmarked with an asterisk.

f) Translate the marginal capital efficiency to the terminal points of the bundle of vectors in the second row, and continue to repeat step e) for each succeeding row up to and including row *m*-1. The solution matrix X^* is $X_{1j}^*, X_{2j}^*, \ldots, X_{m-1,j}^*$ corresponding to a cost of capital $\Delta C_{1j}^* + \Delta C_{2j}^* + \ldots + \Delta C_{m-1,j}^*$ and a maximum net present-value added of $\Delta NPV_{1j}^* + \Delta NPV_{2j}^* + \ldots + \Delta NPV_{m-1,j}^*$.

g) Repeat steps e) and f) using $Ø_{m-1,1} = \Delta R_{m-1,1}/\Delta C_{m-1,1}$ as the new estimate of the marginal capital efficiency, and marking the vectors in each row with extreme terminuses in relation to the line of direction of this new marginal comparison slope with two asterisks. The solution matrix X^{**} is $X_{1j}^{**}, X_{2j}^{**}, \ldots, X_{m-2,j}^{**}$ corresponding to a cost of capital $\Delta C_{1j}^{**} + \Delta C_{2j}^{**} + \ldots + \Delta C_{m-2,j}^{**}$ and a maximum net present-value added of $\Delta NPV_{1j}^{**} + \Delta NPV_{2j}^{**} + \ldots + \Delta NPV_{m-2,j}^{**}$. In this manner, the capital -budgeting schedule can be constructed. It will consist of solution matrices $X^*, X^{**}, \ldots, X^{*(r)}$ with corresponding costs of capital and maximum net present-values added, where *r* is the number of entries in the schedule. The marginal capital efficiency estimates can be selected independently to cover a range of capital costs which the firm contemplates spending. The proof of this solution of the capital-budgeting problem is outlined in Section 1.6 of Chapter One. Problems of estimating marginal capital efficiencies and optimal capital budgeting are illustrated in Chapter One - Exercises.

Chapter One - Exercises

1-1a Project A has nine mutually exclusive alternatives to its ongoing alternative $A_0(\Delta C,\Delta R) = A_0(20,35)$. The $(\Delta C,\Delta R)$ coordinates of the nine alternatives are $A_1(30,42)$, $A_2(27,40)$, $A_3(22,38)$, $A_4(18,32)$, $A_5(16,30)$, $A_6(14,28)$, $A_7(13,26)$, $A_8(20,38)$ and $A_9(24,36)$, where ΔC is present-value input costs and ΔR is present-value output revenues. Determine the best way of doing project A if the marginal capital efficiency ratio, $Ø_m = \Delta R_m/\Delta C_m$, of present-value input costs at the margin of the capital constraint equals 2.5, 2.0 and 1.5 .

The geometric method of solving this problem is to plot the terminal points all ten alternatives in the $(\Delta C,\Delta R)$ plane. Marginal capital efficiency slopes $\tan^{-1}2.5 = 68.2^o$, $\tan^{-1}2.0 = 63.4^o$ and $\tan^{-1}1.5 = 56.3^o$ are now translated parallel to themselves from the upper left region of the $(\Delta C,\Delta R)$ plane towards the terminal points of the alternatives. The first encounter of the marginal capital efficiency slope with a terminal point identifies the alternative which is the best way of doing project A. If two or more terminal points are encountered simultaneously, the alternative with the greatest ΔR coordinate is best.

The problem can also be solved by projecting all ten terminal points on the ΔR-axis using the marginal capital efficiency slopes, $Ø_m$, for the projections. This can be carried out by the formula $\Delta NPV_m\{V\} \equiv \Delta R\{V\}-Ø_m\Delta C\{V\}$ which measures the marginal profitability of each alternative at the margin of the capital constraint. The best way of doing project A are the alternatives which have the greatest marginal profitability for given values of $Ø_m$.

The best way of doing a project may also be determined by the method of binary comparisons in a sequential elimination process called *dynamic programming* as described in Section 1.6 . Sequential elimination can find the best of N alternatives from N-1 binary comparisons. At each binary comparison, the worse alternative is eliminated because it is known not to be the best alternative. After N-1 worse alternatives are eliminated, only the best one is left. The sequential elimination process of binary comparisons is useful to analyze trade offs between absolute profitability and capital efficiency.

1-2a Evaluate ΔR-ΔC and $\Delta R/\Delta C$ for each mutually exclusive alternative of each projects in Table 1B.1 . Select the steepest-slope vector of each project and form a convex envelope in descending order of the vector slopes. Estimate the marginal capital efficiency slope, $Ø_m$, from the smallest vector slope of the convex envelope and calculate the marginal profitability, $\Delta R-Ø_m\Delta C$, for each alternative of each project.

1-2b Budget #1 is formed from steepest-slope vectors of each project in descending order of their slopes which made the convex envelope in Problem **1-2a**. Budget #2 consists of the alternative in each project with the greatest absolute profitability. Budget #3 consists of the alternative in each project with the greatest marginal profitability. The last four columns of the three budgets represent runsums of the in the first four columns. The last row of the last four columns of the three budgets describe the resultant vector of the budgets.

1-2c Resultant vector differences - Show the vector difference between resultants of Budgets #1 and #3 has a capital efficiency (i.e., $(\sum\Delta R_1-\sum\Delta R_3)/(\sum\Delta C_1-\sum\Delta C_3)$ *greater* than $Ø_m$. Show the vector difference between resultants of Budgets #2 and #3 has a capital efficiency (i.e., $(\sum\Delta R_2-\sum\Delta R_3)/(\sum\Delta C_2-\sum\Delta C_3)$ *smaller* than $Ø_m$. It should be noted resultant vectors of budgets are independent of the order in which their individual vectors were summed.

Table 1B.1 - Sample data set of four projects and their mutually exclusive alternatives.

Project A	**ΔR**	**ΔC**	**ΔR-ΔC**	**$\Delta R/\Delta C$**			**$\Delta R - Ø_m\Delta C$**	
A_1	25	14						
A_2	14	6						
A_3	20	10						
A_4	30	18						
Project B	**ΔR**	**ΔC**	**ΔR-ΔC**	**$\Delta R/\Delta C$**			**$\Delta R - Ø_m\Delta C$**	
B_1	38	20						
B_2	43	24						
B_3	32	16						
B_4	48	28						
Project C	**ΔR**	**ΔC**	**ΔR-ΔC**	**$\Delta R/\Delta C$**			**$\Delta R - Ø_m\Delta C$**	
C_1	32	27						
C_2	23	18						
C_3	28	22						
C_4	17	14						
Project D	**ΔR**	**ΔC**	**ΔR-ΔC**	**$\Delta R/\Delta C$**			**$\Delta R - Ø_m\Delta C$**	
D_1	36	26						
D_2	52	38						
D_3	47	34						
D_4	42	30						

Budget #1	**ΔR**	**ΔC**	**ΔR-ΔC**	**$\Delta R/\Delta C$**	**$\sum\Delta R_1$**	**$\sum\Delta C_1$**	**$\sum\Delta R_1$-$\sum\Delta C_1$**	**$\sum\Delta R_1/\sum\Delta C_1$**
Basis	**0**	**0**	**0**	**- - -**	**0**	**0**	**0**	**- - -**
Budget #2	**ΔR**	**ΔC**	**ΔR-ΔC**	**$\Delta R/\Delta C$**	**$\sum\Delta R_2$**	**$\sum\Delta C_2$**	**$\sum\Delta R_2$-$\sum\Delta C_2$**	**$\sum\Delta R_2/\sum\Delta C_2$**
Basis	**0**	**0**	**0**	**- - -**	**0**	**0**	**0**	**- - -**
Budget #3	**ΔR**	**ΔC**	**ΔR-ΔC**	**$\Delta R/\Delta C$**	**$\sum\Delta R_3$**	**$\sum\Delta C_3$**	**$\sum\Delta R_3$-$\sum\Delta C_3$**	**$\sum\Delta R_3/\sum\Delta C_3$**
Basis	**0**	**0**	**0**	**- - -**	**0**	**0**	**0**	**- - -**

Chapter One - Suggested Readings

The books listed here provided sources of information on particular topics of Chapter One. The listed books are also recommended for readers who seek a more detailed understanding of subjects which were neglected or only briefly abstracted in the text.

History of Economic Doctrines
Gide, C. and Rist, C., *A History of Economic Doctrines*, D. C. Heath & Co., 1948.
Heimann, E., *History of Economic Doctrines,* Oxford University Press, 1951.
Lekachman, R., *A History of Economic Ideas*, Harper & Brothers, 1959.
Pasinetti, L. L., *Lectures on the Theory of Production*, Columbia University Press, 1977.
Schumpeter, J. A., *Ten Great Economists*, Oxford University Press, 1951.
Schumpeter, J. A., *History of Economic Analysis*, Oxford University Press, 1954.
Soule, G., *Ideas of the Great Economists,* The New American Library, Viking Press, 1962.

Financial and Managerial Accounting
Brealey, R. and Myers, S., *Principles of Corporate Finance*, McGraw-Hill Book Co., 1984.
Horngren, C. T., and Foster, G., *Cost Accounting, A Managerial Emphasis,* Prentice-Hall, 1987.
Ijiri, Y., *Theory of Accounting Measurement*, American Accounting Association (AAA), 1975.
Ijiri, Y., *Triple-entry Bookkeeping and Income Momentum*, Am Acctg Assn, 1982.
Littleton, A. C., *Structure of Accounting Theory*, Am Acctg Assn, 13th Printing, 1985.
Mason, J. M., *Financial Management of Commercial Banks*, Warren, Gorham & Lamont, 1979.
Mautz, R. K. and Sharaf, H. A., *The Philosophy of Auditing*, AAA, 12th Printing, 1985.
Meigs, W. B., Johnson, C. E., and Meigs, R. F., *Accounting, The Basis for Business Decisions*, McGraw-Hill Book Co., 1977.
Paton, W. A. and Littleton, A. C., *An Introduction to Corporate Accounting Standards*, AAA, 18th Printing, 1985.
Shute, G. E., *Accounting Students and Abstract Reasoning: An Exploratory Study*, AAA, 1979.
Thomas, A. L., *The Allocation Problem in Financial Accounting Theory*, Am Acctg Assn, 1969.
Thomas, A. L., *The Allocation Problem: Part Two*, Am Acctg Assn, 1974.
Zlatkovitch, C. T. et. al., *A Statement of Basic Accounting Theory*, AAA, 6th Printing, 1977.

Dynamic Programming
Beckmann, M. J., *Dynamic Programming of Economic Decisions*, Springer-Verlag, 1968.
Bellman, R., *Dynamic Programming,* Princeton University Press, 1957.
Dreyfus, S. E., *Dynamic Programming and the Calculus of Variations*, Academic Press, 1965.
Howard, R. A., *Dynamic Programming and Markov Processes*, John Wiley and Sons, Inc., 1960.
Nemhauser, G. L., *Introduction to Dynamic Programming*, John Wiley and Sons, Inc., 1967.
Watanabe, S., *Knowing and Guessing*, John Wiley and Sons, Inc., 1969.

Mathematical Inequalities
Beckenbach, E. and Bellman, R., *An Introduction to Inequalities*, Random House, 1961.
Kazarinoff, N. D., *Geometric Inequalities*, Random House, 1961.
Kazarinoff, N. D., *Analytic Inequalities*, Holt, Rinehart and Winston, 1964.

Chapter Two - Net Present Value and Discount Rates

Section 2.1 - Time Preferences and Discount Rates

Time must be utilized as it becomes available. Because time cannot be stored, it must be accumulated by trade-offs of money and physical goods and services. Consequently, the management of time controls the constraints on money and resources. In this respect, time lies at the heart of all economic decision-making.

There is a continuous need to decide how and when projects should be done. It is not a simple matter to recognize different alternatives or to estimate their cash flows. You can buy or lease; make or buy; keep or replace; do earlier or later; substitute machines for labor etc... After recognizing these alternatives, their cash flows and present values must be compared systematically. The problems of recognizing these alternatives are resolved by identifying trade-offs of time, input costs and output revenues. The time value of money is the element that makes comparisons of such alternatives possible.

For example, suppose a company discovers an oil field. If only one well was drilled, it might take 100 years to drain the oil pool. If 50 wells were drilled, it might take only 2 years to drain the oil pool, but the cost of drilling 50 wells would be prohibitive. The best *engineering alternative* depends on trade-offs between increased costs of more wells and increased present-value revenues from earlier oil production.

Engineering alternatives deal with the physical aspects of trade-offs between time, input costs and output revenues. A classic example is that of two types of competing equipment, one with greater input cost which is offset either by better quality and/or quantity output, lower operating cost, or longer lifespan than the other. Owing to the time lags between inputs and outputs, current input costs need to be paid for by borrowing money until sufficient output is sold to pay back the borrowed money with interest. Thus, engineering alternatives are tied to *financial alternatives* through costs of borrowing money which are experienced by the organization as a whole.

It is commonly thought the best alternative has the largest net present-value (NPV). But NPV is not defined until the discount rate is specified. If discount rates are too low, alternatives with large NPV may have output revenues in the distant future. Present values of distant revenues can be reduced with higher discount rates. How high the discount rates should be is an open question in both theory and practice. Let us compare the NPV of alternatives at various discount-rate magnitudes.

Figure 2.1.1 - Insensitivity of net present-value comparisons to positive discount rates.

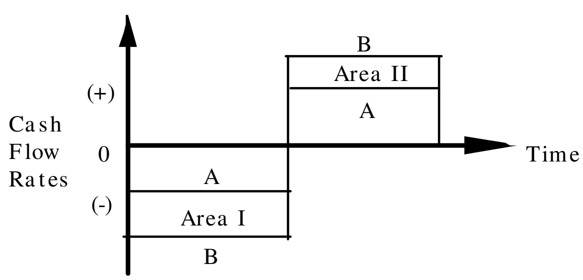


Figure 2.1.1 compares continuous cash-flow alternative A to B whose greater costs (Area I) are larger than the benefit of its increased revenues (Area II) relative to A. This implies that the cumulative cash flows of A, CCF{A}, are *always greater than* CCF{B}. If CCF{A} ≥ CCF{B}, it is proved in Appendix 2A that NPV{A} would be greater than NPV{B} at all positive discount rates.

Let us compare discrete cash-flow alternatives A and B whose end-of-year cash flows are CF{A}(-\$1.00;\$2.50) and CF{B}(-\$1.20;\$2.40). Hence, we calculate CCF{A}(-\$1.00:\$1.50) and CCF{B}(-\$1.20;\$1.20). Since CCF{A} > CCF{B} at all times, it follows NPV{A} > NPV{B} for all positive discount rates. The breakeven discount rate, i_B, at which NPV{A} = NPV{B} is determined from the equation $-1+2.5/(1+i_B) = -1.2+2.4/(1+i_B)$, the solution of which is $(1+i_B) = -0.5$ and $i_B=-1.5$. Hence, NPV{A} > NPV{B} at all discount rates greater than i_B=-150%/year.

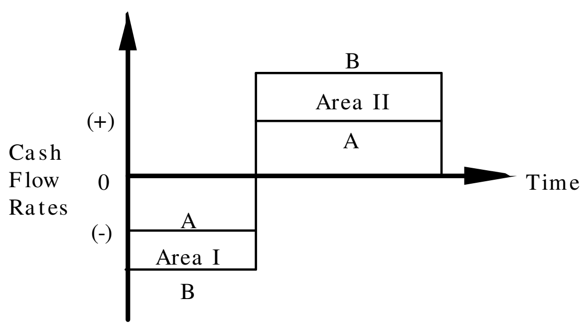
More generally, when net present-value equation $-1+2.5/(1+i) = -1.2+2.4/(1+i)$ is multiplied by (1+i), we obtain net future-value (NFV) equation $-(1+i)+2.5 = -1.2(1+i)+2.4$ at the end of the first year. Both NPV and NFV values for alternatives A and B are listed in Table 2.1.1 below for discount rates ranging from -150% to +150% per year. It is worth noting several results. When i_B=-150%/year, NPV{A}=NPV{B} just as NFV{A}=NFV{B}. For discount rates greater than -150%/year, NPV{A}>NPV{B} just as NFV{A}>NFV{B}. Thus, NPV and NFV comparisons of alternatives are equivalent. Because NPV functions of the discount rate are more complex than NFV functions, the latter may be preferred for analysis.

Table 2.1.1 - NPV and NFV comparisons of alternatives A and B.

Discount Rate	NPV{A}	NPV{B}	NFV{A}	NFV{B}
-150%/yr	-6.0	-6.0	3.0	3.0
-100%/yr	±∞	±∞	2.5	2.4
-50%/yr	4.0	3.6	2.0	1.8
0%/yr	1.5	1.2	1.5	1.2
50%/yr	0.67	0.4	1.0	0.6
100%/yr	0.25	0.0	0.5	0.0
150%/yr	0.0	-0.24	0.0	-0.6

Figure 2.1.2 is similar to Figure 2.1.1, but B is on top of Area II. Since Area I is smaller than Area II, CCF{A} are ***not** always greater than or equal to* CCF{B}. It is proved in Appendix 2A that NPV and NFV comparisons of alternatives would then be sensitive to the magnitude of positive discount rates.

Figure 2.1.2 - Sensitivity of NPV and NFV comparisons to positive discount rates.

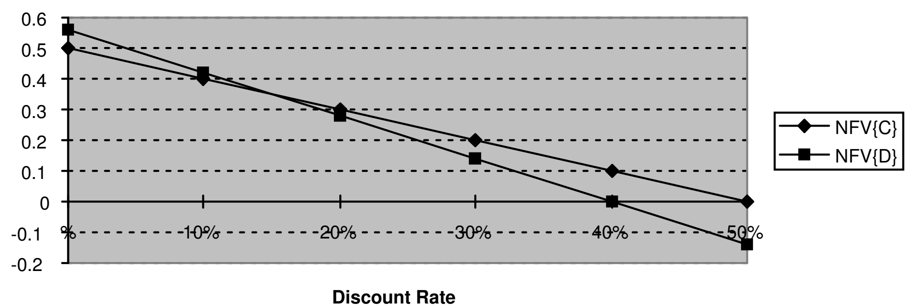


To fix ideas, let us compare alternatives C and D whose end-of-year cash flows are CF{C}(-\$1.00;\$1.50) and CF{D}(-\$1.40;\$1.96). Since CCF{C}(-\$1.00:\$0.50) is *not always greater than* CCF{D}(-\$1.40;\$0.56), NPV and NFV comparisons of C and D are sensitive to positive discount rate magnitudes. The breakeven discount rate, i_B, is determined from the equation $-1+1.5/(1+i_B) = -1.4+1.96/(1+i_B)$, the solution of which is $(1+i_B) = 1.15$ and $i_B = 0.15$. For discount rates greater than 15%/year, NPV{C} > NPV{D}. For discount rates smaller than 15%/year, NPV{C} < NPV{D} (see Table 2.1.2 and Figure 2.1.3).

Table 2.1.2 - NPV and NFV comparisons of alternatives C and D.

Discount Rate	NPV{C}	NPV{D}	NFV{C}	NFV{D}
0%/yr	0.50	0.56	0.50	0.56
10%/yr	0.36	0.38	0.40	0.42
20%/yr	0.25	0.23	0.30	0.28
30%/yr	0.15	0.11	0.20	0.14
40%/yr	0.07	0.00	0.10	0.00
50%/yr	0.00	-0.09	0.00	-0.14

Figure 2.1.3 - NFV{C} versus NFV{D}



Let us now compare two-period alternatives E and F whose end-of-year cash flows are CF{E}(-\$1.00;\$0.60;\$0.72) and CF{F}(-\$2.50;\$1.40;\$1.568). The cumulative cash flows CCF{E}(-\$1.00;-\$0.40;\$0.32) are *not always greater than* CCF{F}(-\$2.50;-\$1.10;\$0.468). Therefore, NPV and NFV comparisons of E and F are sensitive to positive discount rates.

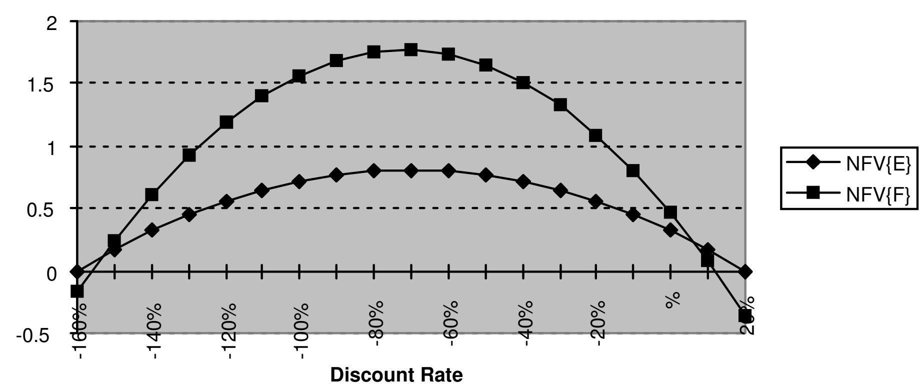
The breakeven discount rate, i_B, at which NPV{E} = NPV{F} is determined from the equation $-1.00 + 0.60/(1+i_B) + 0.72/(1+i_B)^2 = -2.50 + 1.40/(1+i_B) + 1.568/(1+i_B)^2$. The terms on the left side of the equation can be subtracted from both sides of the equation to get NPV{F-E} $= -1.50 + 0.80/(1+i_B) + 0.848/(1+i_B)^2 = 0$.

The NPV{F-E} equation can be multiplied by $(1+i_B)^2$ to get the equation NFV{F-E} $= 1.50(1+i_B)^2 - 0.80(1+i_B) - 0.848 = 0$ for net future values at the end of the second year. The NFV quadratic equation has two solutions, namely, $(1+i_{B1}) = 1.0644$ and $(1+i_{B2}) = -0.5311$, or $i_{B1} = 6.44\%$ and $i_{B2} = -153.11\%$/year (see Appendix 2B). Hence, for discount rates greater than 6.44%/yr, NFV{E} > NFV{F}. For discount rates from -153.11% to 6.44%/yr, NFV{E} < NFV{F}. For discount rates less than -153.11%/yr, NFV{E} > NFV{F}. Similar results are obtained with NPV comparisons (see Table 2.1.3 and Figure 2.1.4).

Table 2.1.3 - NPV and NFV comparisons of alternatives E and F.

Discount Rate	NFV{E}	NFV{F}	NPV{E}	NPV{F}
-160%/yr	0.00	-0.17	0.00	-0.48
-140%/yr	0.32	0.61	2.00	3.80
-120%/yr	0.56	1.19	14.00	29.70
-100%/yr	0.72	1.57	∞	∞
-80%/yr	0.80	1.75	20.00	43.70
-60%/yr	0.80	1.73	5.00	10.80
-40%/yr	0.72	1.51	2.00	4.19
-20%/yr	0.56	1.09	0.88	1.70
0%/yr	0.32	0.47	0.32	0.47
20%/yr	0.00	-0.35	0.00	-0.24

Figure 2.1.4 - NFV{E} versus NFV{F}



The graphs in Figure 2.1.4 are parabolas as explained in Appendix 2B. The data of Tables 2.1.1, 2.1.2 and 2.1.3 are more easily understood when presented in the form of charts or graphs such as Figures 2.1.3 and 2.1.4 . Graphical presentations are readily generated from spreadsheet data in the Chart Wizard program of Microsoft Excel as follows:

Two Preliminary Steps of Chart Wizard in Excel - (1) Select the worksheet data such as the first three columns of Table 2.1.3 to be displayed in Chart Wizard. (2) Click the Chart Wizard button. The cursor changes to a cross-hair and a chart symbol. Click anywhere on the worksheet where you want the chart to be placed. The chart location can be moved afterwards and its size can be adjusted.

Five-step Procedure of Chart Wizard in Excel.

1. Chart Wizard menu asks the user to verify the range of the workseet area. [Next]
2. Select 1 of 15 chart-type choices. [Next]
3. Select 1 of 10 styles for chart-type of step 2. [Next]
4. Sample chart review for variables in Rows or Columns until satisfactory. [Next]
5. Add title of chart, names of variables, and change legend colors. [Finish]

Section 2.2 - Net Present Value and Engineering Discount Rates

The measurement of marginal capital efficiency has led engineers to focus on discount rates at which the net present-value, NPV, and net future-value, NFV, of the least profitable alternative is zero. By definition, the *positive* discount rate at which the NPV and NFV of an alternative is zero is called its *internal rate of return* (IRR). Independent project alternatives are ranked in descending order of their internal rates of return until the cut off of available funds. The IRR of the worst accepted or best rejected alternative is defined as the *minimum attractive rate of return*, *MARR*, which is the discount rate used to determine the NPV of all alternatives of the organization.

The concepts of MARR and marginal capital efficiency have similarities, but also basic differences. Both concepts subscribe to the *theory of opportunity costs* as a framework of economic definitions and measurements. Opportunity cost is defined as the benefit foregone, or opportunity lost, from the least profitable increment of input costs that was not undertaken because of capital constraints. Opportunity costs arise in a world of economic scarcity where available funds and resources are not enough to satisfy all the wants and needs of society. If funds and resources were unlimited, all investments would be undertaken at their own expense, and not at the cost of foregone benefits of rejected alternatives. Opportunity cost is the most fundamental concept of economics because there would be no need to economize in a world with limitless funds and resources.

Cost in economics always means opportunity cost. But engineers, accountants and economists treat opportunity costs very differently. Engineers measure opportunity costs by net present values discounted at MARR. Accountants deal with costs that can be identified on a cash or accrual basis which can be audited in financial statements. Accounting costs differ widely from an economist's opportunity costs if they fail to reflect the highest-value use of available resources. Economists generally look at foregone benefits of projects which were rejected because of limited resources, whereas accountants ignore best rejected projects because they are not recorded in financial statements.

Measurements of opportunity costs cannot be more accurate than the precision of their definition. Opportunity costs occur only with increments or decrements of input costs under an effective capital constraint. More specifically, Ricardo's marginal principle states the necessary and sufficient condition for a group of investments to earn the greatest amount of money from a given cost of investment is that *any* accepted investment makes more money for its capital cost than *any* investment with the same capital cost which has been rejected due to the capital constraint. Assuming accurate forecasts of input costs, output revenues and market rates of interest, measurements of the marginal output/input ratio, $Ø_m$, in the range of the capital constraint were proved to satisfy Ricardo's marginal principle in Chapter One. The major differences between the marginal output/input ratio, $Ø_m$, and conventional engineering measurements of net present values discounted at MARR are:

1. The increment or decrement of input costs under an effective capital constraint may not be the same as the input costs of the least profitable accepted project whose NPV equals zero when discounted at MARR. It is the task of economic analysis to determine the marginal output/input ratio, $Ø_m$, rather than to find the least profitable project by ranking a typical set of projects until available funds are exhausted. From the viewpoint of economics, all projects do equally important things. The object of the organization as a whole is to find the increment or decrement of input costs which maximize its overall NPV discounted at market rates of interest rather than to maximize the NPV of each project by selecting alternatives with the greatest NPV discounted at MARR.

2. The available funds of an organization may not be an effective capital constraint. If investment opportunities require *less* funds than an organization has available, then the excess should be invested in money market funds while waiting for new opportunities. If investment opportunities require *more* funds than is available, then debt and equity financing should provide additional capital. The effective capital constraint is the debt and equity funding which is obtained in conjunction with the marginal output/input ratio, $Ø_m$, for determining the best ways of doing each project and the best projects to do.

3. There is considerable room for questioning the significance of internal rate of return as a criterion for ranking alternatives as outlined below.

a. Unequal discounting - Let us consider the internal rates of return (IRR's) of alternatives E and F which are depicted at the intersections of their parabolas with the positive discount-rate axis in Figure 2.1.4 . The IRR's of CF{E}(-$1.00;$0.60;$0.72) and CF{F}(-$2.50;$1.40;$1.568) are 20% and 12%/year respectively. Since IRR{E} > IRR{F}, E is better than F according to the IRR criterion. However, this conclusion was obtained by comparing the cash flows of E and F at different discount rates even though their cash flows occur at the same times and under the same economic conditions.

If cash flows of E and F were discounted with equal rates, then NPV{E} > NPV{F} with discount rates greater than 6.44%/year, and NPV{E} < NPV{F} for discount rates between zero and 6.44%/year, as can be seen from Figure 2.1.4 . Therefore, breakeven discount rate, i_{B1} = 6.44%/year, is needed to compare E and F by the IRR criterion. Since IRR{E-F} = IRR{F-E} = 6.44%/year, the breakeven discount rate, by itself, cannot determine whether E or F is better. Consequently, the cash flows of E and F are evaluated at zero discount where NPV{E}=NFV{E}=0.32 < NPV{F}=NFV{F}=0.468 and MARR is compared to i_{B1} = 6.44%/year. If MARR < 6.44%/year, F is preferred to E because F has a greater undiscounted cash flow than E. If MARR > 6.44%/year, E is preferred to F because E's undiscounted cash flow is smaller than F's. As a result, preferences of the IRR criterion are the same as those of the NPV criterion discounted at MARR (see Chapter Four).

b. Constant returns to scale - Internal rates of return are independent of the scale of alternatives. For example, CF{E}(-$1.00;$0.60;$0.72) and CF{100E}(-$100;$60;$72) both have IRR's of 20% and -160%/year. As a result, parabolas of CF{E}, CF{100E}, CF{-E} and CF{-100E} all intersect at the same points on the discount-rate axis in Figure 2.1.4 . However, CF{100E} affects the capital constraint much more than CF{E}. The breakeven discount rates of IRR comparisons are sensitive to scale. For example, the breakeven discount rates of CF{F-E} in Section 2.1 are i_{B1} = 6.44% and i_{B2} = -153.11%/year as compared to those of CF{2F-E} which are i_{B1} = 9.94% and i_{B2} = -154.94%/year.

c. Insensitivity to lifespan differences - The IRR criterion may be insensitive to ranking alternatives whose lifespans differ. For example, alternatives A, B and C below have lifespans between one and three years and a common IRR = 20%/year.

CF{A}(-$1.00;$1.20) CF{B}(-$1.00;$0.60;$0.72) CF{C}(-$1.00;$0.60;$0.36;$0.432)

If MARR < 20%/year, NPV{C} > NPV{B} > NPV{A} > 0. But if MARR > 20%/year, 0 > NPV{A} > NPV{B} > NPV{C}. Thus, IRR ranks A, B and C the same regardless of MARR, but NPV ranks A, B and C differently than IRR when MARR ≠ 20%/year. See problem **2-5b**. The breakeven discount rate IRR{B-A} is also 20%/year. Unless the difference in proceeds between A and B at the end of the first year is reinvested at 20%/year for one year, then the two-year NPV comparison of A and B is not equivalent to their IRR comparisons.

d. <u>Degradation of cash flow information</u> - The NPV and NFV functions of the cash flows of multi-period alternatives are polynomials which have as many roots (i.e., internal rates of return) as the duration of the alternatives. The IRR definition for multi-period alternatives characterizes their cash flows by a single positive IRR constant while neglecting all other roots embedded in the cash flow coefficients of their NPV = NFV = 0 equations.

For example, a general two-period alternative has an input of -Z dollars and outputs of (X+Y) dollars the first year and (-XY/Z) dollars the second year. Dividing by Z results in a unitized input and outputs of q_1 = (X+Y)/Z and q_2 = ($-XY/Z^2$) dollars in the next two years. The IRR = i of the unit 2-period alternative CF{U_2}(-1;q_1;q_2) must satisfy NPV equation

$$-1 + (X+Y)/Z(1+i) - XY/Z^2(1+i)^2 = -[1 - X/Z(1+i)]\ [1 - Y/Z(1+i)] = 0 \qquad ....(2.2.1)$$

Since each factor of equation (2.2.1) must equal zero, we get $1+i_1 = X/Z$ and $1+ i_2 = Y/Z$ where $i_1 = IRR_1$ and $i_2 = IRR_2$. Moreover, $q_1 = (1+i_1)+(1+i_2) = 2+i_1+i_2$ and $q_2 = -(1+i_1)(1+i_2)$. Therefore, the cash flow of a unit 2-period alternative is (-1; X+Y; -XY) which can also be expressed as (-1; q_1; q_2) and (-1; $2+i_1+i_2$; $-(1+i_1)(1+i_2)$). Table 2.2.1 below lists a set of unit-cost, two-period alternatives, all of which have $i_1 = IRR_1$ = 20%/year and -260% < i_2=IRR_2 < 60%/year.

Table 2.2.1 - Unit-cost, two-period alternatives (i_1 = 20%/yr and -260% < i_2 < 60%/yr).

#	Z	X	Y	q_1	q_2	i_1	i_2	NPV@0%/yr	NPV@30%/yr
A	-1	1.2	-1.6	-0.4	1.92	20%	-260%	$0.52	-$0.172
B	-1	1.2	-1.2	0.0	1.44	20%	-220%	$0.44	-$0.145
C	-1	1.2	-0.8	0.4	0.96	20%	-180%	$0.36	-$0.124
D	-1	1.2	-0.4	0.8	0.48	20%	-140%	$0.28	-$0.101
E	-1	1.2	-0.0	1.2	0.00	20%	-100%	$0.20	-$0.077
F	-1	1.2	0.4	1.6	-0.48	20%	-60%	$0.12	-$0.053
G	-1	1.2	0.8	2.0	-0.96	20%	-20%	$0.04	-$0.030
H	-1	1.2	1.2	2.4	-1.44	20%	+20%	-$0.04	-$0.006
I	-1	1.2	1.6	2.8	-1.92	20%	+60%	-$0.12	+$0.018

By definition, IRR = 20%/yr ranks A ... I the same regardless of MARR. If MARR is less than 20%/yr, NPV{A} > ... > NPV{I}. If MARR is more than 20%/yr, NPV{I} > ... > NPV{A}. Thus, NPV ranks A ... I differently than IRR when MARR ≠ 20%/yr because NPV keeps all cash flow information intact during the discounting process.

The quadratic formula in equation (2B.7) of Appendix 2B indicates that when the discriminant b^2-4ac is negative, the roots of a quadratic equation will be complex conjugates. Equivalently, in the notation given above, if $(X+Y)^2-4XY = (X-Y)^2 < 0$, then the internal rates of return will be complex conjugates. In this connection, suppose a person deposits $1 in a bank in order to borrow $4 one year later. If the bank requires the loan to be paid back with $5 a year later, what are the internal rates of return of this <u>f</u>inancial <u>t</u>ransaction, FT? The IRRs of the FT cash flow, denoted by CF{FT}(-$1;$4;-$5) must satisfy NPV equation which was factorized in equation (2.2.2) with the help of the quadratic formula.

$$-1 + 4/(1+i) - 5/(1+i)^2 = -[1 - (2+\sqrt{-1})/(1+i)]\ [1 - (2-\sqrt{-1})/(1+i)] = 0 \qquad ....(2.2.2)$$

The solutions of equation (2.2.2) are $1+i_1 = 2+j$ and $1+i_2 = 2-j$ where j denotes the imaginary unit $\sqrt{-1}$. Therefore, $i_1 = 1+j$ and $i_2 = 1-j$. The complex rates of return indicate that the internal rates of return are real variables rather than imaginary constants. We now show that complex rates of return are equivalent to real varying rates of return.

Tables 2.2.2a and 2.2.2b below providean accounting system for EOY (end-of-year) cash flows with BCF (before-cash-flow) balances and ACF (after-cash-flow) balances of a bank using interest rates $i_1 = 1+j$ and $i_2 = 1-j$ respectively during the year.

Table 2.2.2a - Cash Flow Accounting System with Imaginary Bank Interest 1+j per year

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	1+j	$0.00	$1.00	$1.00
1	1+j	2+j	-$4.00	-2+j
2	1+j	$-4+j^2$	$5.00	$0.00

Table 2.2.2b - Cash Flow Accounting System with Imaginary Bank Interest 1-j per year

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	1-j	$0.00	$1.00	$1.00
1	1-j	2-j	-$4.00	-2-j
2	1-j	$-4+j^2$	$5.00	$0.00

Columns (3), (4) and (5) of Tables 2.2.2a and 2.2.2b are now added to get Table 2.2.2c.

Table 2.2.2c - Cash Flow Accounting System with Real Bank Interest i_1 and i_2 per year

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	1	$0.00	$2.00	$2.00
1	i_1	$4.00	-$8.00	-$4.00
2	i_2	-$10.00	$10.00	$0.00

From BCF balances $4.00 and -$10.00, and ACF balances $2.00 and -$4.00, the following two equations are formed which can be solved for real interest rates i_1 and i_2. The results verified in Table 2.2.2d suggest real IRRs may vary just like market rates of interest.

$\$2.00(1+i_1) = \$4.00; \quad i_1 = 1.00$ or 100%/yr
$-\$4.00(1+i_2) = -\$10.00; \quad i_2 = 1.50$ or 150%/yr

Table 2.2.2d - Cash Flow Accounting System with Bank Interest 100% and 150% per year

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	100%	$0.00	$1.00	$1.00
1	100%	$2.00	-$4.00	-$2.00
2	150%	-$5.00	$5.00	$0.00

e. Definitional deficiencies - There are two-period alternatives that have IRRs which are either both positive, negative or complex conjugates as shown in Tables 2.2.1, 2.2.2a-d, and Exercise 2-4a, b, c, d. When alternatives lack a single positive IRR, it is not clear how to rank projects in descending order of the IRR criterion for the determination of MARR. A portion of the cash flows of such alternatives are borrowed at external rates of return (ERR) in order to obtain single positive IRRs (see Chapter Four).

f. Nonadditive rankings - Nonadditive rankings of IRR are shown by alternatives {A}(-$1.00;$1.50) whose IRR{A}=50%/yr, {B}(-$1.20;$1.60) whose IRR{B}=33.33%/yr, and {C}(-$0.90;$1.18) whose IRR{C}=31.11%/yr. Since IRR{A} > IRR{B} > IRR{C}, we would expect IRR{A+B} > IRR{A+C} if IRR rankings were additive. However, {A+B}(-$2.20;$3.10) has IRR{A+B}=40.91%/year and {A+C}(-$1.90;$2.68) has IRR{A+C}=41.05%/year. Therefore, IRR{A+B}=40.91% < IRR{A+C}=41.05%/year contrary to our expectations It follows that ranking alternatives by the IRR criterion may not satisfy Ricardo's marginal principle.

Section 2.3 - Net Present Value and Financial Discount Rates

The concept of *opportunity cost* is the framework from which financial discount rates are defined. However, accounting models from which financial discount rates are determined often assume that the objective of a firm is to maximize the market value of its stockholder's equity. An optimal ratio of debt to equity capital is assumed to exist as a result of *financial leveraging* from which the market value of the firm is derived. The firm's *cost of capital* is then defined as the minimum rate of return which is required for new investments. Thus, the cost of capital and the minimum rate of return are interchangeable terminologies in the definition of financial discount rates.

Optimal debt to equity ratios are evaluated from four sources of capital formation, namely, (1) the firm's net operating income, (2) short-term debt financing, (3) long-term debt financing and (4) the sale of capital stock. The firm's net operating income and sale of capital stock are coupled together as an expected rate of return of equity investors from dividends, capital gains and retained earnings. These rates of return are weighted by their proportions in the firm's balance sheet, and the resulting financial discount rates are called weighted-average-costs-of-capital (WACC).

The WACC determination of financial discount rates is not adaptable to financial institutions, government agencies and nonprofit organizations whose sources of capital are quite different than those of industrial firms. Discount rates derived from from the WACC model are usually much higher than costs of borrowing money because of high investor expectations and upward adjustments to account for risk and inflation. Consequently, short-term returns are emphasized much more than long-term productivity gains.

Financial discount rates are based on the distinction between debt and equity. Although distinguishing between debt and equity is necessary for financial accounting, it does not serve the needs of managerial accounting. More specifically, it does not matter in the cash flow description of an alternative whether the input costs came from either debt or equity sources. The issue is not so much the *source* of input costs as it is how the input costs are used to generate output revenues. The market value of stockholder's equity is important. But the organization as a whole needs to maximize its net present value in order for the market value of stockholder's equity to be maximized.

It is commonly thought that financial discount rates should be greater than market rates of interest in order to pay for risks undertaken by equity investors. But risks of equity investors are not the only kinds of risks that need to be evaluated. Debt financing requires greater rates of interest for loans with a greater risk of default. Engineering risks of not completing the input requirements of an alternative are much different than the marketing risks of not selling its output above cost. Similarly, risks of inflation do not equally affect the input costs and output revenues of alternatives. Consequently, lumping risk, inflation and investor expectations into financial discount rates is a hindrance to the best possible use of available engineering and financial information in economic forecasting.

Financial discount rates are based on *investment opportunity costs* estimated from net operating incomes and costs of borrowed money in financial accounting statements. The proposed discount-rate model is based on *borrowing opportunity costs* estimated from future costs of borrowing money at market rates of interest. At present, the determination of discount rates, both before and after income taxes, are based on investment rather than borrowing opportunity costs. This artificially increases discount rates which obscure the future by emphasizing early returns as opposed to long-term productivity gains.

To fix ideas on the differences between various discount-rate models, let us first neglect income taxes and suppose the marginal investment of a firm is a single-period alternative with a rate of return of 15%/year. Part of the 15%/year marginal investment could be financed by borrowing money from a bank at 10%/year interest and the remaining input cost would come from equity sources. The firm's *opportunity cost* is defined as the benefit foregone by *not* undertaking the 15%/year investment and *not* borrowing any of its input cost from a bank at 10%/year interest. The NPV of the investment could be discounted at a) the 15%/year *internal rate of return* of the total cash flow (TCF), b) the 10%/year *internal rate of return* of debt cash flow (DCF), or c) the *internal rate of return* of the remaining equity cash flow (ECF).

In order to discuss the three discount-rate concepts, suppose the total cash flow consists of an end-of-year (EOY) input of $1 and yields an output of $1.15 one year later for an internal rate of return (IRR) of 15% per year (i.e., -1.00 + 1.15/(1+i) = 0 or 1+i = 1.15 and i = 15%/year). The $1 input cost is financed with 50% equity and 50% debt from a bank which lends money at an interest rate of 10% per year. The $0.50 bank loan is paid back with $0.50*1.10 = $0.55 one year later, leaving a residual of $1.15-$0.55 = $0.60 for equity investors whose IRR is 20% per year (i.e., -0.50 + 0.60/(1+i) = 0 or 1+i = 0.60/0.50 = 1.20 and i = 20%/year). The NPV of the total, debt and equity cash flows are listed in Table 2.3.1 discounted at annual percent rates (APR) of 15%, 10% and 20% per year respectively.

Table 2.3.1 - NPV of total, debt and equity cash flows of a one-period marginal investment.

EOY Cash Flows	APR	NPV@15%/yr	NPV@10%/yr	NPV@20%/yr
TCF(-$1;$1.15)	15%	0	$.045	-$.042
DCF(-$.50;$.55)	10%	-$.022	0	-$.042
ECF(-$.50;$.60)	20%	$.022	$.045	0

APR 15%/year - The NPV{TCF}(-$1;$1.15) of the total cash flow discounted at its 15%/yr internal rate of return (IRR) results in NPV{TCF} = 0. When alternatives are required to have a greater IRR and NPV than the marginal investment, then IRR = 15%/yr acts as a constraint called the minimum attractive rate-of-return (MARR), and NPV = 0 discounted at MARR acts as joint constraint. These MARR and NPV constraints on the firm leave much room for questioning because there is a 15%/year rate of return on the $0.50 equity which has a net yield of $0.50*0.15 = $0.075, and a 15%-10%=5%/year rate of return on the $0.50 debt which has a net yield of $0.50*0.05 = $0.025 . Discounting at 15%/yr is also misleading because NPV{ECF} = $.022 is equal and opposite to NPV{DCF} = -$.022 even though equity and debt have net earnings of $0.075 and $0.025 respectively.

APR 20%/year - The NPV{ECF}(-$.50;$.60) of equity cash flow discounted at its 20%/year internal rate of return (IRR) results in NPV{ECF} = 0. The 20%/year discount rate results in NPV{TCF} = NPV{DCF} = -$.042 for both the total and debt cash flows. When alternatives are required to have a greater IRR and NPV than the equity cash flow, then the marginal investment under consideration could not be undertaken. Such IRR and NPV constraints are used to take into account risk, inflation and equity investor expectations.

APR 10%/year - The NPV{DCF}(-$.50;$.55) of the debt cash flow discounted at its 10%/year internal rate of return (IRR) results in NPV{DCF} = 0. The 20%/year discount rate results in NPV{TCF} = NPV{ECF} = $.045 for both the total and equity cash flows. The marginal capital efficiency, $Ø_m$, of the marginal investment is $\Delta R_m/\Delta C_m$ = (1.15/1.10)/1 = 1.045 . Differences between alternative ways of doing each project discounted at 10%/year could then be compared to $Ø_m$ = 1.045 to determine the best way of doing each project.

Let us generalize the results of Table 2.3.1 by two managerial accounting identities.

$$TCF(i) \equiv DCF(i) + ECF(i) \quad \text{for end-of-year } i = 0, 1, 2, ... \qquad ....(2.3.1)$$

$$NPV\{TCF\} \equiv NPV\{DCF\} + NPV\{ECF\} \quad \text{at all discount rates} \qquad ....(2.3.2)$$

The application of identities (2.3.1) and (2.3.2) to single-period marginal investments with an internal rate of return of w%/year is described in Exercise 2-3a at the end of this chapter. The firm could borrow a fraction x of the unit investment by debt financing at y%/year. The net present value of the marginal investment could be discounted at a) the *internal rate of return* of w%/year of the total cash flow, TCF, b) the *internal rate of return* of y%/year of the debt cash flow, DCF, or c) the *internal rate of return* of the equity cash flow, ECF. The end-of-year (EOY) cash flows satisfying equation (2.3.1) is described in Table 2.3.2 which is followed by NPV discounted at annual percentage rates, APR, satisfying equation (2.3.2). The marginal capital efficiency of the firm is $Ø_m = (1+w)/(1+y)$.

Table 2.3.2 - Generalized TCF, DCF and ECFs of one-period marginal investments.

EOY	TCF	=	DCF	+	ECF
0	-1	=	-x	+	-1+x
1	1+w	=	x(1+y)	+	(1+w)-x(1+y)

APR/yr	NPV{TCF}	=	NPV{DCF}	+	NPV{ECF}
0%	w	=	xy	+	w-xy
y%	(w-y)/(1+y)	=	0	+	(w-y)/(1+y)
w%	0	=	x(y-w)/(1+w)	+	x(w-y)/(1+w)
$\frac{w-xy}{1-x}$ %	$\frac{x(y-w)}{1+w-x(1+y)}$	=	$\frac{x(y-w)}{1+w-x(1+y)}$	+	0

Table 2.3.2 shows that discounting at the IRR of the debt cash flow, APR = y%/year, NPV(TCF) = NPV(ECF) independently of the fraction x of the input cost by debt financing. Therefore, when discounting at the cost of borrowing money, NPV(ECF) always equals NPV(TCF) regardless of the fraction of the unit input cost of the marginal investment that is borrowed. Since money is borrowed against the assets of the organization as a whole, the fraction of the input cost that is borrowed is indeterminate.

The IRR of the equity cash flow is (w-xy)/(1-x). Thus, when w=0.15, y=0.10 and x=0.50 as in Table 2.3.1, then (w-xy)/(1-x) = 0.20 or 20%/year. Keeping w and y fixed as x is increased from 0.50 to 0.75 results in (w-xy)/(1-x) = 0.30 or 30%/year. Again keeping w and y fixed as x is increased from 0.75 to 0.90 results in (w-xy)/(1-x) = 0.60 or 60%/year. Because IRR{ECF} is so sensitive to indeterminate fractions of debt financing, marginal investments are usually evaluated entirely with equity funding which results in NPV{ECF} = 0.

Although the IRR of the marginal investment may be insensitive to its lifespan, the marginal capital efficiency, $Ø_m$, is dependent on the lifespan. In this regard, let us construct a sequence of one-year, two-year, etc. marginal investments with IRR = 15%/year according to the model outlined in Exercise 2-5a at the end of this chapter. Thus, A(-$1.00;$1.15) has $Ø_m\{A\}$ = 1.045 discounted at 10%/year cost of borrowing money; B(-$1.00;$0.575;$0.66125) has $Ø_m\{B\}$ = 1.0692 discounting at 10%/year; and C(-$1.00;$0.575;$0.330625;$0.3802188) has $Ø_m\{C\}$ = 1.1375 discounting at 10%/year. This suggests the marginal capital efficiency of marginal investments should be evaluated by discounting at the cost of borrowing money.

In general, a firm borrows money for projects that earn more than banks charge for their loans. If the firm in Table 2.3.1 increased its debt financing from 50% to 75%, the APR of equity investors would increase from 20% to 30%/year. But the net return to the equity investors decreases from \$.60-\$.50 = \$.10 to \$.325-\$.25 = \$.075 while the net return to the bank increases from \$.55-\$.50 = \$.05 to \$.825-\$.75 = \$.075 . To avoid using up investment opportunities for the bank's benefit, the firm should borrow as little as possible.

Let us now determine discount rates from a two-period marginal investment with an internal rate of return of 15%/year. The firm could finance part of the two-period 15%/year investment by borrowing money from a bank at 10%/year interest. Tables 2.3.3 and 2.3.4 below use 50% and 75% debt financing respectively with ±0.1% approximations of the cash flows and APR discount rates.

Table 2.3.3 - NPV of total, debt and equity of a two-period marginal investment.

EOY Cash Flows	APR	NPV@15%/yr	NPV@10%/yr	NPV@20%/yr
TCF{-\$1;\$.615;\$.615}	15%	0	\$.067	-\$.060
DCF{-\$.50;\$.288;\$.288}	10%	-\$.032	0	-\$.060
ECF{-\$.50;\$.327;\$.327}	20%	\$.032	\$.067	0

Table 2.3.4 - NPV of total, debt and equity of a two-period marginal investment.

EOY Cash Flows	APR	NPV@15%/yr	NPV@10%/yr	NPV@30%/yr
TCF{-\$1;\$.615;\$.615}	15%	0	\$.067	-\$.163
DCF{-\$.75;\$.432;\$.432}	10%	-\$.048	0	-\$.163
ECF{-\$.25;\$.183;\$.183}	30%	\$.048	\$.067	0

<u>APR 15%/year</u> - The NPV{TCF}(-\$1;\$.615;\$.615) of <u>t</u>otal <u>c</u>ash <u>f</u>low (TCF) discounted at its 15%/yr IRR results in NPV{TCF} = 0. As previously mentioned, the IRR of the marginal investment is called the <u>m</u>inimum <u>a</u>ttractive <u>r</u>ate-of-<u>r</u>eturn (MARR). Discounting at MARR results in equal and opposite values of NPV for the <u>d</u>ebt and <u>e</u>quity <u>c</u>ash <u>f</u>lows (DCF) and (ECF) because of the fundamental equations (2.3.1) and (2.3.2)

$$TCF(i) \equiv DCF(i) + ECF(i) \quad \text{for end-of-year } i = 0, 1, 2, ... \qquad ....(2.3.1)$$

$$NPV\{TCF\} \equiv NPV\{DCF\} + NPV\{ECF\} \quad \text{at all discount rates} \qquad ....(2.3.2)$$

which hold true at all discount rates (see Exercise 2-3a). Since NPV{TCF} = 0 discounted at MARR, it follows from equation (2.3.2) that NPV{DCF} = -NPV{ECF}. Thus, discounting at MARR induces the firm to ignore the benefits it could derive from the debt cash flow. Equations (2.3.1) and (2.3.2) are generalized in Chapter Seven to account for income taxes.

<u>APR 10%/year</u> - The NPV{DCF} = 0 when discounted at its 10%/year IRR. It follows from equation (2.3.2) that NPV{TCF} = NPV{ECF}. The marginal capital efficiency, $Ø_m$, of this two-period, 15%/year IRR marginal investment discounted at the 10%/yr cost of borrowing money is equal to $\Delta R_m/\Delta C_m = 1.067/1 = 1.067$ as compared to $Ø_m = 1.045$ for the one-period, 15%/year IRR marginal investment.

<u>APR 20% & 30%/year</u> - The NPV{ECF} = 0 when discounted at the 20% or 30%/year IRRs of their equity cash flows. It follows from equation (2.3.2) that NPV{TCF} = NPV{DCF} when the total and debt cash flows are discounted at the IRRs of the equity cash flows. Moreover, it is worth noting from Tables 2.3.1 to 2.3.4 that smaller fractions of equity in the total cash flow result in higher IRRs for the equity cash flows. When IRRs of equity cash flows are used for discounting, NPVs of total and debt cash flows are negative.

In summary, the WACC and MARR economic models for determining discount rates are based on *investment opportunity costs* that can be evaluated *internally* either from balance sheets or from ranking the IRR's of a current set of investment opportunities. In contrast, the proposed economic model for determining discount rates is based on *borrowing* *opportunity costs* that can be evaluated *externally* from market rates of interest. At present, the dominant economic models for determining discount rates are based on investment rather than borrowing opportunity costs, resulting in high discount rates which focus on shorter-term results as opposed to longer-term productivity gains.

Determining discount rates from given opportunity cost models is inseparable from the problem of formulating the opportunity cost model itself. Chapter One formulates an opportunity cost model consisting of the marginal capital efficiency, $Ø_m$, which is obtained from the last increment of input costs in the range of a given capital constraint. The convex-envelope method of optimal economic decision-making uses the costs of borrowing money for discounting the input costs and output revenues of every alternative. Financial discount rates that are used instead of market rates of interest are *costs of capital* whose average consists of debt and equity sources of capital in the firm's balance sheet. Engineering discount rates that are used instead of market rates of interest consist of internal rates of return of the worst accepted or best accepted investment at the cut off of available funds.

Opportunity costs determined by the IRR criterion have a different meaning than the economic theory of opportunity costs. Opportunity costs occur at the last increment of input costs subject to an effective capital constraint which is not the same as the last ranked project to be accepted. The last increment of input costs may occur between alternatives of the same project. We must first determine the best way of doing each project before the projects are ranked. Lastly, the last ranked project by IRR cannot be proved to make more money for its cost than any rejected project according to Ricardo's marginal principle.

The utility of the IRR definition is not so much as a ranking criterion as it is with specifying interest rates in financial loans. Borrowers and lenders have equal and opposite inputs and outputs in financial loans. Therefore, they have the same internal rates of return and they explicitly agree on a single real positive internal rate of return of the loan. But this does not change the fact that borrowers use loans to earn higher rates of return than lenders are charging as indicated by debt and equity differences in financial accounting.

When projects are ranked in descending order of their internal rates of return, the best rejected project is called a *do-nothing alternative*. Do-nothing alternatives are often confused with *ongoing alternatives* which are defined only for existing projects to separate the effects of past from present decision-making. A do-nothing alternative always has ΔNPV = 0 discounted at MARR. In contrast, the net present-value of an ongoing alternative may be positive, negative or zero when discounted at constant or varying market rates of interest.

This writer believes that the greater emphasis on investment opportunity costs as compared to borrowing opportunity costs distorts the meaning of economic scarcity. This results in major misunderstandings in both private and public sectors of the economy. For these reasons, detailed comparisons of economic decision-making with investment and borrowing opportunity cost models are presented in Chapters Four, Five and Six. In Sections 2.4 to 2.6 which follow, simple examples are given to explore the consequences of using MARR or the cost of borrowing money as discount rates in comparisons of mutually exclusive alternatives.

Section 2.4 - Time Acceleration of Engineering Alternatives

The first example illustrates how discount rates based on investment and borrowing opportunity costs affect the selection of engineering alternatives. The question of the best way to do an existing project includes the possibility that the ongoing alternative is best. Economic decision-making is concerned with challenges from other mutually exclusive alternatives which could replace the ongoing alternative in order to better satisfy the objective of the organization as a whole.

In particular, let us consider a problem of Lorie and Savage ("Three Problems in Rationing Capital," Journal of Business, XXVIII(4), Oct 1955): If nothing was done, an oil well would produce $10,000 worth of oil at the ends of the next two years. An engineer proposes to install a pump costing $1,600 which would produce all $20,000 worth of oil at the end of the first year at which time the pump would have no salvage value. In order to decide if the pump should replace the ongoing alternative, the questions are, "Which alternative is better?" and "At what pump cost would the two alternatives be equally good?".

The Lorie and Savage problem illustrates a choice of realizing the $20,000 output sooner rather than later. Pump alternative CF{P}(-$1,600;$20,000;$0) is the challenger and ongoing alternative CF{Ω}($0;$10,000;$10,000) is the defender. Cumulative Cash Flows of P and Ω are CCF{P}(-$1,600;$18,400;$18,400) and CCF{Ω}($0;$10,000;$20,000). Because CCF{P} is not always larger or smaller than CCF{Ω}, the comparison of NPV{P} and NPV{Ω} is sensitive to the magnitude of the discount rate (see Section 2.1). The discount rates at which NPV{P} equals NPV{Ω} may be found by setting NPV{P-Ω}(-$1,600;$10,000;-$10,000) equal to zero and solving the resulting quadratic equation as shown in Appendix 2B. The results are NPV{P} = NPV{Ω} when i_1 = 25% and i_2 = 400% per year.

In order to show how discount-rate magnitudes affect the comparison of [P] and [Ω], let MARR equal 30% per year and the cost of borrowing money equal 15% per year. Tables 2.4.1 and 2.4.2 have present-value calculations of ΔNPV, ΔR and ΔC for [P] and [Ω] at discount rates of 30% and 15% per year respectively. The results of Tables 2.4.1 and 2.4.2 are plotted below each table in Figures 2.4.1 and 2.4.2 respectively.

Table 2.4.1 - Pump Versus Ongoing Alternatives at a Discount Rate of 30%/Year

² NPV = Net Present-Value	=	² R = Present-Value Output Revenues	-	² C = Present-Value Input Costs
ΔNPV[P] = $13,785	=	$\frac{\$20,000}{1.30} + \frac{\$000}{1.30^2} = \$15,385$	-	$1,600
ΔNPV[Ω] = $13,610	=	$\frac{\$10,000}{1.30} + \frac{\$10,000}{1.30^2} = \$13,610$	-	$000
ΔNPV[P-Ω] = $175	=	$\frac{\$10,000}{1.30} - \frac{\$10,000}{1.30^2} = \$1,775$	-	$1,600

Figure 2.4.1 - Pump Versus Ongoing Alternatives Discounted at 30%/Year

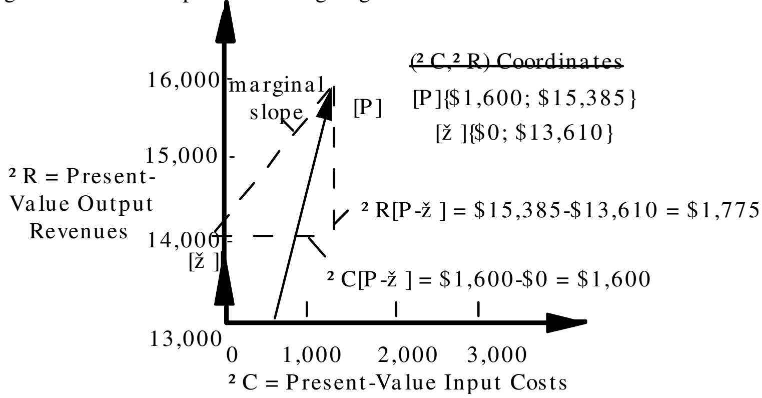
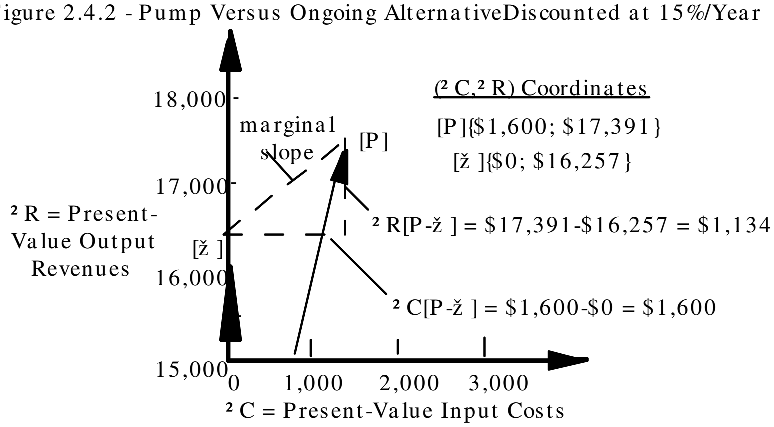


Table 2.4.2 - Pump Versus Ongoing Alternatives at a Discount Rate of 15%/Year

² NPV = Net Present-Value	=	² R = Present-Value Output Revenues	-	² C = Present-Value Input Costs
ΔNPV[P] = $15,791	=	$\frac{\$20,000}{1.15} + \frac{\$000}{1.15^2}$ = $17,391	-	$1,600
ΔNPV[Ω] = $16,257	=	$\frac{\$10,000}{1.15} + \frac{\$10,000}{1.15^2}$ = $16,257	-	$000
ΔNPV[P-Ω] = -$466	=	$\frac{\$10,000}{1.15} - \frac{\$10,000}{1.15^2}$ = $1,134	-	$1,600

Figure 2.4.2 - Pump Versus Ongoing AlternativeDiscounted at 15%/Year



Investment Opportunity Cost of Capital Analysis - The pump and ongoing alternatives are mutually exclusive ways of recovering $20,000 worth of oil. To evaluate the discount-rate sensitivity of NPV to P,Ω differences, CF{P}(-$1.6;$20;$0), CF{Ω}($0;$10;$10}, CCF{P}(-$1.6;$18.4;$18.4) and CCF{Ω}($0;$10;$20} are compared. Since CCF{Ω} is not always greater than CCF{P}, NPV comparisons of P and Ω are affected by discount rates.

The breakeven discount rates are determined by setting ΔNPV{P-Ω} or ΔNPV{Ω-P} equal to zero. The results are ΔNPV{P} = ΔNPV{Ω} when i_1 = 25% and i_2 = 400% per year. At zero discount, ΔNPV{P}=$18,400 < ΔNPV{Ω}=$20,000. Between 25% and 400%/year, ΔNPV{P} > ΔNPV{Ω}. At discount rates greater than 400%/year, ΔNPV{P} < ΔNPV{Ω}. Since 25%/year < MARR=30%/year < 400%/year, it indicates P is better than Ω. The internal rate of return of the cash flow difference {P-Ω} is often referred to as an incremental internal rate of return, ΔIRR. It is a roundabout method of determining the sign of ΔNPV{P-Ω} discounted at MARR = 30%/year. Table 2.4.1 shows ΔNPV{P-Ω} = $175 discounted at 30%/year, indicating P is better than Ω. Thus, determining ΔIRR gives a range of MARR over which the ΔNPV of P is greater than that of Ω.

It is worth mentioning that IRR{P} = 20,000/1,600 = 12.5 or 1,250%/year is smaller than IRR{Ω} which is infinite. Therefore, if IRR is used as a ranking criterion, Ω is better than P. Because the input cost of Ω does not require any cash outlays, it is often called a do-nothing alternative. However, this is not the same as the best rejected alternative from which the minimum-attractive-rate-of-return (MARR) is defined.

Let us now compare challenger P to defender Ω in Figure 2.4.1. The vector difference {P-Ω} is drawn from the terminal point of defender {Ω}($0; $13,610) to the terminal point of challenger {P}($1,600; $15,385) as calculated in Table 2.3.1. The vector difference {P-Ω} is not itself an alternative, but its slope $Ø_{PΩ}$ = ΔR[P-Ω]/ΔC[P-Ω] = $1,775/$1,600 = 1.11 is needed for comparing P to Ω. The marginal capital efficiency ratio of the firm in the range of its capital constraint is $Ø_m$ = 1 because ΔR = ΔC at the MARR discount rate. Since $Ø_{PΩ} > Ø_m$ and $ΔC_{PΩ}$ > 0, we decide by Rule #1 to eliminate defender Ω and keep challenger P. To find the pump cost, X, at which ΔNPV{P}(X; $15,385) = ΔNPV{Ω}($0; $13,610), we get X = $15,385-$13,610 = $1,775. A 45° line drawn at [Ω]{$0;$13,610} would intersect the horizontal line through [P]{$1,600;$15,385} at [P+]{$1,775;$15,385}.

Borrowing Opportunity Cost of Capital Analysis - Table 2.4.2 lists the {ΔC,ΔR} coordinates of vectors [P]{$1,600;$17,391} and [Ω]{$0;$16,257} shown in Figure 2.4.2. The slope of vector difference {P-Ω} is $Ø_{PΩ}$ = ΔR[P-Ω]/ΔC[P-Ω] = $1,134/$1,600 = 0.71 . The marginal capital efficiency ratio of the firm at its capital constraint is $Ø_m$ > 1 when discounting at the 15%/year cost of borrowing money. Since $Ø_{PΩ} < Ø_m$ and $ΔC_{PΩ}$ > 0, Rule #2 indicates we should eliminate challenger P and retain defender Ω. The pump cost, X, at which ΔNPV{P}(X; $17,391) = ΔNPV{Ω}($0; $16,257) is X = $17,391-$16,257 = $1,134. A 45° line drawn at [Ω]{$0; $16,257} intersects the line ΔR=$17,391 at {$1,134; $17,391}.

The borrowing opportunity cost model compares [P] and [Ω] by borrowing the difference in their costs at market rates of interest. Suppose the firm borrows $10,000 at the end of the first year to be paid back with the $10,000 earnings at the end of the second year plus the 15% interest charge of $1,500 for the $10,000 loan. This equal-output comparison of [P] and [Ω] now reduces to a $1,600 pump cost for [P] as opposed to a $1,500 interest charge for [Ω] two years later. Time-shifting $1,500 to two years earlier at a 15% discount rate gives us the present value of $1,500/1.15^2$ = $1,134 which is the most one should pay for [P] with the firm's borrowing opportunity costs. Also, [Ω] provides the option of not borrowing $10,000 at the end of the first year and waiting for better investments.

Section 2.5 - Financial Leveraging

The object of pump alternative [P] is to obtain $10,000 more from the oil well at the end of the first year rather than wait another year for $10,000. This objective could also be met by borrowing $10,000 at the end of the first year to be paid back with $10,000 from the well's second-year production plus $1,500 interest. By adding [B]{$0;$10,000;-$11,500} to [Ω]{$0;$10,000;$10,000}, we get [B+Ω]{$0;$20,000;-$1,500} which serves the same objective as [P]{-$1,600;$20,000;$0}. Since cumulative cash flows CCF[B+Ω]{$0;$20,000;$18,500} are always greater than CCF[P]{-$1,600;$18,400;$18,400}, ΔNPV[B+Ω] is greater than ΔNPV[P] at all discount rates (see Section 2.1).

Let us assume [P]{-$1,600;$20,000;$0} replaces [Ω]{$0;$10,000;$10,000} and then we borrow [B]{$0;$10,000;-$11,500}. To calculate ΔNPV[P+B] discounted at 30% and 15% per year, we calculate ΔNPV[B] at the 30% and 15% discount rates and add these results to those of ΔNPV[P-Ω] found in Tables 2.4.1 and 2.4.2. Discounting at 30% per year, ΔNPV[B] = $888 and ΔNPV[P+B-Ω] = ΔNPV[B]+ΔNPV[P-Ω] = $888 + $175 = $1,063. Discounting at 15% per year, ΔNPV[B] = $0, and ΔNPV[P+B-Ω] = ΔNPV[B]+ΔNPV[P-Ω] = $0-$466 = -$466 so that ΔNPV[P+B-Ω] = ΔNPV[P-Ω] whether or not the borrowing option was exercised.

The increase of ΔNPV[B] from $0 at the 15% discount rate to $888 at the 30% discount rate is caused by borrowing at 15% per year and discounting at 30% per year. This effect is called *financial leveraging*. If [P+B] was compared to [Ω] at the 30% per year discount rate, it increases the ΔNPV of the pump alternative by $888 while the ΔNPV of the ongoing alternative is unchanged. Because financial leveraging could distort present-value comparisons of competing alternatives, special care must be exercised for either direct or indirect forms of borrowing when discount rates differ from the cost of borrowing money.

If the equity of the well production is used as collateral to borrow [B], then the proceeds of the loan increase the amount of funds the company would have available for investing in independent alternatives with internal rates of return of MARR or greater. The calculation of NPV[B] = $888 estimates the net present-value which the company would realize by borrowing [B] at the cost of borrowing money and then investing [B] at the MARR interest rate. In essence, financial leveraging at the MARR discount rate ignores the question of whether [P] or [Ω] is best. Instead, it determines how much the company would benefit if it had more funds to invest in another project at the MARR interest rate.

More generally, it is shown in Appendix 2C that financial leveraging is either positive, zero or negative providing the discount rates are either greater than, equal to, or smaller than the cost of borrowing money, respectively. These results are important for the detection and evaluation of many forms of direct and indirect borrowing which are frequently embedded in engineering and financial alternatives. In particular, Sections 2.6 and 2.7 which follow examine two common problems of financial leveraging, namely, leasing versus purchasing, and debt versus equity funding, in relation to discount rates that differ from the cost of borrowing money.

Section 2.6 - Leasing versus Purchasing as Borrowing Alternatives

In financial accounting, equity financing is posted separately from debt financing on the liability side of the balance sheet. The emphasis here is on who owns the capital, partly for auditing purposes, and partly because creditors and stockholders have different roles in economic decision-making and paying taxes. In the case of leased capital (called *contingency liabilities* in off-the-balance-sheet bookkeeping), the *lessor* is the owner who rents or leases it to a firm (called the *lessee*) who uses it for a period of time specified in the lease agreement.

Whenever it is desired to use equipment, a choice often exists for leasing the equipment as opposed to buying and owning it. If equipment is leased, the lessor owns all rights of that property but assigns the right to use the equipment to a lessee who pays a rental fee for the privilege. If equipment is bought, the purchaser owns all rights of that property. Since the primary concern is for the use of equipment regardless of who owns it, we ask, "What is the best way of acquiring the use of equipment, leasing or purchasing?".

Leasing or purchasing alternatives are mutually exclusive alternatives with a common output. If equipment is leased, the lessee cannot borrow against the leased equipment as collateral because the lessor owns the equipment. In essence, the lessee pays rent from the proceeds of the output and the lessor uses those rental payments to defray the costs of buying and owning the equipment. Thus, leasing is an indirect form of borrowing for the lessee who defers the costs of purchasing equipment by renting its use from the lessor.

If equipment is purchased and not enough funds are available to pay for the equipment, the balance of the purchase price could be borrowed using the equipment as collateral until the loan is repaid later with interest. On the other hand, the rental payments for the use of the equipment needs to be funded directly from output revenues. Thus, there are many leasing and purchasing options that could be exercised to acquire the use of equipment and obtain the same output. In many respects, the issue of leasing versus purchasing is a choice between indirect and direct forms of borrowing. In practical cases, decisions between leasing and purchasing also involve income taxes, residual values, put options and other issues. However, we ignore such issues here and consider only the effects of discounting at interest rates that differ from the cost of borrowing money.

Suppose the choice of leasing versus purchasing in a firm concerns a machine which has no salvage value after a useful life of three years. The firm can either purchase the machine for $24,000 or get a 3-year lease for an annual rental of $10,000 which is payable to the lessor at the beginning of each year. Assume the firm's cost of borrowing money is 15% per year and its minimum attractive rate of return is 30% per year. The questions are, "Which alternative is better?" and "What is the internal rate of return of the lessor who is assumed to have paid $24,000 to purchase the machine?".

Let us denote the leasing alternative by [L]{-$10,000; -$10,000; -$10,000}, and let us denote the purchasing alternative by [P]{-$24,000; $0; $0}. The net present-value of [L] and [P] is shown in Tables 2.6.1 and 2.6.2 using discount rates of 30% per year for MARR and 15% per year for the cost of borrowing money, respectively.

Table 2.6.1 - Leasing Versus Purchasing Alternatives Discounted at 30%/Year.

$$\Delta\text{NPV[L]}: \quad -\$10,000 - \frac{\$10,000}{1.30} - \frac{\$10,000}{1.30^2} = -\$23,609$$

$$\Delta\text{NPV[P]}: \quad -\$24,000 + \frac{\$000}{1.30} + \frac{\$000}{1.30^2} = -\$24,000$$

Table 2.6.2 - Leasing Versus Purchasing Alternatives Discounted at 15%/Year.

$$\Delta\text{NPV[L]}: \quad -\$10,000 - \frac{\$10,000}{1.15} - \frac{\$10,000}{1.15^2} = -\$26,257$$

$$\Delta\text{NPV[P]}: \quad -\$24,000 + \frac{\$000}{1.15} + \frac{\$000}{1.15^2} = -\$24,000$$

Table 2.6.1 shows ΔNPV[L] is $391 *less* than ΔNPV[P] discounting at 30% per year. Table 2.6.2 shows NPV[L] is $2,257 *greater* than NPV[P] discounting at 15% per year. This raises the question "Which answer is correct and why?".

The lessee is seeking the least expensive way of deferring the $24,000 out-of-pocket cost of the machine. One method is to pay $10,000 just like the first payment of [L] and borrow the remaining $14,000 of the purchase price at 15% per year interest to be repaid in two annual installments of $8,612 each. Hence, we get [B]{$14,000; -$8,612; -$8,612} and [P+B]{-$10,000; -$8,612; -$8,612}. Since CCF[P+B]{-$10,000;-$18,612;-$27,224} is always greater than or equal to CCF[L]{-$10,000;-$20,000;-$30,000}, it follows that [P+B] is better than [L] at all discount rates (see Section 2.1).

In order to determine the *internal rate of return* 'i' of the lessor who paid $24,000 to purchase the machine and receives three annual rental payments of $10,000 at the beginning of each year, we must solve equation (2.5.1) below for the unknown 'i'.

$$-\$24,000 = -\$10,000 - \frac{\$10,000}{(1+i)} - \frac{\$10,000}{(1+i)^2} \qquad ....(2.6.1)$$

As explained in Appendix 2B, equation (2.6.1) has two solutions, namely $(1+i_1) = 1.2746$ or $i_1 = 27.46\%$ per year, and $(1+i_2) = -0.5604$ or $i_2 = -156.04\%$ per year. If we ignore i_2, the lessor's internal rate of return is $i_1 = 27.46\%$ per year. The cause of ΔNPV[L] being $391 *less* than ΔNPV[P] in Table 2.6.1 is that the lessor's internal rate of return $i_1 = 27.46\%$ is *less* than the 30% discount rate of the lessee, the consequence of which is to induce the firm to lease rather than purchase the machine. The cause of ΔNPV[L] being $2,257 *more* than ΔNPV[P] in Table 2.6.2 is that the lessor's internal rate of return $i_1 = 27.46\%$ is *more* than the 15% discount rate of the lessee, the consequence of which is to induce the firm to purchase rather than lease the machine.

Moreover, comparing the lessor's 27.46% rate of return to the 30% MARR of the lessee is answering a false question. The lessor does not have the option of charging the lessee with a 27.46% rate of return. The lessee is the one who has the option of either paying the lessor a 27.46% rate of return or paying a financial institution a 15% cost of borrowing money. Leasing versus purchasing alternatives can only be compared by discounting at the cost of borrowing money. Otherwise, the model of investment opportunity costs will serve the purposes of the lessor at the expense of the lessee.

Section 2.7 - Debt versus Equity Financing

An important problem of financial leveraging concerns the choice between debt and equity financing. Ordinarily, we cannot distinguish between borrowed dollars and those that are owned. There is a common responsibility for all dollars. Debt financing requires an organization to mortgage its future revenues and to pledge its assets as security for repaying loans with interest. Consequently, debt financing is an option of last resort. However, if debt and equity financing are compared using discount rates that exceed the cost of borrowing money, then debt financing may appear more profitable than equity financing.

For example, let us consider an investment which costs $10,000 and yields $12,500 one year later. Let us analyze this investment with 0%, 50% and 100% debt financing which correspond to 100%, 50% and 0% equity financing. In Tables 2.7.1 and 2.7.2, these three financing alternatives are compared using a 30%/year MARR discount rate and the 15%/year cost of borrowing money.

Table 2.7.1 - Debt Versus Equity Financing Comparison Discounted at 30%/Year.

Net Present-Value		=	Present-Value Output	-	Present-Value Input
0% Debt Financing:	-385	=	[12,500- 0]/1.30= 9,615	-	10,000
50% Debt Financing:	192	=	[12,500- 5,750]/1.30 = 5,192	-	5,000
100% Debt Financing:	769	=	[12,500-11,500]/1.30 = 769	-	0

Table 2.7.2 - Debt Versus Equity Financing Comparison Discounted at 15%/Year.

Net Present-Value		=	Present-Value Output	-	Present-Value Input
0% Debt Financing:	870	=	[12,500- 0]/1.15 =10,870	-	10,000
50% Debt Financing:	870	=	[12,500- 5,750]/1.15 = 5,870	-	5,000
100% Debt Financing:	870	=	[12,500-11,500]/1.15 = 870	-	0

Table 2.7.1 indicates that as the portion of debt financing increases from 0% to 100%, the net present-value of the $10,000 investment at a discount rate of 30% per year increases from -$385 to $769. This suggests that 100% debt financing should be used for the investment. Then $10,000 more of the company's equity funds would be available for other investments with internal rates of return of 30%/year or higher. This is the rationale of financial leveraging which was previously explained in Section 2.5 .

Table 2.7.2 indicates that as the debt financing portion increases from 0% to 100%, the $870 net present-value of the $10,000 investment remains constant. Hence, discounting at the 15% per year cost of borrowing money is unaffected by accounting distinctions between debt and equity capital. This shows borrowing opportunity costs do not create any artificial differences between debt and equity financing.

Even though debt and equity dollars are indistinguishable, it is commonly thought MARR discount rates provides stockholders with higher rates of return than discounting at the cost of borrowing money. However, high discount rates favor short-term investments and make long-term investments look unprofitable. Such artificial discount rates adversely affects the ability of an enterprise to take advantage of its engineering alternatives.

Section 2.8 - Summary of Chapter Two

Engineering alternatives deal with the physical aspects of the trade-offs between time, input costs and output revenues. Owing to time lags between input and output, engineering alternatives involve *financial alternatives* to pay for input costs until output revenues are realized. Consequently, net present values of input costs and output revenues need to be coordinated with financial alternatives in order to optimize productivity.

It is commonly thought the best alternatives have the largest net present values (NPV). But NPV is not defined until the discount rate is specified. If discount rates are too low, alternatives with large net NPVs may receive output revenues in the distant future. Present values of distant revenues can be reduced with higher discount rates. How high the discount rates should be is an open question in both theory and practice. Chapter Two examines NPVs of alternatives determined at various discount-rate magnitudes.

If the Cumulative Cash Flows CCF of one alternative are *always greater* than those of another, then its NPV is greater than that of the other at all positive discount-rates (Section 2.1). But in many practical cases, the CCF of one alternative is *not always greater* than those of another. Consequently, comparisons of the NPVs of alternatives are often sensitive to discount-rate magnitudes. Breakeven discount rates between alternatives are those rates which equate their NPVs. Single-period alternatives have only one breakeven discount rate, but multi-period alternatives have as many breakeven discount rates as the number of periods in their lifespans. The breakeven discount rates between alternatives are the same whether equating their NPVs or their net future values, NFVs.

The measurement of marginal capital efficiency defined in Chapter One is centered on engineering discount rates in Section 2.2 . By definition, the positive discount rate at which the NPV and NFV of an alternative is zero is called its *internal rate of return* (IRR). Independent project alternatives are ranked in descending order of their IRRs until the cut off of available funds. The IRR of the worst accepted or best rejected alternative is defined as the *minimum attractive rate of return*, *MARR*, which is used for all discounting. It is shown in Section 2.2 that ranking by the IRR criterion is a seriously flawed process.

The concepts of MARR and marginal capital efficiency have similarities, but also basic differences. Both concepts subscribe to the *theory of opportunity costs* which are defined as the benefit foregone, or opportunity lost, from the least profitable increment of input costs that was not undertaken because of capital constraints. But engineers, accountants and economists treat opportunity costs very differently. Engineering opportunity costs are based on net present values discounted at MARR. Accountants deal with costs that can be identified on a cash or accrual basis which can be audited in financial statements. Accounting costs differ from an economist's opportunity costs if they fail to reflect the best use of available resources.

Financial discount rates are discussed in Section 2.3 . Suppose the foregone benefit is an investment opportunity with a rate of return of 15%/year, part of which could be financed by borrowing money from a bank at 10%/year interest. The NPV of the cash flow could be discounted at either a) the *internal rate of return* of the total cash flow (TCF) of the 15%/year investment, b) the 10%/year *internal rate of return* of debt cash flow (DCF), or c) the *internal rate of return* of the remaining equity cash flow (ECF). The three discount-rate possibilities were tested on one- and two-period investment models with various proportions of debt and equity financing. The tests were carried out with equations (2.3.1) and (2.3.2) which holds true in all periods and at all discount rates respectively.

$$TCF(i) \equiv DCF(i) + ECF(i) \quad \text{for end-of-year } i = 0, 1, 2, ... \qquad ....(2.3.1)$$

$$NPV\{TCF\} \equiv NPV\{DCF\} + NPV\{ECF\} \quad \text{at all discount rates} \qquad ....(2.3.2)$$

When discounting at the 15%/year IRR of TCF, ΔNPV{TCF} = 0 and -ΔNPV{DCF} = ΔNPV{ECF}. When discounting at the IRR of ECF, ΔNPV{ECF} = 0 and ΔNPV{TCF} = ΔNPV{DCF} < 0. When discounting at the 10%/year ERR of DCF, ΔNPV{DCF} = 0 and ΔNPV{TCF} = ΔNPV{ECF} > 0 independently of the input-cost fractions that were borrowed. The marginal capital efficiencies, $Ø_m$, of the one- and two-period investment models were 1.045 and 1.067 respectively.

Suppose the foregone benefit is an investment opportunity with a rate of return of 15%/year, part of which could be financed by borrowing money from a bank at 10%/year interest. The firm's *opportunity cost* is defined as the benefit foregone by *not* undertaking the 15%/year investment and *not* borrowing any of its input cost from a bank at 10%/year interest. We identify the best foregone alternative as *an incremental input cost* which is called the *borrowing opportunity cost* of the organization. Conventional models for determining discount rates identify the best foregone alternative as *the best rejected project* whose profitability is called the *investment opportunity cost* of the organization.

Investment and borrowing opportunity cost models result in distinctly different discount rate magnitudes. Both opportunity cost models for determining discount rates simulate the last increment of input costs subject to a capital constraint. The last increment of input costs subject to a capital constraint is not the same as the last project to be accepted or the best rejected project at the cut-off of available funds. From the viewpoint of an economic organization as a whole, all projects do equally important things, and the object of the organization is to undertake the last increment of input costs at the capital constraint which could occur between the alternatives of any project. The capital constraint is not the same as the available funds of an organization which may be greater than the input costs of available investment opportunities or which can be increased by equity and debt financing to fund additional investment opportunities.

The concept of the best rejected project at the cut-off of available funds implies that project alternatives could be ranked in descending order of a single measure of profitability. However, it is shown in Section 1.7 of Chapter One that, in general, each project alternative needs to be characterized with possibly conflicting measures of absolute and relative profitability so that the best rejected project is not determinable by ranking with a single profitability measure.

The financial accounting model for determining discount rates depends upon the weighted-average-cost-of-capital (WACC) from four sources of capital: (1) the firm's net operating income, (2) short-term debt financing, (3) long-term debt financing and (4) sale of capital stock. The firm's net operating income and sale of capital stock are coupled together in the form of an expected rate-of-return (ROR) by equity investors from dividends, capital gains and retained earnings. The rates of return from these sources are then weighted by their proportions on the firm's balance sheet.

Discount rates determined from the WACC model are usually much higher than market rates of interest because of upward adjustments which are made to account for risk, inflation and stockholder expectations. The WACC model for determining discount rates is not practical for financial institutions, nonprofit organizations and government agencies whose sources of capital differ substantially from those of industrial firms.

Investment and borrowing opportunity cost models result in distinctly different discount rate magnitudes. This writer believes that investment opportunity cost models distort the meaning of economic scarcity which gives rise to major misunderstandings among engineers, accountants and economists in both private and public sectors of the economy. Therefore, in Sections 2.4 to 2.7 of this chapter, simple examples are given to bring out the differences of using either MARR or the cost of borrowing money as discount rates in comparisons of mutually exclusive alternatives.

The first example (Section 2.4) concerns time acceleration of engineering alternatives. The ongoing alternative [Ω] consists of an oil well to be depleted by producing $10,000 worth of oil at the end of each year for the next two years. An engineer proposes to install a pump [P] costing $1,600 which would produce all $20,000 worth of oil by the end of the first year. At the 30%/year MARR discount rate, ΔNPV[P] = ΔNPV[Ω] if the pump costs $1,775. At the 15%/year cost of borrowing money, ΔNPV[P] = ΔNPV[Ω] if the pump costs $1,134. Since $1,775 and $1,134 are on opposite sides of the $1,600 pump cost, it follows that the investment opportunity cost model prefers [P] and the borrowing opportunity cost model prefers [Ω].

The second example (Section 2.5) concerns financial leveraging. We use a financial alternative [B] to borrow $10,000 at the end of the first year and pay back the loan one year later with the second-year $10,000 from [Ω] plus 15% interest. Let us now compare [B]{$0;$10,000;-$11,500} plus [Ω]{$0;$10,000;$10,000} to [P]{-$1,600;$20,000;$0}. Because cumulative cash flows CCF[B+Ω]{$0;$20,000;$18,500} are always greater than CCF[P]{-$1,600;$18,400;$18,400}, it follows that ΔNPV[B+Ω] is greater than ΔNPV[P] at all positive discount rates (see Section 2.1).

Suppose we borrow [B]{$0;$10,000;-$11,500} after [P]{-$1,600;$20,000;$0} replaces [Ω]{$0;$10,000;$10,000}. Adding [B]{$0;$10,000;-$11,500} to [P]{-$1,600;$20,000;$0} gives [B+P]{-$1,600;$30,000;-$11,500} as compared to [Ω]{$0;$10,000;$10,000}. In particular, at the 30% and 15% discount rates, ΔNPV[P+B-Ø] equals $1,063 and -$466 respectively. The reason why [B+P] appears much better than [Ω] at the 30% discount rate is because funds are borrowed at 15% per year and discounted at 30% per year. This effect is called *financial leveraging*. More generally, it is shown in Appendix 2C that financial leveraging is either positive, zero or negative when discount rates are either greater than, equal to, or smaller than the cost of borrowing money, respectively.

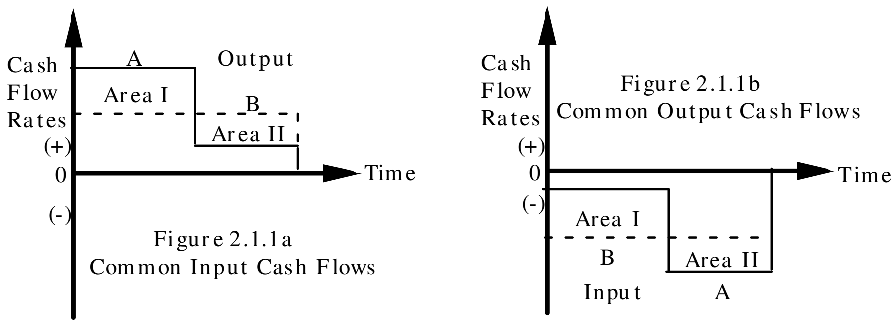
The next example of financial leveraging (Section 2.6) deals with leasing versus purchasing as borrowing alternatives. In this example, the 30%/year MARR discount rate favored leasing and the 15%/year cost of borrowing money favored purchasing. An analysis of these results show that the MARR discount rate enables the lessor to earn much more than the 15% cost of borrowing money and almost as much as the MARR interest rate of the lessee. On the other hand, the 15% per year discount rate compares the lessee's cost of borrowing money to buy the equipment to the indirect cost of borrowing money by leasing.

The last example of financial leveraging (Section 2.7) concerns debt versus equity financing. In this example, the 30%/year MARR discount rate preferred debt to equity financing. This result implies that borrowing money is more profitable than using the company's own money. On the other hand, the 15%/year discount rate shows debt and equity financing are equally profitable. This result substantiates the fact that borrowed and company-owned dollars are indistinguishable. Because of the practical importance of the effects of investment and borrowing opportunity cost models in determining discount rates, Chapters Four, Five and Six treats these issues in much greater detail.

Appendix 2A - Insensitivity of Discount-Rate Magnitudes

In Section 2.1, it was brought out that present-value comparisons of certain mutually exclusive alternatives may be insensitive to the magnitudes of discount rates. In this connection, let us review the present-value comparisons for mutually exclusive alternatives A and B previously shown in Figures 2.1.1(a) and 2.1.1(b) which are reproduced below for convenience.

Figure 2.1.1 - Insensitivity of present-value comparisons to positive discount rates.



In Figure 2.1.1(a), A and B have common input costs so that only output differences between A and B indicated by Areas I and II can determine which one is best. When Area I is greater than Area II, A would have *greater* initial *and* lifetime output values than B. It will be shown below that the net present-value of A will be *greater* than that of B regardless of the magnitudes of discount rates used in present-value calculations.

In Figure 2.11(b), A and B have common output values so that only input differences determine which is best. When Area I is greater than Area II, A would have *smaller* initial *and* lifetime input costs than B. It will be shown below that the present-value profit of A is *greater* than that of B regardless of the magnitudes of discount rates used in present-value calculations. We will now show that this *insensitivity* of discount-rate magnitudes occurs whenever one alternative has greater initial *and* lifetime advantages as compared to other alternatives.

Let $R_A(t)$ and $R_B(t)$ denote the cash flow rates of alternatives A and B at time t. Let $R_1(t)$ equal $R_A(t)$ minus $R_B(t)$ from the start at zero time to the finish at time τ where the differences between $R_A(t)$ and $R_B(t)$ are always positive; and $R_2(t)$ is zero elsewhere. Let $R_2(t)$ equal $R_B(t)$ minus $R_A(t)$ from the start at time τ to the finish at time T where the differences between $R_B(t)$ and $R_A(t)$ are always positive; and $R_1(t)$ is zero elsewhere. The hypothesis is that Area I is greater than Area II. Hence,

$$\int_0^{\tau} R_1(t)dt = \text{Area I} > \text{Area II} = \int_{\tau}^{T} R_2(t)dt \qquad ....(2A.1)$$

In equation (2A.2) below, we express the present values of Areas I and II by discounting their cash flows continuously at the constant discount rate $r > 0$.

$$\int_0^{\tau} R_1(t)e^{-rt}dt = PV[\text{Area I}] > PV[\text{Area II}] = e^{-rt}\int_{\tau}^{T} R_2(t)e^{-rt}dt \qquad ....(2A.2)$$

Equation (2A.2) needs a question mark above the inequality sign in order to indicate we have not yet proven that the present value of Area I is greater than that of Area II at an arbitrary and positive discount rate r. The proof is obtained by invoking the the first mean-value theorem for integrals stated in equation (2A.3) below.

$$\int_a^b f(x)p(x)dx = f(\xi)\int_a^b p(x)dx \qquad ....(2A.3)$$

If p(x) is positive and continuous in [a,b], and f(x) is arbitrary and continuous in [a,b], then there exists a point ξ in [a,b] such that equation (2A.3) is true. (Reference: R. Courant, *Differential and Integral Calculus*, Vol. I, 2nd Edition, Interscience Publishers, New York, pp. 134-35 (1937)).

We may now apply the first mean-value theorem of integral calculus to equation (2A.2). Since $R_1(t)$ and $R_2(t)$ are positive and continuous in the intervals $[0,\tau]$ and $[\tau,T]$, and the exponential e^{-rt} is arbitrary and continuous in the intervals $[0,\tau]$ and $[\tau,T]$, then there exists points t_1 and t_2 in the intervals $[0,\tau]$ and $[\tau,T]$ such that equation (2A.4) is true.

$$e^{-rt_1}\int_0^\tau R_1(t)dt = PV[\text{Area I}] > PV[\text{Area II}] = e^{-rt_2}\int_0^\tau R_2(t)dt \qquad ....(2A.4)$$

Because t_2 is greater than t_1 and r is a positive constant, e^{-rt_1} is greater than e^{-rt_2}. When the left and right hand sides of hypothesis (2A.1) are multiplied by e^{-rt_1} and e^{-rt_2} respectively, it proves that inequality (2A.4) is valid. Therefore, if Area I is greater than Area II, then the present value of A is greater than that of B regardless of the constant discount-rate magnitudes used to calculate their present values.

The following four corollaries are generalizations of the proposition that present-value comparisons of two mutually exclusive alternatives are *insensitive* to constant discount-rate magnitudes if earlier advantages of an alternative are *more* than its later disadvantages:

1. The proposition is still valid if the discount-rate magnitudes are variable. The proof would follow if the exponentials in the integrands of (2A.2) were rewritten as the exponentials $\exp[-\int r(u)du]$ integrated from 0 to τ and from τ to T respectively.

2. When the earlier advantages of one alternative are *less* than its later disadvantages, then present-value comparisons of two mutually exclusive alternatives are *sensitive* to discount-rate magnitudes. The proof follows after changing the direction of inequality (2A.1).

3. When the cash flows and discounting are discrete rather than continuous, the earlier stated propositions are still valid because the discrete conditions can be treated as a limiting case of continuous conditions.

4. Future-value comparisons of two mutually exclusive alternatives are also *insensitive* to discount-rate magnitudes if earlier advantages of an alternative are *more* than its later disadvantages. The proof is obtained when both sides of (2A.2) are multiplied by e^{rT} (or multiplied by $\exp[\int r(u)du]$ integrated from 0 to T if r is a positive variable).

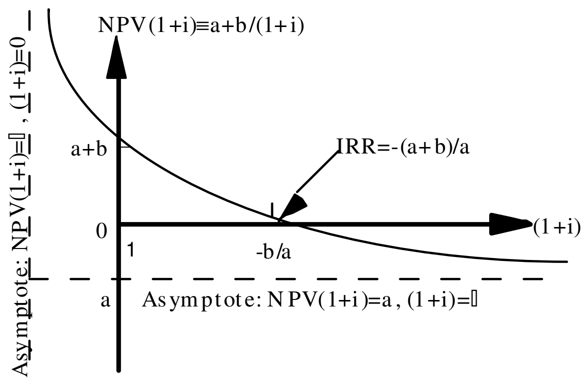
Appendix 2B - Internal Rates of Return of Two-Period Investments

Let us consider a single-period investment of 'a' dollars which returns 'b' dollars one year later. The present value of 'a' dollars input is simply 'a' which we assume is a negative number. Let us assume the output of 'b' dollars is a positive number greater than -a, and its present value is b/(1+i). The net present-value (NPV) of the single-period investment is a function of (1+i) as defined in (2B.1) below.

$$NPV(1+r) \equiv a + b/(1+i) \qquad ....(2B.1)$$

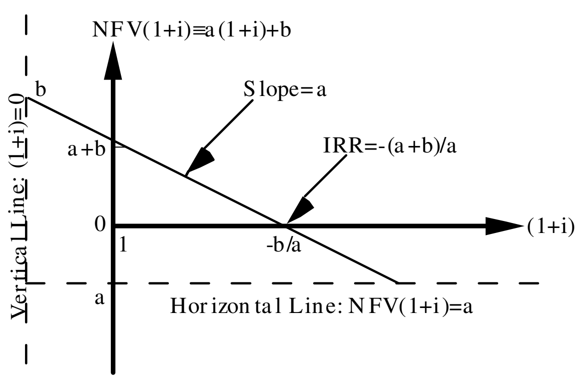
The graph of NPV(1+i) in (2B.1) is a rectangular hyperbola as shown in Figure 2B.1 with asymptotes NPV(1+i) = ∞ when 1+i = 0, and NPV(1+i) = a when 1+i = ∞. The IRR of the investment is at the point where the hyperbola crosses the (1+i)-axis (i.e., where NPV(1+i) = 0 and (1+i) = -(b/a) or i = -(a+b)/a).

Figure 2B.1 - The net present-value function NPV(1+i) of a single-period investment.



Let NPV(1+i) of (2B.1) be multiplied by (1+i) in order to obtain the net future-value function NFV(1+i) ≡ a(1+i) + b of the investment one year later as shown in Figure 2B.2.

Figure 2B.2 - The net future-value function NFV(1+i) of a single-period investment.



The graph of NFV(1+i) in Figure 2B.2 is a straight line with slope equal to a. It also has the same intercepts a+b and -b/a on the NFV(1+i) and (1+i) axes as the NPV(1+i) intercepts on the NPV(1+i) and (1+i) axes. Since NFV(1+i) = NPV(1+i) = 0 where 1+i = -b/a, it follows that i = -(a+b)/a. Therefore, internal rate of return, denoted by IRR, could be defined as the discount rate 'i' that makes either NFV(1+i) or NPV(1+i) equal to zero.

Two-period investments have two *constant* discount rates which make the net present-value, NPV, and net future-value, NFV, of input and output equal to zero. In general, let us consider a two-period investment of 'a' dollars at the beginning of the first year which returns 'b' and 'c' dollars at the ends of the first and second years respectively. The NPV and NFV of this investment discounted or compounded at interest rate 'i' is a function of (1+i) as defined in equations (2B.2) and (2B.3) below.

$$NPV(1+i) \equiv a + b(1+i)^{-1} + c(1+i)^{-2} \qquad ....(2B.2)$$

$$NFV(1+i) \equiv a(1+i)^2 + b(1+i) + c \qquad ....(2B.3)$$

The graph of NFV(1+i) defined in (2B.3) is a parabola which is symmetrical about the vertical line (1+i) = -b/2a. Assuming 'a' is negative, the parabola has a maximum value of $(b^2-4ac)/(-4a)$ at its line of symmetry as shown in Figures 2B.3. If (b^2-4ac) is greater than zero as shown in Figure 2B.3(a), then the equation NFV(1+i) = 0 has two real roots where the left and right branches of the parabola cross the (1+i)-axis. But if (b^2-4ac) is less than zero as shown in Figure 2B.3(b), then the parabola lies below the (1+i)-axis and the two roots of the equation NFV(1+i) = 0 are complex conjugates. If the parabola is tangent to the (1+i)-axis, then the equation NFV(1+i) = 0 has a double root at the point of tangency.

Figure 2B.3 - The net future-value function (NFV) of a two-period investment.

Figure 2B.3(a): $(b^2-4ac) > 0$

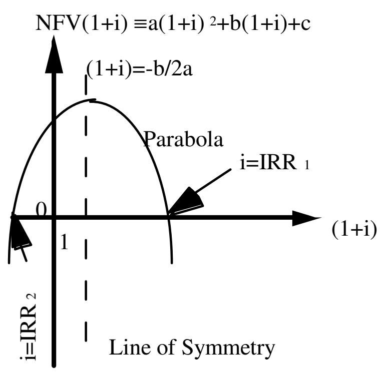
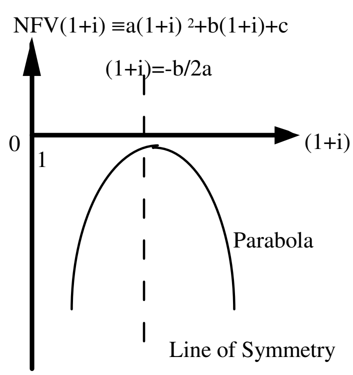


Figure 2B.3(b): $(b^2-4ac) < 0$



When (b^2-4ac) is greater than zero as shown in Figure 2B.3(a), the parabola crosses the (1+i)-axis at two points where NFV(1+i) = 0. The two values of i at these zero-crossings of the parabola define two internal rates of return of a two-period investment, both of which may either be positive, negative, or one positive and one negative as shown in Figure 2B.3(a). The exact values of IRR for a two-period investment can be determined by solving the equation NFV(1+i) = 0. The solution which is outlined below is the well-known quadratic formula which can be found in most textbooks on algebra and analytic geometry. The quadratic formula is valid whether (b^2-4ac) is positive, negative or equal to zero.

Let us divide NFV(1+i) = 0 by 'a' and then both add and subtract $(-b/2a)^2 = b^2/4a^2$ so that the solution to the equation is unchanged.

$$(1+i)^2 + (b/a)(1+i) + c/a + b^2/4a^2 - b^2/4a^2 = 0 \quad ....(2B.4)$$

The first, second and fourth terms of (2B4) form a perfect square as shown in (2B.5).

$$[(1+i) + (b/2a)]^2 + c/a - b^2/4a^2 = 0 \quad ....(2B.5)$$

Multiply the numerator and denominator of c/a by 4a, and transpose the last two terms of (2B.5) to the right-hand side of the equality sign.

$$[(1+i) + (b/2a)]^2 = (b^2-4ac)/4a^2 \quad ....(2B.6)$$

Lastly, taking the square root of both sides of (2B.6) and solving for (1+i) gives us the quadratic formula in (2B.7).

$$(1+i) = \{-b \pm (b^2-4ac)^{1/2}\}/2a \quad ....(2B.7)$$

The investment opportunity cost model assumes only the positive internal rate of return of two-period investments has economic significance. This assumption is based on Descartes rule of signs which indicates two-period investments have at most one positive internal rate of return if successive coefficients a, b and c of (2B.3) have one change of sign. However, coefficients a, b and c of many economically sound two-period investments have two changes of sign in which case there can be two positive, or two negative, or complex conjugate internal rates of return.

Consider the engineering and ongoing alternative [P-Ω]{-$1,600; $10,000; -$10,000} of Section 2.3 . The coefficients of $NFV(1+i) = a(1+i)^2 + b(1+i) + c$ for [P-Ω] are a = -$1,600; b = $10,000; and c = -$10,000 which exhibit two changes of sign. Upon substituting these values for a, b and c in (2B.7), we obtain (1+i) = 1.25 and 5.00, or i_1 = 25% and i_2 = 400% per year. This raises the question of which internal rate of return, if any, has economic significance. Therefore, let us increase the cost of the pump from $1,600 to $1,700 to see if i_1 or i_2 decrease. Upon substituting a = -$1,700; b = $10,000; and c = -$10,000 in (2B.7), we find i_1 *increases* from 25% to 27.7% per year while i_2 *decreases* from 400% to 360.5% per year. Thus, as incredulous as it may sound, only i_2 has economic significance.

Two-period financial leveraging alternatives may have internal rates of return that are complex conjugates. For example, an investor has $1,000 now and wants to invest $3,000 one year later. Therefore, the investor deposits $1,000 with a bank now in order to borrow $3,000 one year later. The investor agrees to pay off the loan with $2,500 one year later. Letting a = -$1,000; b = $3,000; and c = -$2,500; we get (b^2-4ac) = -$1,000,000 which results in internal rates of return that are complex conjugates. But the loan is profitable for the bank who receives a $1,000 interest-free deposit for one year, and then receives 25% interest per year on a one-year loan of $2,000 to the investor.

Other two-period financial leveraging alternatives have two negative internal rates of return. For example, based on a (3,4,5) right triangle, suppose an investor deposits $5,000 with a bank now in order to borrow $3,000+$4,000 = $7,000 one year later. The investor agrees to pay off the loan with ($3,000)($4,000)/($5,000) = $2,400 one year later. Letting a = -$5,000; b = $7,000; and c = -$2,400; we get i_1 = -20% and i_2 = -40% per year, even though the transaction is again quite profitable for the bank.

Appendix 2C - Financial Leveraging

Many firms cannot raise as much capital as they would like to undertake engineering alternatives. All sources of capital have a cost which can be either direct as say a loan with a borrowing opportunity cost, or indirect as say a new project with an investment opportunity cost. In the attempt to minimize the cost of capital, a distinction is often made between the effects of debt and equity capital in the capital structure of a company. Financial leveraging is measured by the ratio of debt capital to the total capital employed by a company, and it is frequently debated as a major factor affecting the cost of capital to a company. The controversy is further complicated by differences in income taxes on corporation profits, dividends distributed for equity capital, and interest paid for debt capital.

Aside from these issues, there is room for questioning whether the distinction between debt and equity capital has an important affect on the cost of capital to a company because dollars of debt or equity are indistinguishable. Moreover, a company's capital is invested in all its activities and it is not meaningful to say that any one operation was financed by either debt or equity funds. In Sections 2.4, 2.5 and 2.6, we have shown that when the proceeds of a two-period loan are discounted at a specific rate which is greater than the cost of borrowing money, then equity capital appears more costly than debt capital. Appendix 2C generalizes these results by showing that equity capital appears more costly (or less costly) than debt capital whenever the proceeds of loans are discounted at rates which are greater than (or less than) the cost of borrowing money.

Let $R_1(t)$ denote the rate of cash inflow from a loan at time t. During the lending period from time zero to time τ, $R_1(t)$ is positive and zero elsewhere. Let $R_2(t)$ denote the rate of cash outflow for paying back the loan with interest at time t. During the payback period from time τ to time T, $R_2(t)$ is negative and zero elsewhere. Since the payback of the loan includes interest payments, we require that

$$\int_0^\tau R_1(t)dt = \text{Area I} < \text{Area II} = \int_\tau^T R_2(t)dt \qquad \text{....(2C.1)}$$

Without loss of generality, let us assume that the constant interest rate 'r' is the cost of borrowing money. Consequently, the difference between the present values of Areas I and II must equal zero when their cash flows are discounted continuously at interest rate r.

$$\int_0^\tau R_1(t)e^{-rt}dt - e^{-r\tau}\int_\tau^T R_2(t)e^{-rt}dt = 0 \qquad \text{....(2C.2)}$$

Let λ represent how much the actual discount rate 'r+λ' differs from the cost of borrowing money 'r'. The present-value difference between Areas I and II discounted at the rate 'r+λ' is

$$\int_0^\tau R_1(t)e^{-(r+\lambda)t}dt - e^{-(r+\lambda)\tau}\int_\tau^T R_2(t)e^{-(r+\lambda)t}dt = 0 \ \text{ if } \lambda = 0 \qquad \text{....(2C.3)}$$

Upon applying the first mean-value theorem of integral calculus to (2C.3), we get

$$e^{-(r+\lambda)t_1}\int_0^\tau R_1(t)dt - e^{-(r+\lambda)t_2}\int_\tau^T R_2(t)dt \geq 0 \ \text{ if } \lambda \geq 0 \qquad \text{....(2C.4)}$$

Since $e^{-(r+\lambda)t_1} > e^{-(r+\lambda)t_2}$ if $\lambda > 0$, and $e^{-(r+\lambda)t_1} < e^{-(r+\lambda)t_2}$ if $\lambda < 0$, it follows that the present value of a loan is greater than (or less than) zero if the discount rate is greater than (or less than) the cost of borrowing money.

Appendix 2D - Algebraic Inequalities

Algebraic inequalities are widely used for comparing alternatives. The following list distinguishes between inequalities that are always true and those that are sometimes true.

1. Transitivity - Theorem 1.0: If $a > b$ and $b > c$, then $a > c$.
   Example 1.0: If $3 > -2$ and $-2 > -3$, then $3 > -3$.
   Theorem 1.1: If $a \geq b$ and $b \geq c$, then $a \geq c$;
   $a = c$ if and only if $a = b$ and $b = c$.
2. Addition - Theorem 2.0: If $a > b$ and $c > d$, then $a+c > b+d$.
   Example 2.0: If $3 > 2$ and $-2 > -3$, then $3-2 > 2-3$.
   Theorem 2.1: If $a > b$ and c is any real number, then $a+c > b+c$.
   Example 2.1: If $3 > 2$ and $c = -3$, then $3-3 > 2-3$.
   Theorem 2.2: If $a \geq b$ and $c \geq d$, then $a+c \geq b+d$;
   $a+c = b+d$ if and only if $a = b$ and $c = d$.
3. Subtraction - Theorem 3.0: If $a > b$ and $c > d$, then $a-c > b-d$ *is not always true*.
   Example 3.00: If $8 > 7$ and $4 > 2$, then $8-4 > 7-2$ *is not true*.
   Example 3.01: If $8 > 6$ and $5 > 4$, then $8-5 > 6-4$ *is true*.
   Theorem 3.1: If $a > b$ and $c > d$, then $a-d > b-c$ *is always true*.
   Example 3.1: If $8 > 7$ and $4 > 2$, then $8-2 > 7-4$ *is true*.
   Theorem 3.2: If $a \geq b$ and $c \geq d$, then $a-d \geq b-c$;
   $a-d = b-c$ if and only if $a = b$ and $c = d$.
4. Multiplication-Theorem 4.0: If $a > b > 0$ and $c > d > 0$, then $ac > bd$ *is always true*.
   Example 4.0: If $8 > 7 > 0$ and $4 > 2 > 0$, then $8*4 > 7*2$ *is true*.
   Example 4.01: If $-7 > -8$ and $-2 > -4$, then $-7*(-2) > -8*(-4)$ *is not true*.
   (i.e., conditions $a > b > 0$ and $c > d > 0$ cannot be dropped).
   Theorem 4.1: If $a > b > 0$ and $c < 0$, then $ac < bc$.
   Example 4.1: If $5 > 4 > 0$ and $c = -1$, then $-5 < -4$.
   Theorem 4.2: If $a \geq b > 0$ and $c \geq d > 0$, then $ac \geq bd$;
   $ac = bd$ if and only if $a = b$ and $c = d$.
5. Division - Theorem 5.0: If $a > b > 0$ and $c > d > 0$, then $a/c > b/d$ *is not always true*.
   Example 5.00: If $8 > 7$ and $4 > 2$, then $8/4 > 7/2$ *is not true*.
   Example 5.01: If $8 > 6$ and $5 > 4$, then $8/4 > 6/5$ *is true*.
   Theorem 5.1: If $a > b > 0$ and $c > d > 0$, then $a/d > b/c$ *is always true*.
   Example 5.1: If $8 > 7$ and $4 > 2$, then $8/2 > 7/4$ *is true*.
   Theorem 5.2: If $a = b = 1$ and $c > d > 0$, then $1/d > 1/c$.
   Example 5.2: If $a = b = 1$ and $4 > 2 > 0$, then $1/2 > 1/4$.
   Theorem 5.3: If $a \geq b > 0$ and $c \geq d > 0$, then $a/d \geq b/c$;
   $a/d = b/c$ if and only if $a = b$ and $c = d$.
6. Power/Roots-Theorem 6.0: If $a > b > 0$ and m, n are positive integers, then $a^{m/n} > b^{m/n}$.
   Example 6.0: If $1.44 > 1.21 > 0$, $m = 1$, and $n = 2$, then $1.44^{1/2} > 1.21^{1/2}$.
   Theorem 6.1: If $a > b > 0$ and m, n are positive integers, then $a^{-m/n} < b^{-m/n}$.
   Example 6.1: If $1.44 > 1.21 > 0$, $m = 1$, and $n = 2$, then $1.44^{-1/2} < 1.21^{-1/2}$.
   Theorem 6.2: If $a \geq b > 0$, m is a nonnegative number and n is a positive number, then $a^{m/n} \geq b^{m/n}$ and $a^{-m/n} \leq b^{-m/n}$; $a^{m/n} = b^{m/n}$ and $a^{-m/n} = b^{-m/n}$ if and only if either $a = b$ or $m = 0$.

Chapter Two - Exercises

2-1a Simple Discount Rates and Simple Interest Rates - Lending institutions often charge interest in advance, called simple discounts. However, different rates of interest should be specified depending upon whether the interest charges are to be paid at the start or finish of the loan period. For example, suppose a person borrows 'b' dollars for 't' years at a simple discount rate of 'r_D' per year. The bank gives the borrower $a = b(1-r_Dt)$ dollars now for which the bank will receive 'b' dollars in return 't' years later when the loan of 'b' dollars is repaid. If the same person borrows 'a' dollars for 't' years at a simple interest rate of 'r_I' per year, the bank expects to be repaid $b = a(1+ r_It)$ dollars at the end of the loan period. It is important for the borrower to determine how simple discount rate 'r_D' is related to simple interest rate 'r_I' for the amounts 'a' and 'b' that the borrower receives and pays at the start and finish of the loan period. Solving for 'a', we get $b(1-r_Dt) = b/(1+ r_It)$ which can be solved for r_D in terms of r_I, or for r_I in terms of r_D as follows:

$$r_D = \frac{r_I}{1 + r_I t} \text{ and } r_I = \frac{r_D}{1 - r_D t} \quad \text{for } r_I,\ r_D,\ t > 0 \qquad ....(2Ex.1)$$

It follows that r_D is *smaller* than r_I. Thus, r_D = 10%/year is equivalent to r_I = 11.1%/year, and r_I = 10%/year is equivalent to r_D = 9.09%/year if 't' equals one year.

2-1b A bank charges 9% interest in advance (i.e., 9%/year simple discount) for loans of one year or less. How much does a borrower receive from the bank upon signing a 6-month promissory note for $2,000? *Ans.* $b(1-r_Dt)$ = $2,000[1-(.09)(6/12)] = $1,910.

2-1c What rate of simple interest does the borrower of Problem **2-1b** pay? *Ans.* $r_I = r_D/(1-r_Dt)$ = .09/[1-(.09)(6/12)] = .09424 or 9.424%/year.

2-1d How much should the promissory note of Problems **2-1b** and **2-1c** be in order for the borrower to receive $2,000 now? Ans. $b = a(1+ r_It)$ = $2,000[1+(.09424)(6/12)] = $2,094.24, or $b = a/(1-r_Dt)$ = $2,000/[1-(.09)(6/12)] = $2,094.24 .

2-2a Net Future- and Present-Value Functions of Two-Period Investments - In Appendix 2B, we defined the net future-value function NFV(1+i) of two-period investments {a, b, c} in equation (2B.3) as follows:

$$\text{NFV}(1+i) \equiv a(1+i)^2 + b(1+i) + c \qquad ....(2B.3)$$

Similarly, the present value-profit function PVP(1+i) of two-period investments {a, b, c} is

$$\text{NPV}(1+i) \equiv a + b(1+i)^{-1} + c(1+i)^{-2} \qquad ....(2B.3')$$

Thus, by definition, $\text{NPV}(1+i) \equiv (1+i)^{-2}\text{NFV}(1+i)$. When r = 0, then NPV(1+i) = NFV(1+i) = a+b+c. Any interest rate $i \neq 0$ that makes NPV(1+i) = 0 also makes NFV(1+i) = 0 and vice versa. Consequently, the internal rates of return of two-period investments {a, b, c} can be determined either from NFV(1+i) or NPV(1+i), whichever is more convenient.

Although the zero-crossing properties of NPV(1+i) and NFV(1+i) bring out points which they have in common, the remaining points on the graphs of NPV(1+i) and NFV(1+i) are significantly different. In particular, NPV(1+i) and NFV(1+i) generally have their maximum (or minimum) values at different interest rates. More specifically, the graph of NFV(1+i) is a parabola which is symmetrical about the vertical line (1+i) = -b/2a, and the NFV(1+i) coordinate of the vertex of the parabola is given by $[b^2-4ac]/[-4a]$. But the graph of NPV(1+i) is asymmetric, and its maximum (or minimum) value is where (1+i) = -2c/b, and the NPV(1+i) coordinate of the stationary point is given by $[b^2-4ac]/[-4c]$.

2-2b Consider the investment {a, b, c} = [P-Ø]{-$1,600; $10,000; -$10,000} derived from Section 2.3 data. Determine the undiscounted values of NPV(1+i) and NFV(1+i) (i.e., at zero discount, or i = 0). *Ans.* PVP(1+i) = FVP(1+i) = a+b+c = -$1,600.

2-2c Determine the internal rates of return of the investment in Problem **2-2b**. *Ans.* (See equation (2B.7)) FVP(1+i) = 0 when (1+i) = $\{-b \pm (b^2-4ac)^{1/2}\}/2a$ = 1.25, 5.00 or i_1 = 25% and i_2 = 400% per year. Also NPV(1+i) = 0 when i_1 = 25% and i_2 = 400% per year.

2-2d What is the maximum negative value of coefficient 'a' in the investment of Problem **2-2b** which will make PVP(1+i) = FVP(1+i) = 0 at discount rates of 30% and 15% per year? Ans. NPV(1.30) = 0 = a + $10,000/1.30 - $10,000/$1.30^2$; a = -$1,775.

NFV(1.30) = 0 = (a)(1.30^2) + ($10,000)(1.30) - $10,000; a = -$1,775.

NPV(1.15) = 0 = a + $10,000/1.15 - $10,000/$1.15^2$; a = -$1,134.

NFV(1.15) = 0 = (a)(1.15^2) + ($10,000)(1.15) - $10,000; a = -$1,134.

2-2e Which interest rates maximize NPV(1+i) and NFV(1+i) for the investment of Problem **2-2b**, and what are their maximum values? *Ans.* The maximum NPV(1+i) occurs when (1+i) = -2c/b = 20,000/10,000 = 2, or i = 100% per year. Therefore, the maximum NPV(1+i) = $[b^2-4ac]/[-4c]$ = $900. The maximum NFV(1+i) occurs when (1+i) = -b/2a = 10,000/3,200 = 3.125, or i = 212.5% per year which is the average of i_1 = 25% and i_2 = 400% per year. The maximum NFV(1+i) = $[b^2-4ac]/[-4a]$ = $5,625.

2-3a (Section 2.3 - Financial Discount Rates) The general problem of determining discount rates from opportunity cost models is exemplified by a firm where the marginal increment of input costs is a single-period investment with a rate of return of w%/year. The firm could finance a fraction x of this investment by borrowing money at y%/year interest. The present value of the input costs and output revenues could be discounted at either a) the *internal rate of return* of the w%/year total cash flow, TCF, b) the *internal rate of return* of the y%/year debt cash flow, DCF, or c) the *internal rate of return* of the equity cash flow, ECF. The end-of-year, EOY, cash flow description is given in the table below, and is followed by another table which evaluates ΔNPV according to equation (2.3.1) at various APR discount rates. The marginal output/input ratio of the firm is $Ø_m$ = (1+w)/(1+y).

EOY	TCF	=	DCF	+	ECF
0	-1	=	-x	+	-1+x
1	1+w	=	x(1+y)	+	(1+w)-x(1+y)

APR%/yr	ΔNPV{TCF}	=	ΔNPV{DCF}	+	ΔNPV{ECF}
0	w	=	xy	+	w-xy
y	(w-y)/(1+y)	=	0	+	(w-y)/(1+y)
w	0	=	x(y-w)/(1+w)	+	x(w-y)/(1+w)
$\frac{w-xy}{1-x}$	$\frac{x(y-w)}{1-x+w-xy}$	=	$\frac{x(y-w)}{1-x+w-xy}$	+	0

2-3b Suppose the bank loan of $0.50 to the firm in Table 2.3.1 at 10%/year interest represents the opportunity cost model of the bank. If the depositors who represent the bank's debt cash flow supplied 90% of the loan at 6%/year interest, determine the marginal capital efficiency, $Ø_m$, of the bank.

2-3c Assume the remaining 10% of the loan represents the equity cash flow of the bank. Determine the internal rate of return of the bank's equity investors and construct a table based on equation (2.3.1) in which the TCF, DCF and ECF of the bank are discounted at their respective internal rates of return. (see Tables 2.3.1 to 2.3.4)

2-4a (Section 2.2 - Engineering Discount Rates) Many alternatives lack a single positive IRR constant. Nevertheless, these alternatives may have economic significance because borrowers pay lenders more than they receive over a delayed time pattern. More specifically, let us consider three numbers (X,Y,Z) where Z > 0. Suppose person A has Z dollars now but needs X+Y dollars one year later. Therefore, A deposits Z dollars with bank B who agrees to loan X+Y dollars to A one year after the deposit. To complete the transaction, A must pay B an amount XY/Z dollars one year after receiving the X+Y loan.

Prove that the Z+XY/Z payments of A are greater than the X+Y loan of B under the following conditions: (1) X,Y > Z > 0, (2) Z > X,Y > 0, (3) Z > X > 0 > Y, and (4) $(X-Y)^2 < 0$ so that X and Y are complex conjugates a±bj where j=√-1. Hint: Divide Z+XY/Z > X+Y by Z and transpose (X+Y)/Z to the left-hand side of the inequality which can be factored into a product [1-(X/Z)][1-(Y/Z)] which must be greater than zero under the four specified conditions.

2-4b Prove that the transaction between A and B has two negative internal rates of return i_1 and i_2 if Z > X,Y > 0. Also prove the transaction between A and B has two positive internal rates of return i_1 and i_2 if 0 < Z < X,Y. Hint: Set the NFV function of (2B.3) equal to zero and solve by the quadratic formula. In general, $Z[(1+i_1)+(1+i_2)] = X+Y$ and $Z[(1+i_1)*(1+i_2)] = XY/Z$.

2-4c Determine the internal rates of return in **2-4a** if (X,Y,Z) = (5000,6000,7000). How much must A pay B at the end of the second year?

2-4d As an alternative to the transactions in **2-4c**, suppose A could invest Z = $7000 at 4% simple interest for one year. Then the proceeds from the $7000 investment are used to reduce the bridge loan which A needs to borrow to make up the X+Y = $11,000 investment. If the bridge loan is borrowed at 15% simple annual interest, how much does A need to pay off the loan one year later? In which way is A better off?

2-5a (Section 2.2 - Engineering Discount Rates) Internal rate of return is insensitive to differences in the scale and lifespan of alternatives. Thus, an investment of $1 which yields $1.50 one year later has the same 50%/year IRR as $2 which yields $3.00 one year later. The insensitivity of IRR to differences in lifespans of alternatives is less obvious. In this regard, let us construct a sequence of one-year, two-year, etc. alternatives which have the same IRR as follows. Suppose the one-year alternative has an input of $1 which yields 1+w dollars one year later for an IRR of w%/year. The two-year alternative has an input of $1 which yields (1+w)/2 dollars one year later and $(1+w)^2/2$ dollars two years later for an IRR=w%/yr. The three-year alternative has an input of $1 which yields (1+w)/2 dollars one year later, $(1+w)^2/4$ dollars two years later, and $(1+w)^3/4$ dollars three years later for an IRR=w%/yr, and so on.

2-5b Construct a sequence of three 30%/year IRR alternatives A, B and C with lifespans of one to three years. Show that the incremental internal rates of return, ΔIRR, of cash-flow differences between pairs of these alternatives is also 30%/year. Show that discounting below 30%/year gives ΔNPV{C} > ΔNPV{B} > ΔNPV{A} > 0 while discounting above 30%/year gives 0 > ΔNPV{A} > ΔNPV{B} > ΔNPV{C}. Assuming MARR = 30%/year and the cost of borrowing money is 10%/year, calculate the marginal capital efficiency, $Ø_m$, of alternatives A, B and C. By definition, $Ø_m$ is the ratio of the present value of the output to the present value of the input, discounted at the cost of borrowing money.

Chapter Two - Suggested Readings

The principles of the internal rate of return concept were first laid down in Irving Fisher's famous work "The Rate of Interest (New York: Macmillan Co., 1907). These ideas may also be found in Fisher's later and better known work "The Theory of Interest" (New York: Macmillan Co., 1930). The books listed here cover the concept of investment opportunity costs appearing in the technical literature on engineering and managerial economics. As far as this author is aware, the economic literature does not distinguish between operational definitions of investment and borrowing opportunity costs, which is a major emphasis of this book. The general concept of opportunity costs is not included in this list of suggested readings because it appears almost everywhere in the literature on macroeconomics and microeconomics.

Au, T. and Au, T. P., *Engineering Economics for Capital Investment Analysis,* Allyn & Bacon Inc., 1983.

Barish, N. N., and Kaplan, S., *Economic Analysis for Engineering and Managerial Decision-Making*, McGraw-Hill Book Co., 1978.

Blank, L.T. and Tarquin, A. J., *Engineering Economy*, 4th Ed., McGraw-Hill Book Co., 1998.

Bussey, L. E., *The Economic Analysis of Industrial Projects*, Prentice-Hall, 1978.

DeGarmo, E. P., Sullivan, W. G. and Bontadelli, J. A., *Engineering Economy*, 8th Ed., Macmillan Publishing Co., 1988.

Eschenbach, T. G., *Engineering Economy: Applying Theory to Practice*, R.D.Irwin, Inc., 1995.

Fabrycky, W. J., Thuesen, G. J. and Verma, D., *Economic Decision Analysis*, 3rd Ed., Prentice-Hall, Inc., 1998.

Fleischer, G. A., *Engineering Economy - Capital Allocation Theory*, Wadsworth, Inc., 1984.

Garrett, D. E., *Chemical Engineering Economics*, Van Nostrand Reinhold, 1989.

Grant, E. L., Ireson, W. G. and Leavenworth, R. S., *Principles of Engineering Economy*, 8th Ed., John Wiley & Sons, 1990.

Humphreys, K. K., *Jelen's Cost and Optimization Engineering*, 3rd Ed., McGraw-Hill Book Co., 1991.

Kleinfeld, I. H., *Engineering & Managerial Economics*, Holt, Rinehart and Winston, 1986.

Levy, H. and Sarnat, M., *Capital Investment & Financial Decisions*, 5th Ed., Prentice-Hall, Inc., 1994.

Morris, W. T., *Engineering Economy*, Richard D. Irwin, Inc., 1960.

Newnan, D. G., *Engineering Economic Analysis*, 4th Ed., Engineering Press, Inc., 1991.

Park, C. S., *Contemporary Engineering Economics*, 2nd Ed., Addison-Wesley, 1997.

Riggs, J. L. and West, T. M., *Engineering Economics*, 3rd Ed., McGraw-Hill Book Co., 1986.

Smith, G. W., *Engineering Economy - Analysis of Capital Expenditures*, 4th Ed., Iowa State University Press, 1987.

Steiner, H. M., *Engineering Economic Principles*, 2nd Ed., McGraw-Hill Companies, 1996.

Swalm, R. O., and Lopez-Leautaud, J. L., *Engineering Economic Analysis: A Future Wealth Approach*, John Wiley & Sons, 1984.

Thuesen, H. G., Fabrycky, W. J. and Thuesen, G. J., *Engineering Economy*, 7th Ed., Prentice-Hall, 1988.

White, J. A., Agee, M. H. and Case, K. E., *Principles of Engineering Economic Analysis*, 3rd Ed., John Wiley & Sons, 1989.

Young, D., *Modern Engineering Economy*, John Wiley & Sons, Inc., 1993.

Chapter Three - Accounting for Time Equivalences

Section 3.1 - Financial and Managerial Accounting

Every economic organization has two forms of accounting, namely, *financial accounting* which deals with the past and *managerial accounting* which plans for the future.

In financial accounting, revenues and expenses are aggregated monthly, quarterly, semiannually or annually without adjusting for the time value of money during an accounting period. All data are recorded on an *as is* basis within each accounting period to provide reliable information for auditing purposes, paying taxes and *external* reporting to creditors and investors in the accepted language of the business world. Owing to the needs of outsiders, financial accounting provides a timely history of past data that are grounded in fact. The use of "generally accepted accounting principles" by certified public accountants (CPAs) may be the best assurance of the accuracy of the data presented.

In managerial accounting, sometimes called *cost accounting*, there is a detailed dating of causally related revenues and expenses of project alternatives that may span many accounting periods. This type of accounting serves the purposes of *internal* reporting to managers who choose best courses of action, implement planning decisions, and evaluate the performance that provides feedback for improving results. Here the primary concerns are controlling when and how to do a project and at what cost; and how much the timing, quality and quantity of the output affect its pricing, marketability, and profitability. Because alternatives often span large periods of time, the time value of money is needed to determine the best alternatives.

In order to illustrate the time value of money in problems of economic decision-making, let us use the ΔNPV criterion to rank 4 mutually exclusive alternatives which cost $1,000 each and yield end-of-year (EOY) cash flows (CF) and cumulative cash flows (CCF) as listed in Table 3.1.1 below.

Table 3.1.1 - Cash flows of mutually exclusive alternatives A_1, A_2, A_3, A_4 costing $1,000 each.

EOY	CF{A_1}	CCF{A_1}	CF{A_2}	CCF{A_2}	CF{A_3}	CCF{A_3}	CF{A_4}	CCF{A_4}
1	$300	-$700	$400	-$600	$400	-$600	$400	-$600
2	$300	-$400	$350	-$250	$350	-$250	$350	-$250
3	$300	-$100	$300	$50	$300	$50	$300	$50
4	$300	$200	$300	$350	$250	$300	$200	$250
5	$300	$500	$300	$650	$200	$500	$150	$400

By inspecting Table 3.1.1 row-by-row, we find the CCF{A_2} are always greater than or equal to CCF{A_1}, CCF{A_3} or CCF{A_4}. Consequently, comparing A_2 to A_1, A_3, or A_4 has the same characteristics as those in Figure 2.1.1(a). Therefore, the net present-value ΔNPV{A_2} is greater than ΔNPV{A_1}, ΔNPV{A_3} or ΔNPV{A_4} for all positive discount rates.

Similarly, the *cumulative cash flows* CCF{A_3} are always greater than or equal to CCF{A_1} and CCF{A_4}. Again, the comparison of A_3 to A_1 or A_4 has the same characteristics as those in Figure 2.1.1(a). Therefore, the net present-value ΔNPV{A_3} is greater than ΔNPV{A_1} or ΔNPV{A_4} for all positive discount rates. Only the ΔNPV comparison of A_1 and A_4 is needed to complete the rankings of the four alternatives.

The *cumulative net revenues* CCF{A_1} are not always greater than or equal to CCF{A_4}. Since CCF{A_1} < CCF{A_4} from EOY 1 to EOY 4, but CCF{A_1} > CCF{A_4} at EOY 5, the ΔNPV comparisons of A_1 and A_4 depend upon discount rate magnitudes. The discount rate, i, which makes ΔNPV{A_1} = ΔNPV{A_4} is determined by setting ΔNPV{A_1-A_4} = 0. By methods which are explained later in this chapter, we find i = 16.94% per year. At all discount rates above 16.94%, ΔNPV{A_1} < ΔNPV{A_4}, and at all nonnegative discount rates below 16.94%, ΔNPV{A_1} > ΔNPV{A_4}. Thus, at a 20% discount rate, ΔNPV{A_1}= -$102.82 < ΔNPV{$A_4$}= -$93.27, and at a 10% discount rate, ΔNPV{A_1}= $137.24 > ΔNPV{$A_4$}= $108.03 .

More generally, economic decision-making for any set of mutually exclusive or independent alternatives can always be carried out by time-shifting their revenues and expenditures to equivalent values at a common point of time where proper comparisons can be made. It is also essential for budgeting purposes that the system of accounting for time equivalences should be compatible with both financial and managerial accounting. In order to account for time equivalences in a manner which is compatible with current systems of financial and managerial accounting, the remainder of this chapter is organized as follows:

<u>Section 3.2 - Cash Flows and Time Equivalences</u>: The definitions of cash flows and present and future time equivalences are explained within the context of an accounting system designed to serve the purposes of both financial and managerial accounting. The proposed system of accounting can handle both constant and variable interest rates, and both discrete and continuous cash flow patterns. The accounting system is also designed to be carried out with scientific pocket calculators or microcomputer spreadsheets.

<u>Section 3.3 - Simple and Compound Interest</u>: The definitions of simple and compound interest are explained within the framework of the same accounting system described in Section 3.2 . The concept of the focal date property is introduced as a relationship between cash flows and time equivalences whereby an agreement between a lender and borrower on the outstanding balance of a loan at the end of one period of time is equivalent to a tacit agreement between the lender and borrower about the loan balance at the ends of all other periods of time.

<u>Section 3.4 - Nominal and Effective Interest Rates</u>: The definitional relationship of nominal and effective interest rates are generalized here to accommodate various compounding frequencies and cash flow periods encountered in financial and managerial accounting. Commonplace confusions between nominal and effective interest rates are underscored in this section. The applications of the focal date property are extended from an agreement between a lender and a borrower concerning the loan balance at discrete points of time to an agreement between the lender and borrower at all points of time.

<u>Sections 3.5 through 3.8 - Constant Interest Rate Formulas</u>: When interest rates are constant and cash flows have simple discrete or continuous patterns, then constant interest rate formulas and tables (Appendix A) are presented here. In particular, constant interest rate formulas are developed in Section 3.5 for *Single Payment Cash Flows*, in Section 3.6 for *Uniform Series Cash Flows*, in Section 3.7 for *Arithmetic Gradient Cash Flows,* and in Section 3.8 for *Geometric Gradient Cash Flows*. These constant interest rate formulas are all derived using the system of accounting developed in Section 3.2.

<u>Section 3.9 - Summary</u>: The material of this chapter and important definitions are summarized here. For easy reference, the constant interest rate formulas are listed for both discrete and continuous cash flows in Appendix 3A. Basic financial functions of microcomputer spreadsheet programs are presented in Appendix 3B.

Section 3.2 - Cash Flows and Time Equivalences

The input and output of every alternative can be characterized by cash flows which are functions of time. When the cumulative cash flows of one alternative are not always greater than or equal to those of another, then the cash flow function of each alternative needs to be reduced to an equivalent value at a common point of time in order to make a proper comparison. The equivalent values of cash flow functions at a given point of time are determined in an accounting system whose mechanics will now be discussed.

When money is borrowed, not only does it need to be repaid, but an additional amount, called *interest*, also needs to be paid for depriving the lender from using the money. The rate at which interest needs to be paid depends on both the amount of money borrowed and the period of time the loan is outstanding. The ratio of the amount of interest to be paid to the amount of the outstanding loan per unit period of time is called the *interest rate*.

Let us consider an accounting system for a 4-year loan of $1,000 at the annual interest rates shown in column (2) of Table 3.2.1 for the years indexed in column (1). End-of-period *cash flows* are all defined as amounts of money that change hands across an interface which separates a system from its surroundings. Either the lender or borrower may be identified as the system. Cash flowing out of a system to its surroundings is defined as negative, and cash flowing into a system from its surroundings is defined as positive.

In the following tables, the lender is identified as the system. All end-of-year (EOY) cash flows are listed in cells $(4)_0$, $(4)_1$, $(4)_2$, $(4)_3$ and $(4)_4$ of column (4), where the cell subscripts refer to the End Of Year indexed in column (1). Thus, $(4)_1$ indicates a -$1,000.00 outflow from the lender to the borrower at the end of Year 0, and $(4)_4$ indicates a $1,286.18 inflow to the lender from the borrower at the end of Year 4. The remaining column and row entries of Table 3.2.1 are generated sequentially in time as described below.

Table 3.2.1 - Cash Flow Accounting for Future Time Equivalences (Forward in Time)

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	7.00%	$0.00	-$1,000.00	-$1,000.00
1	5.00%	-$1,050.00	$0.00	-$1,050.00
2	6.00%	-$1,113.00	$0.00	-$1,113.00
3	7.00%	-$1,190.91	$0.00	-$1,190.91
4	8.00%	-$1,286.18	$1,286.18	$0.00

Initially, we set the BCF (Before-*succeeding*-Cash-Flows) Balance $(3)_0$ = $0.00. The EOY Cash Flow $(4)_0$ = -$1,000.00 is given, and after adding $(4)_0$ to BCF Balance $(3)_0$, we obtain ACF (After-*preceding*-Cash-Flows) Balance $(5)_0$ = -$1,000.00 . During the first year, the interest rate is $(2)_1$ = 5.00%. Interest owed on ACF Balance $(5)_0$ held during the first year is included in BCF Balance $(3)_1$ = -$1,050.00 by multiplying $(5)_0$ = -$1,000.00 with the factor 1.05 which is one plus the 5% interest rate in $(2)_1$ expressed as a decimal fraction. After adding EOY Cash Flow $(4)_1$ = $0.00 to BCF Balance $(3)_1$ = -$1,050.00, we get ACF Balance $(5)_1$ = -$1,050.00 . This process is continued until we reach ACF Balance $(5)_4$ = $0.00. The procedure for determining future time equivalences of any set of cash flows and interest rates such as those in columns (2) and (4) of Table 3.2.1 are conveniently summarized by the following ABC accounting rules.

ABC Rules to Evaluate Future Time Equivalences of Cash Flows $(4)_0$ to $(4)_n$:

A. $(5)_0 = (3)_0 + (4)_0$: Initialize the account by setting BCF Balance $(3)_0$ = <u>$0.00</u>. It follows that $(5)_0 = (3)_0 + (4)_0$ by rule A. (same as Rule C)

B. $(3)_{k+1} = (1+i_{k+1}) \cdot (5)_k$: Continue accounting by multiplying the derived value of $(5)_0$ with $(1+i_1)$ expressed as a decimal to get $(3)_1$ by rule B.

C. $(5)_{k+1} = (3)_{k+1} + (4)_{k+1}$: Add $(4)_1$ to $(3)_1$ to get $(5)_1$ by rule C. Use $(5)_1$ to get $(3)_2$ by rule B. Add $(4)_2$ to $(3)_2$ to get $(5)_2$ by rule C, and so on until the ACF Balance $(5)_n$ of the last period is obtained in rule C.

The *time equivalence* of any set of cash flows is defined as their value at a given point of time. Thus, the cash flow $(4)_0$ = -$1,000.00 of Table 3.2.1 has negative time equivalences $(3)_1$ = -$1,050.00, $(3)_2$ = -$1,113.00, $(3)_3$ = -$1,190.91, and $(3)_4$ = -$1,286.18 at the ends of successive years. The loan $(4)_0$ = -$1,000.00 and its successive time equivalences $(3)_1$, $(3)_2$, $(3)_3$ and $(3)_4$ are depicted in Figure 3.2.1 by solid and dashed arrows respectively which are all negative. Because the borrower tacitly agrees to the time equivalence of $(3)_4$, the lender receives a payment of $(4)_4$ = $1,286.18 which is shown as a positive solid arrow in Figure 3.2.1, thereby making $(5)_4$ = $0.00 which wipes out the debt.

Arrow Conventions

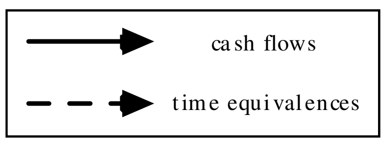
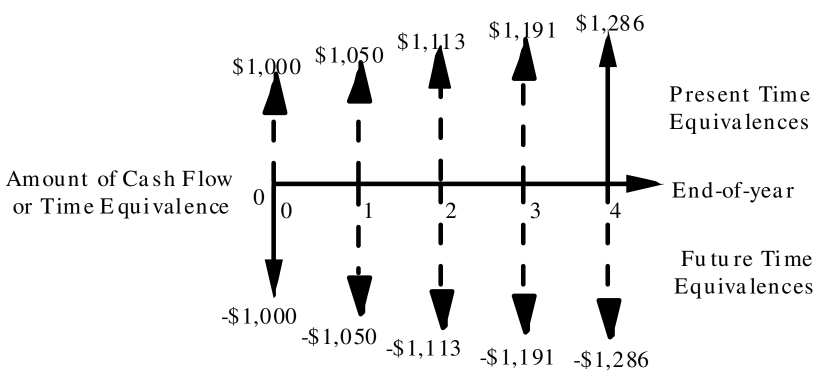


Figure 3.2.1 - Future and Present Cash Flows and Time Equivalences of Tables 3.2.1 & 3.2.2



The future time equivalents of the -$1,000.00 loan in Table and Figure 3.2.1 are amounts owed to the lender at the end of each year in the absence of any payments from the borrower. If the borrower pays back the loan sooner than four years, then the payments would be equal and opposite to the amounts owed to the lender at the end of each year. In order to determine present time equivalences of the borrower's $1,286.18 payment at the end of four years, the following A'B'C' accounting rules were used to generate Table 3.2.2 .

A'B'C' Rules to Evaluate Present Time Equivalences of Cash Flows $(4)_n$ to $(4)_0$:

A'. $(3)_n = (4)_n + (5)_n$:	Initialize by setting the last BCF (Before-*preceding*-Cash-Flows) Balance $(5)_n$ equal to $0.00. ACF (After-*succeeding*-Cash-Flows) Balance $(3)_n = (4)_n + (5)_n$ by Rule A'. (same as Rule C')
B'. $(5)_{n-1} = (3)_n/(1+i_n)$:	Continue accounting by dividing $(3)_n$ with $(1+ i_n)$ expressed as a decimal to get ACF Balance $(5)_{n-1}$ by rule B'.
C'. $(3)_{n-1} = (4)_{n-1} + (5)_{n-1}$:	Add $(4)_{n-1}$ to $(5)_{n-1}$ to get $(3)_{n-1}$ by rule C'. Use $(3)_{n-1}$ to get $(5)_{n-2}$ by rule B'. Add $(4)_{n-2}$ to $(5)_{n-2}$ to get $(3)_{n-2}$ by rule C', and so on until ACF Balance $(3)_0$ is obtained in rule B'.

Table 3.2.2 - Cash Flow Accounting for Present Time Equivalences (Backward in Time)

(1)EOY	(2)Interest/Year	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance
0	7.00%	$0.00	-$1,000.00	$1,000.00
1	5.00%	$1,050.00	$0.00	$1,050.00
2	6.00%	$1,113.00	$0.00	$1,113.00
3	7.00%	$1,190.91	$0.00	$1,190.91
4	8.00%	$1,286.18	$1,286.18	$0.00

The present time equivalences of $(4)_4$ = $1,286.18 in Table 3.2.2 are $(3)_4$ = $1,286.18, $(3)_3$ = $1,190.91, $(3)_2$ = $1,113.00, $(3)_1$ = $1,050.00 and $(5)_0$ = $1,000.00 at the time of the original loan. Payment $(4)_4$ and its time equivalences $(3)_4$, $(3)_3$, $(3)_2$, $(3)_1$ and $(5)_0$ are shown in Figure 3.2.1 by solid and dashed arrows which are all positive. Because $(3)_4$, $(3)_3$, $(3)_2$, $(3)_1$ and $(5)_0$ of Table 3.2.2 are equal and opposite to $(3)_4$, $(3)_3$, $(3)_2$, $(3)_1$ and $(4)_0$ of Table 3.2.1, it shows the borrower and lender were in basic agreement at the end of each year about the amount of the outstanding loan. This example brings out a fundamental property of accounting for cash flows and time equivalences called the *focal date property* as explained below.

Definition of the Focal Date Property - When the time equivalent of a set of cash flows is equal and opposite to the time equivalent of another set of cash flows at any focal point of time, then their time equivalents are equal and opposite at all other points of time. The focal date property is clearly validated in Figure 3.2.1 at the ends of each year. But the focal date property has not yet been demonstrated when end-of-year cash flows occur between the beginning and end of a loan, or when cash flows occur between the ends of a year. The application of the focal date property to end-of-year cash flows occurring between the beginning and end points of a loan will be discussed in Section 3.3 entitled "Simple and Compound Interest". The focal date property will then be applied to cash flows occurring between the ends of a year in Section 3.4 entitled "Nominal and Effective Interest Rates".

Section 3.3 - Simple and Compound Interest

Simple interest is defined as the interest which accrues during a period of time that is based on the interest rate multiplied by the outstanding balance at the beginning of the period without converting any interest into principal during the period. For example, instead of paying $1,286.18 at the end of four years, suppose the borrower of the previous example pays $50, $60, $70, and $80 at the end of each year in time to avoid compounding of interest into principal. At the end of the fourth year, an additional $1,000 is paid in $(4)_4$ to wipe out the loan as shown by $(5)_4$ = $0.00 in Tables 3.3.1 and 3.3.2 . Each simple interest payment is called a return on the lender's investment, and the repayment of the original loan at the end of the fourth year is called a return of the lender's investment.

Table 3.3.1 - Cash Flow Accounting for Future Time Equivalences

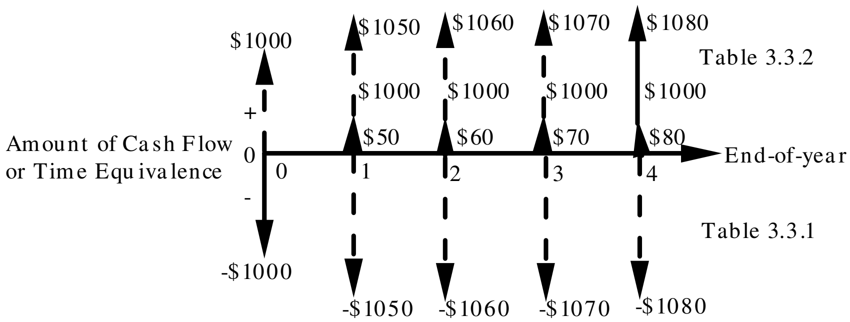
(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	7.00%	$0.00	-$1,000.00	-$1,000.00
1	5.00%	-$1,050.00	$50.00	-$1,000.00
2	6.00%	-$1,060.00	$60.00	-$1,000.00
3	7.00%	-$1,070.00	$70.00	-$1,000.00
4	8.00%	-$1,080.00	$1,080.00	$0.00

Table 3.3.2 - Cash Flow Accounting for Present Time Equivalences

(1)EOY	(2)Interest/Year	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance
0	7.00%	$0.00	-$1,000.00	$1,000.00
1	5.00%	$1,050.00	$50.00	$1,000.00
2	6.00%	$1,060.00	$60.00	$1,000.00
3	7.00%	$1,070.00	$70.00	$1,000.00
4	8.00%	$1,080.00	$1,080.00	$0.00

Since $(3)_4$, $(3)_3$, $(3)_2$, $(3)_1$ and $(5)_0$ in Table 3.3.2 are equal and opposite to $(3)_4$, $(3)_3$, $(3)_2$, $(3)_1$ and $(4)_0$ in Table 3.3.1, it follows that simple interest satisfies the focal date property at the end of every year. Figure 3.31 depicts positive ACF Balances $(3)_4$, $(3)_3$, $(3)_2$ and $(3)_1$ of Table 3.3.2 and negative BCF Balances $(3)_4$, $(3)_3$, $(3)_2$ and $(3)_1$ of Table 3.3.1. The negative BCF Balances of Table 3.3.1 are time equivalents of the outstanding loan before the lender receives payments from the borrower. The positive ACF Balances of Table 3.3.2 represent time equivalents owed to the lender after the borrower's cash flow payments.

Figure 3.3.1 - Simple Interest Cash Flow Diagram of Tables 3.3.1 and 3.3.2



Until the last payment, simple interest provides a return on but no return of the lender's investment. In general, accrued interest is charged at the end of each accounting or *compounding period*. If the accrued interest equals the end-of-period cash flow, it is called simple interest. But if the accrued interest differs from the end-of-period cash flow, then the difference is converted into the outstanding balance on which interest is charged in the next compounding period. These *compound interest* charges are incorporated in the ABC and A'B'C' accounting rules in order to handle problems of both simple and compound interest.

For example, suppose the borrower wants to repay the $1,000 loan with four equal end-of-year payments. The amount 'A' of the four payments can be determined by first finding the present value of four one-dollar end-of-year payments as shown in Table 3.3.3.

Table 3.3.3 - Cash Flow Accounting for Present Time Equivalences

(1)EOY	(2)Interest/Year	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance
0	7.00%	$3.4680	$0.00	$3.4680
1	5.00%	$3.6414	$1.00	$2.6414
2	6.00%	$2.7999	$1.00	$1.7999
3	7.00%	$1.9259	$1.00	$0.9259
4	8.00%	$1.0000	$1.00	$0.00

Since four one dollar installments starting at end of the first year would pay off $3.4680 of present debt, four installments of A = $1,000.00/3.4680 = $288.35 are needed to pay off $(5)_0$ = $1,000.00 of present debt in Table 3.3.2 . This is verified in Table 3.3.4 below.

Table 3.3.4 - Cash Flow Accounting for Present Time Equivalences

(1)EOY	(2)Interest/Year	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance
0	7.00%	$1,000.00	$0.00	$1,000.00
1	5.00%	$1,050.00	$288.35	$761.65
2	6.00%	$807.35	$288.35	$519.00
3	7.00%	$555.33	$288.35	$266.98
4	8.00%	$288.35	$288.35	$0.00

Because of the focal date property, we could use ABC instead of A'B'C' accounting rules to find that four one dollar end-of-year payments are equivalent to $4.4605 at the end of four years. We now have an alternate way of determining that four equal payments A = $1,286.18/4.4605 = $288.35 will pay off the loan balance $(3)_4$ = $1,286.18 in Table 3.2.2 .

A constant interest rate is often provided as an alternative to varying interest rates. For example, the borrower in Tables 3.3.1 and 3.3.4 may be offered a constant rate of interest which is an average that is defined from cash flows that satisfy the focal date property. Constant interest rates can be readily determined from spreadsheet calculations of internal rates of return (IRR) as described in Appendices 3B and 4B. For example, the internal rates of return of the cash flow patterns in Tables 3.3.1 and 3.3.4 are 6.422279% and 5.96345% per year, respectively, as verified in Tables 3.3.5 and 3.3.6.

Table 3.3.5 - Accounting for Future Time Equivalences of Table 3.3.1 Cash Flows

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	6.422279%	$0.00	-$1,000.00	-$1,000.00
1	6.422279%	-$1,064.22	$50.00	-$1,014.22
2	6.422279%	-$1,079.36	$60.00	-$1,019.36
3	6.422279%	-$1,084.83	$70.00	-$1,014.83
4	6.422279%	-$1,080.00	$1,080.00	$0.00

Table 3.3.6 - Accounting for Present Time Equivalences of Table 3.3.4 Cash Flows

(1)EOY	(2)Interest/Year	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance
0	5.96345%	$0.00	-$1,000.00	$1,000.00
1	5.96345%	$1,059.63	$288.35	$771.28
2	5.96345%	$817.28	$288.35	$528.93
3	5.96345%	$560.47	$288.35	$272.12
4	5.96345%	$288.35	$288.35	$0.00

Tables 3.3.5 and 3.3.6 show that a constant or average interest rate depends on the timing of the cash flows as well as the varying interest rates. Let us now determine the effective interest rates when the timing of cash flows differs from the compounding of interest rates.

Section 3.4 - Nominal and Effective Interest Rates

Nominal interest rates, **i**, are defined as interest-to-principal ratios per nominal time period of months, years, etc. Nominal interest rates based on one-year time periods are called Annual Percentage Rates (APR). Let $\mathbf{m \geq 1}$ denote the number of times interest is compounded into principal in the nominal time period. Therefore, **i/m** is a ratio of interest compounded into principal at the end of one-mth of the nominal time period.

Effective interest rate, $\mathbf{i_E}$, is defined as an interest rate which is compounded once at the end of *effective time interval*, $\mathbf{t_E}$, measured in nominal time period units. The product of compounding rate **m** and time interval $\mathbf{t_E}$ represents the number of interest-to-principal compoundings of ratio **i/m** in effective time interval $\mathbf{t_E}$.

Assuming constant interest rates, nominal and effective interest rates must satisfy equation (3.4.1) due to the frequencies of compounding in the nominal and effective time intervals. Equation (3.4.1) obeys two basic rules of dimensional analysis, namely, (1) quantities can be added only if they have the same dimensions. (i.e., ratios of interest rates to compounding rates such as $i_E/1$ or i/m can be added to the number 1 only because both terms are dimensionless), and (2) quantities can be exponentiated only if both the base and exponent are dimensionless (i.e., the base [1+(i/m)] can be exponentiated by mt_E because both are dimensionless).

$$1+(i_E/1) \equiv [1+(i/m)]^{mt_E} \text{ and } 1+(i/m) \equiv [1+(i_E/1)]^{1/mt_E} \quad ...(3.4.1)$$

Notation	Dimension
i = interest-to-principal ratio per nominal time period	1/T = per month, year, ...
m = number of compoundings per nominal time period	1/T = per month, year, ...
i/m = interest-to-principal ratio of 1/mth nominal time period	Ø = dimensionless
i_E = effective interest rate of the effective time interval	1/T = per month, year, ...
t_E = effective time interval in nominal time period units	T = months, years, ...
mt_E = number of compoundings in the effective time interval	Ø = dimensionless

Equation (3.4.1) defines effective interest rates, i_E, for time intervals, t_E, that could differ from nominal time periods. For example, a borrower gets a $1,000 loan at 10% APR compounded semiannually to be paid back at the end of two years. Table 3.4.1 lists nominal and effective interest rates at the ends of semiannual time periods (EOY/2).

Table 3.4.1 - Nominal and Effective Interest Rate Equivalences

(1)EOY/2	(2) (i/m)	(3) t_E	(4) mt_E	(5) $1+(i_E/1) \equiv [1+(i/m)]^{mt_E}$
0	5.0%	0.0	0.0	$1.00000 = [1.05]^{2*0.0}$
1	5.0%	0.5	1.0	$1.05000 = [1.05]^{2*0.5}$
2	5.0%	1.0	2.0	$1.10250 = [1.05]^{2*1.0}$
3	5.0%	1.5	3.0	$1.15763 = [1.05]^{2*1.5}$
4	5.0%	2.0	4.0	$1.21551 = [1.05]^{2*2.0}$

Effective interest rates i_E can be determined for two-year time periods, t_E=2, when the 10% APR is compounded semiannually. Thus, $1+(i_E/1) \equiv [1+(i/m)]^{mt_E} = [1+(0.10/2)]^{2*2} = 1.21551$ or the effective interest rate is i_E = 21.551% per two years. Conseqauently, five semiannual time periods EOY/2 in Table 3.4.2 can be condensed into two rows of two-year time periods EO2Y in Table 3.4.3 in order to obtain results at the end of two years.

Table 3.4.2 - Semiannual Accounting for Future Time Equivalences

(1)EOY/2	(2)Interest Rate	(3)BCF Balance	(4)EOY/2 Cash Flow	(5)ACF Balance
0	5.00%	$0.00	-$1,000.00	-$1,000.00
1	5.00%	-$1,050.00	$0.00	-$1,050.00
2	5.00%	-$1,102.50	$0.00	-$1,102.50
3	5.00%	-$1,157.63	$0.00	-$1,157.63
4	5.00%	-$1,215.51	$1,215.51	$0.00

Table 3.4.3 - Biennial Accounting for Future Time Equivalences

(1)EO2Y	(2)Interest Rate	(3)BCF Balance	(4)EO2Y Cash Flow	(5)ACF Balance
0	21.551%	$0.00	-$1,000.00	-$1,000.00
1	21.551%	-$1,215.51	$1,215.51	$0.00

Another application of equation (3.4.1) is to determine effective interest rate i_E for cash flows that occur before interest is compounded into principal. For example, consider a person who owns a $1,000 bond which pays 10% APR semiannually. Three months after receiving the $50 semiannual dividend, the owner sells the bond at par (i.e., for $1,000) plus accrued interest. We need to determine effective interest rate $\mathbf{i_E}$ for an effective time interval $\mathbf{t_E}$ which is one-fourth of the one-year nominal time period. To solve this problem, we identify the parameters on the right hand side of equation (3.4.1) as follows:

$i = 10\%$ interest-to-principal ratio per one-year nominal time period
$m = 2$ compoundings per one-year nominal time period
$i/m = 10\%/2 = 5\%$ interest-to-principal ratio per compounding
t_E = effective time interval of 1/4 of the one-year nominal time period
$mt_E = 2/4 = 1/2$ compounding in the effective time interval

Substituting in equation (3.4.1) gives $1+(i_E/1) \equiv [1+(i/m)]^{mt_E} = [1+(0.10/2)]^{2*(1/4)} = 1.024695$ or the effective interest rate is $i_E = 2.4695\%$ for three months. Therefore, accrued *compound* interest would be $\$1{,}000 \cdot 0.024695 = \$24.695$. However, in commercial practice, the bondholder is paid accrued *simple* interest which is $\$1{,}000*[(90/360)*0.10] = \$1{,}000*0.025 = \$25.00$ for three months (i.e., 90 days of a 360-day year) or 30.5 cents more than accrued compound interest. In general, between the start and finish of a compounding period, accrued simple interest is larger than accrued compound interest.

Interest rates can be applied continuously between compoundings by increasing the compounding frequency. As the number of compoundings, m, increases without bound, the interest-to-principal ratio per compounding, i/m, approaches zero. The limit of $[1+(i/m)]^m$ as m becomes infinite equals e^i denoted as e^r in equation (3.4.2). Exponent r is an interest rate of the nominal time period which applies continuously during an effective time interval of one nominal time period and exponential e = 2.71828... is the base of natural logarithms.

$$1+(i_E/1) \equiv \lim_{(m\to\infty)} \{[1+(i/m)]^m\}^{t_E} = e^{r \cdot t_E} \text{ and } i_E = e^{r\,t_E} - 1 \qquad ...(3.4.2)$$

Effective time interval, t_E, is a continuous time measurement in nominal time period units so that exponent rt_E is dimensionless. The subscript of t_E indicates that its time interval corresponds to the time period of effective interest rate i_E, both of which are measured in nominal time period units. Taking natural logarithms of equation (3.4.2) gives

$$r \cdot t_E \equiv \ln[1+(i_E/1)] \qquad ...(3.4.3)$$

Let us consider the effective interest rate i_E at ends of t_E = 1, 2, 3 ... nominal time periods in which interest rate 'i' is compounded m times per nominal time period. Hence, $1+(i_E/1) = [1+(i/m)]^{m\cdot 1}$ at the end of t_E = 1 nominal time period, $1+(i_E/1) = [1+(i/m)]^{m\cdot 2}$ at the end of t_E = 2 nominal time periods and $1+(i_E/1) = [1+(i/m)]^{m\cdot n}$ at the end of t_E = n nominal time periods. Effective interest rates $1+(i_E/1) = [1+(i/m)]^{m\cdot n}$ for n = 1, 2, ... may be evaluated as $1+(i_E/1) = e^{r\cdot t_E}$ with accrued compound interest for continuous values of t_E by letting 'm' increase without bounds.

Suppose the $1,000 bond which pays 10% APR semiannually is sold at par three months after receiving a $50 dividend. The problem is to determine the accrued compound interest. First we determine $1+i_E = [1+(0.10/2)]^2 = 1.10250$ for a one-year nominal time period from equation (3.4.1). Then we determine the continuous nominal interest rate r = ln[1.10250] = 0.0975803 from equation (3.4.3). Finally, with t_E = 0.25 nominal time periods, accrued compound interest is $\$1,000i_E = \$1,000[e^{0.0975803*0.25}-1] = \$24.695$ by equation (3.4.2) which is the same as the previous result.

By equation (3.4.2), let us determine the number of years, t_E, required for a present amount to double when the nominal annual interest rate of r% is compounded continuously.

$$1+(i_E/1) = 2 = e^{r\cdot t_E};\ r\cdot t_E = \ln 2 = 0.69315.. \approx 0.70;\ t_E \approx 0.70/r \text{ or } 70/r\% \qquad ...(3.4.4)$$

The relationship $r\cdot t_E \approx 70/r\%$ is known as the Rule of 70. It estimates how many years for a quantity to double when it is compounded continuously at r% per year. Thus, an amount compounded continuously at a 10% per year will double in about 7 years.

The frequency of compounding has little effect on the Rule of 70. For example, the number of years, t_E, required for a quantity to double when it is compounded m times per year at a nominal interest rate of i% per year can be determined from equation (3.4.1). After determining t_E, we evaluate $i\%t_E$ in Table 3.4.4 to show how m and i% affect the Rule of 70. Taking natural logarithms of equation (3.4.1) and solving for t_E gives

$$t_E = \ln[1+(i_E/1)]/\ln[1+(i/m)]^m = \ln[2]/\ln[1+(i/m)]^m \qquad ...(3.4.5)$$

For example, suppose i = 10% per year and m = 2 compoundings per year. Substituting $[1+(i_E/1)] = 2$, $[1+(i/m)]^m = [1+0.05]^2 = 1.1025$ into equation (3.4.5), we find t_E = ln2/ln1.1025 = 7.103 years and $i\%\cdot t_E = 10\%\cdot t_E = 71.03$ as shown in Table 3.4.4 . Therefore, an amount deposited at a 10% APR compounded semiannually will double in 7.103 years which differs little from the 7 years approximated by the Rule of 70.

Table 3.4.4 - Rule of 70 Variations - Values of $i\%t_E$ as a function of i% and m.

i%/m	m = 1	m = 2	m = 4	m = 12	m = 52	m = ∞
5%	71.03	70.18	69.75	69.46	69.35	69.31
10%	72.73	71.03	70.18	69.60	69.38	69.31
15%	74.39	71.88	70.61	69.75	69.41	69.31
20%	76.04	72.73	71.03	69.89	69.45	69.31
25%	77.66	73.56	71.46	70.03	69.48	69.31

Constant interest-rate formulas for discrete end-of-period cash flows have equivalent formulations for continuous cash flows and compounding, and conversely, as explained in Sections 3.5 to 3.9 . Cash flows are almost never purely discrete or continuous. Comparing results of economic decision-making to financial accounting statements requires continuous cash flow estimates of engineering, marketing and finance to have discrete formulations.

Section 3.5 - Time Equivalences of Single Payment Cash Flows

The time equivalences of a single payment cash flow is concerned with finding present or future values of a single cash flow occurring at a given point of time. Consider a case where a lady deposits $1,000 in a bank that pays a nominal interest rate of 12% per annum which is compounded semiannually. She wishes to withdraw her balance at the end of two years. The problem is to find the future time equivalent of the lady's deposit and the present time equivalent of the bank's payment to the lady two years hence as depicted in Figure 3.5.1 where the lady depositor is identified as the system.

Figure 3.5.1 - Cash Flow and Time Equivalence of a Single Deposit and Bank Payment

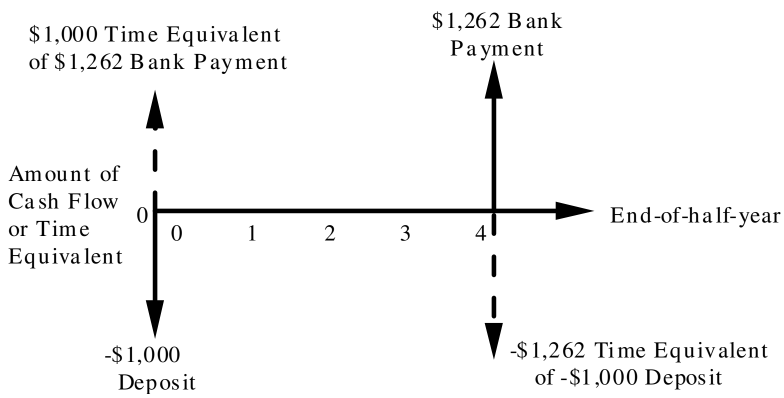


Table 3.5.1 shows the time equivalences of the -$1,000 deposit every half year for two years using the ABC accounting rules. The effective interest rate i_E per t_E = half-year compounding period (CP) is $1+(i_E/1) \equiv [1+(r/m)]^{m/2} = [1+(.12/2)]^{2/2}$, we find $i_E = 0.06$ or 6%.

Table 3.5.1 - ABC Accounting for 2-year Future Time Equivalences of a -$1,000 Deposit

(1)EOCP	(2)Interest/CP	(3)BCF Balance	(4)EOCP Cash Flow	(5)ACF Balance
0	6%	$0.00	-$1,000.00	-$1,000.00
1	6%	-$1,060.00	$0.00	-$1,060.00
2	6%	-$1,123.60	$0.00	-$1,123.60
3	6%	-$1,191.02	$0.00	-$1,191.02
4	6%	-$1,262.48	$1,262.48	$0.00

At the end of two years, the amount in the lady's account is $(3)_4$ = -$1,262.48 . When the bank pays the lady $(4)_4$ = $1,262.48 two years from now, we can verify that the present value of that payment is $(5)_0$ = $1,000 as shown in Table 3.5.2 below.

Table 3.5.2- A'B'C' Accounting for 2-year Present Time Equivalences of a $1,262 Payment

(1)EOCP	(2)Interest/CP	(3)ACF Balance	(4)EOCP Cash Flow	(5)BCF Balance
0	6%	$0.00	-$1,000.00	$1,000.00
1	6%	$1,060.00	$0.00	$1,060.00
2	6%	$1,123.60	$0.00	$1,123.60
3	6%	$1,191.02	$0.00	$1,191.02
4	6%	$1,262.48	$1,262.48	$0.00

As a generalization of Table 3.5.1, let us determine the future time equivalences F after n compounding periods of a single payment P at a constant interest rate of i% per compounding period as shown in Table 3.5.3 below.

Table 3.5.3 - ABC Accounting System for Future Time Equivalences F of Single Payment P

(1)EOCP	(2)Interest/CP	(3)BCF Balance	(4)EOP Cash Flow	(5)ACF Balance
0	i%	$0.00	P	$P(1+i)^0$
1	i%	$P(1+i)^1$	$0.00	$P(1+i)^1$
2	i%	$P(1+i)^2$	$0.00	$P(1+i)^2$
.	.	.	.	.
n-1	i%	$P(1+i)^{n-1}$	$0.00	$P(1+i)^{n-1}$
n	i%	$P(1+i)^n$	$0.00	$F=P(1+i)^n$

The formula for future time equivalent F at the end of n compounding periods of a single payment P which earns i% interest per compounding period is

$$F = P(1+i)^n \equiv P(F/P, i, n) \quad ....(3.5.1)$$

The functional notation of the ASEE (American Society for Engineering Education) for (3.5.1) is $(F/P,i,n) \equiv (1+i)^n$ which represents the future value F per dollar of P at the end of n periods where the interest rate is i% per period. Both (F/P,i,n) and $(1+i)^n$ are dimensionless because F and P are both expressed in dollars; interest rate i is based on one compounding per period so that i/1 is dimensionless; and number of periods n is a dimensionless integer. The quantity (F/P,i,n) can be found in Appendix A for various values of i and n, or it can be evaluated from its formula using a scientific pocket calculator or microcomputer spreadsheet.

The functional notation (F/P,i,n) is very useful for stating questions in shorthand because of the apparent cancellation of the P's in the product P(F/P,i,n). For example, the problem of finding the future value of a $1,000 deposit two years from now when the nominal interest rate of 12% per year is compounded semiannually could be stated as

$$\text{Find } F = P(F/P,i,n) = \$1{,}000(F/P,6\%,4) = \$1{,}000 \cdot (1.06)^4 = \$1{,}262.48$$

The same results that are obtained by discrete compounding using equation (3.5.1) could also be obtained by continuous compounding using equation (3.5.2)

$$F = Pe^{rt} \equiv P(F/P,r,t) \quad ....(3.5.2)$$

where the dimensionless exponent rt is uses the same units of time that i_E is measured. Applying equation (3.4.1), we can solve for the effective interest rate i_E per year from $1+(i_E/1) \equiv [1+(r/m)]^m = [1+(.12/2)]^2$ which gives $i_E = 0.1236$ or 12.36%/year. Since $i_E = 0.1236$ is the effective interest rate per year, it follows that t is a continuous variable also measured in years and r is a continuous annual interest rate which can be determined from equation (3.4.3) as follows:

$$r \cdot t \equiv \ln[1+(i_E/1)] = \ln[1.1236] = 0.1165378 \text{ or } r = 11.65378\%/\text{year}$$

Finally, we can verify the results of Table 3.5.1 and equation (3.5.1) for integer values of n when the future time equivalent F is calculated from equation (3.5.2) with r = 11.65378% per year and t is expressed in years as follows:

$F = Pe^{rt} = \$1,000.00e^{0.1165378 \cdot 0.0} = \$1,000.00 \cdot (1.00000) = \$1,000.00$
$F = Pe^{rt} = \$1,000.00e^{0.1165378 \cdot 0.5} = \$1,000.00 \cdot (1.06000) = \$1,060.00$
$F = Pe^{rt} = \$1,000.00e^{0.1165378 \cdot 1.0} = \$1,000.00 \cdot (1.12360) = \$1,123.60$
$F = Pe^{rt} = \$1,000.00e^{0.1165378 \cdot 1.5} = \$1,000.00 \cdot (1.19102) = \$1,191.02$
$F = Pe^{rt} = \$1,000.00e^{0.1165378 \cdot 2.0} = \$1,000.00 \cdot (1.26248) = \$1,262.48$

In order to generalize the results of Table 3.5.2, let us determine the present time equivalent P of a future single payment F occurring n compounding periods later at a constant interest rate of i% per compounding period as shown in Table 3.5.4 below.

Table 3.5.4 - A'B'C' Accounting for Present Time Equivalent P of Future Single Payment F

(1)EOP	(2)Interest/Period	(3)ACF Balance	(4)EOP Cash Flow	(5)BCF Balance
0	i%	$F(1+i)^{-n}$	\$0.00	$P=F(1+i)^{-n}$
1	i%	$F(1+i)^{-(n-1)}$	\$0.00	$F(1+i)^{-(n-1)}$
2	i%	$F(1+i)^{-(n-2)}$	\$0.00	$F(1+i)^{-(n-2)}$
.	.	.	.	.
n-1	i%	$F(1+i)^{-1}$	\$0.00	$F(1+i)^{-1}$
n	i%	$F(1+i)^{0}$	F	\$0.00

Therefore, the present time equivalent P of a future single payment F that occurs n periods later when money earns interest at the constant rate of i% per period is

$$P = F(1+i)^{-n} \equiv F(P/F, i, n) \qquad ....(3.5.3)$$

The ASEE functional notation for equation (3.5.3) is $(P/F,i,n) \equiv (1+i)^{-n}$ which represents the present value P per dollar of F at the end of n compounding periods during which time the interest rate is i% per compounding period. Suppose we want to determine the present time equivalent of a \$1,262.48 future cash flow that occurs two years later when 12% interest per year is compounded semiannually. This question can now be asked in the form:

Find $P = F(P/F,i,n) = \$1,262.48(P/F,6\%,4) = \$1,262.48(1.06)^{-4} = \$1,000.00$

The same results that are obtained by discrete compounding using equation (3.5.3) could also be obtained by continuous compounding using equation (3.5.4)

$$P = Fe^{-rt} \equiv F(P/F,r,t) \qquad ....(3.5.4)$$

where the dimensionless exponent rt is uses the same units of time that i_E is measured. Applying equation (3.4.1), we previously found i_E = 11.65378%/year. Since i_E = 11.65378% is the effective interest rate per year, it follows that t is a continuous variable also measured in years. Application of equation (3.4.3) with t = 1 verifies that r = 11.65378% is the continuous annual interest rate. Finally, the present time equivalent P is calculated from equation (3.5.4) with r = 11.65378% per year and t = 2 years as follows:

$P = Fe^{-rt} = \$1,262.48e^{-0.1165378 \cdot 2} = \$1,262.48*(0.7920937) = \$1,000.00$

Quantities (F/P,i,n), (F/P,r,t), (P/F,i,n) and (P/F,r,t) are listed in Appendix A for various values of i, n, r and t. They can also be calculated from equations (3.5.1) to (3.5.4) with scientific pocket calculators or microcomputer spreadsheets.

Section 3.6 - Time Equivalences of Uniform Series Cash Flows

Some typical examples with uniform series cash flows include monthly payments made on automobiles, repayments of bank loans in equal installments, etc. Consider a case of a lady winning a $10,000 lottery prize. Under the rules of the lottery, the present value of the $10,000 prize is placed in a bank from which she could withdraw $2,500 at the end of each year for four years as depicted in Figure 3.6.1 below where the lady is identified as the system. However, she decides to accumulate each $2,500 in a bank which pays an annual interest rate of 10% compounded quarterly. To calculate the lady's balance at the end of four years, we use the ABC accounting rules of Section 3.2 in Table 3.6.1. We first calculate the effective annual rate of interest from equation (3.4.1) as follows:

$$1+i_E \equiv [1+(r/m)]^m = [1+(0.10/4)]^4 = 1.1038129 \qquad ....(3.4.1)$$

Figure 3.6.1 - Uniform Series Cash Flows and Future Time Equivalence Diagram

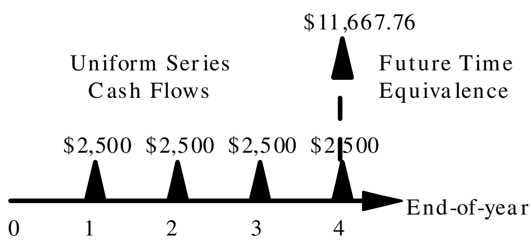


Table 3.6.1 - ABC Accounting for Future Time Equivalences of a Uniform Series of Deposits

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	10.38129%	$0.00	$0.00	$0.00
1	10.38129%	$0.00	$2,500.00	$2,500.00
2	10.38129%	$2,759.53	$2,500.00	$5,259.53
3	10.38129%	$5,805.54	$2,500.00	$8,305.54
4	10.38129%	$9,167.76	$2,500.00	$11,667.76

Thus, the balance in her account from the $10,000 lottery prize is $11,667.76 at the end of four years . This example can be generalized as shown in Table 3.6.2 from which the formulas for the future time equivalence F of a uniform series of cash flows A are derived.

Table 3.6.2 - Sinking Fund Accumulation F from Uniform Series of Cash Flows A

(1)EOP	(2)Interest/Period	(3)BCF Balance	(4)EOP Cash Flow	(5)ACF Balance
0	i%	$0.00	$0.00	$(A/i)[(1+i)^0-1]$
1	i%	$(A/i)[(1+i)^1-(1+i)]$	A	$(A/i)[(1+i)^1-1]$
2	i%	$(A/i)[(1+i)^2-(1+i)]$	A	$(A/i)[(1+i)^2-1]$
.	.	.	.	.
n-1	i%	$(A/i)[(1+i)^{n-1}-(1+i)]$	A	$(A/i)[(1+i)^{n-1}-1]$
n	i%	$(A/i)[(1+i)^n-(1+i)]$	A	$F=(A/i)[(1+i)^n-1]$

The future-value time equivalence F of a uniform series cash flows A in Table 3.6.2 is called a *sinking fund accumulation*. The constant amounts A are deposited at a constant interest rate i% per period. If the kth deposit A was time-shifted to the end of the nth period, the value would be $A(F/P,i,n-k) = A(1+i)^{n-k}$, and the sum of all n time-shifted deposits would be

$$F = \sum_{k=1}^{n} A(1+i)^{n-k} = (A/i)[(1+i)^n - 1] \equiv A(F/A,i,n) \quad ....(3.6.1)$$

Equation (3.6.1) can now be used to obtain the balance in the lady's account at the end of four years as follows:

$$F = \$2{,}500(F/A,10.38129\%,4) = (2{,}500/0.1038129)[1.1038129^4-1] = \$11{,}667.76$$

The functional notation $(F/A,i,n) \equiv [(1+i)^n - 1]/i$ represents how $1 deposited periodically will accumulate at the end of the nth period. Quantity $(F/A,i,n) \equiv [(1+i)^n - 1]/i$ is dimensionless because F and A are expressed in dollars, the interest rate i is based on one compounding per period so that i/1 is dimensionless, and the exponent n is a dimensionless integer.

The reciprocal $(A/F,i,n) \equiv i/[(1+i)^n - 1]$ represents the amount A that needs to be deposited periodically for n periods in order to accumulate a $1 sinking fund at the end of the nth period when the interest rate is i% per period as shown in equation (3.6.2) below. Both (F/A,i,n) and (A/F,i,n) can be found in Appendix A for various values of i and n, or they can be evaluated directly from their formulas with a scientific pocket calculator. When $i = 0$ or $n = \infty$, it can be shown $(F/A,0,n) = n$, $(A/F,0,n) = 1/n$, $(F/A,i,\infty) = \infty$ and $(A/F,i,\infty) = 0$.

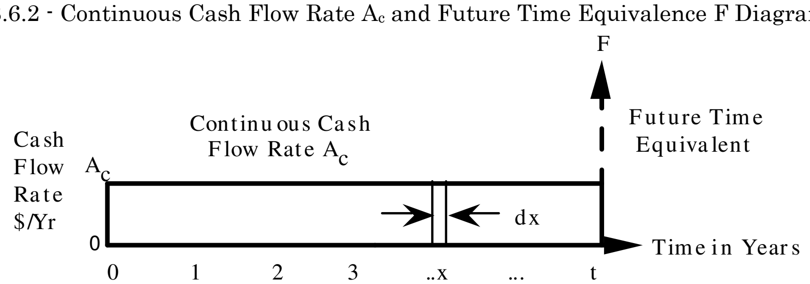
$$A = Fi/[(1+i)^n - 1] \equiv F(A/F,i,n) \quad ....(3.6.2)$$

An example of using equation (3.6.2) is to determine the amount of four uniform end-of-year deposits that are needed to accumulate a sinking fund of $11,667.76 at the end of four years when the interest rate is 10% per year, compounded quarterly. The answer is

$$A = Fi/[(1+i)^n - 1] = \$11{,}667.76 \cdot 0.1038129/[1.1038129^4-1] = \$2{,}500$$

Let us now consider a situation as depicted in Figure 3.6.2 where money is deposited continuously at the rate A_c dollars per year into a sinking fund where interest is compounded continuously at the nominal rate of r% per year. The problem is to determine the future time equivalent F accumulated in the sinking fund at the end of t years.

Figure 3.6.2 - Continuous Cash Flow Rate A_c and Future Time Equivalence F Diagram



During the infinitesimal time interval dx at time x, an amount $A_c dx$ dollars will be deposited into the sinking fund. If the infinitesimal deposit $A_c dx$ was time-shifted forwards from time x to time t, its value would be $A_c dx[e^{r(t-x)}]$. Upon summing elements $A_c dx[e^{r(t-x)}]$ from x=0 to x=t, future time equivalent F of the sinking fund at the end of t years would be

$$F = \int_0^t A_c e^{r(t-x)}\,dx = (A_c/r)[e^{rt} - 1] \equiv A_c(F/A_c,r,t) \quad ....(3.6.3)$$

Despite similarities of $(F/A_c,r,t) \equiv [e^{rt} - 1]/r$ and $(F/A,i,n) \equiv [(1+i)^n - 1]/i$, their dimensions differ significantly. Both $(F/A_c,r,t)$ and F/A_c have dimensions of time because they are ratios of F in dollars to A_c in dollars per unit of time. More specifically, $(F/A_c,r,t) \equiv [e^{rt} - 1]/r$ has dimensions of time because the numerator $[e^{rt} - 1]$ is dimensionless but the denominator r is an interest rate whose reciprocal has dimensions of time. On the other hand, both (F/A,i,n) and F/A are dimensionless, denoted by Ø, because they are ratios of F($) to A($). Moreover, both i/1 and n are dimensionless. In particular, n is an integer which represents both a number of interest periods and a number of distinct cash flows.

The parameters of equation (3.6.3) can be estimated from the discrete data of Figure and Table 3.6.1 so that the results will be the same as those of equation (3.6.1). For this purpose, suppose the lady's lottery winnings are accumulated at a continuous rate of $\$A_c$ per year for four years in a bank that pays an effective annual interest rate i_E = 10.38129% by compounding continuously at its nominal annual interest rate r which can be determined from i_E = 10.38129% using equation (3.4.3) as shown below. The continuous rate of bank deposits A_c can be determined using equation (3.6.9) so that the sinking fund accumulations F_t of equation (3.6.3) in her account would be the same as F_n of equation (3.6.1).

$$r \cdot 1 = \ln[1+(i_E/1)] = \ln[1.1038129] = 0.0987705 \qquad ....(3.4.3)$$

$$A_c = Ar/i_E = \$2,500(0.0987705/0.1038129) = \$2,378.57/yr \qquad ....(3.6.9)$$

Upon substituting A_c = \$2,378.57, r = 9.87705% and t = 0,1,2,3,4 years in equation (3.6.3); and A = \$2,500, i_E = i = 10.38129% and n = 0,1,2,3,4 years in equation (3.6.1), we get

$$F_t = A_c(F/A_c,r,t) = (A_c/r)[e^{rt} - 1] \quad (3.6.3) \qquad F_n = A(F/A,i,n) = (A/i)[(1+i)^n - 1] \quad (3.6.1)$$
$$F_0 = (2,379/0.09877)[e^{0.09877 \cdot 0} - 1] = \$0.00 \qquad F_0 = (2,500/0.1038)[(1.1038)^0 - 1] = \$0.00$$
$$F_1 = (2,379/0.09877)[e^{0.09877 \cdot 1} - 1] = \$2,500.00 \qquad F_1 = (2,500/0.1038)[(1.1038)^1 - 1] = \$2,500.00$$
$$F_2 = (2,379/0.09877)[e^{0.09877 \cdot 2} - 1] = \$5,259.53 \qquad F_2 = (2,500/0.1038)[(1.1038)^2 - 1] = \$5,259.53$$
$$F_3 = (2,379/0.09877)[e^{0.09877 \cdot 3} - 1] = \$8,305.54 \qquad F_3 = (2,500/0.1038)[(1.1038)^3 - 1] = \$8,305.54$$
$$F_4 = (2,379/0.09877)[e^{0.09877 \cdot 4} - 1] = \$11,667.76 \qquad F_4 = (2,500/0.1038)[(1.1038)^4 - 1] = \$11,667.76$$

It is physically impossible to continuously deposit \$2,378.57/year and continuously compound at the interest rate r = 9.87705%/year over the four-year period. But it gives the same results as depositing \$2,500 at the ends of four years in a bank that pays 10% interest per year compounded quarterly. Because continuous cash flows and compounding would be equivalent quarterly to discrete cash flows and quarterly compounding, equation (3.6.3) may be more convenient than equation (3.6.1) for accounting purposes.

The reciprocal of $(F/A_c,r,t)$ is $(A_c/F,r,t) \equiv r/[e^{rt} - 1]$ as shown in equation (3.6.4). It represents the annual rate of depositing money continuously for t years to accumulate a \$1 sinking fund when the nominal annual interest rate is r% compounded continuously. If r = 0 and t = ∞, we get $(F/A_c,0,t) = t$, $(A_c/F,0,t) = 1/t$, $(F/A_c,r,\infty) = \infty$, and $(A_c/F,r,\infty) = 0$.

$$A_c = F(A_c/F,r,t) \equiv Fr/[e^{rt} - 1] \qquad ...(3.6.4)$$

Equation (3.6.4) can be used to determine the continuous cash flow rate A_c needed for a sinking fund accumulation F = \$11,667.76 at the end of t = 4 years when the nominal interest rate r = 9.87705% per year is compounded continuously. The answer is as before

$$A_c = F(A_c/F,r,t) \equiv Fr/[e^{rt} - 1] = \$11,667.76 \cdot 0.0987705/[e^{0.0987705 \cdot 4} - 1] = \$2,378.57/yr$$

The present time equivalent P of a series of future cash flows is called *capital recovery*. In Figure and Table 3.6.3, we depict the present value of a uniform series of $2,500 end-of-year payments at 10% interest per year, compounded quarterly.

Figure 3.6.3 - Uniform Series Cash Flows and Present Time Equivalence Diagram

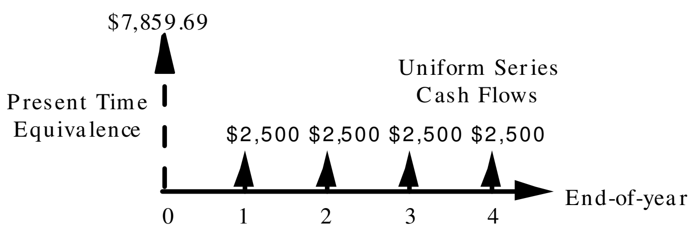


Table 3.6.3 - A'B'C' Accounting for Present Time Equivalence of a Uniform Series of Deposits

(1)EOP	(2)Interest/Period	(3)ACF Balance	(4)EOP Cash Flow	(5)BCF Balance
0	10.38129%	$7,859.69	$0.00	$7,859.69
1	10.38129%	$8,675.63	$2,500.00	$6,175.63
2	10.38129%	$6,816.74	$2,500.00	$4,316.74
3	10.38129%	$4,764.88	$2,500.00	$2,264.88
4	10.38129%	$2,500.00	$2,500.00	$0.00

Thus, the capital recovery (or lumpsum value) of four $2,500 end-of-year lottery prizes is P = $7,859.69 . This example is generalized in Table 3.6.4 below from which the formula for the capital recovery of a uniform series of cash flows is derived.

Table 3.6.4 - Present Capital Recovery P of a Uniform Series of Cash Flows A

(1)EOP	(2)Interest/Period	(3)ACF Balance	(4)EOP Cash Flow	(5)BCF Balance
0	i%	$(A/i)[1-(1+i)^{-n}]$	$0.00	$P=(A/i)[1-(1+i)^{-n}]$
1	i%	$(A/i)[(1+i)-(1+i)^{-n+1}]$	A	$(A/i)[1-(1+i)^{-n+1}]$
2	i%	$(A/i)[(1+i)-(1+i)^{-n+2}]$	A	$(A/i)[1-(1+i)^{-n+2}]$
.	.	.	.	.
n-1	i%	$(A/i)[(1+i)-(1+i)^{-1}]$	A	$(A/i)[1-(1+i)^{-1}]$
n	i%	$(A/i)[(1+i)-(1+i)^{0}]$	A	$0.00

If the kth cash flow A of Table 3.6.4 was time-shifted to the end of the zeroth period, the value would be $A(P/F,i,k) = A(1+i)^{-k}$, and the sum of all n time-shifted cash flows would be

$$P = \sum_{k=1}^{n} A(1+i)^{-k} = (A/i)[1-(1+i)^{-n}] \equiv A(P/A,i,n) \qquad ....(3.6.5)$$

Equation (3.6.5) represents the present time equivalent P of n end-of-period cash flows A when the interest rate is i% per period. Functional notation $(P/A,i,n) \equiv [1-(1+i)^{-n}]/i$ represents the capital recovery per dollar of A. Thus, the capital recovery P of four $2,500 end-of-year cash flows at an interest rate of 10% per year compounded quarterly is

$$P = A(P/A,i,n) = A[1-(1+i)^{-n}]/i = 2500[1-(1.1038129)^{-4}]/0.1038129 = \$7,859.69$$

The value P = $7,859.69 obtained from equation (3.6.5) can be derived from F = $11,667.76 obtained from equation (3.6.1) by the relationship F = P(F/P,i,n) of equation (3.5.1), i.e.,

$$F = P(F/P,i,n) = \$7,859.69(1+i)^{n} = \$7,859.69(1.1038129)^{4} = \$11,667.76$$

The reciprocal $(A/P,i,n) \equiv i/[1-(1+i)^{-n}]$ represents the uniform amount A that needs to be deposited at the ends of n future periods for each lumpsum dollar of present capital recovery P as shown in equation (3.6.6). Both (P/A,i,n) and (A/P,i,n) are listed in Appendix A for various values of i and n, or they may be computed with a pocket calculator. If i = 0 or n = ∞, it can be shown that (P/A,0,n) = n, (A/P,0,n) = 1/n, (P/A,i,∞) = 1/i and (A/P,i,∞) = i.

$$A = Pi/[1-(1+i)^{-n}] \equiv P(A/P,i,n) \qquad ...(3.6.6)$$

The value of A = \$7,859.69•(A/P,10.38129%,4) = \$2,500 obtained from equation (3.6.6) is the same as A = \$11,667.76•(A/F,10.38129%,4) = \$2,500 obtained from equation (3.6.2) because $A \equiv P(A/P,i,n) \equiv F(P/F,i,n)(A/P,i,n) \equiv F(A/F,i,n)$.

From the viewpoint of economic analysis, an important algebraic identity which is useful for subdividing uniform series cash flows into a return <u>on</u> and <u>of</u> the investment is

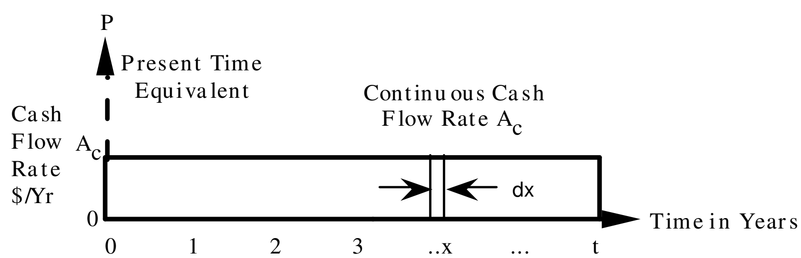
$$(A/P,i,n) \equiv i + (A/F,i,n)$$

For example, a 10.38129% return <u>on</u> investing \$7,859.69 is \$815.94 per year. If \$815.94 is set aside each year from the \$2,500 annual return, then \$2,500 - \$815.94 = \$1,684.06 can be placed in a sinking fund at 10.38129% interest per year for four years in order to accumulate the return <u>of</u> the \$7,859.69 investment as shown below.

$$F = \$1,684.06(F/A,10.38129\%,4) = (1,684.06/0.1038129)[1.1038129^{4}-1] = \$7,859.69$$

Let us now consider Figure 3.6.4 concerning the present time equivalent P of money that will be deposited continuously for t years at the constant rate of A_c dollars per year when interest is compounded continuously at the nominal rate of r% per year.

Figure 3.6.4 - Continuous Cash Flow Rate A_c and Present Time Equivalence P Diagram



During the infinitesimal time interval dx at time x, an amount $A_c dx$ dollars will be deposited. If the infinitesimal deposit $A_c dx$ was time-shifted backwards from time x to time 0, its value would be $A_c dx[e^{-rx}]$. Upon summing time-shifted infinitesimal deposits $A_c dx[e^{-rx}]$ from time x=0 to time x=t, the present lumpsum capital recovery P would be

$$P = \int_0^t A_c e^{-rx}\,dx = (A_c/r)[1-e^{-rt}] \equiv A_c(P/A_c,r,t) \qquad ....(3.6.7)$$

where $(P/A_c,r,t) \equiv [1-e^{-rt}]/r$. The dimensions of $(P/A_c,r,t)$ and $(P/A,i,n)$ are the same as those of $(F/A_c,r,t)$ and $(F/A,i,n)$. In particular, the dimension of $(P/A_c,r,t)$ is time, and $(P/A,i,n)$ is dimensionless. Also, t may not be an integer but n must be an integer.

The relationship between equations (3.6.7) and (3.6.5) for common values of t and n can be illustrated with the $10,000 lottery prize example by continuously depositing A_c = $2,378.57 per year for next four years in a bank whose nominal interest rate is r = 9.87705% per year compounded continuously. After t = 0, 1, ... , 4 years have elapsed, the present value of the remaining years of deposits calculated with equation (3.6.7) should equal those calculated with equation (3.6.5) and listed in Table 3.6.3.

$P_{4-t} = A_c(P/A_c,r,t) = (A_c/r)\ [1-e^{-rt}]$ (3.6.7) $P_{4-n} = A(F/A,i,n) = (A/i)\ [1-(1+i)^{-n}]$ (3.6.5)

$P_4 = (2,379/0.09877)[1-e^{-0.09877\cdot 4}] = \$7,859.69$ $P_4 = (2,500/0.1038)[1-(1.1038)^{-4}] = \$7,859.69$
$P_3 = (2,379/0.09877)[1-e^{-0.09877\cdot 3}] = \$6,175.63$ $P_3 = (2,500/0.1038)[1-(1.1038)^{-3}] = \$6,175.63$
$P_2 = (2,379/0.09877)[1-e^{-0.09877\cdot 2}] = \$4,316.74$ $P_2 = (2,500/0.1038)[1-(1.1038)^{-2}] = \$4,316.74$
$P_1 = (2,379/0.09877)[1-e^{-0.09877\cdot 1}] = \$2,264.88$ $P_1 = (2,500/0.1038)[1-(1.1038)^{-1}] = \$2,264.88$
$P_0 = (2,379/0.09877)[1-e^{-0.09877\cdot 0}] = \$0.00$ $P_0 = (2,500/0.1038)[1-(1.1038)^{-0}] = \$0.00$

If the lady exercised her lumpsum option at the present time and deposited the $7,859.69 in the bank for four years, her balance would be $11,667.76 which is the same as the future value of the four years of A_c = $2,378.57 deposits per year without any withdrawals.

$$F = P(F/P,r,t) = \$7,859.69 * e^{0.09877\cdot 4} = \$11,667.76$$

The reciprocal of $(P/A_c,r,t) \equiv [1-e^{-rt}]/r$ is $(A_c/P,r,t) \equiv r/[1-e^{-rt}]$ which represents the cash flow rate that needs to be deposited continuously for t years in order to get a present time equivalent of $1 when the nominal annual interest rate is r% compounded continuously. If $r = 0$ and $t = \infty$, $(P/A_c,0,t) = t$, $(A_c/P,0,n) = 1/t$, $(P/A_c,r,\infty) = 1/r$, and $(A_c/P,r,\infty) = r$.

$$A_c = Pr/[1-e^{-rt}] = P(A_c/P,r,t) \qquad ...(3.6.8)$$

The value $A_c = \$7,859.69\cdot(A_c/P,9.87705\%,4) = \$2,378.57$ from equation (3.6.8) is the same as $A_c = \$11,667.76\cdot(A_c/F,9.531\%,4) = \$2,378.57$ from equation (3.6.4) because of the relationship $A_c \equiv P(A_c/P,r,t) \equiv F(P/F,r,t)(A_c/P,r,t) \equiv F(A_c/F,r,t)$.

From the viewpoint of economic analysis, an important algebraic identity which is useful for subdividing continuous cash flow rates into a return <u>on</u> and <u>of</u> the investment is

$$(A_c/P,r,t) \equiv r + (A_c/F,r,t)$$

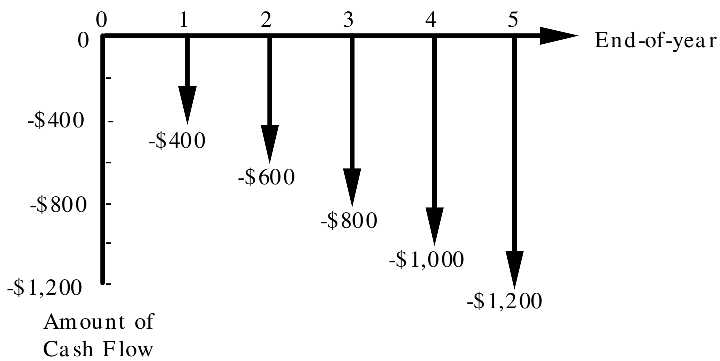
For example, a 9.87705% return <u>on</u> investing $7,859.69 is $776.31 per year. If $776.31 is subtracted continuously from the $2,378.57 annual cash flow rate, then the remainder $2,378.57 - $776.31 = $1,602.26 can be placed in a sinking fund at 9.87705% interest per year compounded continuously for four years in order to accumulate the return <u>of</u> the $7,859.69 investment as shown below.

$$F = \$1,602.26(F/A_c,9.87705\%,4) = (1,602.26/0.0987705)[e^{0.0987705*4} - 1] = \$7,859.69$$

Section 3.7 - Time Equivalences of Arithmetic Gradient Cash Flows

A series of discrete cash flows which change by a constant amount per period is called a *discrete arithmetic gradient*. Let us consider a firm which has annual truck maintenance that costs $400 at the end of the first year and increases $200 per year as shown in Figure 3.7.1. We need to determine the future, present, and uniform series time equivalents of five years of maintenance costs if the interest rate is 12%/year compounded semiannually. The cash flow accounting system will first be used to solve this problem.

Figure 3.7.1 - Discrete Arithmetic Gradient Cash Flows of a Truck Maintenance Problem



In order to use annual instead of semiannual compounding in the cash flow accounting system, the 12% nominal annual interest rate compounded semiannually must be converted to an effective annual interest rate i_E by means of equation (3.4.1) as follows:

$$1+i_E = [1+(r/m)]^m = [1+(0.12/2)]^2 = 1.1236 \text{ or } i_E = 12.36\%/\text{yr.}$$

Table 3.7.1 - Future Time Equivalence F of an Arithmetic Gradient Cash Flow Series

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	12.36%	$0.00	$0.00	$0.00
1	12.36%	$0.00	-$400.00	-$400.00
2	12.36%	-$449.44	-$600.00	-$1,049.44
3	12.36%	-$1,179.15	-$800.00	-$1,979.15
4	12.36%	-$2,223.77	-$1,000.00	-$3,223.77
5	12.36%	-$3,622.23	-$1,200.00	F=-$4,822.23

Table 3.7.2 - Present Time Equivalence P of an Arithmetic Gradient Cash Flow Series

(1)EOY	(2)Interest/Year	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance
0	12.36%	-$2,692.71	$0.00	P=-$2,692.71
1	12.36%	-$3,025.53	-$400.00	-$2,625.53
2	12.36%	-$2,950.04	-$600.00	-$2,350.04
3	12.36%	-$2,640.51	-$800.00	-$1,840.51
4	12.36%	-$2,068.00	-$1,000.00	-$1,068.00
5	12.36%	-$1,200.00	-$1,200.00	$0.00

Table 3.7.3 - Future Time Equivalence of a Uniform Cash Flow Series of $1 per Year

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	12.36%	$0.00	$0.00	$0.00
1	12.36%	$0.00	$1.00	$1.00
2	12.36%	$1.12	$1.00	$2.12
3	12.36%	$2.39	$1.00	$3.39
4	12.36%	$3.80	$1.00	$4.80
5	12.36%	$5.40	$1.00	$6.40*

*More exactly (F/A,12.36%,5) = 6.3984441

Table 3.7.4 - Present Time Equivalence of a Uniform Cash Flow Series of $1 per Year

(1)EOY	(2)Interest/Year	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance
0	12.36%	$3.57	$0.00	$3.57*
1	12.36%	$4.01	$1.00	$3.01
2	12.36%	$3.39	$1.00	$2.39
3	12.36%	$2.68	$1.00	$1.68
4	12.36%	$1.89	$1.00	$0.89
5	12.36%	$1.00	$1.00	$0.00

*More exactly (P/A,12.36%,5) = 3.5728578

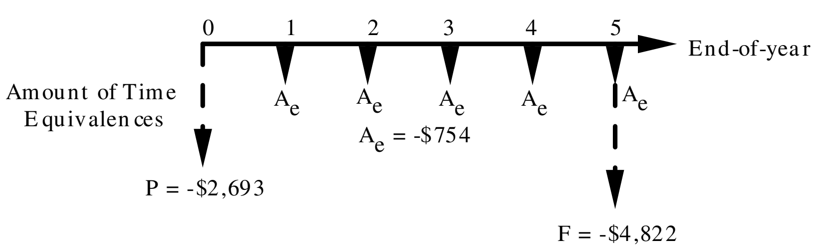
Thus, F = -$4,822.23 from Table 3.7.1 and P = -$2,692.71 from Table 3.7.2 . The uniform series time equivalent A_e can be obtained by solving either the equation A_e = F(A/F,12.36%,5) or the equation A_e = P(A/P,12.36%,5) as follows:

A_e = F(A/F,i,n) = F/(F/A,i,n) or A_e = -$4,822.23/6.3984441 = -$753.66

A_e = P(A/P,i,n) = P/(P/A,i,n) or A_e = -$2,692.71/3.5728578 = -$753.66

Figure 3.7.2 shows the cash flow accounting system solutions for the future, present, and uniform series time equivalences of the discrete arithmetic gradient cash flow problem.

Figure 3.7.2 - Time Equivalences of Figure 3.7.1 Cash Flows



This example is generalized as shown in Table 3.7.5 below to derive the formula for the future time equivalent F of a discrete arithmetic gradient. We define a series of discrete cash flows A which change by a constant amount G per period as a discrete arithmetic gradient. Thus, the series of n successive end-of-period deposits A+0G, A+1G, A+2G, ... , A+(n-1)G starting at the end of the first period as shown in Table 3.7.5 is a discrete arithmetic gradient. The future time equivalent F of this series at the end of n periods at an interest rate of i% per period is derived in Table 3.7.5 using ABC accounting rules.

Table 3.7.5 - Sinking Fund Accumulation F from Arithmetic Gradient Cash Flows

(1)EOP	(2)Interest/Period	(3)BCF Balance	(4)EOP Cash Flow	(5)ACF Balance
0	i%	$0.00	$0.00	$(A+\frac{G}{i})[F/A,i,0]-\frac{0G}{i}$
1	i%	$(A+\frac{G}{i})[F/A,i,1]-\frac{1G}{i}$ -A-0G	A+0G	$(A+\frac{G}{i})[F/A,i,1]-\frac{1G}{i}$
2	i%	$(A+\frac{G}{i})[F/A,i,2]-\frac{2G}{i}$ -A-1G	A+1G	$(A+\frac{G}{i})[F/A,i,2]-\frac{2G}{i}$
3	i%	$(A+\frac{G}{i})[F/A,i,3]-\frac{3G}{i}$ -A-2G	A+2G	$(A+\frac{G}{i})[F/A,i,3]-\frac{3G}{i}$
.	.	.	.	.
n	i%	$(A+\frac{G}{i})[F/A,i,n]-\frac{nG}{i}$ -A-(n-1)G	A+(n-1)G	$F=(A+\frac{G}{i})[F/A,i,n]-\frac{nG}{i}$

The k^{th} cash flow element A+G(k-1) has a value $[A+G(k-1)](1+i)^{n-k}$ at the end of period n, and the value of all n elements as given in equation (3.7.1) is the same as F in Table 3.7.5 .

$$F = \sum_{k=1}^{n}[A + G(k-1)](1+i)^{n-k} = [A+\frac{G}{i}](F/A, i, n) - \frac{nG}{i} \qquad ....(3.7.1)$$

The present time equivalent P of a discrete arithmetic gradient could be derived either with the A'B'C' accounting rules, or by summing present values $[A+G(k-1)](1+i)^{-k}$ of all n elements. However, it is easier to multiply the value of F in equation (3.7.1) by (P/F,i,n) as follows:

$$P = F(P/F,i,n) = [A+\frac{G}{i}](P/A,i,n) - \frac{nG}{i}(P/F,i,n) \qquad ...(3.7.2)$$

Similarly, the uniform series time equivalent A_e of a discrete arithmetic gradient could be obtained by multiplying the values of either F or P of equations (3.7.1) or (3.7.2) by (A/F,i,n) or (A/P,i,n) respectively as follows:

$$A_e = F(A/F,i,n) = P(A/P,i,n) = [A+\frac{G}{i}] - \frac{nG}{i}(A/F,i,n) \qquad ...(3.7.3)$$

Equations (3.7.2) and (3.7.3) were derived with the identities (P/F,i,n)(F/A,i,n) ≡ (P/A,i,n), (F/A,i,n)(A/F,i,n) ≡ 1, (P/A,i,n)(A/P,i,n) ≡ 1, and (A/P,i,n)(P/F,i,n) ≡ (A/F,i,n). These identities are suggested by multiplying the functional symbols (P/F)(F/A) = (P/A), (F/A)(A/F) = 1, (P/A)(A/P) = 1, and (A/P)(P/F) = (A/F), and they greatly facilitate the breakdown of complex cash flow questions into component parts.

To illustrate the use of equations (3.7.1), (3.7.2), and (3.7.3) for the previous truck maintenance problem, we have A = -$400, G = -$200/year, i = 12.36%/year, and n = 5. Upon substituting these values of A, G, i and n in equations (3.71), (3.72) and (3.73), we obtain the same results as before.

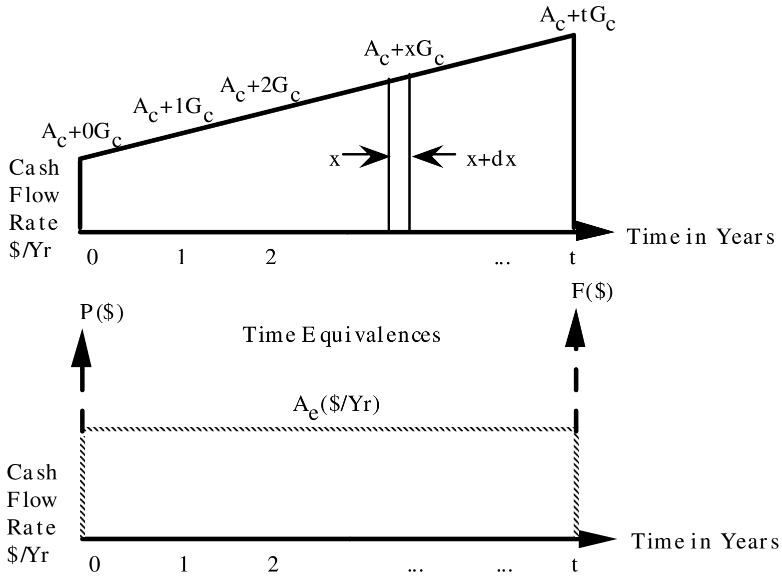
$$F = [A+\frac{G}{i}](F/A,i,n) - \frac{nG}{i} = [-400-\frac{200}{0.1236}]\ (F/A,12.36\%,5) + \frac{5*200}{0.1236} = -\$4,822.23$$

$$P = F(P/F,i,n) = -\$4,822.23\ (P/F,12.36\%,5) = -\$2,692.71$$

$$A_e = -\$4,822.23\ (A/F,12.36\%,5) = -\$2,692.71(A/P,12.36\%,5) = -\$753.66$$

Let us now evaluate the time equivalences of a continuous cash flow rate A_c which changes by a constant amount G_c per period, called a *continuous arithmetic gradient*, as shown in Figure 3.7.3 . On an annual basis, the dimensions of continuous arithmetic gradient parameters are A_c(\$/yr) and G_c(\$/yr^2) as opposed to A(\$) and G(\$/yr) for discrete arithmentic gradients. If the nominal annual interest rate r is compounded continuously, the problem is to determine the future, present and uniform time equivalences of a continuous arithmetic gradient.

Figure 3.7.3 - Cash Flow Rates and Time Equivalences of Continuous Arithmetic Gradients



In time interval dx at time x, an infinitesmal amount $(A_c+G_cx)dx$ dollars will be deposited. If the $(A_c+G_cx)dx$ deposit was time-shifted from time x to time t, its value would be $(A_c+G_cx)dx[e^{r(t-x)}]$. Upon summing $(A_c+G_cx)dx[e^{r(t-x)}]$ from x=0 to x=t, the future time equivalent F of the continuous arithmetic gradient at the end of t years would be

$$F = \int_0^t (A_c+G_cx)\,e^{r(t-x)}dx = [A_c+\frac{G_c}{r}](F/A_c,r,t) - \frac{tG_c}{r} \qquad ....(3.7.4)$$

Equation (3.7.4) consists of time-shifting <u>forward</u> each infinitesimal deposit $(A_c+G_cx)dx$ from time x to its time equivalent $(A_c+G_cx)e^{r(t-x)}dx$ at time t. If $(A_c+G_cx)dx$ was time-shifted <u>backward</u> from time x to its time equivalent $(A_c+G_cx)e^{-rx}dx$ at time zero, the integral would represent the present time equivalent P of the capital recovery at time zero. The same value of P could be obtained more simply by multiplying the value of F already obtained in equation (3.7.4) by (P/F,r,t) as shown in equation (3.7.5) below.

$$P = F(P/F,r,t) = \left[A_c+\frac{G_c}{r}\right](P/A_c,r,t) - \frac{tG_c}{r}(P/F,r,t) \qquad ....(3.7.5)$$

Similarly, the uniform time equivalent A_e of continuous arithmetic gradients could be obtained by multiplying either F or P of equations (3.74) or (3.75) by $(A_c/F,r,t)$ or $(A_c/P,r,t)$ respectively as follows:

$$A_e = F(A_c/F,r,t) = P(A_c/P,r,t) = [A_c+(G_c/r)] - [tG_c/r](A_c/F,r,t) \qquad ....(3.7.6)$$

Forecasting from continuous physical models and financial data is more conveniently carried out with continuous cash flows and compounding which could be partitioned later into financial accounting periods with discrete cash flows and compounding. This suggests that the accuracy of continuous forecasts which are subsequently converted to a discrete form would be greater than forecasting discrete cash flows and compounding directly from continuous physical models and financial data.

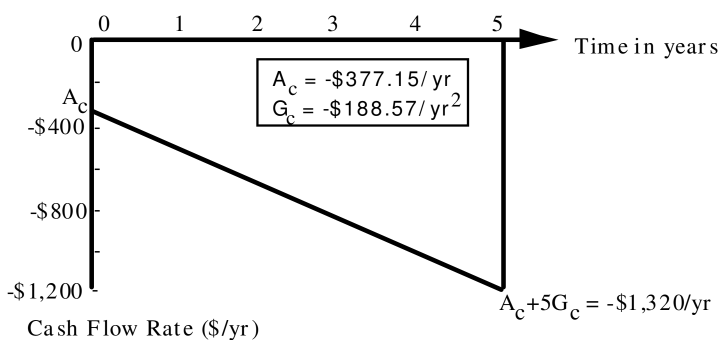
Therefore, it is important to understand the relationship between the parameters r, t, A_c and G_c of equations (3.7.4), (3.7.5) and (3.7.6) for continuous compounding and cash flows and the parameters i_E, n, A and G of equations (3.7.1), (3.7.2) and (3.7.3) for discrete compounding and cash flows. The relationship between r and i_E is given by equation (3.4.3) and they are both measured as a reciprocal of the same units of time. Since the product rt is dimensionless, t must also be measured in the same units of time. However, t is a continuous variable with dimensions of time, and n is an integer variable which is dimensionless. In order for the cash balances to have the same values when t and n have corresponding values, parameters r, A_e and G_e must obey the following relationships.

$$r \cdot 1 = \ln[1+(i_E/1)] = \ln[1.1236] = 0.1165378 \qquad ....(3.4.3)$$

$$A_c = Ar/i_E = -\$400(0.1165378/0.1236) = -\$377.14504/yr$$

$$G_c = Gr/i_E = -\$400(0.1165378/0.1236) = -\$188.57252/yr^2$$

Figure 3.7.4 - Continuous Arithmetic Gradients of the Truck Maintenance Problem



The foregoing relationships between continuous and discrete parameters for common values of t and n can be illustrated for the truck maintenance problem whose future time equivalents F_n were analyzed for integer values of n in Figure, Table and equation (3.7.1). When equation (3.7.4) for future time equivalents F_t is formulated in continuous parameters corresponding to the discrete parameters of equation (3.7.1), then F_t is given by equation (3.7.7) below which will make F_t equal F_n for common values of t and n.

$F_t = [A_c+(G_c/i_E)](F/A_c,r,t) - tG_c/r$ ....(3.7.7)

F_0 = [-377.15-(188.57/0.1236)](F/A_c,0.11654,0)+0*188.57/0.11654 = -$0.00
F_1 = [-377.15-(188.57/0.1236)](F/A_c,0.11654,1)+1*188.57/0.11654 = -$400.00
F_2 = [-377.15-(188.57/0.1236)](F/A_c,0.11654,2)+2*188.57/0.11654 = -$1,049.44
F_3 = [-377.15-(188.57/0.1236)](F/A_c,0.11654,3)+3*188.57/0.11654 = -$1,979.15
F_4 = [-377.15-(188.57/0.1236)](F/A_c,0.11654,4)+4*188.57/0.11654 = -$3,223.77
F_5 = [-377.15-(188.57/0.1236)](F/A_c,0.11654,5)+5*188.57/0.11654 = -$4,822.23

The foregoing relationships between continuous and discrete parameters for common values of t and n can also be illustrated for the truck maintenance problem whose present time equivalents P_n were analyzed for integer values of n in Figure, Table and equation (3.7.2). When equation (3.7.5) for present time equivalents P_t is formulated in continuous parameters corresponding to the discrete parameters of equation (3.7.2), then P_t is given by equation (3.7.8) below which will make P_t equal P_n for common integer values of t and n.

$P_t = F_5*(P/F,r,5\text{-}t) - F_t$ ....(3.7.8)

P_0 = -$4,822.23*(P/F,0.11654,5-0) + $0.00 = -$2,692.71
P_1 = -$4,822.23*(P/F,0.11654,5-1) + $400.00 = -$2,625.53
P_2 = -$4,822.23*(P/F,0.11654,5-2) + $1,049.44 = -$2,350.04
P_3 = -$4,822.23*(P/F,0.11654,5-3) + $1,979.15 = -$1,840.51
P_4 = -$4,822.23*(P/F,0.11654,5-4) + $3,223.77 = -$1,068.00
P_5 = -$4,822.23*(P/F,0.11654,5-5) + $4,822.23 = -$0.00

The uniform time equivalent A_e of continuous arithmetic gradients could be obtained by multiplying either F or P of equations (3.74) or (3.75) by (A_c/F,r,t) or (A_c/P,r,t) as follows:

A_e = -$4,822.23*($A_c$/F,0.11654,5) = -$2,692.71*(A_c/P,0.11654,5) = -$753.66

When i = 0, equations (3.7.1) and (3.7.2) become F = P = nA + n(n-1)G/2, and equation (3.7.3) becomes A_e = A + (n-1)G/2. Thus, the undiscounted cost in the discrete step-function model of Figure 3.7.1 is -5*$400-5*4*$200/2 = -$4,000 and the average undiscounted cost is -$400-(5-1)*$200/2 = $800.

When r = 0, equations (3.7.4) and (3.7.5) become F = P = $tA_c + t^2G_c/2$, and equation (3.7.6) becomes $A_e = A_c + tG_c/2$. Thus, the undiscounted cost in the continuous model of Figure 3.7.4 is the area of a trapezoid which is -5*$377.15-25*$188.57/2 = -$4,242.88 and the average cost is -$377.15-5*$188.57/2 = -$848.58 .

Section 3.8 - Time Equivalences of Geometric Gradient Cash Flows

A series of cash flows which change by a constant proportion per period is called a *discrete geometric gradient*. As an example, let us consider a firm which has annual truck maintenance that costs $500 at the end of the first year which increases 20% per year as shown in Figure 3.8.1 below. The problem is to determine the future, present, and uniform series time equivalents at 12% per year interest compounded annually for five years of maintenance costs. The cash flow accounting system will first be used to solve this problem.

Figure 3.8.1 - Discrete Geometric Gradient Cash Flows of a Truck Maintenance Problem

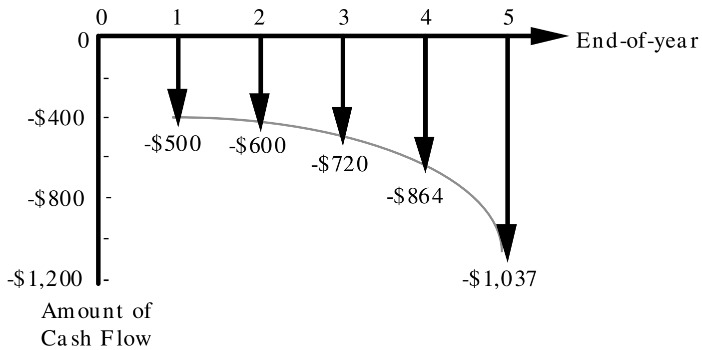


Table 3.8.1 - Future Time Equivalence F of a Geometric Gradient Cash Flow Series

(1)EOP	(2)Interest/Period	(3)BCF Balance	(4)EOP Cash Flow	(5)ACF Balance
0	12%	$0.00	$0.00	$0.00
1	12%	$0.00	-$500.00	-$500.00
2	12%	-$560.00	-$600.00	-$1,160.00
3	12%	-$1,299.20	-$720.00	-$2,019.20
4	12%	-$2,261.50	-$864.00	-$3,125.50
5	12%	-$3,500.56	-$1,036.80	F=-$4,537.36

Table 3.8.2 - Present Time Equivalence P of a Geometric Gradient Cash Flow Series

(1)EOP	(2)Interest/Period	(3)ACF Balance	(4)EOP Cash Flow	(5)BCF Balance
0	12%	-$2,574.62	$0.00	P=-$2,574.62
1	12%	-$2,883.58	-$500.00	-$2,383.58
2	12%	-$2,669.61	-$600.00	-$2,069.61
3	12%	-$2,317.96	-$720.00	-$1,597.96
4	12%	-$1,789.71	-$864.00	-$925.71
5	12%	-$1,036.80	-$1,036.80	$0.00

Table 3.8.3 - Future Time Equivalence of a Uniform Cash Flow Series of $1 per Year

(1)EOP	(2)Interest/Period	(3)BCF Balance	(4)EOP Cash Flow	(5)ACF Balance
0	12%	$0.00	$0.00	$0.00
1	12%	$0.00	$1.00	$1.00
2	12%	-$1.12	$1.00	$2.12
3	12%	-$2.37	$1.00	$3.37
4	12%	-$3.78	$1.00	$4.78
5	12%	-$5.35	$1.00	$6.35*

*More exactly (F/A,12%,5) = 6.3528474

Table 3.8.4 - Present Time Equivalence of a Uniform Cash Flow Series of $1 per Year

(1)EOP	(2)Interest/Period	(3)ACF Balance	(4)EOP Cash Flow	(5)BCF Balance
0	12%	$3.60	$0.00	$3.60*
1	12%	$0.00	$1.00	$3.04
2	12%	$3.40	$1.00	$2.40
3	12%	$2.69	$1.00	$1.69
4	12%	$1.89	$1.00	$0.89
5	12%	$1.00	$1.00	$0.00

*More exactly (P/A,12%,5) = 3.6047762

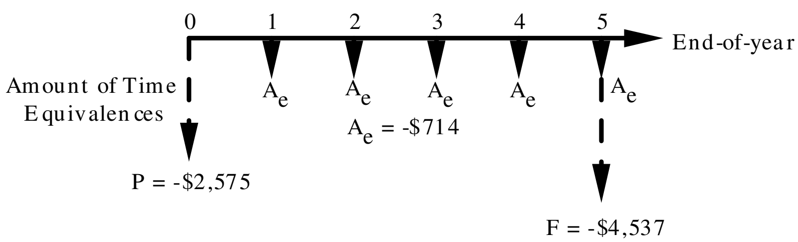
Thus, F = -$4,537.36 from Table 3.8.1, and P = -$2,574.62 from Table 3.8.2 . The uniform series time equivalent A_e can be obtained by solving either the equation $A_e = F(A/F,12\%,5)$ or the equation $A_e = P(A/P,12\%,5)$ as follows:

$$A_e = F(A/F,i,n) = F/(F/A,i,n) \quad \text{or} \quad A_e = -\$4{,}537.36/6.3528474 = -\$714.22$$

$$A_e = P(A/P,i,n) = P/(P/A,i,n) \quad \text{or} \quad A_e = -\$2{,}574.62/3.6047762 = -\$714.22$$

Figure 3.8.2 shows the cash flow accounting system solutions for the future, present, and uniform series time equivalences of the discrete geometric gradient cash flow problem.

Figure 3.8.2 - Time Equivalences of Figure 3.8.1 Cash Flows



This example is generalized in Table 3.8.5 to derive the formula for the future time equivalence F of a discrete geometric gradient. A series of discrete cash flows A which change by a constant proportion g per period as a discrete geometric gradient. The series $A(1+g)^0$, $A(1+g)^1$, $A(1+g)^2$, ... , $A(1+g)^{n-1}$ starting at the end of the first period in Table 3.8.5 is a discrete geometric gradient. The future time equivalent F of this series at the end of n periods when interest is i% per period is derived in Table 3.8.5, where $1+i^* \equiv (1+i)/(1+g)$.

Table 3.8.5 - Sinking Fund Accumulation F from Geometric Gradient Cash Flows

(1)EOP	(2)Interest/Period	(3)BCF Balance	(4)EOP Cash Flow	(5)ACF Balance
0	i%	$0.00	$0.00	$A(1+g)^{0-1}[F/A,i^*,0]$
1	i%	$A(1+i)(1+g)^{0-1}[F/A,i^*,0]$	$A(1+g)^0$	$A(1+g)^{1-1}[F/A,i^*,1]$
2	i%	$A(1+i)(1+g)^{1-1}[F/A,i^*,1]$	$A(1+g)^1$	$A(1+g)^{2-1}[F/A,i^*,2]$
3	i%	$A(1+i)(1+g)^{2-1}[F/A,i^*,2]$	$A(1+g)^2$	$A(1+g)^{3-1}[F/A,i^*,3]$
.	.	.	.	.
n	i%	$A(1+i)(1+g)^{n-2}[F/A,i^*,n-1]$	$A(1+g)^{n-1}$	$F=A(1+g)^{n-1}[F/A,i^*,n]$

The k^{th} cash flow element $A(1+g)^{k-1}$ has a value $A(1+g)^{k-1}(1+i)^{n-k}$ at the end of period n, and the sum of all n elements as given in equation (3.8.1) is the same as F in Table 3.8.5 .

$$F = \sum_{k=1}^{n} A(1+g)^{k-1}(1+i)^{n-k} = A(1+g)^{n-1}[F/A,i^*,n] \qquad ....(3.8.1)$$

The present time equivalent P of geometric gradients could be derived either with the A'B'C' accounting rules, or by summing present values $A(1+g)^{k-1}(1+i)^{-k}$ of all n elements as follows:

$$P = \sum_{k=1}^{n} A(1+g)^{k-1}(1+i)^{-k} = A(1+g)^{-1}[P/A,i^*,n] \qquad ....(3.8.2)$$

The uniform series time equivalent A_e can be obtained by multiplying the value of F or P from equation (3.8.1) or (3.8.2) by (A/F,i,n) or (A/P,i,n) respectively as follows:

$$A_e = A(1+g)^{n-1}(F/A,i^*,n)(A/F,i,n) = A(1+g)^{-1}(P/A,i^*,n)(A/P,i,n) \qquad ...(3.8.3)$$

It should be noted that when i = g, then i* = 0 and (F/A,0,n) = (P/A,0,n) = n.

To illustrate the use of equations (3.8.1), (3.8.2), and (3.8.3) for the previous truck maintenance problem, we have A = -\$500, g = 20%/year, i = 12%/year, and n = 5. Upon substituting these values of A, g, i and n in equations (3.8.1), (3.8.2) and (3.8.3), we obtain the same results as before.

$$1+i^* \equiv (1+i)/(1+g) = (1.12)/(1.20) = 0.9333333 \text{ or } i^* = -6.666667\%$$

$$F = A(1+g)^{n-1}(F/A,i^*,n) = -\$500(1.20)^4 (F/A,-6.667\%,5) = -\$4{,}537.36$$

$$P = A(1+g)^{-1}(P/A,i^*,n) = -\$500(1.20)^{-1} (P/A,-6.667\%,5) = -\$2{,}574.62$$

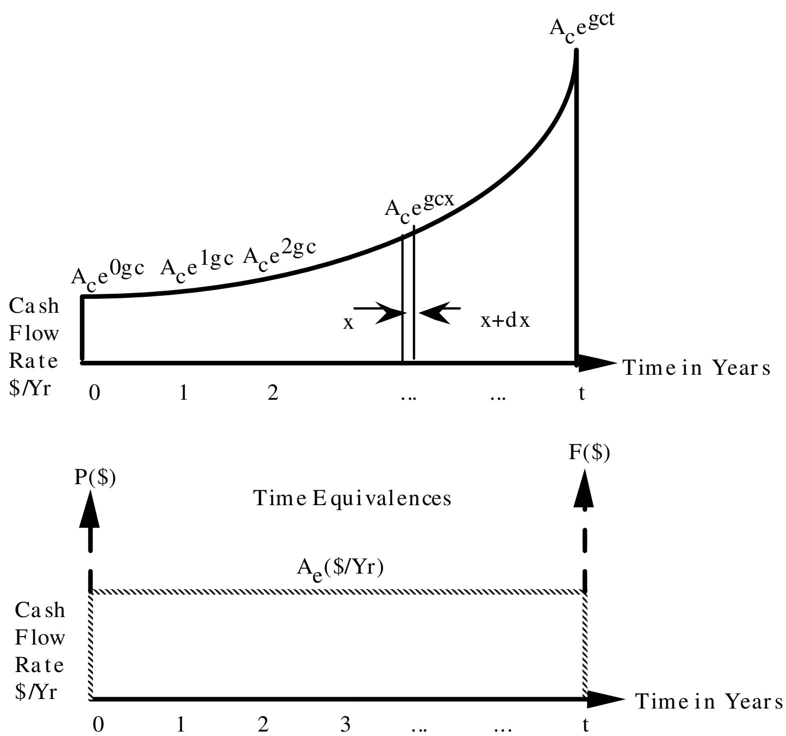
$$A_e = -\$4{,}537.36\,(A/F,12\%,5) = -\$2{,}574.62\,(A/P,12\%,5) = -\$714.22/yr$$

Let us now evaluate the time equivalences of a continuous cash flow rate A_c which changes by a constant proportion g_c per period, called a *continuous geometric gradient*, as shown in Figure 3.8.3 below. On an annual basis, the dimensions of continuous geometric gradient parameters are A_c(\$/yr), g_c(1/yr), r(1/yr) and t(yr) in contrast to A(\$), g(Ø), i(Ø) and n(Ø) for discrete geometric gradients. Let us now determine the future, present and uniform time equivalences at continuous compound interest for continuous geometric gradients.

During the time interval dx at time x, an infinitesmal amount $A_c e^{g_c x}dx$ dollars will be deposited for interest accumulation. If deposit $A_c e^{g_c x}dx$ was time-shifted <u>forward</u> from time x to time t, its value would be $A_c e^{g_c x}dx[e^{r(t-x)}]$. Upon summing $A_c e^{g_c x}dx[e^{r(t-x)}]$ from x=0 to x=t, the future time equivalent F of the sinking fund at the end of t years would be

$$F = \int_0^t A_c e^{g_c x} e^{r(t-x)}dx = A_c e^{g_c t}\left[\frac{e^{(r-g_c)t}-1}{r-g_c}\right] \equiv A_c e^{g_c t}(F/A_c, r-g_c, t) \qquad ....(3.8.4)$$

Figure 3.8.3 - Cash Flow Rates and Time Equivalences of Continuous Geometric Gradients



If the infinitesmal cash flow element $A_c e^{g_c x} dx$ was time-shifted backward from time x to time zero, its value would be $A_c e^{g_c x} e^{-rx} dx$. Upon summing the elements $A_c e^{g_c x} e^{-rx} dx$ from x=0 to x=t, the present time equivalent P of the capital recovery would be

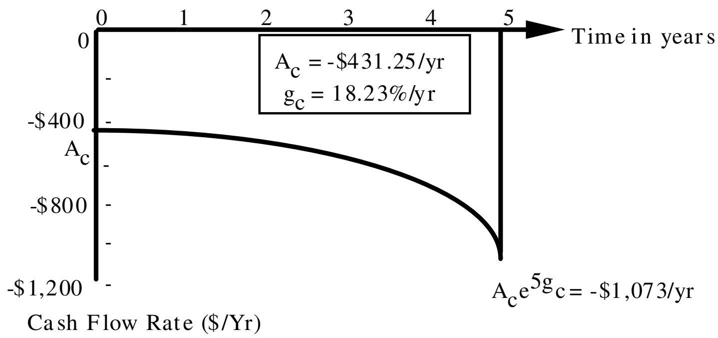
$$P = \int_0^t A_c e^{g_c x} e^{-rx}\, dx = A_c \left[\frac{1 - e^{-(r-g_c)t}}{r - g_c} \right] \equiv A_c(P/A_c, r - g_c, t) \qquad ....(3.8.5)$$

The uniform series time equivalent A_E of the continuous geometric gradient cash flow series could be obtained by multiplying either F or P of equations (3.8.4) or (3.8.5) by $(A_c/F,r,t)$ or $(A_c/P,r,t)$ respectively as follows:

$$A_e = A_c e^{g_c t}\,(F/A_c,r\text{-}g_c,t)(A_c/F,r,t) = A_c(P/A_c,r\text{-}g_c,t)(A_c/P,r,t) \qquad ...(3.8.6)$$

The relationship between continuous parameters r, g_c and A_c of equations (3.8.4), (3.8.5) and (3.8.6), and discrete parameters i, g and A of equations (3.8.1), (3.8.2) and (3.8.3) for common values of t and n can be illustrated with the geometric-gradient truck maintenance problem shown in Figure 3.8.1. For this purpose, r = ln(1+i) = ln1.12 = 0.1133287, g_c = ln(1+i) = ln1.12 = 0.1823216, $A_c = A(r\text{-}g_c)/(i\text{-}g)$ = -$431.245/year which grows 18,23216%/year to $A_c e^{g_c t}$ = -$1,073.08 at the end of 5 years as shown in Figure 3.8.4 below.

Figure 3.8.4 - Example of Continuous Geometric Gradient Cash Flow Rates



We are given A_c = -\$431.245/year, g_c = 18.23216%/year, r = 11.33287%/year and t = 5 years. Upon substitution in equations (3.8.4), (3.8.5) and (3.8.6), we get

$F = A_c e^{g_c t}(F/A_c, r-g_c, t)$ for $r \neq g$; $F = A_c e^{g_c t}(F/A, 0, t) = tA_c e^{g_c t}$ for $r = g$ ....(3.8.4)
$F = -\$431.245e^{0..1823216 \cdot 5}(F/A_c, -6.899929\%, 5) = -\$4,537.78$ (-\$4,537.36 with (3.8.1))

$P = A_c(P/A, r-g_c, t)$ for $r \neq g$; $P = A_c(P/A_c, 0, t) = tA_c e^{g_c t}$ for $r = g$ ....(3.8.5)
$P = -\$431.245(P/A_c, -6.899929\%, 5) = -\$2,574.86$ (-\$2,574.62 with (3.8.2))

$A_e = A_c e^{g_c t}(F/A_c, r-g_c, t)(A_c/F, r, t) = A_c(P/A_c, r-g_c, t)(A_c/P, r, t)$ ....(3.8.6)
$A_e = -\$4,537.78(A_c/F, 11.33287\%, 5) = -\$2,574.86(A_c/P, 12\%, 5) = -\$674.58$

A_e = -\$714.22 with equation (3.8.3)

Engineers often forecast the operations of projects in the form of continuous physical models to which economic forecasts of prices and interest rates are added. For this reason, let us compare continuous to discrete geometric gradient forecasts before discounting. The undiscounted cost in the continuous case is represented by the area above the exponential curve in Figure 3.8.4 which is

$P = A_c(P/A_c, -g_c, t) = -\$431.245(P/A_c, -18.23216\%, 5) = -\$3,520.32$

The undiscounted cost in Figure 3.8.1 is -\$500 -\$600 -\$720 -\$864 -\$1,036.80 = -\$3,720.80 . The same result could be obtained for the discrete case by setting i = 0 in either equation (3.8.1) or (3.8.2). The difference between -\$3,520.32 - (-\$3,720.80) = \$200.48 is relatively minor considering the accuracy of cash flow forecasts of continuous and discrete geometric gradients. Because revenues and expenses generally occur during the year and year end, continuous rather than discrete cash flow forecasts are often preferred for managerial accounting.

More generally, when i = 0, equations (3.8.1) and (3.8.2) become $F = P = A(F/A, g, n)$ and equation (3.8.3) becomes $A_e = A(F/A, g, n)/n$. When r = 0, equations (3.84) and (3.85) become $F = P = A_c(F/A_c, g_c, t)$ and equation (3.8.6) becomes $A_e = A_c(F/A_c, g_c, t)/t$.

Section 3.9 - Equivalent Discrete and Continuous Cash Flows and Compounding

Discrete interest-rate formulas are based on end-of-period cash flows which earn interest just before the cash flows of the next end-of-period. Time-equivalent balances are not defined between the ends of compounding periods. Continuous interest-rate formulas define the time-equivalent balances both between and at ends of compounding periods.

<u>Section 3.5</u> - Single Payment Cash Flows
$F_n= P(F/P,i,n) = P(1+i)^n$ eqn(3.5.1). Table 3.5.1 example: P=-\$1,000; i=12.36%/yr.
$F_{0.0}$=-\$1,000; $F_{0.5}$=-\$1,060; $F_{1.0}$=-\$1,123.60; $F_{1.5}$=-\$1,191.02; $F_{2.0}$=-\$1,262.48 .
$F_t= P(F/P,r,t) = Pe^{rt}$ eqn(3.5.2). Ex: P=-\$1,000; r = ln(1+i) = ln1.1236 = 0.1165378 .
Continuous compound amount equation: $Pe^{rt} = -\$1{,}000e^{0.1165378t}$ for t in years.
Continuous compound rate: r = 0.1165378 per year.
$F_{0.0}$=-\$1,000; $F_{0.5}$=-\$1,060; $F_{1.0}$=-\$1,123.60; $F_{1.5}$=-\$1,191.02; $F_{2.0}$=-\$1,262.48 .

<u>Section 3.6</u> - Uniform Series Cash Flows
$F_n= A(F/A,i,n) = (A/i)[(1+i)^n -1]$ eqn(3.6.1). Table 3.6.1: A=\$2,500; i=10.38129%/yr.
F_0=\$0.00; F_1=\$2,500; F_2=\$5,259.53; F_3=\$8,305.54; F_4=\$11,667.76 .
$F_t= A_c(F/A_c,r,t) = (A/r)[e^{rt}-1]$ eqn(3.6.3). Ex: r= ln1.1038129=.0987705 .
Continuous annual cash flow rate: A_c=Ar/i=\$2,378.57 .
Continuous compound rate: r = 0.0987705 per year.
F_0=\$0.00; F_1=\$2,500; F_2=\$5,259.53; F_3=\$8,305.54; F_4=\$11,667.76 .

<u>Section 3.7</u> - Arithmetic Gradient Cash Flows

Table 3.7.1 -Future Time Equivalence F of Arithmetic Gradient Cash Flows.

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	12.36%	<u>\$0.00</u>	\$0.00	\$0.00
1	12.36%	\$0.00	-\$400.00	-\$400.00
2	12.36%	-\$449.44	-\$600.00	-\$1049.44
3	12.36%	-\$1179.15	-\$800.00	-\$1979.15
4	12.36%	-\$2223.77	-\$1000.00	-\$3223.77
5	12.36%	-\$3622.23	-\$1200.00	-\$4822.23

$F_n= G(F/G,i,n) = (G/i)[(F/A,i,n)-n]$ equation (3.7.1) when A=0.
Table 3.7.1: A=-\$400, G=-\$200/year and i = 12.36%/year.
F_0=\$0.00; F_1=-\$400.00; F_2=-\$1049.44; F_3=-\$1979.15; F_4=-\$3223.77; F_5=-\$4822.23 .
$F_t= [A_c+(G_c/i)](F/A_c,r,t) - tG_c/r$ equation (3.7.7).
Example: r = ln1.1236 =0.1165378; A_c=Ar/i=-\$377.15; G_c=Gr/i=-\$188.57 .
Continuous cash flow rate equivalent: A_c+G_ct = -\$377.15 -\$188.57t for t in years.
Continuous sinking fund equivalent: $(F/A_c,r,t) = (r/i)\ (F/A,i,n)$ when t=n and r=ln(1+i).
F_0=\$0.00; F_1=-\$400.00; F_2=-\$1049.44; F_3=-\$1979.15; F_4=-\$3223.77; F_5=-\$4822.23 .

<u>Section 3.8</u> - Geometric Gradient Cash Flows
$F_n= A(1+g)^{n-1}[F/A,i^*,n]$ where $1+i^* = (1+i)/(1+g)$ equation (3.8.1).
Table 3.8.1: A=-\$500; g = 20%/year and i = 12%/year.
$1+i^* = 1.12/1.2 = 0.9333333$; $i^* = (i-g)/(1+g) = -0.0666667$
F_0=\$0.00; F_1=-\$500; F_2=-\$1160; F_3=-\$2019.20; F_4=-\$3125.50; F_5=-\$4537.36 .
$F_t= A_ce^{g_ct}[F/A_c,r-g_c,t]$ equation (3.8.4) where $A_c = A(r-g_c)/(i-g)$; $g_c = \ln(1+g)$; $r = \ln(1+i)$.
Example from Table 3.8.1: g_c = ln1.2 = 0.1823216; r = ln1.12 = 0.1133287;
$(r-g_c)$ = -0.0689929; (i-g) = -0.08; $A_c= A(r-g_c)/(i-g)$ = -\$500*0.8624113 = -\$431.21 .
Continuous cash flow rate equation: $A_ce^{g_ct} = -\$431.21e^{0.1823216t}$ for t in years.
Continuous compound rate: r = 0.1133287 per year.
F_0=\$0.00; F_1=-\$500; F_2=-\$1160; F_3=-\$2019.20; F_4=-\$3125.50; F_5=-\$4537.36

Section 3.10 - Summary of Chapter Three

This chapter introduces accounting for time equivalences by explaining the roles of *financial* and *managerial* accounting. In financial accounting, there is a simple aggregation of receipts and expenditures within each accounting period and it is mainly used for *external* reporting to government, stockholders and creditors. Managerial accounting involves forecasting causally-related revenues and expenses of alternative ways doing projects that may span many accounting periods. It is mainly used for *internal* reporting to management for purposes of planning present and future operations of an organization.

Cash flows are defined as money that change hands across an interface that separates a system from its surroundings. Cash flow into a system is defined as positive, and cash flow out of a system is defined as negative. Two cash flows at different points of time can be compared by determining their equivalent values at a common point of time. The value of a set of cash flows at a given point of time is known as its time equivalence. The focal date property states that if time equivalences of two sets of cash flows are equal and opposite at one point of time, then their time equivalences are equal and opposite at all other points of time. A general system of accounting is developed in Section 3.2 for determining the future time equivalences (ABC Rules) and present time equivalences (A'B'C' Rules) of any set of discrete cash flows at either constant or varying interest rates.

When repaying a loan (Section 3.3), the borrower pays the lender the borrowed money plus an additional amount called interest. The amount of interest is proportional to the amount of the loan and the period of time in which the loan is outstanding. An interest rate is defined as an amount of interest per unit amount of outstanding loan per unit of time. When periodic cash flow payments to the lender exactly offset the interest charges at the end of each period so that the outstanding loan balance is unchanged, the cash flow payments are known as simple interest. Payments of simple interest until the end of the loan are called a return on the lender's investment, and the last payment which pays back the original loan is called a return of the lender's investment. When cash flow payments differ from the interest charges accrued during each period, then the difference is used to adjust the outstanding loan balance at the end of each period. Cash flow payments which change the outstanding balance at the end of each period are known as compound interest.

Applications of ABC and A'B'C' rules of accounting can be substantially shortened when interest rates are constant and the time interval between successive cash flows is longer than the compounding period (Section 3.4). If an interest rate applies to a nominal time interval such as a year, it is known as a nominal interest rate, or an annual percentage rate (APR). If an interest rate is compounded only once in the cash flow period, it is called an effective interest rate. The relationship between nominal and effective interest rates enables the accounting system to be condensed by replacing many compoundings in a cash flow period with a single compounding. This simplification is especially useful when nominal interest rates are compounded continuously and cash flows are continuous.

Constant interest rate formulas are developed for discrete or continuous cash flows that are either single payments (Section 3.5), uniform series (Section 3.6), arithmetic gradients (Section 3.7), or geometric gradients (Section 3.8). Engineers usually forecast the operations of projects in the form of continuous physical models to which economic forecasts of prices and interest rates are added. For this reason, continuous rather than discrete cash flow forecasts are preferred for managerial accounting.

Appendix 3A - Constant Interest Rate Formulas

Dimension Notations: Ø = dimensionless; n[Ø] = dimensionless integer; T = time in months, years, ... ; 1/T = per month, year, ... ; $/T = dollars per month, year, ...

1. Single Payments (Section 3.5)
   (a) Discrete Compounding: Dimension F($), P($), i(Ø) and n(Ø).
   $F = P(1+i)^n \equiv P(F/P,i,n)$ Eqn (3.5.1)
   $P = F(1+i)^{-n} \equiv F(P/F,i,n)$ Eqn (3.5.3)
   $(F/P,i,n) \equiv 1/(P/F,i,n)$; $(F/P,0,n) \equiv (P/F,0,n) \equiv 1$.
   (b) Continuous Compounding: Dimension F($), P($), r(1/T) and t(T).
   $F = Pe^{rt} \equiv P(F/P,r,t)$ Eqn (3.5.2)
   $P = Fe^{-rt} \equiv F(P/F,r,t)$ Eqn (3.5.4)
   $(F/P,r,t) \equiv 1/(P/F,r,t)$

2. Uniform Series (Section 3.6)
   (a) Discrete Cash Flows and Compounding: F($), P($), A($), i(Ø) and n[Ø] integer.
   $F = (A/i)[(1+i)^n - 1] \equiv A(F/A,i,n)$ Eqn (3.6.1) (Sinking Fund)
   $A = Fi/[(1+i)^n - 1] \equiv F(A/F,i,n)$ Eqn (3.6.2)
   $(F/A,i,n) \equiv 1/(A/F,i,n)$; $(F/A,0,n) = n$; $(F/A,i,\infty) = \infty$, and $(A/F,i,\infty) = 0$.
   $P = (A/i)[1 - (1+i)^{-n}] \equiv A(P/A,i,n)$ Eqn (3.6.5) (Capital Recovery)
   $A = Pi/[1 - (1+i)^{-n}] \equiv P(A/P,i,n)$ Eqn (3.6.6)
   $(P/A,i,n) \equiv 1/(A/P,i,n)$; $(A/P,i,n) \equiv i + (A/F,i,n)$
   $(P/A,0,n) = n$; $(A/P,0,n) = 1/n$; $(P/A,i,\infty) = 1/i$; and $(A/P,i,\infty) = i$.
   (b) Continuous Cash Flows and Compounding: F($), P($), A_c($/T), r(1/T) and t(T).
   $F = (A_c/r)[e^{rt} - 1] \equiv A_c(F/A_c,r,t)$ Eqn (3.6.3) (Sinking Fund)
   $A_c = Fr/[e^{rt} - 1] \equiv F(A_c/F,r,t)$ Eqn (3.6.4)
   $(F/A_c,r,t) \equiv 1/(A_c/F,r,t)$; $(F/A_c,0,t) = t$; $(F/A_c,r,\infty) = \infty$, and $(A_c/F,r,\infty) = 0$.
   $P = (A_c/r)[1 - e^{-rt}] \equiv A_c(P/A_c,r,t)$ Eqn (3.6.7) (Capital Recovery)
   $A_c = Pr/[1 - e^{-rt}] \equiv P(A_c/P,r,t)$ Eqn (3.6.8)
   $(P/A_c,r,t \equiv 1/(A_c/P,r,t)$; $(A_c/P,r,t) \equiv r + (A_c/F,r,t)$
   $(P/A_c,0,t) = t$; $(A_c/P,0,t) = 1/t$; $(P/A_c,r,\infty) = 1/r$; and $(A_c/P,r,\infty) = r$.

3. Arithmetic Gradients (Section 3.7)
   (a) Discrete Cash Flows and Compounding: F($), P($), A($), G($/T), {G/i}($) and n[Ø].
   $F = [A+(G/i)](F/A,i,n) - nG/i$ Eqn (3.7.1)
   $P = [A+(G/i)](P/A,i,n) - (nG/i)\,(P/F,i,n) = F(P/F,i,n)$ Eqn (3.7.2)
   $A_e = [A+(G/i)] - (nG/i)\,(A/F,i,n) = F(A/F,i,n) = P(A/P,i,n)$ Eqn (3.7.3)
   When $i = 0$, $F = P = nA + n(n-1)G/2$ and $A_e = A + (n-1)G/2$
   (b) Continuous Cash Flows & Compounding: F($), P($), A_c($/T), G_c($/T^2), r(1/T), t(T).
   $F = [A_c+(G_c/r)](F/A_c,r,t) - (tG_c/r)$ Eqn (3.7.4)
   $P = [A_c+\frac{G_c}{r}](P/A_c,r,t) - \frac{tG_c}{r}(P/F,r,t) = F(P/F,r,t)$ Eqn (3.7.5)
   $A_e = [A_c+\frac{G_c}{r}] - \frac{tG_c}{r}(A_c/F,r,t) = F(A_c/F,r,t) = P(A_c/P,r,t)$ Eqn (3.7.6)
   When $r = 0$, $F = P = tA_c + t^2G_c/2$ and $A_e = A_c + tG_c/2$

4. Geometric Gradients (Section 3.8)
   (a) Discrete Cash Flows and Compounding: F($), P($), A($), {g/1}(Ø), {i/1}(Ø) and n[Ø].

$$1+i^* \equiv \frac{1+i}{1+g}$$

$$F = A(1+g)^{n-1}\,(F/A,i^*,n) \qquad \text{Eqn (3.8.1)}$$

$$P = A(1+g)^{-1}\,(P/A,i^*,n) \qquad \text{Eqn (3.8.2)}$$

$$A_e = A(1+g)^{n-1}\,(F/A,i^*,n)(A/F,i,n) = A(1+g)^{-1}\,(P/A,i^*,n)(A/P,i,n) \qquad \text{Eqn (3.8.3)}$$

When $i = 0$, then $F = P = A(F/A,g,n)$ and $A_e = A(F/A,g,n)/n$.
When $i = g$, then $i^* = 0$ and $(F/A,0,n) = (P/A,0,n) = n$.

   (b) Continuous Cash Flows & Compounding: F($), P($), A_c($/T), g_c(1/T), r(1/T), t(T).

$$F = A_c e^{g_c t}\,(F/A_c,r-g_c,t) \qquad \text{Eqn (3.8.4)}$$

$$P = A_c(P/A_c,r-g_c,t) \qquad \text{Eqn (3.8.5)}$$

$$A_e = A_c e^{g_c t}\,(F/A_c,r-g_c,t)(A_c/F,r,t) = A_c(P/A_c,r-g_c,t)(A_c/P,r,t) \qquad \text{Eqn (3.8.6)}$$

When $r = 0$, then $F = P = A_c(F/A_c,g_c,t)$ and $A_c = A_c(F/A_c,g_c,t)/t$.
When $r = g$, then $r-g_c = 0$ and $(F/A_c,0,t) = (P/A_c,0,t) = t$.

Appendix 3B - ABC and A'B'C' Spreadsheet Calculations

The ABC and A'B'C' rules of accounting for time equivalences in Section 3.2 provide a dynamic bookkeeping system that is readily carried out in real time with microcomputer spreadsheet calculations. Different formats and syntax are used with various programs and types of microcomputers, but the same logic is used in the computations of virtually all microcomputer spreadsheets. The following two examples illustrate how microcomputer spreadsheet calculations would be implemented to generate Tables 3.2.1 and 3.2.2 in Section 3.2 of Chapter Three.

Concerning the ABC rules generated in Table 3B.1 below, the column headings of Table 3.2.1 are filled in cells A1 through E1. The column entries of cells A2 to A6, B2 to B6, and D2 to D6 are filled in from the data given in Table 3.2.1 . The ABC rules are initiated by placing zero in cell C2. The formula C2 + D2 is entered for cell E2 which then shows the amount ($1,000.00). Cell E2 is then selected and its formula is copied into cells E3 to E6. As a formula is copied into successive cells, it is automatically updated according to the cell location. Consequently, the formula in cell E3 is C3 + D3, the formula in cell E4 is C4 + D4, and so on. We then enter the formula E2*(1+B3) in cell C3, and the formula in C3 is copied into cells C4 to C6. This completes Table 3B.1 with the proper numerical values in the cells of columns C and E.

Concerning the A'B'C' rules generated in Table 3B2 below, the column headings of Table 3.22 are filled in cells A1 through E1. The column entries of cells A2 to A6, B2 to B6, and D2 to D6 are filled in from the data given in Table 3.22 . The A'B'C' rules are initiated by placing zero in cell E6. The formula E6 + D6 is entered for cell C6 which then shows the amount $1,286.18. Cell C6 is then selected and its formula is copied into cells C5 to C2. As a formula is copied into successive cells, it is automatically updated according to the cell location. Consequently, the formula in cell C5 is E5 + D5, the formula in cell C4 is E4 + D4, and so on. We then enter the formula C6/(1+B6) in cell E5, and the formula in E5 is copied into cells E4 to E2. This completes Table 3B.2 with the proper numerical values in the cells of columns C and E.

Table 3B.1 - Spreadsheet Calculations for ABC Rules of Table 3.2.1

	A	B	C	D	E
1	(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
2	0	7.00%	$0.00	($1,000.00)	($1,000.00)
3	1	5.00%	($1,050.00)	$0.00	($1,050.00)
4	2	6.00%	($1,113.00)	$0.00	($1,113.00)
5	3	7.00%	($1,190.91)	$0.00	($1,190.91)
6	4	8.00%	($1,286.18)	$1,286.18	$0.00

Table 3B.2 - Spreadsheet Calculations for A'B'C' Rules of Table 3.2.2

	A	B	C	D	E
1	(1)EOY	(2)Interest/Year	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance
2	0	7.00%	$0.00	($1,000.00)	$1,000.00
3	1	5.00%	$1,050.00	$0.00	$1,050.00
4	2	6.00%	$1,113.00	$0.00	$1,113.00
5	3	7.00%	$1,190.91	$0.00	$1,190.91
6	4	8.00%	$1,286.18	$1,286.18	$0.00

The ABC accounting rules may be used to determine the future value of any set of uniformly-spaced discrete cash flows with either constant or varying interest rates. Referring to Table 3B.1 for arbitrary cash flows spanning four periods, the procedure is initiated by placing zero in cell C2 and determining the future value at the end of the fourth period in cell E6. In order to get a time average of the interest rates weighted by the dollar amounts of the cash flows, it is necessary to make the cash flow at the end of the fourth period equal and opposite to the BCF Balance in cell C6 so that the entry in column E6 would equal zero. The adjusted set of cash flows may then be used as input to the internal-rate-of-return (IRR) function of microcomputer spreadsheets in order to obtain the desired time average of the specified interest rate(s) in column B. As described in Appendix 4B, the IRR function of microcomputer spreadsheets requires the user to guess an initial interest rate in the neighborhood of the answer. This initial guess speeds up the iterative process of the IRR function, and it enables one to obtain local answers when multiple rates of return exist. Some spreadsheets require the initial cash flow to be negative, in which case it may be necessary to multiply the set of cash flows by minus one in order to satisfy this requirement.

The A'B'C' accounting rules may be used to determine the present value of any set of uniformly-spaced discrete cash flows with either constant or varying interest rates. Referring to Table 3B.2 for arbitrary cash flows spanning four periods, the procedure is initiated by placing zero in cell E6 and determining the present value at the end of the zeroth period in cell C2. In order to get a time average of the interest rates weighted by the dollar amounts of the cash flows, it is necessary to make the cash flow at the end of the zeroth period equal and opposite to the BCF Balance in cell E2 so that the entry in column C2 would equal zero. The adjusted set of cash flows may then be used as input to the internal-rate-of-return (IRR) function of microcomputer spreadsheets in order to obtain the desired time average of the specified interest rate(s) in column B.

Appendix 3C - Spreadsheet Calculations of IRR

Internal rates of return (IRR) are major tools of breakeven and sensitivity analyses which frequently arise in the evaluation of engineering and financial alternatives. Because there is no convenient direct method of determining internal rates of return from given cash flow streams, it is usually necessary to calculate IRR by laborious trial and error methods. Spreadsheet calculations of IRR offer a programmed approach to trial and error methods which saves considerable amounts of time and effort as shown in the following examples.

Example 1 - A $3,000 investment was returned $1,000 per year over a period of four years. What was the internal rate of return of the investment? In functional notation, the problem is, Solve the equation $3,000 = $1,000(P/A,i%,4) for 'i'. Without the aid of a spreadsheet, one would scan the interest tables to find (P/A,12%,4) = 3.037 and (P/A,15%,4) = 2.855 . Consequently, 'i' must lie between 12% and 15% per year, and the exact answer must be found by interpolation.

The spreadsheet approach consists of listing {-3000; 1000; 1000; 1000; 1000} in consecutive row or column cells of the spreadsheet. Then an answer cell is chosen. The programmed function IRR is then key punched into the answer cell together with the range of cell addresses of the cash flow and an initial guess of the solution. Different formats are used to code the IRR function for various programs and types of microcomputers. For example, if the cash flow is listed in cells A1 to A5 and the initial guess of the solution is 10% per year, the user would enter @IRR(0.1,A1..A5) in the Lotus 1-2-3 program for IBM compatible computers, and the user would enter =IRR(A1:A5,0.1) in the Excel program for Macintosh computers. Upon execution of the program, the result in the answer cell is i = 0.12589832 or 12.589832% per year.

Example 2 - The Lorie and Savage problem in Section 2.3 has multiple rates of return. If nothing was done, an oil well would produce $10,000 worth of oil at the end of each year for the next two years. An engineer proposes to install a pump costing $1,600 which would produce all $20,000 worth of oil at the end of the first year, at which time the pump would have no salvage value. The "incremental" cash flow between the ongoing and pump alternatives is {-$1,600; $10,000; -$10,000}. As explained in Appendix 2B, the internal rates of return of the incremental cash flow is 25% and 400% per year.

Although this problem is easily solved by means of the quadratic formula, it is instructive to follow the spreadsheet approach to the solution. Suppose we enter {-1600, 10000, -10000} in cells A1 to A3 of the spreadsheet. If we now enter 100% per year as an initial guess of the solution, the programmed IRR function will not be able to give an answer. The reason for this misbehavior is that the slope of the PVP(Ø-P) curve is horizontal at i = 100% per year. The programmed IRR function depends upon the slope of the PVP curve at the initial guess in order to converge to the answer, and a horizontal slope will not permit the program to converge. However, if the initial guess was 80% per year (or less), the answer would converge to 25% per year; and if the initial guess was 120% per year (or more), the answer would converge to 400% per year.

Example 3 - Constant Equivalents of Variable Rates of Return - IRR calculations are useful for finding constant equivalents of fluctuating rates of return for random sequences of deposits and withdrawals. For example, let us determine constant equivalents of the variable interest rates shown in column (2) of Tables 3C.1 and 3C.2 below.

Table 3C.1 - Cash Flow Accounting for Future Time Equivalences of Random Deposits

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	7.00%	$0.00	$1,000.00	$1,000.00
1	5.00%	$1,050.00	$50.00	$1,100.00
2	6.00%	$1,166.00	$60.00	$1,226.00
3	7.00%	$1,311.82	$70.00	$1,381.82
4	8.00%	$1,492.37	$1,080.00	$2,572.37

Table 3C.2 - Cash Flow Accounting for Present Time Equivalences of Random Deposits

(1)EOY	(2)Interest/Year	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance
0	7.00%	$2,000.00	$1,000.00	$1,000.00
1	5.00%	$1,050.00	$50.00	$1,000.00
2	6.00%	$1,060.00	$60.00	$1,000.00
3	7.00%	$1,070.00	$70.00	$1,000.00
4	8.00%	$1,080.00	$1,080.00	$0.00

In order to find the equivalent rate of return for the future value of $2,572.37 in Table 3C.1, let us change the last cash flow of $1,080 to -$1,492.37 as if the depositor withdrew the last balance. Spreadsheet calculations would then show IRR = 6.555697% per year. The IRR solution is verified in Table 3C.3.

Table 3C.3 - Cash Flow Accounting for Future Time Equivalences of Random Deposits

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	6.555697%	$0.00	$1,000.00	$1,000.00
1	6.555697%	$1,065.56	$50.00	$1,115.56
2	6.555697%	$1,188.69	$60.00	$1,248.69
3	6.555697%	$1,330.55	$70.00	$1,400.55
4	6.555697%	$1,492.37	$1,080.00	$2,572.37

In order to find the equivalent rate of return for the present value of $2,000.00 in Table 3C.2, let us change the first cash flow of $1,000 to -$1,000 as if the depositor was loaned the value of the future deposits. Spreadsheet calculations would then show IRR = 6.422279% per year which is verified in Table 3C.4.

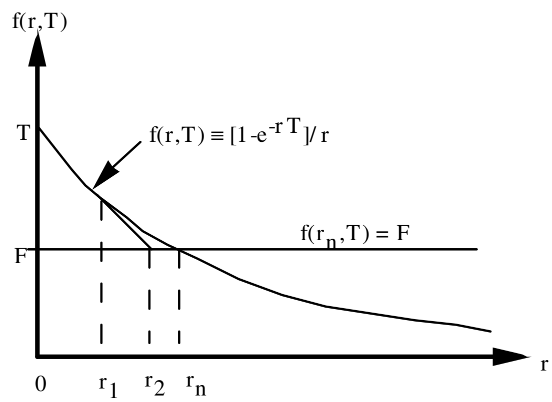
Table 3C.4 - Cash Flow Accounting for Present Time Equivalences of Random Deposits

(1)EOY	(2)Interest/Year	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance
0	6.422279%	$2,000.00	$1,000.00	$1,000.00
1	6.422279%	$1,064.22	$50.00	$1,014.22
2	6.422279%	$1,079.36	$60.00	$1,019.36
3	6.422279%	$1,084.83	$70.00	$1,014.83
4	6.422279%	$1,080.00	$1,080.00	$0.00

Example 4 - Newton-Raphson Iteration Method - The built-in IRR functions of microcomputer spreadsheets assume that input cash flows are discrete and periodic. In cases where cash flows are continuous, or if cash flows are discrete but aperiodic, the built-in IRR functions are inapplicable. As explained below, other spreadsheet solutions encapsulate the Newton-Raphson iteration method in order to determine the IRR of investments whose cash flows are either continuous or discrete and aperiodic. The Newton-Raphson iteration method will be presented here by an example of investing $6,000 which returns $1,000 per year continuously for 20 years. The problem is to determine the discount rate, r, which makes the present value of the returns equal to the $6,000 investment (i.e., Solve the equation $6,000 = $1,000•($P/A_c$,r%,20) for r).

Let $f(r,T) \equiv (P/A_c,r\%,T) \equiv [1-e^{-rT}]/r$. The problem is to determine the unknown discount rate, r_n, which makes $f(r_n,T) = F$ when $F = 6$ and $T = 20$. We could estimate r_n from interest tables of $(P/A_c,r\%,T)$ at the back of this book. Suppose the first estimate of r_n is r_1, and we want a closer estimate r_2 to the unknown value of r_n shown in Figure 3C.1.

Figure 3C.1 - Newton-Raphson Iteration Method



Let us calculate $f(r_1,T)$ which gives us a point $[r_1,f(r_1,T)]$ on the curve of $f(r,T)$ versus r where a tangent is drawn until it intersects horizontal line $f(r_n,T)= F$ at the point $[r_2,f(r_n,T)]$. The slope of the tangent at $[r_1,f(r_1,T)]$ equals the first derivative $f'(r_1,T)$ given in (3C.1).

$$f'(r_1,T) = \frac{f(r_1,T) - F}{r_1 - r_2} \quad \text{or} \quad r_2 = r_1 - \frac{f(r_1,T) - F}{f'(r_1,T)} \qquad \text{....(3C.1)}$$

Equation (3C.1) provides us with an estimate r_2 that depends only on r_1 but which is much closer to r_n. The next estimate of r_n is r_3 which can be calculated from equation (3C.1) by replacing $[r_1,f(r_1,T)]$ and $f'(r_1,T)$ with $[r_2, f(r_2,T)]$ and $f'(r_2,T)$ and so on. The iterations converge to r_n if the second derivative $f''(r,T)$ has the same sign between r_1 and r_n.

For continuous functions of r such as $(P/A_c,r,T) = f(r,T)$, the first derivative $f'(r,T)$ can be calculated analytically as follows:

$$f'(r,T) = \frac{rTe^{-rT} - [1 - e^{-rT}]}{r^2} = \frac{Te^{-rT} - f(r,T)}{r} \qquad \text{....(3C.2)}$$

Let us now determine the value of r_n that satisfies the equation $(P/A_c,r\%,20) = 6$. The headings of the spreadsheet are filled in the first row of Table 3C.1 below. Then enter 1 in A2, $r_1 = 0.1$ in B2, T = 20 in C2, F = 6 in D2, $f(r_1,T)$ = (1-EXP(-B2*C2))/B2 in E2, and $f'(r_1,T)$ = (C2*EXP(-B2*C2) - E2)/B2 in F2. Upon entering the formulas for $f(r_1,T)$ and $f'(r_1,T)$ in E2 and F2, the values 8.6466472 and -59.39942 will appear in their cells. On row 3, enter A2+1 in A3, and r_2 = B2 - (E2-D2)/F2 in B3 whose value becomes 0.1445568 . We can now select and copy from C2 to F2 onto C3 to F3 for the second iteration. Subsequent iterations are obtained by selecting and copying from A3 to F3 onto A4 to F4, A5 to F5 etc. until r = .159853 when D6 and E6 become sufficiently close in value. The spreadsheet copying process updates cells in formulas except those with $ prefixes.

Table 3C.1 - Newton-Raphson Iteration to Solve $(P/A_c, r\%, 20) = 6$ for 'r'.

	A	B	C	D	E	F
1	Iteration	r	T	F	f(r,T)	f'(r,T)
2	1	0.1	20	6	8.6466472	-59.39942
3	2	.1445568	20	6	6.5336736	-37.51751
4	3	.1587814	20	6	6.0349129	-32.74662
5	4	.1598476	20	6	6.0001739	-32.42114
6	5	.159853	20	6	6.	-32.41951

For cash flow patterns, discrete or continuous, whose present-value functions f(r,T) are not easily differentiated with respect to r, the first derivative f'(r,T) is approximated numerically by the secant line between adjacent values of r as shown in equation (3C.3). The Newton-Raphson iteration method is modified accordingly in equation (3C.4).

$$f'(r_1,T) = \frac{f(r_1,T) - f(r_2,T)}{r_1 - r_2} \qquad \text{....(3C.3)}$$

$$r_3 = r_1 - (r_1 - r_2)\frac{f(r_1,T) - F}{f(r_1,T) - f(r_2,T)} \qquad \text{....(3C.4)}$$

Iterations of equation (3C.4) use two of the three values of r_1, r_2 and r_3 for which $f(r_1,T)$, $f(r_2,T)$ and $f(r_3,T)$ come closest to the value of F. Many spreadsheet solutions encapsulate Newton-Raphson iterations with equations (3C.3) and (3C.4).

<u>Example 5</u> - IRR Calculations for Bond Purchases - If a bond is purchased on an interest payment date, the purchaser obtains the right to receive future interest payments plus the redeemable value of the bond either on the date that it is callable or the date of its maturity. Suppose \$920 is paid on July 1, 1991 to purchase a \$1,000 bond with 9 1/2% interest payable semiannually on January and July 1st of each year until it is redeemable at par on January 1, 2014. The purchaser receives the first two interest payments of 0.095•(\$1,000)/2 = \$47.50 on January 1 and July 1, 1992 as well as the next 21 years for a total of 44 interest payments. On January 1, 2014, the purchaser receives the 45th interest payment of \$47.50 plus the \$1,000 bond redemption value. The semiannual rate of return of this bond purchase is readily obtained from the built-in IRR function of a spreadsheet by entering -920 in cell A1, 47.50 in cells A2 to A45, and 1047.50 in cell A46. The spreadsheet solves the equation 920 = 47.50(P/A, i, 45) + 1000(P/F, i, 45) to give i = 5.2143%.

Suppose \$920 is paid on September 20, 1991 to purchase the bond described above. Eighty-one days (or 81/180 = 0.45 of a half-year) have elapsed since the July 1, 1991 interest payment. The buyer must pay accrued *simple* interest of 0.45•\$47.50 = \$21.38 plus the \$920 cost of the bond. The semiannual rate of return, i, is now found where the function f(i) ≡ -941.38*(P/F,i,0.45) + 47.50*(P/A,i,45) + 1000*(P/F,i,45) equals zero. Although the \$941.38 payment was shifted to a semiannual period by multiplying with (P/F,i,0.45), the unknown numerical quantity -941.38*(P/F,i,0.45) cannot be entered into a cell of the spreadsheet. Therefore, the Newton-Raphson iteration method is used outside the framework of the built-in IRR function. The derivative f'(i) can be evaluated analytically, i.e.: $f'(i) = -941.38*(-0.45)*(1+i)^{-1.45} - 47.50i^{-2}[1-(1+i)^{-45}] + 47.50i^{-1}[45*(1+i)^{-46}] + 1000*(-45)*(1+i)^{-46}$. Consequently, the solution of this problem was determined by the Newton-Raphson iteration method described in Example 4. *Ans*. i = 5.2137%.

6. Solving Transcendental Equations by Microsoft Excel Solver - The Newton-Raphson iteration method can be encapsulated into spreadsheet programs using either analytical or numerical derivatives. In particular, one may install an add-in feature in Microsoft Excel called **Solver** which appears in the **Tools** menu. The procedure for solving transcendental equations with Microsoft Excel Solver is illustrated below for the problem described in Example 4.

In Example 4, a $6,000 investment is made which returns $1,000 per year continuously for 20 years. The problem is to determine the discount rate, r, which makes the present value of the returns equal to the $6,000 investment (i.e., Solve the equation $6,000 = $1,000•$(P/A_c,r\%,20)$ for r). For this purpose, we let $f(r) \equiv (P/A_c,r\%,20) \equiv [1-e^{-20r}]/r$. We now need to determine the unknown discount rate, r, which makes f(r) = 6.

Before the Tools menu is activated, the formula for f(r) needs to be placed in a cell which is later referred to as the Target Cell in the Solver Parameters dialogue box. Suppose cell C2 is where the formula (1-EXP(-20*B2))/B2 for f(r) is entered. Cell B2 represents the variable r for which we are solving. A guess of, say, 0.1 is entered in cell B2 to serve as a starting point for Solver. Then Excel updates cell C2 to a numerical value of 8.64664717 .

At this point, the **Tools** menu is activated and the **Solver** option is selected. With the cursor in the Target Cell field of the Solver Parameters dialogue box, click on cell C2 and **C2** will appear in the Target Cell field. With the cursor in the By Changing Cells field of the Solver Parameters dialogue box, click on cell B2 and **B2** will appear in the Changing Cells field. The value 6 which f(r) must equal is entered in the Equal To field of the Solver Parameters dialogue box.

After entering values in the Target Cell, Changing Cells and Equal To fields of the Solver dialogue box, click on the **Solve** button. Unless the solution diverges or does not exist, the variable cell (in our case B2) will contain the solution 0.15985316, and the target cell (in our case C2) will contain the Equal To value 5.99999337 which was reached in the solution. After finding the solution, a Solver Results dialogue box appears in which Excel gives the option to keep and obtain reports about the solution.

Reference: *A Guide to Microsoft Excel for Scientists and Engineers* by Bernard V. Liengme, copublished by John Wiley & Sons, Inc., New York, 1997.

Chapter Three - Exercises

Sections 3.2 and 3.3

3-1a A bank which pays interest rates compounded annually as indicated in column (2) of Table I, receives deposits(+) and withdrawals(-) from a company according to column (4) of Table I. Using ABC accounting rules and Table I, determine the ACF Balance (end-of-year 4) of the company's cash flows.

3-1b Using ABC accounting rules, determine an equivalent 4-year uniform series of cash flows for the company's ACF Balance at the end of four years and verify the result.

3-1c Replace the $8,000 cash flow at end-of-year 4 with the negative of the BCF Balance in the 4th year, and determine the equivalent rate of return of the resulting cash flow series using a spreadsheet IRR function. Verify that the equivalent rate of return would yield the same ACF Balance at the end of 4 years. (see Example 3 of Appendix 3C)

3-1d Using A'B'C' accounting rules and Table II, determine the present value of the company's cash flows (i.e., the ACF Balance at the end of year zero).

3-1e Using A'B'C' accounting rules, determine an equivalent 4-year uniform series of cash flows for the company's ACF Balance at the end of year zero and verify the result.

3-1f Replace the $7,000 cash flow at end-of-year 0 with the negative of the BCF Balance in the zeroth year, and determine the equivalent rate of return of the resulting cash flow series using a spreadsheet IRR function. Verify that the equivalent rate of return would yield the same ACF Balance at the end of year zero. (see Example 3 of Appendix 3C)

(1)EOY	(2)Interest/Yr	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance	Table I
0	7.4%	$0.00	$7,000.		
1	8.6%		-$5,000.		
2	9.8%		$9,000.		
3	10.2%		-$6,000.		
4	8.7%		$8,000.		

(1)EOY	(2)Interest/Yr	(3)ACF Balance	(4)EOY Cash Flow	(5)BCF Balance	Table II
0	7.4%		$7,000.		
1	8.6%		-$5,000.		
2	9.8%		$9,000.		
3	10.2%		-$6,000.		
4	8.7%		$8,000.	$0.00	

Sections 3.4, 3.5 and 3.6

A student buys a car with a $7,000 loan to be paid off in 4 years with equal end-of-month installments at a 10.5% Annual Percentage Rate (APR). If the student fails to make a monthly payment, the interest due that month is added to the principal amount of the outstanding loan. If the monthly payment exceeds the interest due that month, the excess is used to reduce the principal amount of the outstanding loan. These rules define monthly financing or monthly compounding of the 10.5% APR.

3-2a Determine the effective interest rate i_E per cash flow period of one month of the nominal interest rate r = 10.5% per year. (Eqn (3.4.1))

3-2b Determine the future amount F of the outstanding loan at the end of 4 years if the student fails to make any of the 48 monthly payments. (eqn (3.5.1) F = P(F/P,i,n))

3-2c Determine the monthly payments A if the the student makes all 48 installments of the loan payments on time. (eqn (3.6.2) A = F(A/F,i,n))

3-2d Determine the monthly payments A in **3-2c** using equation (3.6.6) A = P(A/P,i,n).

3-2e Determine the loan balance F_n after the first n = 22 monthly payments. $F_n = P(F/P,i,n) - A(F/A,i,n)$ from equations (3.5.1) and (3.6.1)

3-2f Determine the loan balance P_{48-n} before the last 48-n = 26 monthly payments based on equation (3.6.5) $P_{48-n} = A(P/A,i,48-n)$

3-2g Determine the nominal interest rate r compounded continuously during a cash flow period of one month using equation (3.4.3) $r*t = \ln(1+i_E/1)$ with t = 1 month.

3-2h Determine the continuous cash flow rate A_c for paying off the $7,000 loan with continuous compounding at nominal interest rate r found in 3-2g. Use the relationship $F = A(F/A,i,n) = A_c(F/A_c,r,t)$ obtained by setting equation (3.6.1) equal to (3.6.3). When integer n equals continuous time variable t measured in cash flow periods (i.e., months), the relationship above reduces to $A_c = rA/i$.

3-2i Determine the loan balance F_t after the first t = 22 monthly payments using the relationship $F_t = P(F/P,r,t) - A_c(F/A_c,r,t)$ from equations (3.5.2) and (3.6.3) with continuous cash flows and compounding. Compare the results with **3-2e**.

3-2j Determine the loan balance P_T before the last 48-t = 26 monthly payments based on equation (3.6.7) $P_{48-t} = A_c(P/A_c,r,48-t)$ with continuous cash flows and compounding. Compare the results with **3-2f**.

3-3a A person deposits $1,200 at the end of the first and second years in a bank which pays 9.6% interest per year compounded quarterly. Find the balance accumulated at the end of five years.

3-3b The balance is withdrawn in five equal annual installments at the ends of the sixth, seventh, eighth, ninth, and tenth years. Determine the amount of each annual withdrawal.

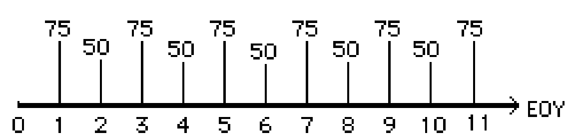
3-4a You borrow $50,000 in the form of a variable-rate mortgage to be repaid in equal monthly amounts over the next 25 years. The initial interest rate is 10.20% per year compounded monthly. How much are the monthly payments during the first year?

3-4b After the eighteenth payment, the bank adjusts the interest rate to 9.60% per year. How much would the monthly payments be now?

3-4c Before you undertook the variable-rate mortgage, you had the option of taking a fixed-rate mortgage of $50,000 to be repaid in equal monthly amounts for 25 years at an interest rate of 9.25% per year. If 2 points are required to obtain this fixed-rate mortgage (i.e., the 2 point stipulation means that 2% is added in advance to the $50,000 principal amount of the mortgage), how much would the monthly payments be for the fixed-rate mortgage?

3-5a Determine the present-value at end of year zero of the series of end-of-year cash flows in the time diagram below when i = 10.40% per year compounded quarterly.

3-5b Determine the future-value at the end of year 12 of the series of end-of-year cash flows in the time diagram below when i = 10.40% per year compounded quarterly.



3-6a A new plant needs a boiler that can be fired *continuously* either by natural gas, fuel oil, or coal. The installation cost would be $15,000 for natural gas; $55,000 for fuel oil; and $150,000 for coal. If natural gas is used, the fuel cost rate will be $5,000/yr more than fuel oil. If coal is used, the fuel cost rate will be $10,000/yr less than fuel oil. Assume 7% nominal annual interest compounded *continuously*, a 20-year period of analysis, and no net salvage value. Which fuel is most economical?

3-6b At what interest rate are the present-value costs of natural gas and coal equal? (see Appendix 3C, Example 4)

3-6c When will the 8%/year discounted costs of natural gas and coal be equal?

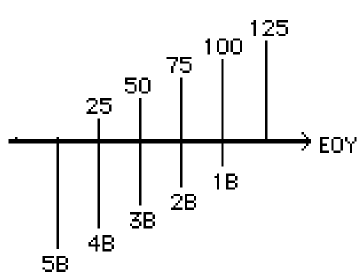
3-7a Pacific Bell (PacTT) 9 7/8 '16 bonds were bid 106 on 9/13/91. This means buyers bid 106% of the bond's $1,000 face value in order to buy it. The bond pays semiannual dividends of 9.875% per year on Feb. and Aug. 15th each year until maturity on Feb. 15, 2016 when the bondholder would receive the semiannual dividend plus the bond's $1,000 face value. Determine the current yield or annual percent return (APR) defined as the percent ratio of one-year dividend income to the current price of the bond. Find the cost of a bond bought 9/13/91 including accrued simple interest since the Aug. 15, 1991 dividend plus $15 broker commission per bond.

3-7b The bonds could be called Feb. 15, 1996 at 105 (i.e., $1,050/bond). Determine the yield to call (i.e., the effective semiannual rate of return) based on the 9/13/91 cost of a bond. (see Appendix 3C, Example 5)

3-8a A coal mine is expected to produce a net end-of-year income of $75,000 for each of the next 20 years. At the end of 20 years, the coal will be exhausted and the mine site would then have an estimated value of $150,000. A prospective buyer wishes to earn 12% per year compounded quarterly on his initial investment for 20 years. For this purpose, a sinking fund is being set up into which uniform annual payments will be made. The sinking fund earns 9%/year compounded quarterly and will return the initial investment 20 years from now. What is the most the buyer should pay for the mine?

<u>Sections 3.7 and 3.8</u>

3-9a Assume the value of the 5 negative end-of-year cash flows in the time diagram below equals the value of the 5 positive cash flows when i = 10.60% per year compounded semiannually. Determine B.



3-10a The initial production rate from an oil well is worth $150,000 per year which declines *continuously* at the rate of 7% per year (i.e., g_c = -7%/year). When the production rate declines to $50,000 per year, it is expected the oil well could be sold for $100,000. Determine the present value of the production and the later sale of the oil well when the nominal annual interest rate r = 10.8% is compounded *continuously*. (Appendix 3A - Continuous Geometric Gradients)

3-11a Tuition costs are expected to inflate 8% annually. A student will be entering a four-year program one year from now when the first year's tuition will be $3,900 payable at the beginning of the year. A fund is being set up today to pay the student's tuition for four years. How large must the fund be if the annual interest rate is (A) 7%, (B) 8%, and (C) 9%, compounded quarterly? (Appendix 3A - Discrete Geometric Gradients)

3-12a Service records indicate a machine has initial end-of-year maintenance costs of $600 which increase $50 per year during its 10-year service life. A 10-year service contract is offerred when the machine is purchased. Assuming a 9.6%/yr interest compounded quarterly, determine the net present value of the 10-year service contract.

3-12b In how many years will the maintenance costs discounted at 6%/year interest be equal to the net present value of the 10-year service contract? (Appendix 3A - Discrete Arithmetic Gradients)

Chapter Three - Suggested Readings

Dimensional Analysis

Bridgman, P. W., *Dimensional Analysis,* Yale University Press,1931.

De Jong, F. J., *Dimensional Analysis for Economists*, North-Holland Publishing Co., 1967.

Duncan, W. J., *Physical Similarity and Dimensional Analysis*, Edward Arnold & Co., 1953.

Langhaar, H. L., *Dimensional Analysis and Theory of Models*, John Wiley and Sons, Inc., 1962.

LeCorbeiller, P., *Dimensional Analysis*, Appleton-Century-Crofts, 1966.

Chapter Four - What is the best way of doing each project?

Section 4.1 - Objectives of Economic Decision-Making

The objective of economic decision-making is to select engineering and financial alternatives which maximize the net present-value added (ΔNPV) to an organization as a whole subject to a capital constraint. This objective is a generalization of the principles of maximizing output for a given input and minimizing input for a given output. However, most alternatives do not have either the same input or output in which cases the objective is not well defined. As a result, multiple criteria have been developed in practice which seek to maximize the net present-value objective in different ways. The proposed and conventional methods of satisfying the objective have many similarities, but also important differences. Chapter Four explains both the similarities and differences between the proposed and conventional methods of answering the question, What is the best way of doing each project?

The conventional criteria for determining the best way of doing each project include net present-value discounted at MARR (NPV), equivalent uniform worth (EUW), internal-rate-of-return (IRR), benefit/cost ratio (B/C), payback period (PBP), and average annual percent profit (AAPP). Because multiple criteria appear to satisfy the same objective, there is room to question whether different criteria would determine the same best alternatives. There is also a need to question how best alternatives determined by conventional criteria differ from alternatives selected by the proposed method of economic decision-making.

It is often assumed that if the best way of doing a project was found by one criterion, then that criterion would also be optimized for the economic organization as a whole. Different criteria would then define multiple objectives which implies economic organizations would have more than one economic objective. But when multiple objectives exist, there are times when different objectives give conflicting answers. If multiple objectives always determine the same best alternative, they would serve as the same objective. Therefore, multiple objectives imply conflicts, at least some of the time. Whenever conflicts occur, multiple objectives cannot determine which alternatives are best. Instead, a system of weights must be used to decide how much each objective should be followed.

A firm's *marginal* output/input ratio, $Ø_m$, is proposed here as a single criterion which enables one to consistently select the best way of doing each project in order to maximize the firm's ΔNPV objective subject to a capital constraint. When net present-value discounted at MARR (NPV) is used as a single criterion to determine the best way of doing each project, it is shown in Chapter One that a firm's ΔNPV objective may exceed its capital constraint. When the output/input ratio of individual alternatives is used as a single criterion, the firm's overall capital efficiency may be maximized but not its ΔNPV objective.

Section 4.2 - Measurements of Net Present-Value (NPV)

Net present-value (NPV) of an alternative is defined as the net present-value of its input and output cash flows discounted at the minimum attractive rate of return, MARR. Similarly, net present-value (ΔNPV) of an alternative is defined as the net present-value of the *change* in cash flows of an organization between accepting and not accepting that alternative discounted at the cost of borrowing money. Measurements of NPV and ΔNPV differ with respect to **(1)** defining "do-nothing" and "ongoing" alternatives, **(2)** estimating investment and borrowing opportunity costs, and **(3)** determining the capital constraint.

(1) Defining "do-nothing" and "ongoing" alternatives - The do-nothing alternative is defined in conventional economic decision-making as follows. The choice of not investing in a project is called the *do-nothing alternative*. The foregone benefit of not investing in the worst accepted project would be to invest elsewhere for a minimum attractive rate of return, MARR, which is called the *investment opportunity cost*. The net present-value, NPV, of do-nothing alternatives discounted at MARR is zero just like *null alternatives* which literally do nothing because they consist only of zero cash flows. Hence, in NPV measurements discounted at MARR, it does not matter whether or not do-nothing alternatives are included in the set of mutually exclusive alternatives of each project.

The zero NPV of do-nothing alternatives discounted at MARR has far-reaching consequences in selecting the best way of doing each project. If either cost-increasing or cost-decreasing alternatives have negative NPVs discounted at MARR, then such alternatives would not be accepted. Thus, do-nothing alternatives act as minimum acceptance constraints of both the NPV and IRR criteria. These constraints are thought to be necessary for maximizing the NPV and IRR of the organization as a whole.

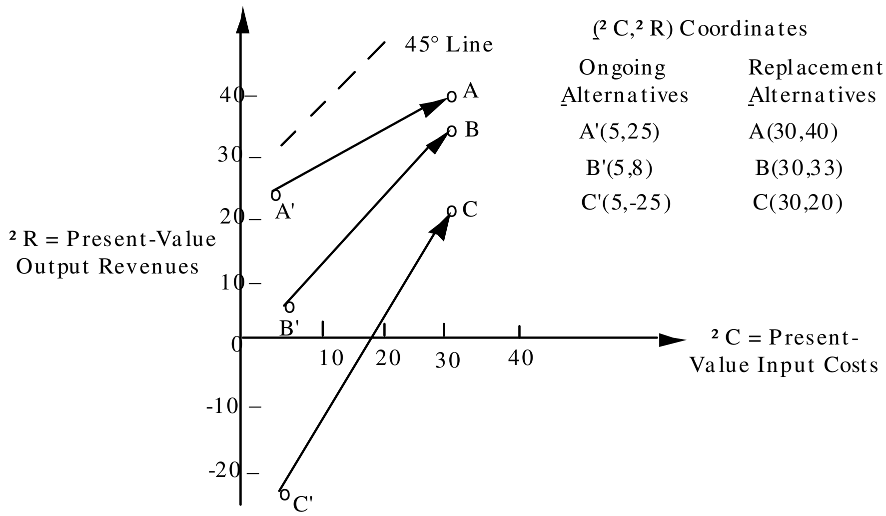
Negative NPV alternatives may be required because of public regulations governing pollution, safety, health, zoning, etc. Because such alternatives are often money-losing propositions, there is even a greater need for optimal economic decision-making in order to cut losses which deplete a firm's resources. An enterprise can benefit as much or more by reducing negative NPV alternatives as it can by increasing positive NPV alternatives. Many private and public projects have negative net present-values because their evaluations do not include intangible or difficultly measureable quantities. However, the comparison of intangible and difficultly measurable quantities which are common to two mutually exclusive alternatives will cancel out and not affect the decision of which one is better.

Unlike do-nothing alternatives, *ongoing alternatives* are defined as continuations of existing projects without changing the effects of past decisions. What you do at present is the outcome of past decisions. Current operating conditions are assumed to result from past decisions. You may not want to change what you do at present in which case ongoing alternatives are continued as planned in the past. But if you want to change what you are doing, then ongoing alternatives must be compared to other mutually exclusive alternatives. An ongoing alternative is only one element in a set of mutually exclusive alternatives of an existing project, and only other alternatives in the set can replace the ongoing alternative.

Replacements of ongoing alternatives can be either cost-increasing or cost-decreasing. When cost-increasing alternatives replace ongoing alternatives, the organization either increases borrowing or decreases lending. Conversely, when cost-decreasing alternatives replace ongoing alternatives, the organization either decreases borrowing or increases lending. In either case, the foregone alternatives of not replacing ongoing alternatives are measured *externally* in terms of *borrowing opportunity costs* at market rates of interest. The net present-value of the ongoing alternative of an existing project discounted at market rates of interest could be either positive, negative or zero.

To compare the consequences of differences between NPV and ΔNPV measurements, let us evaluate replacements of do-nothing and ongoing alternatives as depicted in Figure 4.2.1 which plots present-values of input costs (ΔC) and output revenues (ΔR). Do-nothing alternatives have zero net present-values which lie on a 45° line through the origin and replacements of do-nothing alternatives must lie above the 45° line. In the proposed method of economic decision-making, terminal points of replacements must lie above the marginal comparison slope, $Ø_m$, drawn through the terminal point of the ongoing alternative.

Figure 4.2.1 - Effects of Replacing Do-nothing and Ongoing Alternatives.



Let us first consider replacing do-nothing alternative A"(5,5) with NPV[A"] = \$5-\$5 = \$0 by A(30,40) with NPV[A] = \$40-\$30 = \$10. Therefore, if A replaced A", NPV would increase by NPV[A-A"] = \$10-\$0 = \$10. But replacing do-nothing alternative A" by A actually means A replaces A' which is the ongoing alternative that came from past decisions. Consequently, ΔNPV[A-A'] = (\$40-\$30)-(\$25-\$5) = -\$10. Thus, replacing A' by A decreases NPV by \$10 instead of increasing NPV by \$10 as implied by the virtual replacement of A" by A. Another way of showing challenger A should not replace defender A' is to compare the slope $Ø_{CD} = Ø_{AA'} = (40-25)/(30-5) = 0.6$ which is less than $Ø_m > 1$. It follows that the definition of do-nothing alternative A" causes the negative NPV replacement of A' by A.

If do-nothing alternative B"(5,5) is replaced by B(30,33), we would go from NPV[B"] = \$0 to NPV[B] = \$33-\$30 = \$3 which is an apparent increase of NPV[B-B"] = \$3-\$0 = \$3. But replacing do-nothing alternative B" by B actually means ongoing alternative B' is being replaced by B. Consequently, ΔNPV[B-B'] = (\$33-\$30)-(\$8-\$5) = \$0. Thus, replacing B' by B actually increases input costs by ΔC[B-B'] = \$30-\$5 = \$25 without any increase of NPV and without the supposed \$3 increase in NPV that was indicated by the virtual replacement of B" by B. Another way of seeing that B should not replace B' is that the slope $Ø_{CD} = Ø_{BB'} = (33-8)/(30-5) = 1.0$ is less than $Ø_m > 1$.

Lastly, let us replace do-nothing alternative C"(5,5) by C(30,20). If C replaced C", we would go from NPV[C"] = \$0 to NPV[C] = \$20-\$30 = -\$10 which is an apparent decrease of NPV[C-C"] = -\$10-\$0 = -\$10. Such a decrease of NPV would indicate that C should not replace C". But not replacing do-nothing alternative C" by C actually means that ongoing alternative C'(5,-25) will not be replaced by C(30,20). If ongoing alternative C'(5,-25) with NPV[C'] = -\$25-\$5 = -\$30 was replaced C(30,20) with NPV[C] = \$20-\$30 = - \$10, the net present-value would increase by NPV[C-C'] = -\$10-(-\$30) = \$20. Thus, the previous result that NPV[C-C"] = -\$10 misleads the decision-maker into not replacing C' by C even though such a replacement would increase net present-value by NPV[C-C'] = \$20. Another way of seeing that C should replace C' is that the slope $Ø_{CD} = Ø_{CC'} = (20-(-25))/(30-5) = 1.8$ is probably greater than $Ø_m > 1$. It follows that the definition of do-nothing alternative C" causes a lost opportunity of increasing NPV by \$20 by not replacing C' with C.

The major focus of defining do-nothing alternatives is not so much a choice of not investing in a project as it is to define MARR which is a baseline for the IRR criterion and a discount rate needed to calculate net present-values. Do-nothing alternatives discounted at MARR are determined independently of any specific project, whereas ongoing alternatives are needed to distinguish between past and present decisions which are designed for finding the best way of doing each project.

Conventional economic decision-making treats *sunk costs* which are revenues or expenses that occurred before the present decision as being irrelevant because sunk costs are common to all current alternatives and only *differences* between alternatives can decide which choice is best. But past decisions go beyond sunk costs by determining the ongoing alternative of a project. Although present decisions cannot affect sunk costs, they can entertain cost-increasing or cost-decreasing alternatives which may affect only portions of the ongoing alternative. Consequently, the ongoing alternative can be compared to other mutually exclusive alternatives of the project and common portions of past and present decisions will cancel out in the binary comparisons of project alternatives.

(2) Estimating investment and borrowing opportunity costs - Besides differences between do-nothing and ongoing alternatives, NPV and ΔNPV measurements differ with respect to the discount rates used in present-value calculations. Engineering and financial discount rates used in NPV measurements are based on *investment opportunity costs* estimated from an organization's minimum attractive investments and cost of capital. Discount rates used in ΔNPV measurements are based on *borrowing opportunity costs* estimated externally from costs of borrowing money or market rates of interest.

Engineering discount rates (Section 2.2) are derived from foregone alternatives of minimum attractive investments which are called best rejected projects or *do-nothing alternatives* among other synonyms. Assuming equal IRRs of minimum attractive investments and best rejected projects, their IRR is called the minimum attractive rate of return or (MARR). In theory and practice, it is not clear how to determine MARR. In theory, independent projects are ranked in descending order of their IRRs until reaching the cut-off of the capital constraint. The foregone alternative of last accepted project is the best rejected project or *do-nothing alternative* whose internal rate of return is MARR. The cash flows of all alternatives are then discounted at MARR to determine their net present-values.

Financial discount rates (Section 2.3) use a *cost of capital* approach to determine the rate of return of the minimum attractive investment. The underlying assumption governing this approach is that the sources of capital for an economic organization all come from the debt and equity components of its capitalization. Consequently, a weighted average cost of capital (WACC) is taken of short- and long-term debt interest rates and the expected rates of return of equity investors from dividends, capital appreciation and retained earnings. The operational difficulties of estimating MARR and WACC discount rates from observable data are still a deep-seated problem.

The discount rates used in ΔNPV measurements are based on *borrowing rather than investment opportunity costs* of MARR or WACC discount rates. The foregone alternative of not accepting the minimum attractive investment is the option of reducing the debt of the organization. It was shown in Section 2.3 that the NPV of the minimum attractive investment is independent of the debt fraction of its input costs when discounting at the cost of borrowing money. Borrowing opportunity costs are less than MARR or WACC discount rates because organizations can function in the long run only if their rates of return exceed their costs of borrowing money.

Let us compare the consequences of discount rates derived from investment and borrowing opportunity costs. Suppose MARR or WACC discounting is 30%/year and the cost of borrowing money is 15%/year. An early cash flow alternative ECF{-$100; $130; $10} is compared to a late cash flow alternative LCF{-$100; $10; $160} as two mutually exclusive alternatives. Tables 4.2.1 and 4.2.2 below show the net present-value calculations of the ECF and LCF alternatives at discount rates of 30% and 15% per year respectively.

The cumulative cash flows of ECF and LCF are CCF{ECF}[-$100; $30; $40] and CCF{LCF}[-$100; -$90; $70] which show the NPV comparison of ECF and LCF is sensitive to discount rate magnitudes. Solving the equation NPV{LCF-ECF}[0;-120;150] = 0 for the breakeven discount rate, i, which makes NPV{ECF} = NPV{LCF} gives i = 25%/year. Thus, NPV{ECF} > NPV{LCF} for i > 25%, and NPV{ECF} < NPV{LCF} for i < 25%.

Table 4.2.1 - Early Versus Late Cash Flow Alternatives at a Discount Rate of 30%/Year

	Net Present-value	=	Present-value Output	-	Present-value Input
Early Cash Flow:	$5.92	=	$\frac{130}{1.30} + \frac{10}{1.30^2} = 105.92$	-	100
Late Cash Flow:	$2.37	=	$\frac{10}{1.30} + \frac{160}{1.30^2} = 102.37$	-	100

Table 4.2.2 - Early Versus Late Cash Flow Alternatives at a Discount Rate of 15%/Year

	Net Present-value	=	Present-value Output	-	Present-value Input
Early Cash Flow:	$20.60	=	$\frac{130}{1.15} + \frac{10}{1.15^2} = 120.60$	-	100
Late Cash Flow:	$29.68	=	$\frac{10}{1.15} + \frac{160}{1.15^2} = 129.68$	-	100

Table 4.2.1 shows NPV[ECF-LCF] = $5.92-$2.37 = $3.55 which makes ECF *better than* LCF at the 30%/year discount rate. Table 4.2.2 shows NPV[ECF-LCF] = $20.60-$29.68 = -$9.08 which makes ECF *worse than* LCF at the 15%/year discount rate. As explained in Appendix 2A, future cash flows are heavily discounted by the 30%/year MARR discount rate which makes the first-year ECF advantage of $130-$10 = $120 appear greater than the second-year LCF advantage of $160-$10 = $150. Future cash flows are discounted only moderately by the 15%/year cost of borrowing money which enables the second-year LCF advantage of $150 appear greater than the first-year ECF advantage of $120.

Because 30% and 15%/year discount rates give opposite choices for ECF and LCF, let us explain the 30%/year choice by the option of reinvesting the first-year ECF advantage of $120 at 30%/year which is denoted by INV_{30}{0; -$120, $156}. Adding INV_{30} to ECF gives [ECF+INV_{30}]{-$100; $10; $166} which is clearly better than LCF{-$100; $10; $160}. Conversely, let the first-year ECF advantage of $120 be used to exercise the option of reducing the outstanding debt of 15%/year which is denoted by $DEBT_{15}${0; -$120, $138}. Adding $DEBT_{15}$ to ECF gives [ECF+$DEBT_{15}$]{-$100; $10; $148} which is clearly worse than LCF{-$100; $10; $160}. Thus, whether ECF is better or worse than LCF depends upon excercising the first-year options of the $120 ECF advantage for either reinvestment at 30%/year or debt reduction at 15%/year, respectively.

(3) Determining the capital constraint - The determination of the capital constraint is carried over from problems of estimating investment and borrowing opportunity costs. Engineering interpretations of investment opportunity costs leads one to the estimation of the capital constraint in the form of a minimum attractive rate of return (MARR) for the IRR of each alternative and a zero net present-value (NPV) for the cash flow of each alternative discounted at MARR. Owing to the possibility that discounting at MARR may allow alternatives with large net present-values to have output revenues which are received in the distant future, another constraint is usually placed on the minimum payback period (PBP) that output revenues can be received to cover the input costs of each alternative.

Financial accounting interpretations of investment opportunity costs make a strong distinction between debt and equity capital because of differences in their risks and rewards. Thus, debt capital has relatively little risks and smaller fixed rewards, whereas equity capital has more risks which are compensated with greater variety of rewards. By gaining access to both debt and equity capital, the overall cost of capital can be lowered. Up to a limit, financial leveraging derived from debt capital can potentially increase the market value of equity capital without unduly increasing its risks. Assuming organizations have an optimal financing mix of debt and equity capital which maximizes the market value of equity capital, the net present-value of the cash flow of projects with an average risk profile should be discounted by a weighted average cost of capital (WACC) described above.

The proposed application of borrowing opportunity costs for managerial accounting does not distinguish between debt and equity capital before income taxes because it was shown in Section 2.3 that the NPV of project alternatives does not depend on the debt fraction of their input costs when discounting at the before-tax cost of borrowing money. Likewise, in Chapter Seven, the distinction between debt and equity capital in after-tax evaluations is irrelevant when discounting at the after-tax cost of borrowing money. The proof in Section 1.6 shows that the marginal capital efficiency (or the marginal output/input ratio), $Ø_m$, of an economic organization is the sole determinant of its capital constraint.

Section 4.3 - Measurements of Equivalent Uniform Worth (EUW)

Equivalent uniform worth EUW{A} is defined as a uniform series of cash flows over the lifespan of alternative A which has an equivalent worth as NPV{A} or NFV{A} discounted at the same interest rate. Therefore, EUW{A} = NPV{A}(A/P,i,n) = NFV{A}(A/F,i,n). If alternatives A and B have equal lifespans, EUW and NPV comparisons of A and B always give the same results. However, alternatives often have unequal lifespans. The ΔNPV measurements are applied uniformly to all members of a set of project alternatives whether or not they have equal lifespans. But when project alternatives have unequal lifespans, conventional NPV measurements require alternatives to have equal lifespans in order for their EUW and NPV comparisons to be equivalent. Two methods of adjusting EUW and NPV comparisons of alternatives with unequal lifespans are (1) the *common-multiple* method, and (2) the *common-analysis-period* method as described below:

The *common-multiple* method coterminates alternatives with unequal lifespans by repeating them often enough to fill a common multiple of their lives. Thus, two alternatives with lives of 3 and 4 years have a least common-multiple life of 12 years. Repeating the 3-year alternative 4 times, and the 4-year alternative 3 times, results in a common 12-year period of comparison. The common-multiple-lifespan method makes continuing-service alternatives with unequal lifespans more comparable. It does not imply alternatives will be replaced by successors having identical input and output characteristics.

The *common-analysis-period* method terminates alternatives with unequal lifespans in a period of time that a project is expected to be in operation. An appropriate time horizon usually reflects either the estimated duration of required service or the shortest life of competing alternatives to account for technological obsolescence. When an alternative is terminated before the end of its service life, it is necessary to estimate the residual value of its operation or the salvage value and abandonment cost upon termination.

We will now show the EUW of alternative A with an n-year lifespan is constant regardless of how many times it is repeated. Suppose NPV{A} = P_1 discounted at an effective interest rate of i% per year. The EUW{A}, denoted by A_1, is given by

$$A_1 = P_1 (A/P,i,n) = \frac{P_1 i}{1-(1+i)^{-n}} \qquad \text{....(4.3.1)}$$

If the alternative is repeated at the end of its n-year life, then the net present-value of the twice repeated alternative is $P_2 = P_1[1+(1+i)^{-n}]$, and the equivalent uniform worth of the twice repeated alternative, denoted by A_2, can be determined from equation (4.3.2).

$$A_2 = P_2(A/P,i,2n) = P_1[1+(1+i)^{-n}]\frac{i}{1-(1+i)^{-2n}} \qquad \text{....(4.3.2)}$$

The denominator $[1-(1+i)^{-2n}]$ of (4.3.2) can be factored into $[1+(1+i)^{-n}][1-(1+i)^{-n}]$, the first factor of which cancels the factor immediately after P_1. Consequently, equation (4.3.2) reduces to

$$A_2 = P_1[1+(1+i)^{-n}](A/P,i,2n) = \frac{P_1 i}{1-(1+i)^{-n}} = A_1 \qquad \text{....(4.3.3)}$$

If the alternative is repeated three times at the end of its n-year life, the net present-value is is $P_3 = P_1[1+(1+i)^{-n}+(1+i)^{-2n}]$, and the equivalent uniform worth, denoted by A_3, can be determined from equation (4.3.4).

$$A_3 = P_3(A/P,i,3n) = P_1[1+(1+i)^{-n}+(1+i)^{-2n}]\frac{i}{1-(1+i)^{-3n}} \qquad \text{....(4.3.4)}$$

The denominator $[1-(1+i)^{-3n}]$ can be factored into $[1+(1+i)^{-n}+(1+i)^{-2n}][1-(1+i)^{-n}]$, the first factor of which cancels the factor immediately after P_1. Consequently, equation (4.3.4) reduces to

$$A_3 = P_1[1+(1+i)^{-n}+(1+i)^{-2n}](A/P,i,3n) = \frac{P_1 i}{1-(1+i)^{-n}} = A_1 \qquad \text{....(4.3.5)}$$

Thus, $A_R = A_1$, and the equivalent uniform *series* worth of an alternative is constant regardless of the number of of times, R, the alternative is repeated. Similarly, it can be shown the equivalent uniform *continuous* worth of an alternative is independent of the number of its identical replacements.

Although EUW is constant regardless of how many times an alternative is repeated, NPV increases as a geometric progression with each repetition as given in equation (4.3.6).

$$P_R = P_1 \frac{1-(1+i)^{-Rn}}{1-(1+i)^{-n}} \qquad \text{....(4.3.6)}$$

If alternatives have unequal lifespans and the EUW of each alternative is evaluated from NPV with individual (A/P,i,n) factors, then EUW and NPV criteria may give opposite results. But if EUW of differences between two alternatives is evaluated with an (A/P,i,n) factor whose parameter, n, is the lifespan of the cash flow difference between the two alternatives, then EUW and NPV criteria give the same results without invoking the common-multiple-lifespan or common-analysis-period assumptions. For example, let us compare alternatives X and Y costing $400 and $500, and whose lifespans are 4 and 5 years, as shown in Table 4.3.1 . The EUW and NPV criteria in Table 4.3.2 are discounted from 0% to 90% per year.

Table 4.3.1 - Cash Flow Descriptions of Alternatives X and Y.

End of Year	0	1	2	3	4	5	IRR%
Alternative X	-400	500	375	250	125	0	94.14
Alternative Y	-500	500	400	300	200	100	74.94
Difference Y-X	-100	0	25	50	75	100	26.66

Table 4.3.2 - EUW and NPV Comparisons of Alternatives X and Y.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
i%	NPV{X}	NPV{Y}	NPV{Y-X}	EUW{X}	EUW{Y}	EUW[{Y}-{X}]	EUW{Y-X}
0	$850.00	$1,000.00	$150.00	$212.50	$200.00	-$12.50	$30.00
10	$637.67	$709.21	$71.54	$201.17	$187.09	-$14.08	$18.87
20	$482.04	$504.69	$22.65	$186.21	$168.76	-$17.45	$7.58
30	$364.07	$354.81	-$9.26	$168.06	$145.68	-$22.38	-$3.80
40	$272.12	$241.21	-$30.91	$147.15	$118.52	-$28.63	-$15.19
50	$198.77	$152.67	-$46.10	$123.85	$87.91	-$35.94	-$26.54
60	$139.09	$82.05	-$57.04	$98.48	$54.42	-$44.06	-$37.83
70	$89.73	$24.58	-$65.15	$71.35	$18.51	-$52.84	-$49.06
80	$48.29	-$22.98	-$71.27	$42.70	-$19.41	-$62.11	-$60.20
90	$13.08	-$62.92	-$76.00	$12.75	-$59.01	-$71.76	-$71.27

i < 26.66%/year:	NPV{Y-X} > 0;	EUW[{Y}-{X}] < 0	EUW{Y-X} > 0
i > 26.66%/year:	NPV{Y-X} < 0;	EUW[{Y}-{X}] < 0	EUW{Y-X} < 0

Comparisons of alternatives X and Y in columns (2), (3) and (4) of Table 4.3.2 show NPV{Y} > NPV{X} or NPV{Y-X} > 0 discounting at i < 26.66%/year, and NPV{Y} < NPV{X} or NPV{Y-X} < 0 discounting at i > 26.66%/year. When discounting at i < 26.66%/year, the NPV results are not consistent with those of the EUW criterion in columns (5), (6) and (7) which are based on the individual (A/P,i,4) and (A/P,i,5) factors of X and Y respectively. However, the NPV results in Table 4.3.2 are all consistent with the EUW results in column (8) which is based on the (A/P,i,5) factor of the cash flow difference Y-X.

Only cash flow differences can determine the best mutually exclusive alternative. Therefore, an incremental or marginal cash flow analysis is needed to compare alternatives with equal or unequal lifespans. Consequently, NPV{Y-X} = NPV{Y}-NPV{X} was compared to the zero NPV of do-nothing alternatives. Similarly, EUW{Y-X} = EUW{Y}- EUW{X} was compared to the zero EUW of do-nothing alternatives in column (8) using a common (A/P,i,5) factor based on the lifespan of Y-X. The EUW measurements in columns (5), (6) and (7) annualize the NPVs of X and Y over 4 and 5-year periods which may be repeated indefinitely to keep individual EUW values constant. However, this changes the question of whether X or Y is better once to whether X or Y is better forever.

In the proposed method of determining the best mutually exclusive alternative, X may be chosen as the defender and Y as the challenger. Assuming the cost of borrowing money is 10%/year, we may determine the ($\Delta C,\Delta R$) coordinates of X(400,1037.67) and Y(500,1209.21) from the data of Tables 4.3.1 and 4.3.2 . Alternatives X and Y are compared by drawing vector difference {Y-X} from the terminal point of X to the terminal point of Y. Because X and Y are indivisible, {Y-X}(100,171.54) is not itself an alternative, but its slope $Ø_{YX} = \Delta R_{YX}/\Delta C_{YX} = 171.54/100 = 1.7154$ can be compared to the marginal capital efficiency or marginal output/input ratio $Ø_m = \Delta R_m/\Delta C_m$ of the organization as a whole. If $Ø_{YX} > Ø_m$, then Y should replace X; and if $Ø_{YX} < Ø_m$, then X should be retained.

Because NPV and EUW are equivalent ranking criteria, EUW measurements are mostly used in annualizing cash flow descriptions of alternatives with large cash flows at the start and finish of their lifespans and relatively smaller intermediate cash flow rates. If interest rates are the same from start to finish, then EUW = P(A/P,i,n) - S(A/F,i,n) where P and S represent the start and finish cash flows. But if start and finish interest rates i_P and i_S differ, then EUW = Pi_P + (P-S)(A/F, i_S,n). (see Exercise 3-8a)

Section 4.4 - Measurements of Internal Rate of Return (IRR)

The internal rate of return (IRR) of an alternative is defined in Section 2.2 as the positive discount rate which equates the present values of its input and output. Because of the focal date property (Section 3.2), IRR may also be defined as the interest rate which equates the future values of the input and output. In conventional decision-making, an alternative must have a greater IRR than MARR in order to be acceptable.

Results from IRR and NPV criteria often differ, even with single-period alternatives. For example, let A_1(-\$1.00; \$1.50) and A_2(-\$1.40; \$1.96) be two mutually exclusive alternatives. Since IRR[A_1] = 50%/year and IRR[A_2] = 40%/year, A_1 is best by the IRR criterion. But between discount rates of 0% and 15%/year, A_2 is best by the NPV criterion (i.e., NPV$[A_1]_{0\%}$ = \$0.50 < PVP$[A_2]_{0\%}$ = \$0.56). Because IRR is a rate and NPV has dollar dimensions, they cannot, in general, select the same best alternatives.

In order to get the same results from IRR and NPV criteria, the incremental internal rate of return ΔIRR[A_1-A_2] = ΔIRR[A_2-A_1] of alternatives A_1 and A_2, also called their positive breakeven discount rate, is defined by the IRR of the difference between their cash flows. The ΔIRR criterion requires the ΔIRR of an incremental investment to be greater than MARR in order to accept the more costly investment. Since ΔIRR[A_1-A_2] = ΔIRR[A_2-A_1], incremental investments are defined by ranking alternatives according to increasing initial investment costs or increasing undiscounted net present-values. If ΔIRR is greater than MARR, the more costly alternative is compared to the next ranked alternative. If ΔIRR is smaller than MARR, the less costly alternative is compared to the next ranked alternative. The process of successive elimination is repeated until the best mutually exclusive alternative remains.

The ΔIRR criterion described above can be illustrated with three single-period mutually exclusive alternatives A_1(-\$1.00; \$1.50), A_2(-\$1.40; \$1.96) and A_3(-\$2.10; \$2.73) which are ranked in order of increasing initial investment cost. We may now form Table 4.4.1 of NPV values using MARR discount rates of 9%, 14% and 20% per year. At the 9% discount rate, A_3 is the best alternative; at the 14% discount rate, A_2 is the best alternative; and at the 20% discount rate, A_1 is the best alternative. Table 4.4.1 shows that NPV and IRR criteria are compatible only at the 20% discount rate. The problem now is to show the ΔIRR and NPV criteria are compatible at the 9% and 14% MARR discount rates as well.

Table 4.4.1 - NPV, IRR and ΔIRR measurements for the best mutually exclusive alternative.

Alternative	NPV@ 9%	NPV@ 14%	NPV@ 20%	IRR	ΔIRR[A_k-A_j]
A_1(-$1.00; $1.50)	0.376	0.316	0.250*	50%	[A_1-A_0] = 50%
A_2(-$1.40; $1.96)	0.398	0.319*	0.233	40%	[A_2-A_1] = 15%
A_3(-$2.10; $2.73)	0.405*	0.295	0.175	30%	[A_3-A_2] = 10%

* indicates best alternative at given discount rate; NPV[A_0]$_{0\%}$ = 0.

<u>MARR = 9% per year</u> - A_3 is best by the NPV criterion.
ΔIRR[A_1-A_0] = 50% > MARR = 9%; retain A_1, eliminate A_0.
ΔIRR[A_2-A_1] = 15% > MARR = 9%; retain A_2, eliminate A_1.
ΔIRR[A_3-A_2] = 10% > MARR = 9%; eliminate A_2. A_3 is best by the ΔIRR criterion.

<u>MARR = 14% per year</u> - A_2 is best by the NPV criterion.
ΔIRR[A_1-A_0] = 50% > MARR = 14%; retain A_1, eliminate A_0.
ΔIRR[A_2-A_1] = 15% > MARR = 14%; retain A_2, eliminate A_1.
ΔIRR[A_3-A_2] = 10% < MARR = 14%; eliminate A_3. A_2 is best by the ΔIRR criterion.

<u>MARR = 20% per year</u> - A_1 is best by the NPV and IRR criteria.
ΔIRR[A_1-A_0] = 50% > MARR = 20%; retain A_1, eliminate A_0.
ΔIRR[A_2-A_1] = 15% < MARR = 20%; retain A_1, eliminate A_2.
ΔIRR[A_3-A_2] = 10% < MARR = 20%; eliminate A_3. A_1 is best by the ΔIRR criterion.

The IRR and NPV criteria may be incompatible with mutually exclusive alternatives of two or more periods in which case the ΔIRR method needs additional logic to resolve such problems. Table 4.4.2 illustrates the ΔIRR method with alternatives ECF{-$100;$130;$10} and LCF{-$100;$10;$160}. As explained in Appendix 2B, IRR[ECF] and IRR[LCF] can be determined by setting the future values of their cash flows equal to zero as follows:

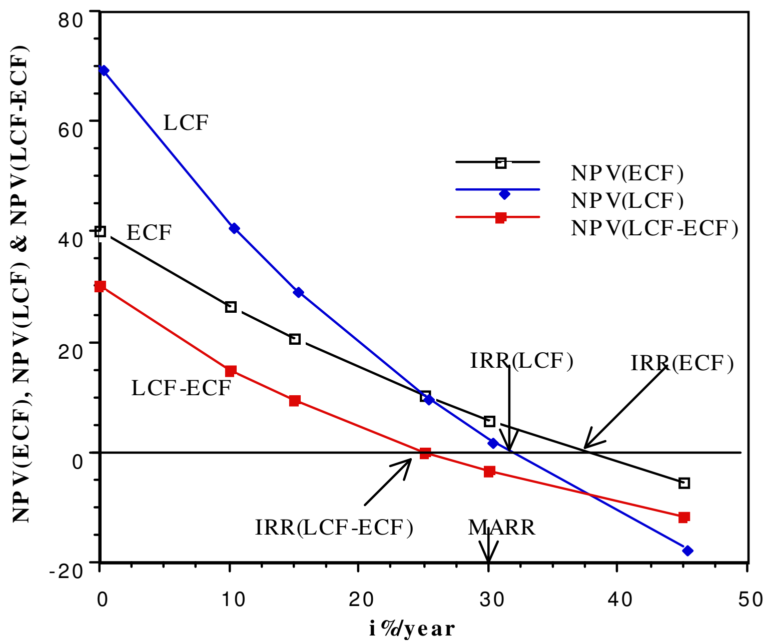
ECF: $-100x^2 + 130x + 10 = 0$; $x = 1.3728, -0.0728$; $i = 37.28\%, -107.28\%$/yr
LCF: $-100x^2 + 10x + 160 = 0$; $x = 1.3159, -1.2159$; $i = 31.59\%, -221.59\%$/yr
LCF-ECF: $-120x + 150 = 0$; $x = 1.25$; $i = 25\%$/yr.

The ECF and LCF equations above were solved for unknown x = (1+i) by the quadratic formula. The equation for LCF-ECF is linear and was solved directly for x = (1+i). Since IRR is defined only for positive roots of polynomial equations, IRR[ECF] = 37.28%/yr and IRR[LCF] = 31.59%/yr. Although IRR[ECF]=37.28% > IRR[LCF]=31.59%, Table 4.4.2 and Figure 4.4.1 show that NPV[ECF] < NPV[LCF] for all discount rates below 25%/year.

Table 4.4.2 - Comparison of NPV and IRR Criteria for the ECF and LCF Alternatives

	i%/year	NPV[ECF]	NPV[LCF]	NPV[LCF-ECF]
	0.00%	$40.00	$70.00	$30.00
	10.00%	$26.44	$41.32	$14.88
	15.00%	$20.60	$29.68	$9.08
IRR[LCF-ECF]	25.00%	$10.40	$10.40	$0.00
	30.00%	$5.92	$2.37	-$3.55
IRR[LCF]	31.59%	$4.57	<u>$0.00</u>	-$4.57
IRR[ECF]	37.28%	<u>$0.00</u>	-$7.82	-$7.82
	45.00%	-$5.50	-$17.00	-$11.50

Figure 4.4.1 - Comparison of NPV and IRR Measurements



The difficulty of ranking ECF and LCF on the basis of incremental investment costs is that both have $100 initial investment costs. Consequently, the incremental investment cost to which ΔIRR is attributable must be defined more generally by ranking the ECF and LCF alternatives according to increasing undiscounted net present-values. Since $NPV[ECF]_{0\%}$ = $40 and $NPV[LCF]_{0\%}$ = $70, the $70-$40 = $30 incremental investment cost from ECF to LCF is assigned to the following comparison of ΔIRR and MARR. It may also be required that the IRR of each mutually exclusive alternative should be greater than MARR.

<u>MARR = 15% per year</u> - $NPV[LCF]_{15\%}$\$29.68 > $NPV[ECF]_{15\%}$=\$20.60 .
LCF is best by the NPV criterion.
ΔIRR[LCF-ECF] = 25% < MARR = 15%; replace ECF, retain LCF.
LCF is best by the NPV and ΔIRR criteria.

<u>MARR = 30% per year</u> - $NPV[ECF]_{15\%}$\$5.92 > $NPV[LCF]_{15\%}$=\$2.37 .
ECF is best by the NPV and IRR criteria.
ΔIRR[LCF-ECF] = 25% < MARR = 30%; replace LCF, retain ECF.
ECF is best by the NPV, ΔIRR and IRR criteria.

Although NPV, ΔIRR, IRR and MARR measurements may be compatible, there are many cases where ΔIRR, IRR and MARR are not well defined. For example, differences in cash flows of *engineering acceleration alternatives* may have two or more positive rates of return. Moreover, cash flows of *financial credit alternatives* may have only complex or negative rates of return as shown in Appendices 4A and 4B. The <u>e</u>xternal-<u>r</u>ate-of-<u>r</u>eturn (ERR) method of handling such ambiguities is explained in Appendix 4C.

The Lorie and Savage problem described in Sections 2.3 and 4.3 is useful to study the compatibility of multiple criteria. An oil well would produce $10,000 worth of oil at the end of each year for the next two years. An engineer proposes to install a $1,600 pump which would produce all $20,000 worth of oil at the end of the first year, at which time the pump would have no salvage value. The choice here is between the ongoing alternative, Ω{$0; $10,000; $10,000}, and the pump alternative, P{-$1,600; $20,000; $0}. This problem is analyzed in Figure 4.4.2 and Table 4.4.3 below.

Figure 4.4.2 - NPV, IRR & ² IRR Measurements for Table 4.4.3

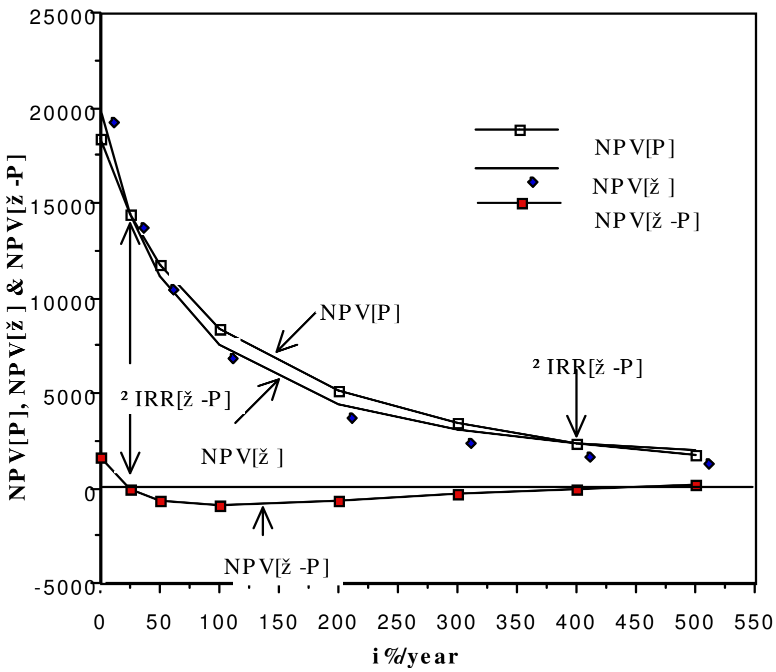


Table 4.4.3 - Comparison of NPV and IRR Measurements for the Lorie and Savage Problem

	i%/yr	NPV[P]	NPV[Ω]	NPV[Ω-P]
	0%	$18,400	$20,000	$1,600
ΔIRR[Ω-P]	25%	$14,400	$14,400	$0
	50%	$11,733	$11,111	-$622
	100%	$8,400	$7,500	-$900
	200%	$5,067	$4,444	-$623
	300%	$3,400	$3,125	-$275
ΔIRR[Ω-P]	400%	$2,400	$2,400	$0
	500%	$1,733	$1,944	$211

We first determine IRR[Ω] = ∞ and IRR[P] = 1,150%/year, each of which is greater than either MARR = 15% per year or MARR = 30% per year. However, ΔIRR[Ω-P] has two positive rates of return, 25% and 400%/year. Figure and Table 4.4.2 show that NPV[Ω-P] is negative within the 25% to 400% range, and it is positive outside that range. Consequently, ΔIRR measurements need to be modified in order to be compatible with the NPV[Ω-P] criterion.

When ΔIRR of two mutually exclusive alternatives has two positive rates of return, then Descartes' Rule of Signs indicates that the differences between the cash flows of the two alternatives have at least two changes of signs. For example, the ongoing and pump alternatives have cash flow series Ω{$0; $10,000; $10,000} and P{-$1,600; $20,000; $0}. The differences of their cash flow series is [Ω-P]{$1,600; -$10,000; $10,000} which has two changes of sign (i.e., from $1,600 to -$10,000 and from -$10,000 to $10,000).

In order to modify ΔIRR measurements from two to one positive rate of return, the cash flow differences between the two alternatives need to be reduced to only one sign change. The cash flow differences between two alternatives can always be reduced to one sign change by 'increments of borrowing' at *external rates of return* (ERR) based on the cost of borrowing money. For example, suppose we borrow $1,600 at an ERR of 15% per year, and repay $1,840 one year later from the $20,000 proceeds of P. Adding B{$1,600; -$1,840; $0} to P{-$1,600; $20,000; $0} gives [B+P]{$0; $18,160; $0}. Since IRR[Ω] = IRR[B+P] = ∞, we determine [Ω-B-P]{$0; -$8,160; $10,000} and ΔIRR[Ω-B-P] = 22.55%/year. Because MARR=15% < ΔIRR=22.55%, it follows that Ω > B+P is compatible with $NPV[\Omega]_{15\%}$=$16,257 > $NPV[B+P]_{15\%}$= $15,791. Since MARR=30% > ΔIRR=22.55%, it follows that Ω<B+P is compatible with $NPV[\Omega]_{30\%}$=$13,609 < $NPV[B+P]_{30\%}$ =$13,969.

Another problem with IRR measurements concerns *financial credit alternatives* where compensating balances are needed to establish lines of credit in banking relationships. For example, suppose a person deposits $1 in a bank now for the purpose of borrowing $3 one year later. If the bank requires a $2.50 pay back of the loan one year later, should the person accept the $3 line of credit when MARR = 30%/yr and the cost of borrowing money is 15%/yr? The negative response to this question can be readily obtained using NPV measurements based on either the 30% or 15%/year discount rates as shown below.

$$\text{NPV @ 30\%/yr:}\quad -1.00 + \frac{3.00}{1.30} - \frac{2.50}{1.30^2} = -\$0.17$$

$$\text{NPV @ 15\%/yr:}\quad -1.00 + \frac{3.00}{1.15} - \frac{2.50}{1.15^2} = -\$0.28$$

The IRR of financial credit alternative, F{-$1; $3; -$2.50}, can be determined by setting its future values at the unknown interest rate i%/yr equal to zero. (see Appendix 2B)

$$-\$1.00(1+i)^2 + \$3.00(1+i) - \$2.50 = 0$$

The solutions of this quadratic equation are (1+i) = 1.5+0.5j and 1.5-0.5j, where j denotes the imaginary unit $\sqrt{-1}$. Therefore, IRR[F] = 0.5+0.5j and 0.5-0.5j. Because IRR[F] are complex conjugates, it shows that IRR is not always measurable by constant discount rates. Moreover, the complex rates of return of IRR[F] are applicable to both borrower and lender, and they are the same from either viewpoint (i.e., IRR{-$1; $3; -$2.50} = IRR{$1; -$3; $2.50}). Appendices 4A and 4B present mathematical methods of interpreting the economic significance of complex and negative rates of return, respectively.

A major assumption of IRR and ΔIRR measurements is that mutually exclusive alternatives can be compared by a single positive rate of return and that all negative and complex conjugate rates of return lack economic significance. However, this assumption is unrealistic in a number of important cases, and it only serves to mask the existence of varying interest rates. For example, consider an investor who needs $15,500 one year later but only has $10,000 at the present time. A bank offers to pay the investor 5%/yr interest on the $10,000 deposit and to charge him 19%/yr on the outstanding balance. From the bank's viewpoint, the cash flow of this financing is F{$10,000; -$15,500; $5,950}. The IRR of this transaction can be determined by setting its future values at the unknown interest rate i%/yr equal to zero.

$$\$10{,}000(1+i)^2 - \$15{,}500(1+i) + \$5{,}950 = 0$$

The solutions of this quadratic equation are (1+i) = 0.85 and 0.70 . Consequently, we have IRR(F) = -15% and -30% per year. Tables 4.4.3 and 4.4.4 use negative rates of return, and Table 4.4.5 uses actual interest rates for positive and negative balances.

Table 4.4.3 - Constant Interest Rate -15%/yr for F{$10,000; -$15,500; $5,950}.

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	-15.00%	$0.00	$10,000.00	$10,000.00
1	-15.00%	$8,500.00	-$15,500.00	-$7,000.00
2	-15.00%	-$5,950.00	$5,950.00	$0.00

Table 4.4.4 - Constant Interest Rate -30%/yr for F{$10,000; -$15,500; $5,950}.

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	-30.00%	$0.00	$10,000.00	$10,000.00
1	-30.00%	$7,000.00	-$15,500.00	-$8,500.00
2	-30.00%	-$5,950.00	$5,950.00	$0.00

Table 4.4.5 - Actual Interest Rates -15%/yr for F{$10,000; -$15,500; $5,950}.

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	5%, 19%	$0.00	$10,000.00	$10,000.00
1	5%, 19%	$10,500.00	-$15,500.00	-$5,000.00
2	5%, 19%	-$5,950.00	$5,950.00	$0.00

Internal rate of return calculations are a major tool of breakeven analyses which can be used for averaging variable interest rates. For example, Tables 3.2.1 and 3.3.1 illustrate two ways of repaying a $1,000 loan in four years at interest rates which vary from 5% to 8% per year. The IRR of the two cash flow series of loan repayments in Tables 3.2.1 and 3.3.1 can be determined by spreadsheet calculations explained in Appendix 3C. In this manner, we determined IRR{-$1,000; $0; $0; $0; $1,286.18} = 6.4941%/year and IRR{-$1,000; $50; $60; $70; $1,080} = 6.4223%/year which are verified in Tables 4.46 and 4.47 below.

Table 4.4.6 - Constant Interest Rate Loan Repayments of Table 3.2.1

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	6.4941%	$0.00	-$1,000.00	-$1,000.00
1	6.4941%	-$1,064.94	$0.00	-$1,064.94
2	6.4941%	-$1,134.10	$0.00	-$1,134.10
3	6.4941%	-$1,207.75	$0.00	-$1,207.75
4	6.4941%	-$1,286.18	$1,286.18	$0.00

Table 4.4.7 - Constant Interest Rate Loan Repayments of Table 3.3.1

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	6.4223%	$0.00	-$1,000.00	-$1,000.00
1	6.4223%	-$1,064.22	$50.00	-$1,014.22

2	6.4223%	-$1,079.36	$60.00	-$1,019.36
3	6.4223%	-$1,084.83	$70.00	-$1,014.83
4	6.4223%	-$1,080.00	$1,080.00	$0.00

In summary, internal-rate-of-return, IRR, measurements enable the best mutually exclusive alternative in a set to be determined without specifying discount rates. The concept of incremental-internal-rate-of-return, ΔIRR, was introduced in order to obtain the same results as NPV discounted at MARR. The ΔIRR, IRR and MARR criteria are not well defined when alternatives have multiple positive, negative or complex rates of return in which cases the ambiguities can be resolved by cash flow modifications based on external-rates-of-return, ERR.

Section 4.5 - Measurements of Benefit/Cost Ratio (B/C)

Conventional economic decision-making uses measurements of benefit/cost ratio (B/C) for selecting mutually exclusive alternatives mostly in the public sector of the economy. The major reasons why public and private project alternatives are handled differently are (a) public works are not intended for profit, (b) benefits of public projects are not easily quantified in monetary terms, (c) public works are not subject to public taxation, and (d) ownership of public projects is not directly connected to the source of capital funding.

The differences outlined above do not justify using different decision-making criteria for public and private projects for the following reasons: (a) Although public works are not intended for profit, all private, public and nonprofit enterprises seek economic efficiency by minimizing input for a given output and maximizing output for a given input. (b) Although some benefits of public projects are not easily quantified in monetary terms, there are also quantifiable tangibles in need of economic decision-making. (c) Taxes do not need to be taken into account for projects that are not taxed. (d) Public projects are not only funded by public taxation, but also by borrowing in private capital markets. Since the differences between private and public projects do not have economic significance, we will only discuss how B/C differs from the NPV criterion in selecting mutually exclusive alternatives under the concepts of investment opportunity costs and do-nothing alternatives.

The interest rate used for discounting benefits and costs in public projects is defined as a "social discount rate" which represents the internal rate of return of minimally acceptable public projects. The social discount rate is conceptually similar to but smaller than the MARR of industrial firms because financial risks are often smaller and tax benefits are usually greater in public than in private projects. However, in practice, the social discount rate used for public projects frequently approximates the market rates of interest which is incurred when money is borrowed or debt is refinanced for public projects.

The benefit/cost ratio of an investment is defined as the ratio of its present-value output revenues to its present-value input costs. Since the dollar dimensions of the numerator and denominator cancel, the B/C ratio is dimensionless. Alternatives are acceptable only if their B/C ratio is greater than one. Upon subtracting one numerically from the B/C ratio, the result is (B-C)/C which is a ratio of net present-value to present-value cost. Therefore, B/C ratios greater than one are equivalent to positive net present-values.

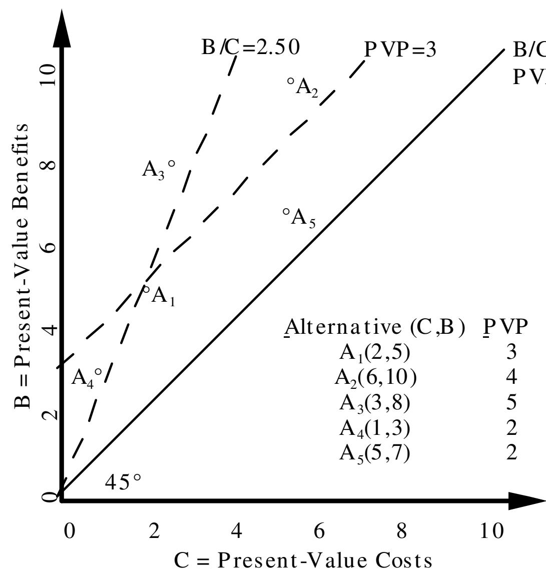
Despite similarities between B/C and NPV criteria, there are also differences. The B/C ratio is dimensionless, but NPV has dollar dimensions. Therefore, the B/C and NPV yardsticks are not commensurate. For example, let $A_1(2,5)$, $A_2(6,10)$, $A_3(3,8)$, $A_4(1,3)$ and

$A_5(5,7)$ be five mutually exclusive alternatives whose (C,B) coordinates denote their present-value costs and benefits respectively as shown in Table and Figure 4.5.1 .

Table 4.5.1 - B/C and PVP Selections of Mutually Exclusive Alternatives

Alternative i	B	C	B/C	B-C
$A_1(2,5)$	5	2	2.5	3
$A_2(6,10)$	10	6	1.67	4
$A_3(3,8)$	8	3	2.67	5
$A_4(1,3)$	3	1	3	2
$A_5(5,7)$	7	5	1.4	2

Figure 4.5.1 - B/C Ratio and NPV Selections of Mutually Exclusive Alternatives



Let us compare $A_1(2,5)$ to the four other alternatives by drawing a 45° line through A_1 so that all points above (below) the 45° line have larger (smaller) NPV than $NPV[A_1(2,5)] = B_1-C_1 = 3$. Let us draw another line from the origin through A_1 so that all points above (below) the line have larger (smaller) B/C ratios than $B/C[A_1(2,5)] = B_1/C_1 = 2.5$.

Compatibility of B/C and NPV Criteria

If we compare $A_1(2,5)$ to alternatives in the same regions of the graph as either $A_3(3,8)$ or $A_5(5,7)$, the B/C and NPV criteria would select the same alternatives. More specifically, $A_3 > A_1$ because both $B/C[A_3]=2.67 > B/C[A_1]=2.50$ and $NPV[A_3]=5 > NPV[A_1]=3$. Likewise, $A_5 < A_1$ because both $B/C[A_5]=1.40 < B/C[A_1]=2.50$ and $NPV[A_5]=2 < NPV[A_1]=3$.

Incompatibility of B/C and NPV Criteria

If we compare $A_1(2,5)$ to alternatives in the same regions of the graph as either $A_2(6,10)$ or $A_4(1,3)$, the B/C and NPV criteria would select opposite alternatives. More specifically, B/C finds $A_2 < A_1$ because $B/C[A_2]=1.67 < B/C[A_1]=2.50$, but NPV finds $A_2 > A_1$ because $NPV[A_2]=4 > NPV[A_1]=3$. Likewise, B/C finds $A_4 > A_1$ because $B/C[A_4]=3.00 > B/C[A_1]=2.50$, but NPV finds $A_4 < A_1$ because $NPV[A_4]=2 < NPV[A_1]=3$.

Compatibility of ΔB/ΔC and NPV Criteria

In order for B/C and NPV criteria to make the same selections between pairs of alternatives, the *"incremental" benefit/cost ratio* (ΔB/ΔC) is defined as $Ø_{CD} \equiv (B_C-B_D)/(C_C-C_D)$ which is the slope of the line between challenger $A_C(C_C,B_C)$ and defender $A_D(C_D,B_D)$. Assuming marginal comparison slope $Ø_m = 1$, we may use the five decision rules in Exercise **1-1a** to determine the best mutually exclusive alternative in Table 4.5.1 as follows.

Comparing challenger A_2 to defender A_1, $Ø_{21} = (10-5)/(6-2) = 5/4$ for which A_1 is replaced and A_2 is retained as a defender (Rule #1). Comparing challenger A_3 to defender A_2, $Ø_{32} = (8-10)/(3-6) = 2/3$ for which A_2 is replaced and A_3 is retained as a defender (Rule #4). Comparing challenger A_4 to defender A_3, $Ø_{43} = (3-8)/(1-3) = 5/2$ for which A_4 is replaced and A_3 is retained as a defender (Rule #2). Comparing challenger A_5 to defender A_3, $Ø_{53} = (7-8)/(5-3) = -1/2$ for which A_5 is replaced and A_3 is retained as a defender (Rule #1).

In the absence of any other challengers, A_3 is the best mutually exclusive alternative of Table 4.5.1 by both the ΔB/ΔC and NPV criteria. But this result could be obtained from the NPV criterion alone. Moreover, the marginal comparison slope $Ø_m = 1$ differs significantly from the marginal capital efficiency or output/input ratio which must be greater than one in order to satisfy the NPV objective of an economic organization subject to a capital constraint.

Section 4.6 - Measurements of Payback Period (PBP)

The *payback period* (PBP) is defined as the period of time required for undiscounted output net revenues (i.e., revenues minus expenses) to payback the undiscounted input costs of an investment before net profit is generated. The alternative which pays for its input costs soonest (i.e., the shortest PBP) ranks highest in comparison with other alternatives. Dimensionally, PBP(yr) and IRR(1/yr) are reciprocals so that minimizing PBP has the same effect as maximizing IRR.

The PBP yardstick is not a measure of profitability in the same sense as NPV, EUW, IRR and B/C ratio because PBP only measures how soon there is a return of the investment without any return on the investment. Because PBP ignores the economic consequences after the payback period, it may be incompatible with other criteria of economic decision-making. Instead, PBP measurements reflect a natural concern for the period of time needed to recoup first investment costs. The popularity of PBP as an investment criterion stems from its simplicity of calculation and intuitively logical meaning. Other reasons for PBP popularity are that its calculation is helpful in establishing lines of credit and arranging debt repayment schedules. Investors are also accustomed to using PBP measurement in the form of *price-earnings ratio (P/E)* which is defined as the ratio of the price to annual earnings per share of common stock. The P/E ratio of a stock measures the number of years required for current annual earnings per share to 'payback' the current price per share of stock.

Section 4.7 - Measurements of Average Annual Percent Profit (AAPP)

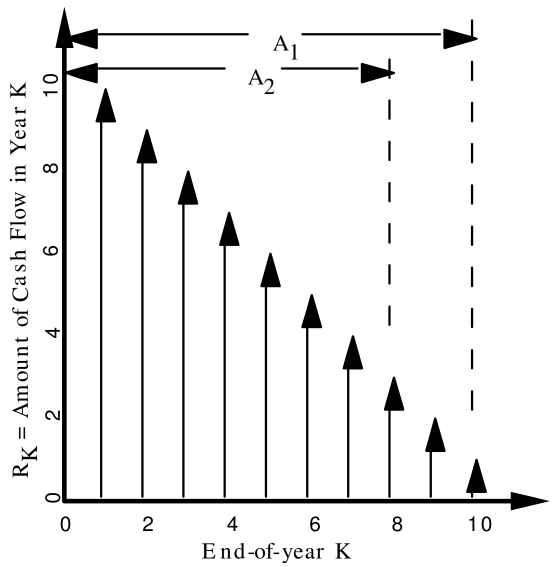
A measure of profitability used in extractive industries (i.e., mining of minerals and petroleum) is the *average annual percent profit* (AAPP) which is defined as the percentage ratio of the sum of annual profits to the sum of annual outstanding capitals. The annual profits are obtained by amortizing capital cost C of the project each year according to the fraction of the lifetime net income generated during that year. Although AAPP(1/yr) and IRR(1/yr) have the same dimensions, their relationship to NPV is quite different.

For example, let us compare two mutually exclusive alternatives A_1 and A_2, each costing \$10. Alternative A_1 has net incomes starting at \$10 at the end of the first year and declining \$1 per year for 10 years as shown in Figure 4.7.1 below. Alternative A_2 has the same net incomes but it only lasts 8 years. Obviously, NPV of A_1 is greater than that of A_2 at all positive discount rates. But AAPP measurements lead to the opposite conclusion. Before comparing A_1 and A_2, it is convenient to derive a simple expression for the AAPP of an investment with a lifespan of N years. By definition, we have

$$AAPP \equiv \frac{P_1 + P_2 + ... + P_N}{C_1 + C_2 + ... + C_N} \qquad ....(4.7.1)$$

where P_K is the amount of profit in year K, and C_K is the amount of outstanding capital during year K for K = 1, 2, ... , N. However, P_K and C_K are not directly observable quantities, except for $C_1 \equiv C$ which is the first cost of the investment. The amounts of cash flow R_K in year K are observable quantities because they represent revenues minus expenses other than depreciation or depletion. The sum of R_K from K = 1 to K = N is denoted by R. Since the original cost of the investment is C, the total profit P of the venture is equal to R-C. Therefore, the sum of P_K from K = 1 to K = N in the numerator of (4.7.1) which represents P must also be equal to R-C.

Figure 4.7.1 - Measurements of AAPP for Mutually Exclusive Alternatives



In order to determine the annual profit P_K in year K, we need to subtract from R_K the amount of capital investment cost that was depleted that year. By definition of AAPP, the amount of depletion in year K is the fraction R_K/R of the original capital cost C. Therefore,

$$P_K \equiv R_K - (R_K C/R) \qquad ....(4.7.2)$$

$$C_K \equiv C - \sum_{J=1}^{K-1} R_J C / R \qquad ....(4.7.3)$$

Equation (4.7.4) below shows that the sum of outstanding capitals C_K from K = 1 to K = N in the denominator of (4.7.1) is equal to $C\tau$ where τ denotes the center of gravity of the cash flow diagram in Figure 4.7.1 above.

$$\sum_{K=1}^{N} C_K = \sum_{K=1}^{N}\left(C - \sum_{J=1}^{K-1} R_J C / R\right) = \frac{C}{R}\sum_{K=1}^{N}\left(R - \sum_{J=1}^{K-1} R_J\right) = \frac{C}{R}\sum_{K=1}^{N}\sum_{J=K}^{N} R_J = \frac{C}{R}\sum_{J=1}^{N} JR_J \equiv C\tau \qquad ....(4.7.4)$$

We may now evaluate AAPP for A_1 and A_2 from the simple formula AAPP = $(R-C)/C\tau$.

AAPP(A_1): R = \$55, C = \$10, τ = 4.00 years, AAPP(A_1) = 112.50%/yr
AAPP(A_2): R = \$52, C = \$10, τ = 3.69 years, AAPP(A_2) = 113.75%/yr

Thus, AAPP measurements indicate A_1 is worse, instead of better than A_2. Although AAPP and NPV measurements are incompatible in this example, IRR and NPV measurements are compatible in this example because IRR(A_1) = 88.75%/yr and IRR(A_2) = 88.67%/yr.

Section 4.8 - Summary of Chapter Four

The objective of economic decision-making is to select engineering and financial alternatives in order to maximize the net present-value added (ΔNPV) to the organization as a whole subject to a capital constraint. The proposed method of satisfying this objective is based on a single criterion of the organization's marginal capital efficiency or present-value output/input ratio. Conventional methods of economic decision-making are carried out with multiple criteria for determining preferred engineering and financial alternatives. The criteria used in the proposed and conventional methods of economic decision-making have many similarities, but also important differences. Chapter Four explains both the similarities and differences between the proposed and conventional criteria with respect to answering the question, What is the best way of doing each project?

The conventional criteria for determining preferred alternatives include net present-value (NPV), equivalent uniform worth (EUW), internal rate of return (IRR), benefit/cost ratio (B/C), payback period (PBP), and average annual percent profit (AAPP). Because multiple criteria are used in selecting alternatives, the question arises whether different criteria prefer the same or different alternatives. The use of multiple criteria imply that they select different alternatives at least some of the time. When conflicts occur, multiple criteria cannot determine preferred alternatives and weights are needed to determine how much each criterion should be followed. Moreover, there is a need to investigate the differences between the preferred alternatives determined by conventional criteria and those selected by the proposed method of economic decision-making.

Net Present-Value (NPV) (Section 4.2) is the major criterion used in conventional economic decision-making to answer the question "What is the best way of doing each project?". Despite the outward similarity of NPV and the proposed criterion of selecting alternatives, conventional NPV measurements differ from those in Section 1.5 in three fundamental respects, namely, in the **(1)** definition of "do-nothing" and "ongoing" alternatives, **(2)** estimation of investment and borrowing opportunity costs, and **(3)** determination of the capital constraint.

Both *do-nothing and ongoing alternatives* are defined as doing nothing to change a current set of operating conditions. Conventional do-nothing alternatives assume that the forgone benefit of not investing in the best rejected project at the margin of an organization's capital constraint would be to invest elsewhere in a do-nothing alternative whose internal rate of return represents an *investment opportunity cost* or a minimum attractive rate of return (MARR). The NPV of each alternatives discounted at MARR must be positive because the NPV of do-nothing alternatives are zero.

Each project has a set of mutually exclusive alternatives which has one *ongoing alternative* that is defined as a continuation of the project without changing the effects of past decisions. Ongoing alternatives are the first defenders against the challenges from other mutually exclusive alternatives in their sets for satisfying the objective of the organization as a whole under a capital constraint. The proposed method of economic decision-making assumes that the forgone benefit of not undertaking cost-increasing or cost-decreasing project alternatives would show up in changes of *borrowing opportunity costs* (i.e., either decreasing debt or increasing liquidity) of the enterprise as a whole. Therefore, cash flows of all alternatives are discounted at the cost of borrowing money, and the discounted value of the cash flows in ongoing alternatives may be positive, negative or zero.

Conventional NPV measurements do not distinguish between input cost ΔC and output revenue ΔR components of net present value. Consequently, two mutually exclusive alternatives may have the same NPV measurement but vastly different capital efficiency or output/input ratios $\Delta R/\Delta C$. However, it was shown in Chapter One, Section 1.7, that these $\Delta R/\Delta C$ ratios must be compared to the marginal output/input ratio, $Ø_m$, in order to select cost-increasing and cost-decreasing alternatives which would maximize the net present-value added, ΔNPV, to the enterprise as a whole under a capital constraint.

Measurements of *equivalent uniform worth* (EUW) (Section 4.3) are used to complement NPV measurements when alternatives have unequal lifespans. The cash flows of alternatives with unequal lifspans are modified to equal lifespans with common-multiple-lifespan or common-analysis-period assumptions in order for NPV and EUW criteria to give equivalent results. However, cash flow modifications are not needed for NPV and EUW to be compatible if NPV is converted to EUW measurements by multiplication with an (A/P,i,n) factor where 'n' s the lifespan of the difference in cash flows between the two alternatives.

It is shown that EUW measurements systematically favor alternatives with shorter lifespans by annualizing their NPV over shorter periods of time with individual (A/P,i,n) factors. An example is given of mutually exclusive alternatives X and Y which have the same costs and incomes but Y lasts one year longer than X. Although NPV{Y} is greater than NPV{X} at discounts below 26.66%/year, EUW measurements show the exact opposite. Instead of answering the question of whether X or Y is better once, EUW measurements answer the question of whether X or Y is better forever. In Section 1.5, ΔNPV measurements evaluate alternatives directly with equal or unequal lifespans.

The *internal rate of return* (IRR) of an alternative (Section 4.4) is defined as the positive discount rate at which the present value of its input and output are equal and its NPV is zero. Measurements of IRR and NPV are frequently incompatible, especially when alternatives with early cash flows are compared to alternatives with larger cash flows that occur later. In order to make IRR and NPV measurements compatible, the *incremental internal rate of return* (ΔIRR) is defined as the positive discount rate which equates the NPV of the difference in cash flows between two alternatives.

Since $\Delta IRR[A_1\text{-}A_2] = \Delta IRR[A_2\text{-}A_1]$, it is necessary in applications of ΔIRR to determine whether $[A_1\text{-}A_2]$ or $[A_2\text{-}A_1]$ entail incremental investment costs. For this reason, alternatives are ranked according to either increasing initial investment costs or increasing undiscounted net present-values. If ΔIRR is greater than MARR, the less costly alternative is eliminated and the more costly alternative is compared to the next ranked alternative. If ΔIRR is smaller than MARR, the more costly alternative is eliminated and the less costly alternative is compared to the next ranked alternative. The process of successive elimination is repeated until the best mutually exclusive alternative remains.

In cases of *engineering acceleration alternatives* and *financial credit alternatives*, ΔIRR often yields multiple positive, complex or negative rates of return which can be converted into single positive internal rates of return by means of an *external rate of return* (ERR) (see Appendix 4C). Significant economic alternatives with complex and negative internal rates of return are explained in Appendices 4A and 4B.

The *benefit/cost ratio* (B/C) of an alternative (Section 4.5) is defined as the ratio of its present-value output to its present-value input. The B/C measurements are largely used to select alternatives in the public sector of the economy. The "social" discount rate used for such purposes is closer in concept to borrowing rather than investment opportunity costs. The B/C and NPV often give opposite results. To resolve this conflict, the *incremental benefit/cost ratio* (ΔB/ΔC) is defined as the ratio of the difference between the present-value output and present-value input of mutually exclusive alternatives. Measurements of ΔB/ΔC are compatible with NPV measurements providing the five decision-making rules described in Chapter One are used in conjunction with a marginal capital efficiency $Ø_m = 1$.

The *payback period* (PBP) of an investment (Section 4.6) is defined as the period of time required for output incomes (i.e., revenues minus expenses) to payback its first costs. Dimensionally, PBP(yr) and IRR(1/yr) are reciprocals so that minimizing PBP has the same effect as maximizing IRR. The PBP yardstick is not a measure of investment efficiency in the same sense as NPV, EUW, IRR and B/C ratio because PBP only measures how soon there is a return <u>of</u> the investment without regard to the return <u>on</u> the investment. Because PBP ignores the economic consequences after the payback period, it may be incompatible with other measures of investment efficiency. The popularity of the PBP criterion stems from its simplicity of calculation and intuitively logical meaning.

The *average annual percent profit* (AAPP) of an investment (Section 4.7) is defined as the percentage ratio of the sum of annual profits to the sum of annual outstanding capitals. The derivation of AAPP is based on amortizing the capital cost of an investment each year according to the fraction of the total net income produced during that year. Although AAPP(1/yr) and IRR(1/yr) have the same dimensions, they differ in their incompatibility with NPV measurements. An example is given of mutually exclusive alternatives A_1 and A_2 where $NPV[A_1] > NPV[A_2]$ and $IRR[A_1] > IRR[A_2]$ but $AAPP[A_1] < AAPP[A_2]$.

Appendix 4A - Complex Rates of Return

In Section 4.4, we presented a problem in which a person deposits $1 in a bank now for the purpose of borrowing $3 one year later. If the bank requires the person to pay back the loan with $2.50 a year later, what are the internal rates of return of these transactions? The IRR of this financial credit loan, denoted by F{-$1; $3; -$2.50}, can be found by setting the net future-value of the cash flows at the unknown interest rate r%/yr equal to zero as follows:

$$-1.00(1+r)^2 + 3.00(1+r) - 2.50 = 0 \qquad \text{....(4A.1)}$$

The solutions of equation (4A.1) are $(1+r) = 1.5+0.5j$ and $1.5-0.5j$ where j denotes the imaginary unit $\sqrt{-1}$. Therefore, $r_1 = 0.5+0.5j$ and $r_2 = 0.5-0.5j$. The complex rates of return indicate interest rates are variable rather than constant. We now present a method of showing that complex rates of return are equivalent to real varying rates of return.

Tables 4A.1 and 4A.2 below provide two cash flow accounting systems for the future time equivalences of a bank using interest rates $r_1 = 0.5+0.5j$ and $r_2 = 0.5-0.5j$ respectively.

Table 4A.1 - Cash Flow Accounting System for Future Time Equivalences of the Bank

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	0.5+0.5j	$0.00	$1.00	$1.00
1	0.5+0.5j	1.50+0.50j	-$3.00	-1.50+0.50j
2	0.5+0.5j	$-2.25+0.25j^2$	$2.50	$0.00

Table 4A.2 - Cash Flow Accounting System for Future Time Equivalences of the Bank

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	0.5-0.5j	$0.00	$1.00	$1.00
1	0.5-0.5j	1.50-0.50j	-$3.00	-1.50-0.50j
2	0.5-0.5j	$-2.25+0.25j^2$	$2.50	$0.00

Columns (3), (4) and (5) of Tables 4A.1 and 4A.2 are now added to get Table 4A.3 below.

Table 4A.3 - Cash Flow Accounting System for Future Time Equivalences of the Bank

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	1	$0.00	$2.00	$2.00
1	i_1	$3.00	-$6.00	-$3.00
2	i_2	-$5.00	$5.00	$0.00

From BCF balances of $3.00 and -$5.00, and ACF balances of $2.00 and -$3.00, we form the following two equations which can be solved for variable interest rates i_1 and i_2. The results are verified in Table 4A.4 below.

$\$2.00(1+i_1) = \$3.00;\quad i_1 = 0.50$ or 50%/yr
$-\$3.00(1+i_2) = -\$5.00;\quad i_2 = 2/3$ or 66.67%/yr

Table 4A.4 - Cash Flow Accounting System for Future Time Equivalences of the Bank

(1)EOY	(2)Interest/Year	(3)BCF Balance	(4)EOY Cash Flow	(5)ACF Balance
0	100.00%	$0.00	$1.00	$1.00
1	50.00%	$1.50	-$3.00	-$1.50
2	66.67%	-$2.50	$2.50	$0.00

Appendix 4B - Negative Rates of Return

In Appendix 4B, we will explain the meaning of negative rates of return for one and two-period financial credit loans. The internal rate of return (IRR) of an investment is defined as the *constant* discount rate which makes the net present- or future-value of its input and output equal to zero. Consider a single-period investment of \$1.00 which returns \$0.80 a year later. The present value of \$1.00 input is -\$1.00, and the present value of \$0.80 output discounted at an interest rate of -20% per year is also \$1.00 (i.e., \$0.80/(1-0.20) = \$1.00). Hence, the IRR of the \$1.00 investment is -20% per year. If the \$0.80 output was reinvested in a single-period investment with IRR = -20% per year, the output from the reinvestment would be 80% of the \$0.80 input (i.e., \$0.80*(1-0.20) = \$0.64). Reinvesting N times would result in an output of $(1\text{-}0.20)^N$ from the initial investment of \$1.

Two-period financial credit loans complicate the definition of internal rates of return because there could be two *negative* discount rates which could make the future value of input and output equal to zero. For example, consider an investor who deposits one dollar in a bank in order to borrow \$1.40 one year later. The loan is paid off with \$0.48 one year later. Thus, the bank makes a net (undiscounted) profit of \$0.08 from this loan.

In order to find the unknown interest rate 'r' which will make the net future-value of input and output equal to zero, we have to satisfy the following equation:

$$-\$1.00(1+r)^2 + \$1.40(1+r) - \$0.48 = 0 \qquad ....(4B.1)$$

Solving for (1+r) by the quadratic formula (2B.7), we get two negative solutions, namely, r_1 = -20% and r_2 = -40%. If the investor deposited \$1.40 in the preceding year under the same credit conditions (i.e., 1.40*{-1.00: 1.40; -0.48}), and the outputs were redeposited N times under the same credit conditions, then as shown in Table 4B.1 below, the amount of the deposit in the Nth period would be $(1\text{-}0.20)^N+(1\text{-}0.40)^N$ (i.e., for N=3: $(0.80)^3+(0.60)^3$ = \$0.728; for N=4: $(0.80)^4+(0.60)^4$ = \$0.5392; which behaves as a -20% rate of return for N=∞).

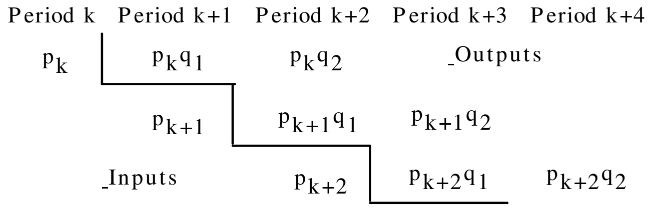
Period 1	Period 2	Period 3	Period 4	Period 5
1.40	1.96	-0.672		
	1.00	1.400	-0.4800	
		0.728	1.0192	-.34944
			0.5392	.75488
				.40544

Table 4B.1 - Reinvesting with negative rates of return.

More generally, let $i_1 \equiv 1-r_1$ and let $i_2 \equiv 1-r_2$, where r_1 and r_2 are negative rates of return between zero and -100%. Then we can construct two-period financial credit loans, denoted by $F\{-1,\ i_1+i_2,\ -i_1i_2\}$ with negative rates of return r_1 and r_2 whose deposits of the previous two outputs would be $i_1^N + i_2^N$ at the end of the Nth period of the successive loans. Assuming $i_1 > i_2$, then deposits will decline at the rate of i_1% as N increases indefinitely.

Both positive and negative IRR's of alternatives affect their reinvestment potential which is evaluated as follows. Let p_k be the input at the end of the kth year, the outputs from which are p_kq_1 and p_kq_2 at the ends of years k+1 and k+2. If outputs $p_{k+1}q_1$ and p_kq_2 equal input p_{k+2} as shown in Figure 2.2.1, then difference equation (2.2.3) is valid for all k.

Figure 2.2.1 -



$$p_{k+2} - p_{k+1}q_1 - p_kq_2 = 0 \quad \text{for all } k \qquad \text{....(2.2.3)}$$

The general solution of difference equation (2.2.3) is $p_k = C_1(X/Z)^k + C_2(Y/Z)^k$ where C_1 and C_2 are arbitrary constants which can be determined by specifying two consecutive values of p_k. Hence, let us specify $p_0 = 2$ and $p_1 = (X+Y)/Z = q_1$. Upon solving the resulting simultaneous equations of (2.2.3), we get $C_1 = C_2 = 1$. Therefore, a particular solution of (2.2.3) is

$$p_k = (X/Z)^k + (Y/Z)^k \quad \text{for all } k \qquad \text{....(2.2.4)}$$

If Z = 1, then $p_k = X^k + Y^k$ which represents the input at the end of the kth year from reinvesting the preceding two outputs from the unit-cost investment. Reinvestments p_k from p_0=2 to p_8 are listed in Table 2.2.2 below for the set of alternatives in Table 2.2.1 .

Table 2.2.2 - Reinvestment potential $p_k = X^k + Y^k$ for alternatives in Table 2.2.1 .

#	X	Y	p_0	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
A	1.2	-1.6	2	-0.4	4.00	-2.37	8.63	-8.00	19.76	-23.26	47.25
B	1.2	-1.2	2	0.0	2.88	0.00	4.15	0.00	5.97	0.00	8.60
C	1.2	-0.8	2	0.4	2.08	1.22	2.48	2.16	3.25	3.37	4.47
D	1.2	-0.4	2	0.8	1.60	1.66	2.10	2.48	2.99	3.58	4.30
E	1.2	0.0	2	1.2	1.44	1.73	2.07	2.49	2.99	3.58	4.30
F	1.2	0.4	2	1.6	1.60	1.79	2.10	2.50	2.99	3.59	4.30
G	1.2	0.8	2	2.0	2.08	2.24	2.48	2.82	3.25	3.79	4.47

Reinvestments in alternatives D, E, F and G indicate monotonically increasing inputs from p_2 to p_8 because IRR_1 = 20%/year is sufficiently greater than IRR_2 which varies from -140% to -20%/year. Reinvestments in alternatives A, B and C indicate fluctuating inputs from p_2 to p_8 because IRR_1 = 20%/year is not sufficiently large for neglecting the influence of IRR_2 which varies from -260% to -180%/year.

Moreover, the cash flow coefficients of all two-period alternatives equal the sum and product of both rates of return as shown in equation (2.2.2). More generally, multi-period alternatives have as many *constant* IRR discount rates which make ΔNPV=0 as the number of periods in their lifespans, and their cash flow coefficients consist of various combinations of sums and products of all these IRR discount rates. Consequently, cash flow forecasts of multi-period alternatives are degraded by their characterization by a single positive IRR.

Appendix 4C - External Rate of Return (ERR)

In Appendix 2B, we considered two-period investments which had two internal rates of return that were both positive, complex or negative. Since the two-period investments are typical of engineering and financial alternatives that have economic significance, they could not be compared sensibly by the IRR method to other important alternatives which have only a single positive internal rate of return. The *external rate of return* (ERR) is used to convert the cash flows of investments which do not have a single positive internal rate of return into cash flows of investments which do have a single positive internal rate of return.

The ERR method of converting cash flows is based on Descartes' Rule of Signs for polynomials of the n-th degree. Applied to cash flows, the rule is that there may be as many positive values of IRR as there are sign changes in the cash flow. Zero cash flows are ignored in counting the number of sign changes.

First example - The engineering alternative [P-Ø]{-$1,600; $10,000; -$10,000} of Section 2.3 has cash flows that exhibit two changes of sign. The cash flows correspond to the coefficients of equation (2B.7): $NFV(1+r) = a(1+r)^2 + b(1+r) + c$ where a = -$1,600; b = $10,000; and c = -$10,000. The roots of this polynomial equation are $1+r_1 = 1.25$ and $1+r_2 = 5.00$, or $r_1 = 25\%$ and $r_2 = 400\%$ per year which are the two positive internal rates of return permitted by Descartes' Rule of Signs.

Suppose we borrow the -$1,600 cash flow at an external rate of return ERR = 10% per year, and the $1,600 loan is paid back with interest (i.e., $1,600+$160 = $1,760) at the end of the first year. The cash flow is now [P-Ø+B]{$0; $8,240; -$10,000} which has only one sign change and whose internal rate of return is 1+r = 1.2136 or r = 21.36% per year.

Second example - An investor deposits $1,000 with a bank now in order to borrow $3,000 one year later. The investor pays off the loan with $2,500 one year later. . The cash flows exhibit two changes of sign, corresponding to the coefficients of equation (2B.7): $NFV(1+r) = a(1+r)^2 + b(1+r) + c$ where a = -$1,000; b = $3,000; and c = -$2,500. The roots of this polynomial equation are complex conjugates $1+r_1 = 1.5+0.5j$ and $1+r_2 = 1.5-0.5j$, or $r_1 = 0.5+0.5j$ and $r_2 = 0.5-0.5j$ per year, none of which are the two positive internal rates of return permitted by Descartes' Rule of Signs.

Suppose we borrow the -$1,000 cash flow at an external rate of return ERR = 10% per year, and the $1,000 loan is paid back with interest (i.e., $1,000+$100 = $1,100) at the end of the first year. The cash flow is now {$0; $1,900; -$2,500} which has only one sign change and whose internal rate of return is 1+r = 1.3158 or r = 31.58% per year.

Third example - An investor deposits $5,000 with a bank now in order to borrow $3,000+$4,000 = $7,000 one year later. The investor pays off the loan with ($3,000)($4,000)/($5,000) = $2,400 one year later. The cash flows exhibit two changes of sign, corresponding to the coefficients of equation (2B.7): $NFV(1+r) = a(1+r)^2 + b(1+r) + c$ where a = -$5,000; b = $7,000; and c = -$2,400. The roots of this polynomial equation are both negative $1+r_1 = 0.8$ and $1+r_2 = 0.6$, or $r_1 = -20\%$ and $r_2 = -40\%$ per year, none of which are the two positive internal rates of return permitted by Descartes' Rule of Signs.

Suppose we borrow the -$5,000 cash flow at an external rate of return ERR = 10% per year, and the $5,000 loan is paid back with interest (i.e., $5,000+$500 = $5,500) at the end of the first year. The cash flow is now {$0; $1,500; -$2,400} which has only one sign change and whose internal rate of return is 1+r = 1.60 or r = 60% per year.

Chapter Four - Exercises

4-1a An investor needs financing for a $50,000 investment one year later. A bank offers to accept a $10,000 deposit now in order to lend $50,000 one year later. The loan must then paid back to the bank with $44,000 one year later. Determine the NPV of these transactions discounted at 7%/yr, compounded semiannually, and determine two positive internal rates of return of these transactions.

4-1b Alternatively, the investor could earn 7%/yr, compounded semiannually, by buying a $10,000 certificate of deposit. One year later, the balance of the required $50,000 could be obtained as a one-year loan at an interest rate of 12.125%/yr, compounded quarterly. Determine the NPV of these transactions discounted at 7%/yr, compounded semiannually, and determine two positive internal rates of return of these transactions.

4-1c Based on the NPV criterion, should the investor select **4-1a** or **4-1b**? Based on the smaller positive IRR criterion, should the investor select **4-1a** or **4-1b**? Based on the larger positive IRR criterion, should the investor select **4-1a** or **4-1b**?

4-2a An owner of a used car worth $1,800 is making plans for the next 6 years. The car can be used 3 more years with end-of-year operating costs $900 and increasing $200 per year. After 3 years, the used car has an $800 salvage value. Then another $1,800 used car would be bought for another 3 years with the same operating costs and salvage value.

Alternatively, the owner could purchase a new car now for $6,900 with a $2,500 trade-in for the used car (i.e., $700 more than its $1,800 market value). End-of-year operating costs of the new car start at $500 per year and increase 12% per year. At the end of 6 years, the new car has a $1,000 salvage value.

Based on a 6-year planning horizon and an annual interest rate of 10.5% compounded semiannually, determine the present-value costs of the used and new car alternatives and the amount of the difference. Also, determine the equivalent uniform annual costs of the used and new car alternatives and the amount of the difference.

4-2b Prepare a table of i%/year discount versus NPV(old car) and NPV(new car) for i%/year = -10%, -5%,0%, 5%, 10% and 15%. Plot the results of the table in a line graph using a spreadsheet software program.

4-3a Three mutually exclusive alternatives, each with a 5-year lifespan, have end-of-year cash flows as follows:

End-of-year	Alternative A	Alternative B	Alternative C
0	-$1,600	-$1,630	-$1,570
1	$400	$500	$300
2	$400	$450	$350
3	$400	$400	$400
4	$400	$350	$450
5	$400	$300	$500

Using a MARR of 7.50% per year, select the alternative with the best net present-value.

4-3b Determine the benefit/cost ratio of A, B and C to select the best alternative, using MARR = 7.50%/year. Determine the incremental benefit/cost ratio $\Delta B/\Delta C$ of B-A, C-A or C-B in order to select the best alternative, using MARR = 7.50% per year.

4-3c Determine the internal rates of return IRR of A, B and C, and the incremental internal rates of return ΔIRR of B-A, C-A and C-B. Use external rate of return ERR = 7.25%/yr to convert the cash flows of B-A, C-A and C-B to a single positive ΔIRR and select the best alternative. (see Appendix 4C)

4-3d Select the alternative with the best undiscounted payback period.

Chapter Five - Which are the best projects to do?

Section 5.1 - Independent Alternatives

The objective of economic decision-making is to select engineering and financial alternatives which maximize the net present-value added (ΔNPV) to an organization as a whole subject to a capital constraint. The proposed and conventional criteria used to satisfy this objective have many similarities, but also important differences. The proposed and conventional criteria were compared in Chapter Four with respect to sets of mutually exclusive project alternatives in order to determine the best way of doing each project. Chapter Five explains both similarities and differences between the proposed and conventional criteria in regards to independent alternatives in connection with answering the second question, Which are the best projects to do?

As described in Chapter Four, the major criterion of conventional economic decision-making is the maximization of net present-value (NPV) using either MARR or WACC discount rates. The conventional NPV criterion is a measure of absolute profitability which uses zero net present-value as a reference standard. Thus, the best mutually exclusive alternative must have a greater positive NPV than other ways of doing the same project. The conventional NPV criterion is complemented by other criteria which may give equivalent selections of the best way to do each project, namely, equivalent uniform worth (EUW) based on the lifespan of cash flow differences between alternatives, incremental internal rate-of-return (ΔIRR) compared to MARR, and incremental benefit/cost ratio (ΔB/ΔC) measurements. The proposed ΔNPV criterion measures relative profitability by binary comparisons of the ratios of present-value output/input differences between mutually exclusive alternatives to the marginal capital efficiency of the organizationdiscounted at the cost of borrowing money.

In Chapter Five, conventional methods of economic decision-making are explained with respect to answering the second question, Which are the best projects to do? Before answering the second question, it is presumed the first question has already been answered. If the first question has not yet been answered, then the answer to the second question may include inferior ways of doing projects which need to be re-evaluated with respect to finding the best way of doing each project. The proposed method of economic decision-making in Section 1.7 shows that the first and second questions cannot be answered definitively without also answering the third question.

When economic organizations undertake projects because of *prior commitments*, it is still necessary to evaluate alternative ways of doing those projects in order to satisfy their objective. Conventional decision-making sometimes answers the third question "Which projects should be funded?" before answering the first two questions by *economic feasibility constraints* such as rejecting projects with negative NPV's or whose IRR's are below MARR. Conventional economic decision-making also uses NPV, EUW, ΔIRR and ΔB/ΔC criteria to answer all three questions simultaneously. However, in Chapter Five, the answers from conventional and proposed methods of analysis to the second question are considered separately from the first and third questions.

The basic difference between the first two questions is that the second question only involves *adding* alternatives of independent projects whereas the first question only involves *differences* between mutually exclusive alternatives of the same project, only one of which can be an independent alternative. Answers to both questions depend on the effectiveness of capital constraints which conventional criteria do not take into proper account.

Section 5.2 - Measurements of Net Present-Value (NPV)

In Section 4.2, conventional NPV measurements were compared to those of ΔNPV with respect to determining the best way of doing each project. The results indicated that conventional NPV measurements differ from those of ΔNPV in three major respects:

1. Definitions and applications of do-nothing and ongoing alternatives.
2. Estimation of discount rates from investment and borrowing opportunity costs.
3. Answering the question, Which are the best projects to do? separately from What is the best way of doing each project? and Which projects should be funded?.

The first two differences between NPV and ΔNPV measurements were largely explained in Chapter Four in answer to the first question, What is the best way of doing each project? The same two differences between NPV and ΔNPV measurements persist when they are used to determine independent alternatives in answer to the second question, Which are the best projects to do? However, the additivity property of independent alternatives has a cumulative effect which amplifies the first two differences between NPV and ΔNPV measurements because the optimal selection of mutually exclusive and independent alternatives depends upon the effectiveness of capital constraints.

Another important difference between NPV and ΔNPV measurements concerns the processes of ranking versus successive elimination or dynamic programming. Conventional NPV measurements are used to *rank both* mutually exclusive and independent alternatives according to their magnitudes above the zero NPV of do-nothing alternatives. But ΔNPV measurements are not used to rank the elements in a set of mutually exclusive alternatives because the *difference* between any two elements in the set is not an alternative way of doing that project. Instead, the difference between two mutually exclusive alternatives can only be compared to the input costs of other mutually exclusive or independent alternatives.

The use of conventional NPV measurements to rank both mutually exclusive and independent alternatives overlooks the need to account for the effectiveness of capital constraints. The object of measuring NPV for mutually exclusive alternatives is to determine the best alternative according to the NPV criterion. However, applying the NPV criterion by itself assumes unlimited capital. Moreover, in order to take the limitation of capital into account, the profitability of last increment of input costs must be evaluated.

In order to bring out the need for taking the effectiveness of capital constraints into account, suppose A_1(-\$1.00; \$1.50) and A_2(-\$1.40; \$1.96) are two mutually exclusive scale alternatives as presented in Section 4.4 of Chapter Four. The NPV comparisons of A and B at MARR discount rates of 10%, 15% and 20% per year are tabulated below:

$NPV[A_2]_{10\%} = \$0.3818 > NPV[A_1]_{10\%} = \$0.3636$	$NPV[A_2 - A_1]_{10\%} = \$0.0182$
$NPV[A_2]_{15\%} = \$0.3043 = NPV[A_1]_{15\%} = \$0.3043$	$NPV[A_2 - A_1]_{15\%} = \$0.0000$
$NPV[A_2]_{20\%} = \$0.2333 < NPV[A_1]_{20\%} = \$0.2500$	$NPV[A_2 - A_1]_{20\%} = -\$0.0167$

Without a predetermined discount rate, the NPV criterion cannot determine whether A or B is best. But if we specify the MARR discount rate is 10%, 15% or 20% per year, then the NPV criterion indicates A_2 is better than, equal to or worse than A_1 respectively.

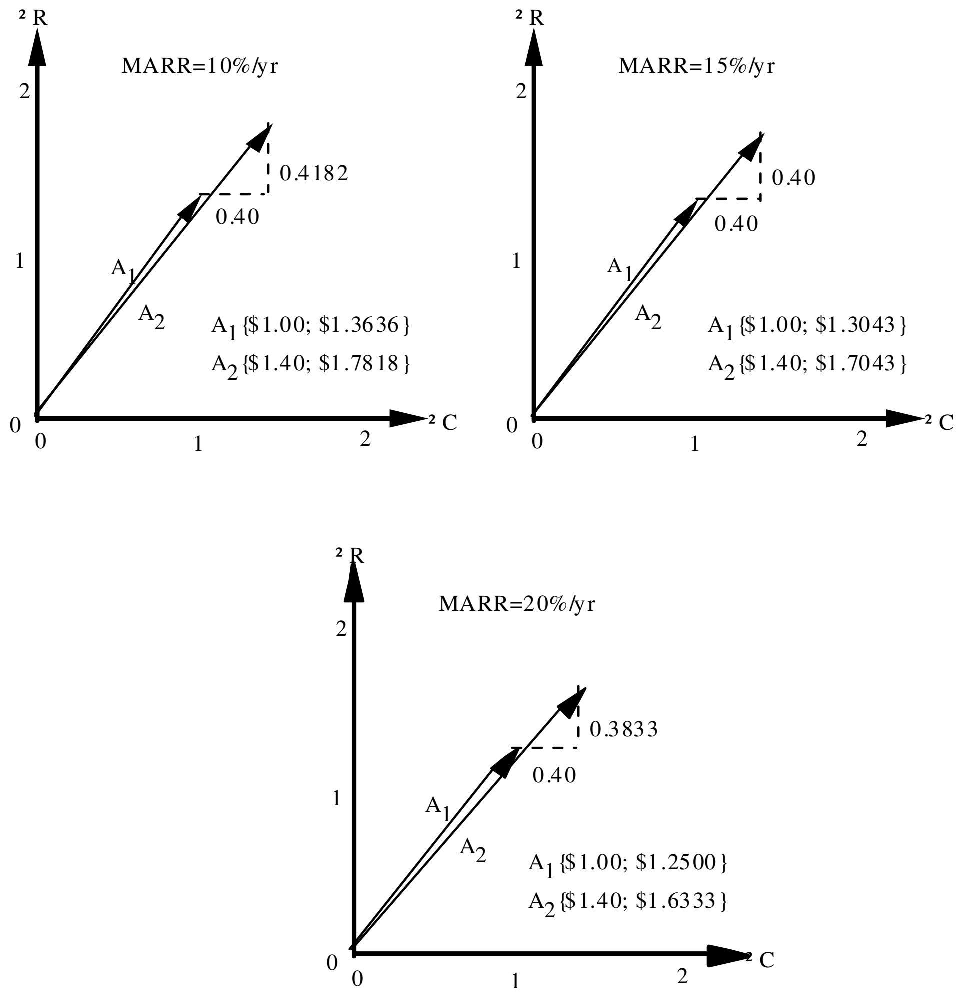
These NPV results can be compared to ΔNPV measurements by stating A_1 and A_2 in terms of their ΔC and ΔR components discounted at 10%, 15% and 20% per year. The ΔNPV analysis shown in Table and Figure 5.2.1 assume a zero-cash-flow ongoing alternative.

Table 5.2.1 - ΔNPV analysis of alternatives A_1 and A_2 at 10%, 15% and 20% discount rates.

$\Delta C[A_1]_{10\%}=\$1.00$; $\Delta R[A_1]_{10\%}=\$1.3636$; $\Delta NPV[A_1]_{10\%}=\Delta R[A_1]_{10\%}-\Delta C[A_1]_{10\%}=\$0.3636$.
$\Delta C[A_1]_{15\%}=\$1.00$; $\Delta R[A_1]_{15\%}=\$1.3043$; $\Delta NPV[A_1]_{15\%}=\Delta R[A_1]_{15\%}-\Delta C[A_1]_{15\%}=\$0.3043$.
$\Delta C[A_1]_{20\%}=\$1.00$; $\Delta R[A_1]_{20\%}=\$1.2500$; $\Delta NPV[A_1]_{20\%}=\Delta R[A_1]_{20\%}-\Delta C[A_1]_{20\%}=\$0.2500$.

$\Delta C[A_2]_{10\%}=\$1.40$; $\Delta R[A_2]_{10\%}=\$1.7818$; $\Delta NPV[A_2]_{10\%}=\Delta R[A_2]_{10\%}-\Delta C[A_2]_{10\%}=\$0.3818$.
$\Delta C[A_2]_{15\%}=\$1.40$; $\Delta R[A_2]_{15\%}=\$1.7818$; $\Delta NPV[A_2]_{15\%}=\Delta R[A_2]_{15\%}-\Delta C[A_2]_{15\%}=\$0.3043$.
$\Delta C[A_2]_{20\%}=\$1.40$; $\Delta R[A_2]_{20\%}=\$1.7818$; $\Delta NPV[A_2]_{20\%}=\Delta R[A_2]_{20\%}-\Delta C[A_2]_{20\%}=\$0.2333$.

Figure 5.2.1 - ΔNPV analysis of alternatives A_1 and A_2 at 10%, 15% and 20% discount rates.



In the absence of capital constraints, the slope of the marginal comparison vector is 45°. Consequently, since $\Delta R[A_2\text{-}A_1]_{10\%}$ = \$0.4182 is larger than $\Delta C[A_2\text{-}A_1]_{10\%}$ = \$0.40, the marginal comparison slope would first touch the terminus of A_2. This would mean that A_2 is better than A_1, and A_2 would then be treated as an independent alternative at the 10% MARR discount rate. Similarly, since $\Delta R[A_2\text{-}A_1]_{20\%}$ = \$0.3833 is smaller than $\Delta C[A_2\text{-}A_1]_{20\%}$ = \$0.40, the marginal comparison slope would first touch the terminus of A_1. This would mean that A_1 is better than A_2, and A_1 would then be treated as an independent alternative at the 20% MARR discount rate. Therefore, the ΔNPV and NPV criteria would select the same alternatives under the unrealistic assumptions that there is no distinction between do-nothing and ongoing alternatives, the same discount rates are used, and an unlimited supply of capital is available.

However, in the real world, there is neither an unlimited supply of investment opportunities at a fixed minimum rate of return nor an unlimited supply of borrowing opportunities at a fixed cost of borrowing money. Most economic organizations cannot borrow nearly as much money as they would like, and most financial institutions do not lend out nearly as much as they would like. In the final exchange equilibrium, economic organizations scale down the capital costs of their investment opportunities and maintain sufficient short-term cash balances in order to meet contingencies and limit the amount of borrowed capital needed to improve projects and start new ventures. Simultaneously, financial institutions and other creditors evaluate the abilities of potential borrowers to repay outstanding loans with interest.

Aside from differences between ΔNPV and NPV analyses with respect to discount rates and do-nothing and ongoing alternatives which have been discussed in Chapter Four, the major difference between the two approaches concerns how they treat the effectiveness of capital constraints. The object of NPV measurements is limited to studying a method of selecting the best mutually exclusive alternatives, but the results are not directly related to a specific objective of the organization as a whole. Consequently, the NPV criterion prematurely determines which mutually exclusive alternatives should become independent alternatives. More specifically, NPV calculations obscure observable cause and effect relationships between output revenues and input costs by direct subtraction of their present values without taking into account the capital constraints.

Conventional NPV measurements serve the multiple purposes of an objective, a criterion and a constraint. The objective of an economic organization is defined as the selection of mutually exclusive and independent alternatives which maximize the NPV of the organization as a whole. The best alternatives which satisfy this objective are those with the largest measure of the NPV criterion. Lastly, any alternative with a nonpositive NPV measurement can neither maximize the NPV objective nor meet the NPV constraint.

In contrast, the ΔNPV analysis distinguishes between the absolute and relative profitability of output revenues and input costs in order to determine a *marginal capital efficiency* which maximizes the net present-value added to the organization as a whole under a given capital constraint. When the slope of the vector difference between two mutually exclusive alternatives is larger (smaller) than the marginal comparison slope, then the alternative with the smaller (larger) present-value input cost is eliminated. In each set of mutually exclusive alternatives, the last element which has not been eliminated is designated as an independent alternative. Consequently, by the systematic elimination of mutually exclusive alternatives on the basis of their vector difference properties, the net present-value added by independent alternatives to an economic organization as a whole is maximized under the given capital constraint.

Section 5.3 - Measurements of Equivalent Uniform Worth (EUW)

The compatibility of EUW and NPV measurements is discussed in Section 4.3 of Chapter Four with respect to mutually exclusive project alternatives. When two alternatives have equal lifespans, EUW and NPV criteria make the same selection of which alternative is better because the same (A/P,i,n) factor multiplies the NPV of each alternative. However, when two alternatives have unequal lifespans, different (A/P,i,n) factors would apply to the NPV of each alternative. Hence, in comparisons of mutually exclusive alternatives with unequal lifespans, EUW and NPV criteria may select different ways of doing each project.

In real circumstances, mutually exclusive alternatives often have unequal lifespans. Therefore, it is difficult to understand why conventional economic decision-making requires mutually exclusive alternatives to have equivalent lifespans by common-multiple-lifespan or common-analysis-period assumptions before they can be compared. The consequences of such forced assumptions are for the EUW criterion to make the same selections of best alternatives as would be made by NPV absolute measurements of profitability, thereby making EUW and NPV criteria equivalent measures of absolute profitability.

However, the only way of determining which mutually exclusive project alternative is best is by evaluating their *cash flow differences*, usually called *incremental analyses*. The lifespan of cash flow differences between two mutually exclusive alternatives is equal to the lifespan of the alternative with the longer cash flow stream. To obtain the EUW of cash flow differences, the same (A/P,i,n) factor multiplies the NPV of the cash flow differences between the two alternatives. Consequently, EUW and NPV criteria are equivalent measures of absolute profitability whether or not mutually exclusive alternatives have equal lifespans.

Independent alternatives can hardly be expected to have equal lifespans. Therefore, one cannot make common-multiple-lifespan or common-analysis-period assumptions to compare independent alternatives in order to make EUW and NPV criteria compatible. The example shown in Tables 4.3.1 and 4.3.2 shows that when EUW measurements annualize the cash flows of individual alternatives, a short-lifespan alternative X with a small NPV can have a greater EUW than a longer-lifespan alternative Y with a larger NPV. The EUW preference for alternatives with short lifespans tends to complement the systematic bias of NPV measurements based on high MARR discount rates and unlimited capital constraints.

The proposed method of determining the best projects to do, independent alternatives may be ranked in descending order of their absolute or relative profitabilities, ΔNPV or ΔR/ΔC. However, individual rankings of independent alternatives that are funded is not important. It is the process of comparing the mutually exclusive alternatives of each project to the marginal capital efficiency or present-value output/input ratio of the organization as a whole that is fundamental for optimal economic decision-making.

For example, let us assume the cost of borrowing money is 10%/year which enables us to determine the (ΔC,ΔR) coordinates of alternatives X(400,1037.67) and Y(500,1209.21) from the data in Tables 4.3.1 and 4.3.2 . Alternatives X and Y are compared by drawing vector difference {Y-X} from the terminal point of X to the terminal point of Y. Because X and Y are indivisible, {Y-X}(100,171.54) is not itself an alternative, but its slope $Ø_{YX} = ΔR_{YX}/ΔC_{YX}$ = 171.54/100 = 1.7154 can be compared to the marginal capital efficiency or marginal output/input ratio $Ø_m = ΔR_m/ΔC_m$ of the organization as a whole. If $Ø_{YX} > Ø_m$, then Y should replace X; and if $Ø_{YX} < Ø_m$, then X should be retained.

Section 5.4 - Measurements of Internal Rate of Return (IRR)

The internal rate of return of an investment, denoted by IRR, is defined in Section 2.2 as the discount rate which equates the present values of its input and output. Because of the focal date property explained in Section 3.2, the IRR of an investment may also be defined as the interest rate which equates the future value of its input and output. Measurements of IRR are very popular because they produce percentage ranking figures that seem easy to understand and which are often useful for comparing financial alternatives. Owing to the differences of dimensions between NPV($) and IRR(1/yr) measurements, their rankings of independent alternatives cannot always be the same.

Ranking independent alternatives by IRR measurements not only determines the best projects to do, but also to determine the discount rate used to calculate other criteria used in conventional economic decision-making. A set of independent alternatives is arranged in descending order according to their *internal rates of return*. Available funds are then rationed among the alternatives according to their rank. The IRR of the alternative at the cut-off of available funds is called MARR (Minimum Attractive Rate of Return) which is the discount rate used for determining the net present-values of all alternatives.

In connection with the model of investment opportunity costs, it is especially important to understand the effects of borrowing on IRR measurements of independent alternatives. For this purpose, let us consider an investment of $1.00 which yields $1.30 one year later. If 80% of the $1.00 investment requirement is borrowed at an interest rate of 15% per year, what is the NPV and IRR of the net equity investment? A cash flow description of this problem is tabulated below.

	End-of-year 0	End-of-year 1	Internal Rate of Return
Total Investment:	-$1.00	$1.30	30%
80% Borrowing:	$0.80	-$0.92	15%
Equity Investment:	-$0.20	$0.38	90%

The equity investment of $0.20 has a profit of $0.38-$0.20 = $0.18 or a 90% rate of return on the $0.20 equity investment. Thus, 80% borrowing only reduces the net profit of the total investment from $0.30 to $0.18 while increasing its IRR from 30% to 90% per year. Similarly, the manager of every project could utilize borrowing to enhance the IRR of the project with only a minor reduction of its NPV rating. Because such financial leveraging distorts ranking procedures and uses up investment opportunities for the benefit of lenders, borrowing is rarely allowed in IRR measurements. But excluding borrowing from IRR measurements seriously affects estimates of capital constraints.

As explained in the discussion of engineering discount rates in Section 2.2, it may not always be feasible to rank independent alternatives in descending order of their IRR measurements because of multiple positive, negative and complex conjugate rates of return. However, when it is feasible to rank independent alternatives by IRR measurements, situations may arise where David Ricardo's marginal principle is violated. In brief, Ricardo's marginal principle states the necessary and sufficient condition for a group of investments to earn the greatest amount of money from a given cost of investment is that the marginal investment (i.e., the lowest ranking of the accepted investments) makes more money for its capital cost than any other investment which has not been accepted. The following example typifies situations where IRR rankings of independent alternatives are inconsistent with Ricardo's marginal principle.

Let us consider a firm which has \$60,000 available for investing in any two of the following three independent alternatives, each of which costs \$30,000:

	End-of-year 0	End-of-year 1	End-of-year 2
Alternative A:	-\$30,000	\$30,000	\$6,000
Alternative B:	-\$30,000	\$18,000	\$18,000
Alternative C:	-\$30,000	\$30,000	\$4,350

Alternatives A and C realize most of their output cash flow at the end of the first year, while B has its output cash flow evenly divided between the ends of the first and second years. The internal rate of return of A, B and C can be determined by setting the future values of their cash flows at the unknown compound factors $x \equiv 1+r$ equal to zero. The resulting quadratic equations may be solved for the unknown $x \equiv 1+r$ by means of the quadratic formula from which the unknown internal rate of return of r%/yr may be determined as follows:

A: $-30{,}000x^2 + 30{,}000x + 6{,}000 = 0$; $x = 1.17082$; IRR(A) = 17.082%/yr
B: $-30{,}000x^2 + 18{,}000x + 18{,}000 = 0$; $x = 1.13066$; IRR(B) = 13.066%/yr
C: $-30{,}000x^2 + 30{,}000x + 4{,}350 = 0$; $x = 1.12849$; IRR(C) = 12.849%/yr

The IRR ranking of the independent alternatives is A > B > C. Since \$60,000 is available, only A and B would be undertaken. Inasmuch as B is the lowest ranking of the two investments, we would expect that A+B should make more money for the \$60,000 capital cost than A+C which includes rejected investment C. We will now show these IRR measurements contradict Ricardo's marginal principle because IRR{A+C} > IRR{A+B}. Moreover, if C had been chosen instead of B as the lowest ranking of the accepted investments, the discount rate would be IRR{C} = 12.849% instead of IRR{B} = 13.066%/yr. This raises questions about determining discount rates at the cutoff of available funds since NPV{A+C} discounted at IRR{C} is greater than NPV{A+B} discounted at IRR{B}. We will now show IRR{A+C} > IRR{A+B}, and that NPV{A+C} discounted at IRR{C} is greater than NPV{A+B} discounted at IRR{B} in contradiction of Ricardo's marginal principle.

A+C: $-60{,}000x^2 + 60{,}000x + 10{,}350 = 0$; $x = 1.15000$; IRR{A+C} = 15.000%/yr
A+B: $-60{,}000x^2 + 48{,}000x + 24{,}000 = 0$; $x = 1.14833$; IRR{A+B} = 14.833%/yr

A+C: $-60{,}000 + 60{,}000(1.12849)^{-1} + 10{,}350(1.12849)^{-2} = \$1{,}295.65 = \text{NPV}\{A+C\}$
A+B: $-60{,}000 + 48{,}000(1.13066)^{-1} + 24{,}000(1.13066)^{-2} = \$1{,}226.49 = \text{NPV}\{A+B\}$

There are two general explanations of inconsistencies which arise between IRR rankings of independent alternatives and Ricardo's marginal principle. When IRR measurements are made on different alternatives, future cash flows occurring at the same time from each alternative are discounted to the present time with different discount factors. However, cash flows which occur at the same time are indistinguishable from one another, and the use of different discount factors creates an unrealistic distinction between simultaneously occurring cash flows. Because high discount factors have little effect on the present values of early incomes or expenses, IRR measurements generally favor alternatives with early cash flows as opposed to alternatives with greater amounts of future income.

Inconsistencies of IRR rankings and Ricardo's marginal principle can also be explained by the reinvestment properties of IRR measurements. For example, let A_1 denote alternative A{-\$30,000; \$30,000; \$6,000} after a single reinvestment in itself at the end of two years, thereby resulting in the following cash flow description:

A_1\{-\$30,000; \$30,000; \$(6,000-30,000); \$30,000; \$6,000\}

We can easily verify that IRR{A_1} = IRR{A} = 17.082%/yr as follows:

$$NPV\{A_1\} = -30{,}000 + \frac{30{,}000}{1.17082} - \frac{24{,}000}{1.17082^2} + \frac{30{,}000}{1.17082^3} + \frac{6{,}000}{1.17082^4} = 0$$

This example can be generalized to show that IRR measurements, just like EUW measurements, remain constant regardless of how many times the alternatives being evaluated undergo identical reinvestments at the end of their lifespans.

However, the reproductive properties of IRR measurements are more extensive than those of EUW measurements. For example, let us reinvest in only one-half of alternative A at the end of the first year. If $A_{1.5}$ denotes alternative A{-\$30,000; \$30,000; \$6,000} after half of it (i.e., {-\$15,000; \$15,000; \$3,000}) is reinvested at the end of the first year, its cash flow description would be

$A_{1.5}$ {-\$30,000; \$(30,000-15,000); \$(6,000+15,000); \$3,000}

Again, we can easily verify that IRR{$A_{1.5}$} = IRR{A_1} = IRR{A} = 17.082%/yr as follows:

$$NPV(A_{1.5}) = -30{,}000 + \frac{15{,}000}{1.17082} + \frac{21{,}000}{1.17082^2} + \frac{3{,}000}{1.17082^3} = 0$$

The phantom reinvestment properties of IRR measurements are devoid of economic significance. They are simply mathematical properties of IRR measurements rather than economic properties of any alternative. The real question is whether an independent alternative should be undertaken once, and not whether it should be undertaken as often as it can be reproduced mathematically. Independent alternatives which rank highest as a result of reproductions spawned by IRR measurements are those with early cash flows which can utilize such phantom reinvestments both sooner and more often than alternatives with later cash flows. This is the same kind of systematic bias that is exerted by NPV and EUW measurements when they are applied with high discount rates such as MARR.

Section 5.5 - Measurements of Benefit/Cost Ratio (B/C)

Measurements of benefit/cost ratio (B/C) were studied in Section 4.5 in order to compare mutually exclusive alternatives in the public sector of the economy. However, aside from governmental taxation, common distinctions between alternatives in the public and private sectors of the economy are largely superficial. We then proceeded to show that B/C and NPV comparisons of mutually exclusive alternatives are frequently incompatible, and that the 'incremental' benefit/cost ratio ($\Delta B/\Delta C$) would remedy this problem.

Different aspects of $\Delta B/\Delta C$ and NPV criteria are involved when ranking independent rather than mutually exclusive alternatives. The private sector of the economy ranks both mutually exclusive and independent alternatives by the NPV criterion, but the B/C criterion is largely used as a constraint in the public sector of the economy. The proposed ΔNPV measurements treat both private and public sector alternatives by comparing their present-value output/input ratio to the marginal capital efficiency which is greater than one. Let us begin by describing the origins of B/C ratio before comparing it to NPV, ΔNPV and $\Delta B/\Delta C$ models for selecting independent alternatives.

The principle of benefit-cost analysis for public projects was institutionalized in the United States by the Flood Control Act of 1936. The underlying idea was that measurements of the benefits of a capital investment project should be greater than those of its costs in order to justify its acceptance. Although the benefit-cost ratio is not the only measure used to compare public projects, the higher the ratio is above unity, the more favorably a project is regarded by government agencies and the United States Congress, especially with respect to planning water resource projects and transportation systems.

Because B/C and NPV rankings of independent alternatives depend on public and private policy, let us examine B/C and NPV measurements from a different viewpoint. In this regard, let us consider three sets of mutually exclusive alternatives depicted by three bundles of vectors with common initial points as shown in Figure 5.5.1 below. The horizontal component of each vector represents the difference between accepting and rejecting its present-value input cost. Similarly, the vertical component of each vector represents the difference between accepting and rejecting its present-value output benefit. Each bundle has a vector of steepest slope, and starting from the origin the steepest-slope vectors are added geometrically in descending order of their slopes to form the convex envelope shown in Figure 5.5.1 .

Figure 5.5.1 - B/C and NPV Selections of Mutually Exclusive and Independent Alternatives

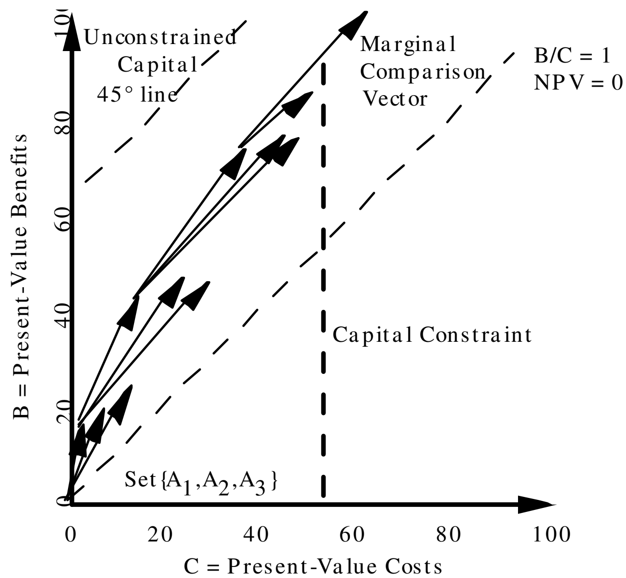


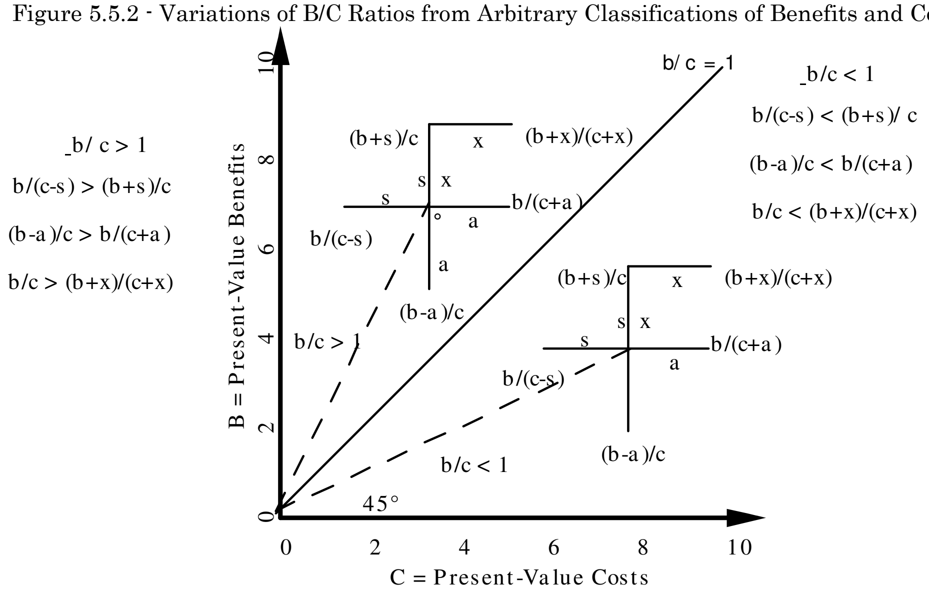
Table 5.5.1 - B/C and NPV Properties of Mutually-Exclusive-Alternative Set 'A'.

Alternative	A_1	A_2	A_3	A_2-A_1	A_3-A_2	A_3-A_1
ΔPV Costs	3.0	5.0	8.0	2.0	3.0	5.0
ΔPV Benefits	15.0	20.0	24.0	5.0	4.0	9.0
Benefit/Cost Ratios	5.0	4.0	3.0	2.5	1.33	1.8
Net Present-Values	12.0	15.0	16.0	3.0	1.0	4.0

Concerning the bundle of vectors A_1, A_2, and A_3 in Figure and Table 5.5.1, the greatest NPV of set 'A' is NPV{A_3} = 16.0 but the greatest B/C of set 'A' is B/C{A_1} = 5.0 . In order to make B/C compatible with NPV measurements, ΔB/ΔC measurements are used in conjunction with $Ø_m$=1.0 and the five decison rules in Exercise 1-1a. Thus, $Ø_{21}$=2.5 > $Ø_m$=1.0 which means defender A_1 is eliminated. Then $Ø_{32}$=2.5 > $Ø_m$=1.0 which means defender A_2 is eliminated. Therefore, NPV{A_3} = 16.0 is greatest which could also be found by translating a 45° line towards the vector bundle until it first touches A_3. But if the capital constraint does not permit all three maximum NPV alternatives to be undertaken, then the marginal comparison slope of, say $Ø_m$ = 1.5, should be used instead of $Ø_m$ = 1.0 for the marginal comparison slope. Consequently, NPV{A_2} = 15.0 would be better than PVP{A_3} = 16.0 under an effective capital constraint.

Measurements of B/C ratios sometimes have problems which do not occur with NPV measurements. These problems are often traced to the famous passage in the United States Flood Control Act of 1936 which stated that the Federal Government should participate in the improvement of navigable waters for flood control purposes if "the benefits to whomsoever they may accrue are in excess of the estimated costs,..". This phrasing was gradually adopted in evaluating many other types of federal and local public works projects both in the United States and other countries of the world. But this phrasing lead to many ambiguities in B/C ratio measurements due to arbitrary classifications of one or more items as either benefits or costs. For example, the present-value 's' of salvageable items from a project can be treated either as an addition to present-value benefits 'b' or as a reduction of present-value costs 'c' as depicted in Figure 5.5.2 below.

Figure 5.5.2 - Variations of B/C Ratios from Arbitrary Classifications of Benefits and Costs



Whether 's' was added to 'b' or subtracted from 'c', NPV would increase from b-c to (b+s)-c = b-(c-s) = b-c+s. However, when b/c > 1, then b/(c-s) > (b+s)/c > b/c; and when b/c < 1, then (b+s)/c > b/(c-s) > b/c. In particular, when secondary and tertiary benefits are included in 's', then projects with b/c < 1 would be enhanced more by adding 's' to 'b' than by subtracting 's' from 'c'. And if 's' is sufficiently large, b/c < 1 could be changed to (b+s)/c > 1.

The variation of the B/C ratio due to the choice of classifying 's' as an added benefit or a reduced cost has important practical applications in projects which have dominant secondary and tertiary benefits. But if the primary purpose of a project is to secure specific benefits rather than salvage values and other secondary and tertiary benefits, then the present-value of salvage 's' should be classified as a reduced cost rather than an added benefit because of the limited availability of input resources.

Similarly, ambiguities of the B/C ratio also occur with different classifications of abandonment costs. For example, the present-value 'a' of abandonment costs of a project can be treated either as a reduction of present-value benefits 'b' or an addition to present-value costs 'c' as depicted in Figure 5.5.2 above. Whether 'a' was added to 'c' or subtracted from 'b', the NPV would decrease from b-c to b-(c+a) = (b-a)-c = b-c-a. Although the subtraction of abandonment cost also decreases the b/c ratio, the decrease of the b/c ratio would be smaller if 'a' was subtracted from 'b' than if it was added to 'c', providing b/c > 1 (i.e., b/c > (b-a)/c > b/(c+a) when b/c > 1). But if b/c < 1, then b/c > b/(c+a) > (b-a)/c. In public projects with prolonged periods of current benefits as well as operating and maintenance costs, it is necessary according to the "Rule of Delta" to subtract costs from benefits on a current basis to obtain net benefits. The effect of this rule is the same as subtracting 'a' from 'b' rather than adding 'a' to 'c'.

Lastly, suppose the present-value benefit and cost of a public project is 'b' and 'c' respectively. As a method of ingratiating legislators with their constituents, local improvements whose present-value benefits and costs are each equal to 'x' are tacked on the original project. The NPV of the project before and after *pork barrelling* is unchanged because b-c = (b+x)-(c+x). However, unlike NPV measurements, the value of the B/C ratio would change because (b+x)/(c+x) is smaller than b/c when b/c > 1, and (b+x)/(c+x) is greater than b/c when b/c < 1. Proof:

$$\frac{b}{c} - \frac{b+x}{c+x} = \frac{cb-cb+bx-cx}{c[c+x]} = \frac{[b-c]x}{c[c+x]} \quad \text{where } c, x > 0$$

Therefore, when b/c > 1 and b-c > 0, then b/c > (b+x)/(c+x); and when b/c < 1 and b-c < 0, then b/c < (b+x)/(c+x). Thus, the incidence of pork barrelling or other increases in scale which are unrelated to the primary purpose of a project would be reflected in changes of B/C ratio measurements but not in NPV measurements. These characteristics of B/C and NPV measurements are depicted in Figure 5.5.2 above for the cases where b/c > 1 and b/c < 1.

In many projects, there are tangible and intangible components whose values are difficult to quantify in monetary terms. When this occurs, $\Delta B/\Delta C$ ratios may be less arbitrary than conventional NPV measurements. For example, suppose the benefit and cost components whose values are difficult to quantify in monetary terms are denoted by β and κ respectively. If we compare the *difference* between two alternatives which have β and κ in common, then the effects of these difficultly quantifiable terms will cancel out in their $\Delta B/\Delta C$ ratio measurements. The same effect is accomplished in ΔNPV analysis by comparing the ongoing alternative to other mutually exclusive alternatives in a set. On the other hand, the techniques described below are being developed to handle nonmonetary values of β and κ which ongoing and its mutually exclusive alternatives do not have in common.

In cases where mutually exclusive alternatives do not share nonmonetary values of β and κ in common with their ongoing alternatives, a *cost-effectiveness* approach may be used in which the effectiveness of input costs are measured in terms of realizing specific outcomes. The outcomes are measured in terms of their physical attributes rather than monetary units. If outcomes vary only with input cost alternatives, then only *differences* between outcomes of the same type need to be evaluated in monetary terms. (see H. M. Levin, "Cost-Effectiveness, A Primer", Sage Publications, California, 1983, and M. S. Thompson, "Benefit-Cost Analysis for Program Evaluation", Sage Publications, California, 1982) For evaluating project alternatives which involve costs or benefits that are hard to quantify in the absence of competitive market prices, discussions on shadow prices in linear programming are useful.

Section 5.6 - Summary of Chapter Five

In Chapter Five, conventional measures of ranking independent alternatives are analyzed with respect to answering the second question, Which are the best projects to do? If the first question was not previously answered, then the answer to the second question may include inferior ways of doing projects which need to be re-evaluated after better ways of doing each project are determined. . The proposed method of economic decision-making in Section 1.7 shows such re-evaluations could change the selections of the best projects to do.

Sometimes the third question is answered first by ruling out projects which are not *economically feasible* on the basis of commonly accepted criteria such as negative NPV or having IRR less than MARR. However, a project may become economically feasible after finding the best way of doing that project. An explanation is needed of why conventional answers to the second question may be inconsistent, and how these inconsistencies would be resolved with the proposed method of ΔNPV measurements.

In Section 5.2, we consider how conventional measurements of net present-values (NPV) differ from those of ΔNPV with respect to ranking independent alternatives. The purpose of ranking independent alternatives is to prioritize their status with respect to the competition for available funds and resources. The rank of every independent alternative is important only because of its potential contribution to the net present-value added to the system as a whole. However, conventional NPV measurements rank mutually exclusive and independent alternatives in the same way, even though such ranking assumes unlimited capital. Besides NPV ranking of independent alternatives not taking capital constraints and efficiencies into proper account, it also differs from ΔNPV analysis with respect to using MARR discount rates derived from the internal rate of return of do-nothing alternatives instead of using discount rates based on the external cost of borrowing money.

Aside from differences between ΔNPV and NPV analyses with respect to discount rates and do-nothing and ongoing alternatives which have been discussed in Chapter Four, the major difference between the two approaches concerns their treatment of the effectiveness of capital constraints. The object of NPV measurements is limited to determining preferred mutually exclusive alternatives, but the resulting preferences are not directly related to the objective of the organization as a whole. Consequently, NPV prematurely determines which mutually exclusive alternatives should become independent alternatives. More specifically, NPV calculations obscure cause and effect relationships between input costs and output revenues by direct subtraction of their present values.

In Section 5.3, measurements of equivalent uniform worth (EUW) are considered for ranking independent alternatives that have unequal lifespans. Although independent alternatives can hardly be expected to have equal lifespans, conventional economic decision-making may require alternatives to have a common-analysis-period by making unrealistic reinvestments, replacements and liquidations. As a consequence, EUW measurements with ($/year) dimensions may may appear equivalent to NPV measurements with ($) dimensions.

As shown in Section 4.3, when EUW measurements are used to annualize NPV measurements, alternatives with short lifespans are amplified more than those with longer lifespans. This exaggeration of short-term alternatives stems from the property of EUW measurements that alternatives with short lifespans are used for reinvestment more often than those with long lifespans. The EUW amplification of short-term investments is consistent with the preferences of conventional NPV measurements based on high discount rates such as MARR.

In Section 5.4, measurements of internal rate of return (IRR) are considered for ranking independent alternatives. It is shown that IRR measurements can produce internally inconsistent rankings of independent alternatives. For example, if A, B, and C are independent alternatives with equal costs and IRR(A) > IRR(B) > IRR(C), then it is possible that IRR(A+C) > IRR(A+B) which is contrary to the expectation that IRR(A+B) > IRR(A+C). This inconsistency creates an ambiguity in determining the minimum attractive rate of return (MARR) which is defined by ranking independent alternatives in descending order of their IRR's until the IRR of the best rejected independent alternative at the cut-off of available funds.

The inconsistencies in IRR rankings of independent alternatives may be explained as follows. Because the IRR's of independent alternatives differ, simultaneous future cash flows from different alternatives are discounted to the present time with different discount factors. Since high discount factors affect future incomes more than earlier incomes, IRR measurements generally favor alternatives with early cash flows as opposed to alternatives with greater amounts of future income.

Another explanation of inconsistencies in IRR rankings of independent alternatives stems from mathematical reinvestment properties of IRR measurements. It can be shown that IRR measurements, just like EUW measurements, remain constant regardless of how many times identical reinvestments are made in alternatives at the ends of their lifespans. Moreover, the IRR of an investment is unchanged even when identical reinvestments are scaled either up or down before or after the end of its lifespan. The IRR measurements of alternatives with early cash flows are thereby provided with forecasts of phantom reinvestment opportunities both sooner and more often than alternatives with later cash flows. This IRR preference for early cash flows is the same bias as NPV and EUW measurements have with high discount rates such as MARR.

Conventional economic decision-making uses *benefit/cost (B/C) analysis* (Section 5.5) only for projects in the public sector of the economy because of requirements by government agencies. Governmental use of B/C ratios is primarily concerned with benefits exceeding costs to provide a minimum level of acceptability of a project. But the higher the B/C ratio is above unity, the more favorably a project is regarded by governmental agencies. Thus, the B/C ratio is used indirectly to rank public projects. The B/C ratio is rarely used to evaluate projects in the private sector where it is thought to be a reformulation of present-value profits which is subject to undesirable variations due to arbitrary classifications of benefits and costs.

The undesirable variations of B/C ratios due to arbitrary classifications of benefits and costs is explained with respect to classifying salvage values 's', abandonment costs 'a', and unrelated marginal projects 'x' which is commonly called pork barrelling. For example, suppose the present-value benefits 'b' and present-value costs 'c' of a project have been established except for one item such as the present-value salvage 's'. It then appears as if 's' could be classified either as an added benefit (b+s) or as a reduced cost (c-s). In either case, the NPV would increase from b-c to b-c+s because (b+s)-c = b-(c-s). Although the addition of salvage value also increases the B/C ratio, the increase of the B/C ratio would be greater if 's' was subtracted from 'c' than it would be if 's' was added to 'b' when b/c > 1.

Similarly, the present value of abandonment costs 'a' could be classified either as a reduced benefit (b-a) or as an added cost (c+a). In either case, NPV would decrease from b-c to b-c-a because (b-a)-c = b-(c+a). Although the subtraction of abandonment cost also decreases the B/C ratio, the decrease of the B/C ratio would be greater if 'a' was subtracted from 'b' than it would be if 'a' was added to 'c' providing b/(c+a) > 1.

Lastly, suppose the present-value benefit and cost of a public project is 'b' and 'c' respectively. As a method of ingratiating legislators with their constituents, local improvements whose present-value benefits and costs are each equal to 'x' are tacked on the original project. The NPV of the project before and after pork barrelling is unchanged because b-c = (b+x)-(c+x). However, unlike NPV measurements, the value of the B/C ratio would change because (b+x)/(c+x) is smaller than b/c when b/c > 1, and (b+x)/(c+x) is greater than b/c when b/c < 1. Thus, the incidence of pork barrelling or other unnecessary increases in the scale of a project would be reflected in changes of B/C ratio but not NPV measurements.

Despite outward similarities of the B/C and capital efficiency ratios, their function in economic decision-making is quite different. The B/C ratio is used in the public sector of the economy primarily as an acceptance criterion for independent alternatives, and little or no provision is made for the capital constraints of an economic organization as a whole.

The ΔNPV analysis distinguishes between the present-values of input costs and output revenues in order to determine a marginal comparison slope which maximizes the net present-value added to the organization as a whole under a given capital constraint. When the slope of vector difference between two mutually exclusive alternatives is larger (smaller) than the slope of the marginal comparison vector, then the alternative with the smaller (larger) present-value input cost is eliminated. The systematic elimination of mutually exclusive alternatives on the basis of their vector differences determines independent alternatives whose slopes must also be greater than the marginal comparison slope in order to maximize the ΔNPV of the organization as a whole under a given capital constraint.

In many private and public projects, there are tangible and intangible components whose values are difficult to quantify in monetary terms. When this occurs, measurements of ΔB/ΔC ratios may be less arbitrary than conventional NPV measurements. For example, suppose the benefit and cost components whose values are difficult to quantify in monetary terms are denoted by β and κ respectively. If we compare the *difference* between two alternatives which have β and κ in common, then the effects of these difficultly quantifiable terms will cancel out in their ΔB/ΔC ratio measurements. The same effect is accomplished in ΔNPV analysis by subtracting the ongoing alternative from each mutually exclusive alternative in a set. The techniques described below are being developed to handle nonmonetary values of β and κ which are not common to ongoing and mutually exclusive alternatives.

In cases where independent alternatives do not share nonmonetary values of β and κ in common with their ongoing alternatives, a *cost-effectiveness* approach may be used in which input costs are measured in terms of realizing specific outcomes. The outcomes are measured in terms of their physical attributes rather than monetary units. If outcomes vary only with input cost alternatives, then only *differences* between outcomes of the same type need to be evaluated in monetary terms. (see H. M. Levin, "Cost-Effectiveness, A Primer", Sage Publications, California, 1983, and M. S. Thompson, "Benefit-Cost Analysis for Program Evaluation", Sage Publications, California, 1982) For evaluating project alternatives which involve costs or benefits that are hard to quantify in the absence of competitive market prices, discussions on shadow prices in linear programming are useful.

Chapter Five - Exercises

5-1a A firm has \$60,000 available for investing in the following three independent alternatives X, Y and Z. Each alternative costs \$30,000 and yields a continuous cash flow rate for $0 < t \leq 20$ years as follows:

$X(t) = \$15{,}000e^{-0.1t}$; $Y(t) = \$5{,}000$; $Z(t) = \$10{,}000e^{-0.2t}$.

Use the constant interest rate tables with continuous cash flows and compounding for ranking X, Y and Z according to the payback period criterion undiscounted.

5-1b Rank X, Y and Z according to the NPV criterion discounted continuously at a 10% nominal annual interest rate.

5-1c Rank X, Y and Z according their individual rates of return (see Example 4 of Appendix 3C).

5-1d Rank X+Y, X+Z and Y+Z according to their overall rate of return (see Example 4 of Appendix 3C).

5-2a In order to control a persistent flooding problem, two plans are proposed:

	Initial Cost (MM\$)	Annual Benefits (MM\$/yr)	Annual Costs (MM\$/yr)
Plan A	25.000	2.250	0.315
Plan B	40.000	3.370	0.280

Each plan is estimated to have a lifespan of 50 years with no expected salvage value or abandonment cost. Determine the preference for plans A or B using the undiscounted payback period criterion calculated to an accuracy of at least three decimal places.

5-2b Using a discount rate of 6.8% per year, which plan would have a greater net present value?

5-2c Each plan has an internal rate of return between 7% and 8% per year. Which plan would be preferred by the internal rate of return calculated to an accuracy of at least three decimal places?

5-2d Treat plans A and B as mutually exclusive alternatives, and determine which plan would be preferred by the incremental benefit/cost ratio criterion using a discount rate of 6.8% per year calculated to an accuracy of at least three decimal places. Show the preference is unchanged whether annual costs reduce annual benefits or annual costs increase the initial cost.

5-2e Treat the annual costs of each plan as a reduction of its annual benefits, and assume each plan is an independent alternative. Determine which plan would be preferred by the benefit/cost ratio criterion using a discount rate of 6.8% per year calculated to an accuracy of at least three decimal places.

5-2f Same as 5-2e, except that the annual costs are added to the initial cost instead of subtracting annual costs from annual benefits.

Note: Additive analyses are essential for comparing independent alternatives. Incremental analyses are only applicable to comparisons of mutually exclusive alternatives.

Chapter Six - Which projects should be funded?

Section 6.1 - Capital-Budgeting Formulation and Solution

The proposed objective of economic decision-making is to select current engineering and financial alternatives in various parts of an economic organization in order to maximize the net present-value added (ΔNPV) to the organization as a whole under given capital constraints. Conventional economic decision-making uses multiple criteria for selecting the preferred engineering and financial alternatives. Chapter Six further explains both the similarities and differences between the proposed and conventional methods of answering the last question, Which projects should be funded?

The answer to the third question is the decision itself. All that has gone before in the decision-making process has led us to this step. The decision is not automatic even though the answers to the previous questions may have indicated the order of preferences. The limited funds that an economic organization has available for capital investments need to be rationed among project alternatives for the purposes of optimizing an objective of the organization as a whole. The selection of preferred alternatives under budgetary constraints is referred to as the *capital-budgeting* problem.

Formulation of the problem - Conventional economic decision-making formulates the capital budgeting problem as selecting projects which maximize present-value profits discounted at MARR for a given capital constraint of the enterprise as a whole. The MARR discount rate is determined internally from the investment opportunity cost of the enterprise estimated either from the internal rate of return of the best rejected project (i.e., the "do-nothing" alternative) or from a weighted average cost of capital WACC. The alternatives are treated independently of one another which presupposes the question "What is the best way of doing each project?" has already been answered by applying conventional criteria such as NPV, EUW, IRR and B/C as well as ΔIRR and ΔB/ΔC which make IRR and B/C measurements compatible with NPV. It is assumed that independent alternatives are indivisible and nonoverlapping, and that each project is continued once it is started.

In the proposed method of deciding which projects should be funded, the problem is formulated as one of maximizing net present-value discounted at the cost of borrowing money for a given capital constraint of the enterprise as a whole. The discount rate is determined externally from the borrowing opportunity cost of the enterprise. Mutually exclusive and independent alternatives are treated together because the questions "What is the best way of doing each project?" and "Which are the best projects to do?" are dependent on each other as well as the capital constraint. It is assumed that mutually exclusive alternatives are indivisible and they overlap with the ongoing alternative which expresses the continuation of each project.

Solution algorithm - Because formulations of the capital-budgeting problem are usually quite involved, a "solution algorithm" is generally needed for the solution. An algorithm is merely a step-by-step recipe that guarantees reaching the correct solution to a complicated problem. Owing to the indivisibilities of mutually exclusive and independent alternatives, the capital-budgeting problem is an integer programming problem whose solution time is much greater than comparable linear programming problems whose decision variables can vary continuously. The solution algorithm in Section 1.7 will permit a rapid scanning of the capital-budgeting problem within the range of the planned capital cost without exhaustively examining the profitability of every possible combination.

Section 6.2 - Capital-Budgeting of Independent Alternatives

The capital-budgeting problem rests on the assumption that only limited funds are available for all projects that an economic organization could undertake. Consequently, there is a problem of selecting from a number of available projects those that will maximize the net return on the capital invested and whose cost is within limits of funds available for investment. The projects considered in Section 6.2 are assumed to be independent alternatives whose indivisibilities force their funding to be a zero-or-one binary choice.

The general method of budgeting indivisible alternatives which are either mutually exclusive or independent is called zero-one *integer programming*. However, conventional applications of integer programming for budgeting mutually exclusive alternatives are based on different assumptions than those used for independent alternatives. More specifically, the differences of underlying assumptions used in conventional capital budgeting for mutually exclusive and independent alternatives concern the treatments of different economic service lives and capital costs.

Independent alternatives rarely have equal economic service lives. The only relationship that exists between independent projects of an economic organization is their joint appearance in the same capital budget. But this is not sufficient reason for requiring shorter-lived independent projects to have their lives equalized to those of longer-lived independent alternatives by cycles of reinvestment. Therefore, if the economic service lives of independent alternatives are unequal, conventional methods of economic analysis do not require their economic service lives to be equalized.

However, in order to make "fair" comparisons in a given set of mutually exclusive alternatives whose economic service lives are unequal, conventional economic analysis assumes that their economic service lives should be equalized so that each alternative could provide the same service. Consequently, conventional economic analysis equalizes all economic service lives in each set of mutually exclusive alternatives by the methods described under Measurements of Equivalent Uniform Worth (EUW) in Section 4.3 of Chapter Four. This writer regards that such equalizations of economic service lives are unrealistic because the economic motivation for different service lives amongst mutually exclusive alternatives overrides the need for mathematical equality.

For example, suppose an investor is planning to build either a two- or three-story apartment house. In a mathematical sense, two or three stories *are not* mutually exclusive alternatives because three-stories overlaps and already includes the two-story alternative. But in an economic sense, two or three stories *are* mutually exclusive alternatives because the investor plans to build either two- or three-stories, but not both. In order to justify the greater cost and time to build three- rather than two stories, we need to determine whether the present-value benefit of the third story is worth more than its extra present-value cost. There is no need to equalize the economic service lives of mutually exclusive alternatives by cycles of reinvestment in order to make "fair" economic comparisons between them.

Conventional economic decision-making assumes mutually exclusive alternatives must be equalized both in economic services and service lives before a proper economic comparison can be made. However, competing alternatives usually increase input costs in order to obtain greater outputs of their economic services and/or service lives. It is the function of marginal economic analysis to determine if the extra benefit is worth the extra cost without distorting the margins of economic comparison by introducing extraneous equalizing investments.

In essence, the distinction between mutually exclusive and independent alternatives is not so much their indivisibilities, as it is their scale alternatives which need to selected by marginal economic analysis under given capital constraints in order to maximize the net present-value added to an economic organization as a whole. To fix ideas on the similarities and differences between conventional and proposed methods of capital budgeting, let us first consider an example of four indivisible independent alternatives as listed in Table 6.2.1 .

Table 6.2.1 - Four Independent Alternatives Subject to Capital Budgeting

Project	ΔR	ΔC	ΔNPV=ΔR-ΔC	ΔR/ΔC
A	30	10	20	3.00
B	15	6	9	2.50
C	36	16	20	2.25
D	44	22	22	2.00

If the independent alternatives were ranked in descending order of the net present-value (ΔNPV=ΔR-ΔC), the results would be D, A or C, and B. However, ΔNPV does not measure the return per dollar of invested capital. Therefore, let us rank the independent alternatives in descending order of their the capital efficiency ratios ΔR/ΔC so that the addition of their vectors would form a convex polygon as described in Section 1.7 . The ΔR/ΔC ranking results in A, B, C and D as shown in the last column of Table 6.2.1 .

Because the four independent alternatives are indivisible, there are 2^4-1 = 15 different capital budget possibilities consisting of distinct combinations of one, two, three or four projects as listed in Table 6.2.2 below. It is worth mentioning that each project represents a vector sum in the (ΔC,ΔR) coordinates of Figure 1.7.1, and that the ABCD vector sum forms a convex polygon by successive additions of those vectors in descending order of their capital efficiency ratios. The underlined combinations of projects represent vector sums of successive combinations of projects in the convex polygon.

Table 6.2.2 - Capital Budgeting of Independent Alternatives Under Capital Constraints

Projects	ΔR	ΔC	ΔNPV=ΔR-ΔC	Ø=ΔR/ΔC
A	30	10	20	3.000
B	15	6	9	2.500
C	36	16	20	2.250
D	44	22	22	2.000
AB	45	16	29	2.813
AC	66	26	40	2.538
AD	74	32	42	2.313
BC	51	22	29	2.318
BD	59	28	31	2.107
CD	80	38	42	2.105
ABC	81	32	49	2.531
ABD	89	38	51	2.342
ACD	110	48	62	2.292
BCD	95	44	51	2.159
ABCD	125	54	71	2.315

The project combinations A, AB, ABC and ABCD are defined as *optimal capital budgets* for the capital constraints of 10, 16, 32 and 54 respectively because there are no project combinations with the same or smaller capital cost that could have as large a total net present-value added. For example, AB has $\Delta C = 16$ and $\Delta NPV = 29$, and project combinations A, B, and C each have $\Delta C \leq 16$ and $\Delta NPV \leq 20$. As a further example, ABC has $\Delta C = 32$ and $\Delta NPV = 49$, and project combinations AC, AD, BC and BD each have $\Delta C \leq 32$ and $\Delta NPV \leq 42$. The proposed integer programming method determines only these four optimal capital budgets without evaluating all 15 project combination possibilities.

On the other hand, conventional integer programming solutions require all 15 capital budget possibilities to be analyzed when arbitrary capital budget constraints are specified. For example, if an arbitrary capital constraint $\Delta C = 25$ is specified, then the maximum $\Delta NPV = 29$ could be obtained either from BC or AB. However, since BC has $\Delta C = 22$ and AB has $\Delta C = 16$, the project combination AB with $\Delta C = 16$ is defined as the optimal capital budget for the arbitrary capital constraint $\Delta C = 25$. Because integer programming solutions lie on a razors edge for indivisible independent alternatives, further considerations of the indivisibilities of independent alternatives and the effectiveness of capital constraints are discussed in the next two sections of Chapter Six.

Section 6.3 - Capital-Budgeting of Mutually Exclusive Alternatives

Even if there are only a moderate number of independent projects and mutually exclusive alternatives for modifying or initiating new projects, the problem of finding the combination of best mutually exclusive alternatives with the greatest net present-value added within a planned range of capital costs is so large that a method of mathematical programming is required for its simplification. For example, if there are only two mutually exclusive alternatives to each one of a hundred nonoverlapping projects, then there exists $(3)^{100} \approx 51.5 \cdot (10)^{45}$ budget choices, since each project may be funded in either one of two ways, or not at all. The programming method which is detailed in Appendix 1B and proved to give a correct solution in Section 1.7, permits a rapid scanning of the net present-value added within the range of the planned capital cost, without exhaustively examining the profitabilities of every possible combination.

The characteristics of the proposed solution algorithm for capital-budgeting problems provide major simplifications and advantages for engineering and financial management. After a number of iterations, the final marginal comparison vector is determined at the given capital constraint with a positive slope that is greater than one for cost-increasing alternatives, and less than one for cost-decreasing alternatives. The slope of the final marginal comparison vectors represent the minimum capital-efficiency or maximum capital-deficiency ratios that are applicable to the vector bundle of each set of mutually exclusive alternatives for projects in all parts of an economic organization. These individual project optimizations are then budgeted as independent alternatives as described in Section 6.2 . Each individual project optimization does not depend on the capital constraint. It is only the cumulation of individual project optimizations that depend on the capital constraint.

The proposed and conventional methods of solving the capital budgeting problem differ in two major respects, namely, marginal economic analysis and effectiveness of capital constraints. In order to maximize the net present-value added for given capital constraints, the proposed solution uses marginal comparison vectors which treat mutually exclusive and independent alternatives in all parts of an enterprise equally. In contrast, conventional solutions use NPV, EUW, ΔIRR, and ΔB/ΔC measurements in pairwise comparisons to determine the best mutually exclusive alternative of each project. Moreover, ΔIRR and ΔB/ΔC measurements are primarily designed to make IRR and B/C ratios compatible with NPV measurements. This use of incremental analysis to justify the use of multiple criteria tends to distort economic comparisons of marginal revenues to marginal costs.

As previously mentioned, conventional optimization of individual project alternatives is insensitive to capital constraints. The choice between alternatives can then be made independently of the sources of funds to be used. However, the capital constraint largely determines the future of each project. Therefore, in order to make optimal use of available capital, it is essential for engineering and financial managment to understand how capital-budgeting decisions are affected by the sources of capital financing.

Section 6.4 - Equity and Debt Sources of Capital Financing

The framework for rationing funds and resources among competing mutually exclusive and independent alternatives relative to the sources of available capital is called capital budgeting. The preparation of a capital budget is usually carried out in a financial accounting environment as described in Chapter Seven. Financial accounting provides a set of periodic balance sheets and income statements that result from a systematic recording of past events. However, managerial accounting should be used for capital budgeting because budget figures are forecasts of the future rather than historical data on past events and all project alternatives need to be evaluated before an optimal capital budget can be finalized.

Financial accounting classifies debt and equity as two distinct external sources of capital. *Debt financing* comes from private and institutional investors and corporate-issued bonds where creditors have explicit guarantees that borrowed funds will be repaid with a specified interest rate within a given time interval. Yields on corporate bonds reflect the market rate of interest as well as the company's credit worthiness in capital markets.

Equity funding comes from stockholders of corporations and owners of private businesses. Because equity holders have no guarantee of seeing their money again, they seek high future returns and capital gains as a reward for their risk. Since dividend yields and future stock prices may have unpredictable variations, the implicit cost of equity capital is very difficult to determine.

The distinction between debt and equity capital has economic significance because of differences in their risks and rewards. Debt capital has relatively little risks and smaller fixed rewards, whereas equity capital has more risks compensated with greater variety of rewards. By gaining access to both debt and equity capital, the overall cost of capital can be lowered. Up to a limit, financial leveraging derived from debt capital can potentially increase the market value of equity capital without unduly increasing its risks. The rationale for discounting cash flows of 'average risk profile' projects with a weighted-average-cost-of-capital (WACC) is based on the assumption that an optimal financing mix of debt and equity capital exists which maximizes the market value of equity capital.

Conventional economic decision-making selects independent alternatives in descending order of either IRR, or NPV discounted at the weighted-average-cost-of-capital (WACC) until the cut-off point of available equity and debt capital. The discount rate is a crucial factor in determining the cut-off point of available capital. For example, suppose the best rejected project has IRR = 20% per year and NPV = 0, and the firm's cost of borrowing money is 10% per year. There is considerable room for questioning why this project was rejected because equity investors could earn 20% per year on their money as well as the differential of 10% per year on the money which could be borrowed to undertake the project.

For this reason, it is commonly thought that WACC discount rates should be tailored only for projects whose risk profile simulates that of the organization as a whole. If the cash flow of a project is certain, then it is argued that the NPV of the project should be discounted at the risk-free interest rate of the cost of borrowing money. Similarly, if a project is characterized by very high risk, then a higher discount rate than WACC should be employed. But these arguments fail to differentiate *financial risks* of varying market rates of interest from *technological risks* of input costs and *marketing risks* of output revenues.

In contrast, the proposed method of economic decision-making would consider funding a project whose IRR = 20% per year providing the ΔNPV of the project discounted at the 10% per year cost of borrowing money was positive and the ΔR/ΔC output/input ratio of the project was greater than the marginal capital efficiency of the firm. The comparison of ΔR/ΔC to the firm's marginal capital efficiency satisfies Ricardo's marginal principle, namely, the necessary and sufficient condition for a group of investments to earn the greatest amount of money from a given cost of investment is that the marginal investment (i.e., the lowest ranking of the accepted investments) makes more money for its capital cost than any other investment which has not been accepted.

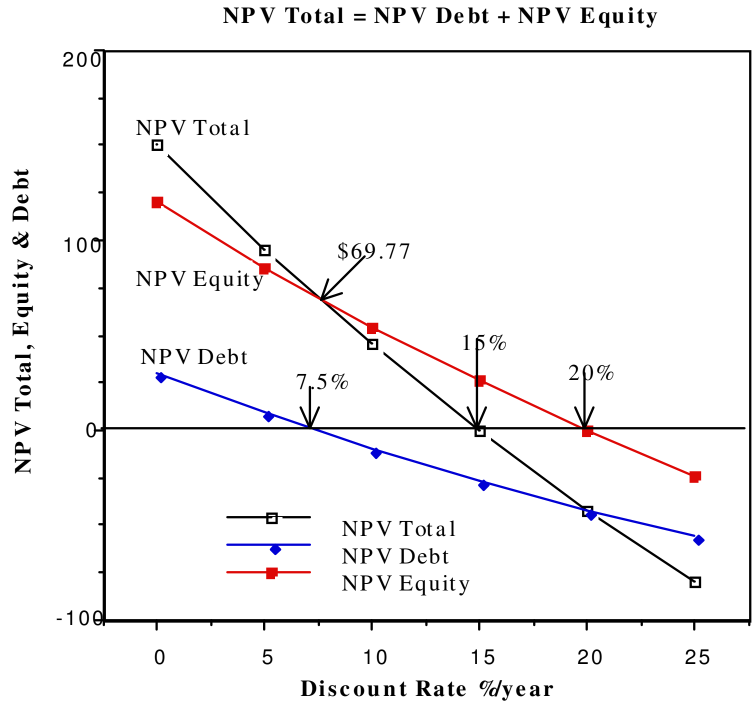
It is commonly believed that discount rates based on investment opportunity costs will insure equity investors of higher returns on their investment. The following example traces the origins of such concepts. An independent alternative costs $1,000 and yields $1,150 one year later. Only $600 of equity capital is available and $400 could be borrowed at 7.5%/year interest as outlined in the table below. The net present values (NPV) of the debt and equity portions of the total investment are calculated as a function of the discount rates in Table 6.4.1 and the results are plotted in Figure 6.4.1 below.

Investment	EOY-0 Cash Flow	EOY-1 Cash Flow	Rate of Return
100% Total	-$1,000	+$1,150	15.0%/year
40% Debt	-$400	+$430	7.5%/year
60% Equity	-$600	+$720	20.0%/year

Table 6.4.1 - Net present-values (NPV) of total, 40% debt and 60% equity investments.

Discount Rate	NPV Total	=	NPV Debt	+	NPV Equity
0.0%/yr	150.00		30.00		120.00
5.0%/yr	95.24		9.52		85.72
7.5%/yr	69.77		0.00		69.77
10.0%/yr	45.45		-9.09		54.54
15.0%/yr	0.00		-26.09		26.09
20.0%/yr	-41.67		-41.67		0.00
25.0%/yr	-80.00		-56.00		-24.00

Figure 6.4.1 - Net present-values (NPV) of total, 40% debt and 60% equity investments.



Let us first examine the consequences of discounting with MARR = 20%/year. Then, according to conventional economic decision-making, the total investment of $1,000 would not be undertaken because it has a net present-value of -$41.67 at the 20%/year discount rate as shown by the point of intersection of the *debt* and *total* NPV curves in Figure 6.4.1. Therefore, the $1,000 total investment would not be undertaken because its 15%/year internal rate of return and its -$41.67 net present-value both fall below the cut-off point of the capital budget.

It is not obvious that the total investment is unprofitable as implied in conventional decision-making. At zero discount, a $600 equity investment yields a profit of $120 one year later, representing a 20% rate of return on the equity investment, equivalent to MARR. Discounting at the 7.5%/year cost of borrowing money, the net present value is $69.77 for both the total investment of $1,000 and the equity investment of $600 as shown by the intersection of the *equity* and *total* NPV curves in Figure 6.4.1.

The reason why the total investment of $1,000 would not be undertaken in conventional decision-making is the opportunity cost of investing $600 in a 15% rate of return project is a foregone benefit of not being able to invest $600 in another project with a rate of return of 20% per year or better. However, investing $600 in the 15% rate of return project on hand is not necessarily mutually exclusive with investing $600 in another unidentified project with a rate of return of 20% per year or better. We cannot rule out the possibility that the bank would be willing to lend $1,000 instead of just $400 if the 15% rate-of-return project was used as collateral. Consequently, the opportunity cost of the $600 equity investment would vanish in the face of this $600 additional borrowing opportunity.

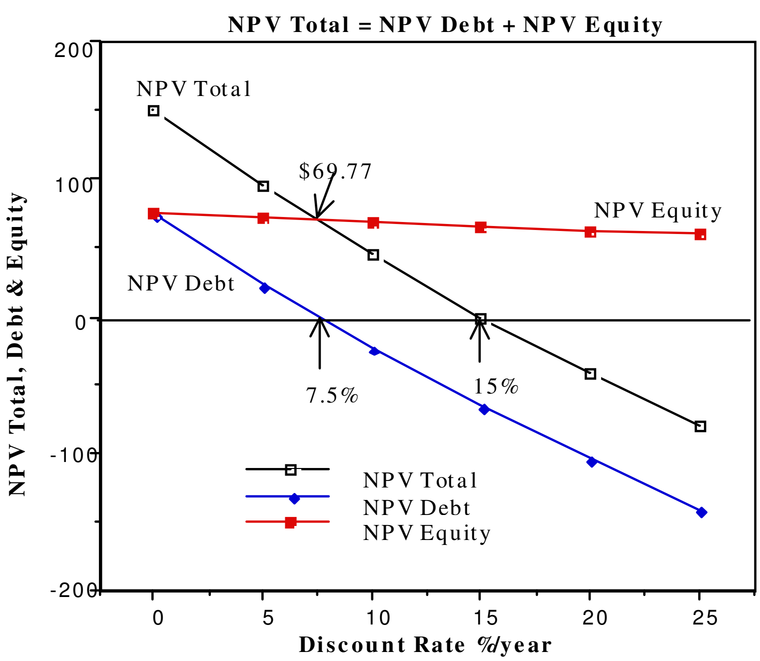
More specifically, suppose the entire $1,000 of the total investment was borrowed at an interest rate of 7.5%/year. The results of undertaking this venture would then be as described below and evaluated in Table and Figure 6.4.2 .

Investment	EOY-0 Cash Flow	EOY-1 Cash Flow	Rate of Return
100% Total	-$1,000	+$1,150	15.0%/year
100% Debt	-$1,000	+$1,075	7.5%/year
0% Equity	$0	+$75	∞%/year

Table 6.4.2 - Net present-values (NPV) of total, 100% debt and 0% equity investments.

Discount Rate	NPV Total	-	NPV Debt	=	NPV Equity
0.0%/yr	150.00		75.00		75.00
5.0%/yr	95.24		23.81		71.43
7.5%/yr	69.77		0.00		69.77
10.0%/yr	45.45		-22.73		68.18
15.0%/yr	0.00		-65.22		65.22
20.0%/yr	-41.67		-104.17		62.50
25.0%/yr	-80.00		-140.00		60.00

Figure 6.4.2 - Net present-values (NPV) of total, 100% debt and 0% equity investments.



Let us now examine the consequences of discounting with MARR = 20%/year. According to conventional economic decision-making, the total $1,000 investment has a net present-value of -$41.67 and an internal rate of return of 15% per year as shown in Table and Figure 6.4.2 . Therefore, the total $1,000 investment would not be undertaken because its -$41.67 net present-value and its 15%/year internal rate of return both fall below the cut-off point of the capital budget. Although the $1,000 debt would be paid off with $1,075 at the end of the first year, its net present-value is -$104.17 at the 20%/year discount rate. Therefore, ignoring the benefits that the 7.5% borrowing opportunity provides to equity investors, the $1,000 total investment would not be undertaken because its 15%/year internal rate of return is below MARR and its net present-value at the MARR discount rate is negative.

However, as a consequence of borrowing the total $1,000 capital requirement, equity investors would have zero cost and an infinite rate of return. At a discount rate of 20%/year, the net present value of the equity investment would be $62.50 which satisfies the present-value-profits criterion. Hence, if the equity capital appeared in the capital budget without the debt investment, it would definitely be undertaken.

The difference in acceptability between the equity and total investments is attributable to the high value of MARR which exaggerates the benefits of substituting debt for equity in the total investment, commonly called *leveraging*. When the $1,000 total investment was borrowed, the rate of return of the equity investment became infinite. This gives the false appearance that borrowing money is more profitable than using your own money. Consequently, borrowing opportunities are usually not included in capital budgets.

The economic artifacts attributable to borrowing money can be eliminated by discounting at the cost of borrowing money. For example, if the $400 capital requirement of the total investment was borrowed at 7.5%/year as in Table 6.4.1, the undiscounted profit of the equity investment would be $120 which has a present value of $69.77. Similarly, if the $1,000 capital requirement of the total investment was borrowed at 7.5%/year as in Table 6.4.2, the undiscounted profit of the equity investment would be $75 which also has a present value of $69.77. The present-value of both the total and equity investments discounted at 7.5%/year is always $69.77 as shown by the intersection of the NPV Total and NPV Equity curves in Figures 6.4.1 and 6.4.2 because the NPV Debt curve has a zero net present-value discounted at 7.5%/year regardless of how much capital was borrowed.

Discounting at the cost of borrowing money to obtain net present-values has important practical advantages over discounting with MARR. In particular, the cost of borrowing money can be readily estimated directly from market rates of interest as compared to the difficulty of estimating MARR either from the IRR of the best rejected project or from the weighted average cost of capital. Secondly, indirect borrowing alternatives such as leasing can be evaluated with engineering alternatives in a set of mutually exclusive alternatives because the net present-value of each alternative discounted at the cost of borrowing money is independent of the amount that was borrowed or the source of capital funding. Finally, as described in Chapter Seven, the cost of borrowing money is an actual cost which can be posted in the accounting system and capital budgets both before and after taxes in the same manner as all other expenses. Thus, discounting with the cost of borrowing money would be compatible with the accounting system and enable borrowing to be included in capital budgets.

Section 6.5 - Summary of Chapter Six

Chapter Six continues to explain both the similarities and differences between the proposed and conventional criteria and objectives in regards to answering the third and final question, Which projects should be funded? The answer to the third question is the decision itself. All that has gone before in the decision-making process has led us to this step. The decision is not automatic even though the answers to the previous questions may have indicated the order of preferences. The limited funds that an economic organization has available for capital investments need to be rationed among project alternatives for the purposes of optimizing an objective of the organization as a whole. The selection of preferred alternatives under budgetary constraints is referred to as the *capital-budgeting* problem.

Formulation of the problem - Conventional economic decision-making formulates the capital budgeting problem as selecting projects which maximize present-value profits discounted at MARR for a given capital constraint of the enterprise as a whole. The MARR discount rate is determined internally from models of investment opportunity costs. The alternatives are usually treated as independent of one another which presupposes the best mutually exclusive alternatives have already been selected by applying NPV, EUW, IRR and B/C criteria as well as ΔIRR and ΔB/ΔC measurements which make IRR and B/C compatible with NPV. It is assumed that independent alternatives are indivisible and nonoverlapping, and that each project is continued once it is started.

The proposed method of formulating the capital budgeting problem is to select projects which maximize net present-value discounted at the cost of borrowing money for a given capital constraint of the enterprise as a whole. The discount rate is determined externally from the borrowing opportunity cost of the enterprise. Mutually exclusive and independent alternatives are treated together because their selection is dependent on each other as well as the capital constraint. Mutually exclusive alternatives are assumed to be indivisible and overlapping with the ongoing alternative which expresses the continuation of each project from past decisions.

Solution algorithm - Because proposed and conventional formulations of the capital-budgeting problem are usually involved, a "solution algorithm" is needed for the solution. An algorithm is a step-by-step recipe that guarantees reaching the correct solution to a complicated problem. Owing to the indivisibilities of mutually exclusive and independent alternatives, the capital-budgeting problem is an integer programming problem whose solution time is much greater than comparable linear programming problems whose decision variables can vary continuously. The solution algorithm in Section 1.7 permits a rapid scanning of the capital-budgeting problem within the range of the planned capital cost without exhaustively examining the profitability of every possible combination.

Section 6.2 deals with capital budgeting independent alternatives which are assumed to be indivisible and nonoverlapping. The general method used in capital budgeting indivisible alternatives which are either mutually exclusive or independent is called *integer programming*. However, unlike the proposed integer programming, conventional applications of integer programming for mutually exclusive alternatives are based on different assumptions than those used for independent alternatives. The differences of integer-programming assumptions for mutually exclusive and independent alternatives concern differences in their economic service lives and capital costs.

Independent alternatives rarely have equal economic service lives and conventional methods of economic analysis do not equalize their economic service lives. However, in order to make "fair" comparisons of mutually exclusive alternatives whose economic service lives are unequal, conventional economic analysis assumes that their economic service lives should be equalized so that each alternative could provide the same service over the same period of time. Such equalizations of economic service lives are unrealistic because the economic motivation for different service lives amongst mutually exclusive alternatives overrides the need for mathematical equality.

For example, consider the alternatives of building either a two- or three-story apartment house at a particular site. In a mathematical sense, two or three stories *are not* mutually exclusive alternatives because three-stories overlaps and already includes the two-story alternative. But in an economic sense, two or three stories *are* mutually exclusive alternatives because if a decision is made to build the apartment house, either two- or three-stories will be built, but not both. It is the function of marginal economic analysis to determine if the extra benefit of the third story is worth its extra cost without distorting the margins of economic comparison by introducing extraneous equalizing investments.

In essence, the distinction between mutually exclusive and independent alternatives is not so much their indivisibilities, as it is their scale alternatives which need to selected by marginal economic analysis under given capital constraints in order to maximize the net present-value added to an economic organization as a whole. To fix ideas on the similarities and differences between conventional and proposed methods of capital budgeting by integer programming, an example is given of four indivisible independent alternatives which represent new projects with differing service lives and capital costs. These four independent alternatives provide $2^4-1 = 15$ distinct capital budget possibilities. But only four of these possibilities are *optimal capital budgets* in the proposed integer programming method because there are no project combinations with the same or smaller capital cost that could have as large a total net present-value added. On the other hand, conventional integer programming solutions require all 15 capital budget possibilities to be analyzed when arbitrary capital budget constraints are specified.

The characteristics of the proposed solution algorithm for capital-budgeting problems provide major simplifications and advantages for engineering and financial management. After a number of iterations, the final marginal comparison vector is determined at the given capital constraint with a positive slope that is greater than one for cost-increasing alternatives, and less than one for cost-decreasing alternatives. The slope of the final marginal comparison vectors represent the minimum capital-efficiency or maximum capital-deficiency ratios that are applicable to the vector bundle of each set of mutually exclusive alternatives for projects in all parts of an economic organization. These individual project optimizations are then budgeted as independent alternatives as described in Section 6.2 . Each individual project optimization does not depend on the capital constraint. It is only the cumulation of individual project optimizations that depend on the capital constraint.

Further considerations of the indivisibilities of mutually exclusive and independent alternatives and the effectiveness of capital constraints are discussed in the next two sections of Chapter Six. Section 6.3 deals with integer programming of mutually exclusive alternatives. The integer programming method which is proposed here will permit a rapid scanning of the net present-value added within the range of the planned capital cost, without exhaustively examining the profitabilities of every possible combination.

This proposed solution of capital-budgeting mutually exclusive alternatives differs from conventional methods of solution in two major respects, namely, marginal economic analysis and effectiveness of capital constraints. In order to to maximize the net present-value added for given capital constraints, the proposed solution uses marginal comparison vectors which treat mutually exclusive and independent alternatives in all parts of an enterprise equally. In contrast, conventional solutions use NPV, EUW, ΔIRR, and ΔB/ΔC measurements in pairwise comparisons to determine the best mutually exclusive alternative of each project. Moreover, ΔIRR and ΔB/ΔC measurements are primarily designed to make IRR and B/C ratios compatible with NPV measurements. This use of incremental analysis to justify the use of multiple criteria tends to distort economic comparisons of marginal revenues to marginal costs.

Section 6.4 considers the effects of equity and debt sources of capital financing on capital budgeting. Conventional economic decision-making selects independent alternatives in descending order of either IRR or NPV discounted at the weighted average cost of capital. The projects are then undertaken in descending order of their rank until the cut-off point of either MARR, or NPV = 0, or available debt and equity capital. Debt financing enables an enterprise to defer payment for project costs until a later time. Therefore, it should be included with equity funding in determining the capital constraint. However, debt financing is based on the cost of borrowing money which is much less than MARR. Consequently, when high MARR rates are used to discount the deferred payments of borrowing, it makes borrowing appear much less expensive than actual out-of-pocket costs. For this reason, all project evaluations usually exclude debt financing when discounting at MARR.

Purchasing versus leasing alternatives can be sensibly evaluated by discounting at the costs of borrowing money which are readily observable either by sample inquiries or public quotations of market rates of interest. The net present-value of an alternative and its ΔR/ΔC output/input ratio could be evaluated independently of the amount borrowed or the source of capital funding. Finally, as described in Chapter Seven, the cost of borrowing money is an actual cost which fits in the accounting system and capital budgets both before and after taxes just like all other expenses.

Chapter Seven - Cash Flow Accounting and Income Taxes

Section 7.1 - Balance Sheets

Financial accounting is the language of the business world. The work of financial accountants is tailored for *external* reports to parties outside an enterprise such as bankers, creditors, investors and government agencies. Financial accountants provide periodic reports in a format of balance sheets and income statements on which the investment community and government agencies can rely. Because the outside world depends heavily upon the accuracy of an organization's financial statements, they are usually audited by independent Certified Public Accountants (CPA) according to "generally accepted accounting principles" established by the Financial Accounting Standards Board (FASB) of the United States.

Financial accounting is the indispensable language of business because it has been adapted over the centuries to serve the purposes of auditing, paying taxes and timely reporting and communication. However, the information contained in balance sheets and income statements are poorly suited for the cash flow descriptions that are needed in economic decision-making. Financial accountants aggregate causally unrelated revenues and expenses into accounting periods that do not fully span the lifetimes of different projects. Consequently, revenues of one accounting period may be caused by expenses of previous accounting periods, and current expenses may be planned for revenues in later accounting periods. For auditing purposes, all revenues and expenses which occur at the beginning or end of an accounting period are recorded on an *as-is basis* without taking the time value of money into account. Because revenues and expenses in the same accounting period usually lack causal connections, financial accountants employ a *matching principle* which states that expenses should be deducted against directly related revenues. As a result, adjustments are required at the ends of accounting period to appropriately reflect the revenue earned and the expenses incurred in an accounting period.

Financial accounting recognizes revenues and expenses either on a *cash basis* or an *accrual basis* of accounting. The cash basis of accounting recognizes revenues and expenses only when the cash related to those revenues and expenses are received and disbursed. Individuals and professionals generally use the cash basis of what they earned in a given time period because the accounting is relatively simple and taxes are payable only on revenues that were actually received. But when large inventories exist, the cash basis of accounting is inappropriate. To keep track of price fluctuations, output is first sold into inventory and then it is bought out of inventory to be sold to final consumers at a later time. Most corporations use the accrual basis of accounting in which revenues are recognized when they are *earned* rather than when they are received, and expenses are recorded when they are *incurred* rather than when they are paid for. This creates a difficulty in the cash flow descriptions of alternatives for economic decision-making.

Depreciation, in a financial accounting sense, is the systematic allocation of the initial cost of an asset against the income derived from its use during its lifetime. If an asset lasts one year or less, its initial cost is deductible as an ordinary expense for reducing taxable income. But if the lifetime of an asset is greater than one year, then its initial cost must be expensed over time according to the depreciation allowances permitted by income tax laws. Because income taxes are major expenses of private enterprises, they must be included in the cash flow descriptions of alternatives for economic decision-making. In this regard, it should be noted that financial accounting methods of reasonable book write-offs of depreciation based on the collective experience of the enterprise as a whole may differ appreciably from actual income tax accounting at the project level.

One of the most important differences between *financial accounting external* reports to outside parties and *managerial accounting internal* reports to managers who plan the future operations of the enterprise concerns the treatment of interest expense at the project level. The managerial definition of cash flow for equity capital expenditures of a project is the actual capital cost plus the interest that the enterprise would have earned on those capital expenditures if less money was borrowed and the project was not accepted. Therefore, the opportunity cost of not accepting a project is defined as the firm's cost of borrowing money. However, financial accountants only take into account the investment opportunity cost of not accepting a project in evaluating project profitability.

Financial accounting deals with past happenings and managerial accounting plans for future events. Nevertheless, the *results* of economic decision-making must finally enter into the balance sheets and income statements of financial accounting for outside communication. But the *processes* of economic decision-making require engineering and financial alternatives and constraints to be defined in a managerial accounting language that enables one to optimize current budget decisions with a cause-and-effect analysis. In order to achieve a better understanding between financial and managerial accounting, it is essential to be familiar with basic financial accounting practices of describing the cash flow of alternatives in the balance sheets and income statements of economic organizations.

Financial accounting provides monetary records of business enterprises in the private sector of the economy on the basis of which income taxes are determined. A basic document of financial accounting is the *Balance Sheet* which is a statement of the financial position of a firm at a given point of time. It balances the firm's assets with claims against those assets according to the equation

$$\text{Assets} = \text{Liabilities} + \text{Owners' Equity} \qquad \text{....(7.1.1)}$$

where "assets" are things of monetary value that the system possesses, and "liabilities" and "owners' equity" tell us who supplied those resources to the business and how much was supplied by either creditors or owners. The assets of the Balance Sheet always equals the total of liabilities and owners' equity because everything that a business owns has been supplied either by creditors or owners. Assets which a business uses but does not own such as rented or leased equipment and buildings are not recorded directly on balance sheets. Instead, rented or leased assets are recorded as *contingency liabilities* in footnotes to the balance sheet.

It is worth mentioning that assets and liabilities on balance sheets of banks and financial intermediaries have a different meaning than those of industrial companies and business enterprises. Loans to investors are treated as bank assets because they are the primary source of bank revenues. The funds of depositors are recorded as bank liabilities. The difference between assets and liabilities is the bank owners' equity which is usually very small in comparison to the bank's liabilities. Therefore, in general, assets are sources of revenues for the system, and liabilities are the sources of revenues that were not supplied by owners.

Assets are usually listed in descending order of liquidity on either the left side or top of the Balance Sheet. The first assets to be listed are current assets consisting of cash, marketable securities, and accounts receivable. The term cash refers to readily available money on hand or on demand deposit at banks where they generally earn low rates of interest. Marketable securities earn higher rates of interest than cash held in banks, but they must be sold, transferred or used as loan collateral for transaction purposes. Accounts receivable are amounts owed to the firm by customers who have not yet paid for past purchases.

Current assets also include inventories. In descending order of liquidity, inventories consist of finished products available for sale, work in progress, and raw materials. The volume of finished products is largely affected by the accounts receivable, and work in progress and raw materials tend to correlate with volume of production. When current assets are reported net of inventories they are called quick assets. A miscellaneous category of current assets is prepaid expenses such as rent and insurance.

Noncurrent assets are largely fixed assets such as land, buildings, machinery, and equipment. Both the original book value of fixed assets and the cumulative depreciation are listed, followed by the current book value of those assets net of depreciation. Other noncurrent assets consist of intangibles such as trademarks, copyrights, patents, and goodwill which are expected to generate revenues in the future. Land, precious metals and some intangible assets are not subject to depreciation because they are considered to have unlimited useful lives. The market value of assets with limited useful lives may either appreciate or depreciate, but they are always depreciated in business accounting because their original cost can be charged off as an expense to reduce taxable income.

Liabilities are outstanding debts owed to creditors for the supply of business assets. Just like assets, liabilities are classified as current or noncurrent in descending order of urgency or priority, and they are usually listed on either the right side or bottom of the Balance Sheet. Current liabilities consist of accounts payable, salaries payable, taxes payable, dividends payable, short-term notes payable, interest payable, and current maturities.

The difference between current assets and current liabilities is one definition of working capital, which is widely used as a measure of a firm's liquidity. A better measure of liquidity than working capital is the current ratio, which is defined as the ratio of current assets to current liabilities. In many industries, current ratios less than 2 are often coupled with difficulties of meeting short-term obligations, because current assets include accounts receivable and inventories which are slowly convertible into cash.

Noncurrent liabilities consist of long-term debt. Long-term notes payable such as mortgages are documents signed by both borrower and lender and secured by real property. Other long-term notes payable such as bonds, debentures, and convertible debentures are nonsecured debt. Long-term notes payable offer greater security than less formal instruments of current liabilities such as accounts receivable that are based on purchase orders of the buyer and invoices of the seller. Frequently, long-term notes payable such as mortgages may be stated "less current portion" which are listed under current liabilities. Long-term leases are stated as contingent liabilities which may be found as footnotes to the Balance Sheet, because businesses have use and possession of leased assets without having any ownership.

Owners' Equity (or Net Worth) is the difference between assets and liabilities, and it usually appears below the Liabilities on either the right side or bottom of the Balance Sheet. Net worth is the portion of the business assets that was supplied by owners, and it comes last in seniority of repayment. The assets and liabilities sections of the balance sheet are essentially the same for corporations as for single proprietorships and partnerships of two or more persons who, as co-owners, carry on a business for profit. However, significant differences appear in the owners' equity section of the balance sheet for corporations as compared to single proprietorships and partnerships.

The most common reason why single proprietorships seek to become partnerships is the lack of sufficient capital to begin or expand a business. Partnerships afford a means of combining the capital and abilities of two or more persons. A partnership can be formed without any legal formalities. When individuals agree to become partners, a partnership is created. A partnership agreement is voluntary. No one can be forced into a partnership or be forced to continue as a partner. Each partner can act as an agent of the partnership, and the partnership is bound by the acts of any partner as long as these acts are within the scope of normal operations. Consequently, each partner is personally responsible for all the debts of the partnership. Because of unlimited partner liability, wealthy persons often prefer participating in corporations rather than entering into partnerships.

The net worth of corporate balance sheets is called Stockholders' Equity, which consists of Capital Stock and Retained Earnings. Capital originally paid in by stockholders is permanent capital not subject to withdrawal. The amount of Capital Stock equals the number of issued shares of stock multiplied by an arbitrary par value per share. The unit of corporate ownership is a share of stock which represents the share of net income to be distributed to stockholders as dividends. Retained Earnings are the accumulation of net incomes that have not been distributed to the stockholders as dividends since the inception of the business. If the company had been unprofitable and had incurred losses since its organization, the amount of Capital Stock would remain fixed but the Retained Earnings would be reported as a Deficit.

The owner's equity section of the balance sheet is especially important in a partnership because the net income is divided among two or more owners. Partnership accounting requires a separate capital account to be maintained for each partner as well as a separate drawing account for each partner. Another distinctive feature of partnership accounting is that each year's net income or loss is divided among the partners in the proportions specified by the Partnership Agreement. Consequently, annual net income or loss flows through partnerships as individual owners without any accumulation in the form of Retained Earnings.

The differences in the owners' equity section of the balance sheet between corporations and single proprietorships or partnerships are significant from the standpoint of income taxes. A corporation's profit is taxed, and if the profit remaining after paying corporate taxes is distributed to stockholders in the form of dividends, then the stockholders must declare those dividends as income on their individual income tax returns. Thus, corporate profit is really taxed twice. Corporations are taxed in the year profits are earned, and stockholders are taxed in the year dividends are received.

Corporate income taxes could be postponed by accumulating profits in the Retained Earnings account. However, United States' tax laws impose additional taxes on corporations "improperly accumulating surplus". This pressures corporations to distribute much of their current earnings as dividends each year. Under certain conditions, corporations with 35 or fewer stockholders can elect to be taxed as partnerships. Such corporations are called Subchapter S Corporations. Ordinarily, net business incomes of partnerships and single proprietorships are subject to individual income taxes in the year they are earned whether the profits are retained in the business or withdrawn by the owners.

Section 7.2 - Income Statements

Income Statements provide the data base for levying income taxes. Financial activities during the time interval between successive balance sheets are documented periodically in the form of Income Statements (also called Profit and Loss Statements). Income Statements provide a list revenues and expenses of enterprises during the accounting period, together with the net income (and Retained Earnings of corporations) at the end of the accounting period according to the equation

$$\text{Revenues} - \text{Expenses} = \text{Net Income, or Net Profits} \quad ....(7.2.1)$$

Income Statements use single-entry bookkeeping for revenues and expenses in each accounting period which are tied in to the Balance Sheets at the beginning and end of the accounting period. This enables successive Balance Sheets to be synchronized, checked and validated. Thus, at the end of each accounting period, we can expand the Balance Sheet equation as follows:

$$\text{Assets} = \text{Liabilities} + \{\text{Beginning Ownership} + \text{Revenues} - \text{Expenses}\} \quad ....(7.2.2)$$

At the top of the income statement is *gross revenues* 'R' which may include both *operating incomes* 'R_O' and *nonoperating incomes* 'R_N' (i.e., $R = R_O + R_N$). Operating incomes R_O refer to the sales of goods and services of the firm. Nonoperating incomes R_N refer to revenues from cash deposits, marketable securities, and gains or losses from sales of assets.

From the gross revenues, we need to subtract various costs and expenses which may also be classified as *operating* and *nonoperating*. The major operating costs are referred to as *cost of goods sold* which include material and energy costs, wages of workers, and interest on debt capital. Other operating costs are called *costs of selling* which cover the cost of advertising and the salaries, commissions and expenses of sales persons. Nonoperating expenses include overhead salaries of managers and administrative personnel who are not directly engaged in production, as well as depreciation and depletion allowances.

In order to calculate income taxes, we do not need to make detailed separations between operating and nonoperating incomes and expenses. The needs of economic decision-making are largely met by financial accountants whose responsibilities for calculating income taxes (as well as auditing and external financial reporting to creditors, investors and government agencies) are carried out by using the following terminology.

The gross revenues 'R' are reduced by operating expenses 'E' (i.e., cost of goods sold and costs of selling) to obtain the *net operating income* 'NOI' = R-E. Then the net operating income has depreciation and depletion allowances 'D' (both of which are classified as nonoperating expenses) subtracted from it to yield *net income before taxes* 'NIBT', as shown in equation (7.2.3). Because the net income before taxes is the basis upon which income tax liabilities are calculated, it is often called *taxable income*.

$$NIBT = R - (E + D) = NOI - D \quad ....(7.2.3)$$

Although depreciation and depletion 'D' on the right side of equation (7.23) are often called expenses, they are not costs which need to be paid for on a cash basis. Because depreciation and depletion are not cash flows, it is useful to define the *before-tax cash flow* 'BTCF' as net income before taxes plus the depreciation and depletion allowances as shown in equation (7.2.4).

$$BTCF = NIBT + D = NOI \quad ....(7.2.4)$$

The taxable income NIBT is now multiplied by the applicable tax rate 't' in order to determine the amount of income taxes 'IncTx' to be paid, as shown in equation (7.2.5).

$$IncTx = t*NIBT = t*[NOI - D] \quad ....(7.2.5)$$

The income taxes are subtracted from NIBT to determine the net income after taxes 'NIAT', as shown in equation (7.2.6).

$$NIAT = [NOI - D] - t[NOI - D] = (1-t)*NIBT \quad ....(7.2.6)$$

When income taxes are paid, they are cash flow. Therefore, the *after-tax cash flow* 'ATCF' is defined as the net income after taxes plus the depreciation and depletion allowances as shown in equation (7.2.7).

$$ATCF = NIAT + D = BTCF - IncTx = (1-t)*NIBT + D \quad ....(7.2.7)$$

The cash flow definitions given above serve the purposes of financial accounting, but more detailed definitions are given in the following sections for managerial accounting purposes.

Within each accounting period, revenues and expenses of the income statement are recorded on an "as is" basis without adjusting for time equivalences within or between accounting periods. When an *accrual basis of accounting* is used, revenues are recorded in the accounting period in which they are earned rather than the period they are collected in cash, and expenses are recorded in the accounting period in which they are incurred rather than the period they are paid. Since accounting income has meaning only with respect to a specific period of time, it must be calculated from revenues and expenses in the same accounting period.

In economic decision-making, the positive economic effect of business revenues should be recognized at the time they are earned, and the negative economic effect of business expenses should be recognized at the time the goods and services are consumed. Therefore, most businesses use the accrual basis of accounting. Alternatively, under the *cash basis of accounting*, revenues are recorded only on receipt of cash, and expenses are recorded only when cash payments are made. Most individuals and professionals such as physicians and lawyers maintain their accounting records on a cash basis. Because the cash basis of accounting ignores uncollected revenues which have been earned and expenses which have been incurred but not paid, the accrual basis of accounting is preferred for economic analysis.

Many short-term revenues and expenses are easily partitioned into accounting periods of a year or a quarter of a year. However, dividing the revenues and expenses of an enterprise into time segments such as a year or quarter of a year distorts the cause-and-effect relationships between expenses and revenues which belong to different accounting periods. Nevertheless, investors, creditors and governments require these periodic measurements of net income and financial position for decision-making purposes and the estimation of taxable income. The need of parties outside an enterprise for annual or quarterly financial statements creates many of the most serious problems in financial accounting.

In particular, the time allocation of long-term expenses such as the *cost of goods sold* and *depreciation allowances* depends upon the selection of inventory flow assumptions and statutory depreciation methods which can have a sizable impact on income statements and balance sheets over the course of many accounting periods. The accounting problems encountered in the timing of these two long-term expenses is presented in Sections 7.3 and

7.4 in order to bring out fundamental issues. A more detailed explanation may be found in selected references on financial accounting at the end of this chapter.

Section 7.3 - Cost of Goods Sold

Most goods cannot be sold as soon as they are produced. Consequently, they are placed in inventory until actual sales are withdrawn from inventory. The cost of goods sold then depends on how the inventory is priced. Because inventory pricing directly affects operating incomes, income taxes and asset evaluation upon which business decisions are based, we need a more detailed explanation of the fundamental equation of inventory accounting.

Cost of Goods Sold = Beginning Inventory + Costs of Added Inventory - Ending Inventory ....(7.3.1)

The beginning inventory consists of all unsold goods that are owned and available for sale at the beginning of the accounting period. The costs of added inventory during the accounting period consist of purchases of raw materials and finished products and the costs of manufacturing and making them available for sale. When the actual inventory of all unsold goods at the end of the accounting period is subtracted from the inventory at the beginning of the accounting period plus the costs of inventory added during the accounting period, we arrive at the cost of goods sold during the accounting period.

Although cost is the primary basis for pricing inventories, situations arise where inventory may properly be valued at less than its cost. If the value of inventory has fallen below cost due to physical deterioration, obsolescence or decline in prices, the loss may be realized in the current period by writing down the accounting value of inventory from cost to a lower value designated as the cost of replacing the inventory at market prices. For this reason, inventory is often priced at cost or market, whichever is less.

However, market selling prices do not always drop when the cost of replacing inventory declines. Consequently, if ending inventory is written down from cost to a lower market figure but the goods are sold in the next period at usual prices, then the write-down will first reflect a fictitious loss and then an exaggerated profit in the next period. If inventory could be sold at prices which yield a normal profit, it appears that inventory should be carried at cost even though current market prices are lower.

During a period of rising prices, the cost of goods available for sale consists of older lower-cost goods as well as higher-cost goods placed in inventory at later times. The *first-in, first out* method of inventory evaluation, commonly referred to as *FIFO*, is based on the assumption that the goods first placed in inventory are the goods that are sold first. The FIFO method of determining the cost of goods sold may be adopted for income tax purposes whether or not actual sales are made from the oldest units in stock. By removing lower cost goods first, the cost of ending inventory is higher and the cost of goods sold is lower. Consequently, goods sold in the current period will result in a relatively larger taxable income and a larger inventory in the current asset section of the balance sheet.

The *last-in, first-out* method of pricing inventories, commonly called *LIFO*, assumes the most recently acquired goods are sold first, and that the ending inventory consists of "old" merchandise first placed in inventory. Such an assumption may be adopted for income tax purposes regardless of which physical units of merchandise are being delivered to customers.

During a period of rising prices, LIFO removes higher cost goods first so that the cost of ending inventory is lower and the cost of goods sold is higher. Consequently, goods sold in the current period will result in a relatively smaller taxable income and a smaller inventory in the current asset section of the balance sheet. If LIFO tax rules are elected for income tax purposes, then generally accepted accounting principles require that this method must also be used in published financial statements.

Besides FIFO and LIFO, there are two other methods of pricing inventory in common use which affect the cost of goods sold and the cost of the ending inventory. If the units in the ending inventory can be identified from specific purchases, they may be priced using the *specific identification method* from the amounts listed on the corresponding purchase invoices. This method of pricing inventory is likely to give meaningful results in the purchase and sale of high-priced distinguishable items such as automobiles, boats and jewelry.

The *average-cost method* of pricing inventory is computed by dividing the total cost by the number of units available for sale. When the average-cost method is used, the cost figure for the ending inventory is influenced by all the various prices paid during the year. This computation gives a weighted-average unit cost which is then applied to the units in the ending inventory. A common criticism of the average-cost method of pricing inventory is that it attaches the same weight to prices at the beginning and end of a year.

The search for the "best" method of inventory valuation is difficult because the inventory figure is used directly in the balance sheet and indirectly as the cost of goods sold in the income statement, and these two financial statements are intended for different purposes. Accountants who favor the LIFO method argue that it provides a realistic measure of operating income in light of current selling prices. During periods of inflation, LIFO has the advantage of avoiding the payment of income taxes on the basis of exaggerated measurements of taxable income. But the use of LIFO during a period of rising prices is apt to produce a balance sheet figure for inventory which is far below current replacement costs. The FIFO method leads to a balance sheet figure for inventory which is more in line with current replacement costs. However, many accountants believe that the use of FIFO during a period of inflation results in reporting of fictitious operating incomes and consequently in paying excessive income taxes.

Section 7.4 - Depreciation Accounting

A major expense in the income statement of private enterprises consists of depreciation and depletion allowances. From *gross revenues* 'R' are subtracted *operating expenses* to arrive at *net operating income*, as shown in equation (7.4.1).

$$\text{Net Operating Income } \underline{NOI} \equiv \text{Gross Revenues } \underline{R} - \text{Operating Expenses } \underline{E} \quad ....(7.4.1)$$

The operating expenses that can be used as deductions from gross revenues are labor, utilities, maintenance, interest, rent, property taxes, sales taxes and excise taxes. However, the cost of property held to produce income for more than one year may not be fully expensed in the year it is purchased. If all capital expenditures could be expensed when purchased, hardly any net operating income would be available for taxation. In a year of large capital acquisitions, the expenses would largely offset all other income, leaving little or no net operating income. Subsequently, when those large capital acquisitions produce goods and services, the net operating income would be very high unless offset again by other capital acquisitions.

In order to stabilize the amount of taxable income, depreciable property is defined as any property, tangible or intangible, which was acquired and is held to produce income, and which has a limited and estimable life exceeding one year. The cost of depreciable property must first be capitalized and then expensed through depreciation allowances over its estimable life for income tax purposes. Depreciation cannot be accumulated for high income years - it must be deducted annually from net operating income to obtain *net income before taxes*.

Net Income Before Taxes NIBT ≡ Net Operating Income NOI - Depreciation D ....(7.4.2)

By definition, net income before taxes is the taxable income which is multiplied by the applicable tax rate to determine the *income taxes* that must be paid. *Net income after taxes* is then obtained by subtracting the income taxes from the net income before taxes.

Income Taxes IncTx ≡ Tax Rate t * Net Income Before Taxes NIBT ....(7.4.3)

Net Income After Taxes NIAT ≡ Net Income Before Taxes NIBT - Income Taxes IncTx ....(7.4.4)

In corporations, net income after taxes belongs to the owners of equity capital. Interest on debt was already taken into account in the operating expenses of equation (7.4.1). Net income after taxes is distributed as cash dividends to preferred and common stockholders with the remainder going to retained earnings. Funds for further capital investment are then derived from accumulated retained earnings, additional equity funding and debt financing. By this process of capital formation and transforming investments into goods and services, business enterprises hope to systematically increase their net incomes before and after taxes.

There are a number of widespread misunderstandings regarding depreciation and depletion allowances that should be cleared up before discussing how they enter into the definition of taxable income in greater detail.

1. Depreciation implies the decrease in asset value over time due to deterioration, wearout, obsolescence and depletion. However, it should be emphasized that it is *cost*, not *value*, that is apportioned in depreciation allowances. The *market value* of a capital asset or its *value to the owner* may actually appreciate in the course of time, or the value may depreciate faster than depreciation allowances permit. Although land, precious metals, money and art may be appreciating or depreciating in value with time, they are not depreciable assets because they are defined as having unlimited lives.

2. Depreciation allowances are not intended for the replacement of depreciated assets. It may be undesirable, uneconomical or even infeasible to replace depreciated assets. In the case of mineral deposits, it is not feasible to replace depleted assets. But depletion allowances are still deductible from net operating income in order to determine the taxable income in any accounting period.

3. It is commonly thought that depreciation allowances are designed to compensate for the physical consumption of capital investments. Most capital assets depreciate during their lifetime because of physical factors such as depletion, deterioration and wear and tear. Consequently, depreciation allowances often correlate with capital consumption. However, depreciation allowances are arbitrary in nature. They cannot be expected to compensate for

technological factors such as obsolescence, for economic factors such as price changes of inputs and outputs, or for valuations of intangible assets. Depreciation allowances are standardized for all competitors in the private sector of the economy. Although depreciation exists in public projects, it is not computed there because income taxes are not applicable.

Depreciation, in a financial accounting sense, is the systematic allocation of the initial cost of an asset against the income derived from goods and services obtained from using the asset during its lifetime. The *book value* of an asset at any time is its first cost minus the cumulative depreciation allowances. In other words, the book value of an asset is that part of its original cost in excess of accumulated depreciation. When a depreciable asset is purchased at a price which is greater than the seller's book value, then the purchaser can elect to step up the cost basis to the purchase price of the property. The cost basis of the asset for income tax purposes is then the cash purchase price plus the cost of making the asset serviceable.

There are many different methods of writing off the cost basis of depreciable assets. Prior to the Economic Recovery Tax Act of 1981 (ERTA), the principal depreciation methods used in the United States required an estimate of the useful lives of depreciable assets and the salvage values of those assets at the end of their useful lives. In light of the difficulties of predicting the useful life and salvage value of durable assets, ERTA(81) provided an Accelerated Cost Recovery System (ACRS) of depreciation to be used for assets placed in service after December 31, 1980. Two major advantages of ACRS depreciation are: (1) depreciation allowances are made using "property class lives" that are less than "actual useful lives"; and (2) salvage values are assumed to be zero. ACRS is mandatory for most tangible depreciable assets placed in service from January 1, 1981 through July 31, 1986.

The modified ACRS (MACRS) rules created by the Tax Reform Act of 1986 (TRA) is now the principal means of writing off the cost basis of depreciable assets. The differences between ACRS and MACRS rules are mainly the class lives of assets and the methods of recovering their costs. The primary objective of both ERTA(81) and TRA(86) is to stimulate growth of the U. S. economy by relying more on market forces and less on income tax considerations for economic decision-making.

The intent of the remainder of this section is to present the major depreciation accounting methods used prior to 1981. Then we will discuss the main features of ACRS and MACRS depreciation accounting methods created by ERTA(81) and TRA(86). For more specific details, the reader is referred to "Depreciation", Internal Revenue Service (IRS) Publication 134, U. S. Government Printing Office. The IRS is the federal agency that collects taxes and issues regulations for the implementation of tax legislation passed by United States Congress.

Section 7.4.1 - Straight-Line (SL) Depreciation Accounting

In the straight-line method prior to 1981, estimates are made of the service life, N, of the asset and its prospective salvage value, S, at the end of its service life. *Straight-line depreciation* assumes a constant annual rate over the lifetime of an asset for the depreciation of its first cost, P, and its ending salvage value, S. Let $DR_K(SL)$ denote the straight-line depreciation rate during the Kth year, and let BV_K denote the book value of the asset at the end of the Kth year. Then $P = BV_0$, $S = BV_N$ and $DR_K(SL) = (BV_0 - BV_N)/N*BV_0$. Since $DR_K(SL)$ is a constant independent of K, the annual depreciation charge $DC_K = -DR_K(SL)*BV_0$ is also constant for each of the N years.

Table 7.4.1 illustrates straight-line depreciation. As an example, let BV_0 = \$5,000, BV_N = \$1,000, N = 4 years and $DR_K(SL) = (BV_0 - BV_N)/N*BV_0$ = \$4,000/4*\$5,000 = 20%/yr. The annual depreciation charge is then $DC_K = DR_K(SL)*BV_0$ = -20.00%*\$5,000 = -\$1,000/yr. Table 7.4.1 is similar to the ABC tables in Chapter Three with several exceptions. The period of depreciation accounting is generally a calendar year. Column (2) is now a depreciation rate rather than an interest rate. Column (3) is the book value of the asset before the depreciation charge (BDC) in column (4), and column (5) is the book value of the asset after the depreciation charge (ADC) in column (4). All of the depreciation rates from $(2)_1$ to $(2)_4$ operate only on the first cost BV_0 in $(5)_0$ to determine the entries in $(4)_1$ to $(4)_4$.

Table 7.4.1 - Straight-Line (SL) Depreciation Accounting for a 4-year Asset Life

(1)EOY	(2)Depr Rate (SL)	(3)BDC Book Value	(4)EOY Depr Charge	(5)ADC Book Value
0	20.00%	\$0.00	\$0.00	\$5,000.00
1	20.00%	\$5,000.00	-\$1,000.00	\$4,000.00
2	20.00%	\$4,000.00	-\$1,000.00	\$3,000.00
3	20.00%	\$3,000.00	-\$1,000.00	\$2,000.00
4	20.00%	\$2,000.00	-\$1,000.00	\$1,000.00

Section 7.4.2 - Sum-of-Years-Digits (SOYD) Depreciation Accounting

This method was first authorized in the United States by the 1954 tax law. Estimates are made of the service life, N, of the asset and its prospective salvage value, S, at the end of its service life. The digits corresponding to the number of years are added together. Since the average number of years of life is (N+1)/2, the sum-of-years-digits is SOYD = N(N+1)/2. The depreciation rate $DR_1(SOYD)$ in the first year is (N/SOYD) times the fraction $(BV_0 - BV_N)/BV_0$ of depreciable cost. In the second year, $DR_2(SOYD)$ is (N-1)/SOYD times the fraction $(BV_0 - BV_N)/BV_0$ of depreciable cost. And in the Kth year, $DR_K(SOYD)$ is (N-K+1)/SOYD times the fraction $(BV_0 - BV_N)/BV_0$ of depreciable cost. The depreciation charge $DC_K = -DR_K(SOYD)*BV_0$ is largest when K = 1 and it is smallest when K = N.

For example, let BV_0 = \$5,000, BV_N = \$1,000 and N = 4 years as in the straight-line depreciation example. Then SOYD = N(N+1)/2 = 4*5/2 = 10. Substituting in the formula $DR_K(SOYD) = ((N-K+1)/SOYD)*(BV_0 - BV_N)/BV_0$, we find $DR_1(SOYD)$ = 32.00%, $DR_2(SOYD)$ = 24.00%, $DR_3(SOYD)$ = 16.00% and $DR_4(SOYD)$ = 8.00% as listed in column (2) of Table 7.4.2 below. The procedure for sum-of-years-digits depreciation accounting in Table 7.4.2 is the same as that used for straight-line depreciation accounting in Table 7.4.1 above.

Table 7.4.2 - Sum-of-Years-Digits (SOYD) Depreciation Accounting for a 4-year Asset Life

(1)EOY	(2)Depr Rate (SOYD)	(3)BDC Book Value	(4)EOY Depr Charge	(5)ADC Book Value
0	32.00%	\$0.00	\$0.00	\$5,000.00
1	32.00%	\$5,000.00	-\$1,600.00	\$3,400.00
2	24.00%	\$3,400.00	-\$1,200.00	\$2,200.00
3	16.00%	\$2,200.00	-\$800.00	\$1,400.00
4	8.00%	\$1,400.00	-\$400.00	\$1,000.00

Section 7.4.3 - Declining-Balance (DB) Depreciation Accounting

The contribution of assets to income is usually much greater in the early years of life than in later years. During the final years of their lives, many assets are inefficiently used below their capacity and they require much more maintenance. The failure to replace, retire or refurbish such assets in a more timely manner not only causes industry to be less competitive, but it also results in the collection of less income taxes by the government. For these reasons, several ways of making more rapid write-offs in the early years were introduced for income tax purposes in the United States in 1954. One of these accelerated depreciation methods was the so-called double-rate declining-balance method.

In the double-rate declining-balance (DDB) method, the depreciation rate was computed as 200% divided by the estimated life in years. This depreciation rate was applied each year as a *fixed percentage* to the *remaining book value* of the asset instead of its first cost. When using DDB depreciation, the salvage value is *not* used to determine either the depreciation rate or the book value. These features are illustrated in Table 7.4.3 below. The fixed percentage for a 4-year asset life is 200%/4 = 50%/year as shown in column (2). This depreciation rate is applied to book values $(3)_1$ to $(3)_4$ in order to obtain depreciation charges $(4)_1$ to $(4)_4$.

Table 7.4.3 - Double-rate Declining-Balance (DDB) Depreciation for a 4-year Asset Life

(1)EOY	(2)Depr Rate (DDB)	(3)BDC Book Value	(4)EOY Depr Charge	(5)ADC Book Value
0	50.00%	$0.00	$0.00	$5,000.00
1	50.00%	$5,000.00	-$2,500.00	$2,500.00
2	50.00%	$2,500.00	-$1,250.00	$1,250.00
3	50.00%	$1,250.00	-$625.00	$625.00
4	50.00%	$625.00	-$312.50	$312.50

Because the declining-balance method cannot reach a book value of zero within a finite lifespan, it is permissible to start with the declining-balance method and to finish with the straight-line method. In this regard, let us determine the time to switch from DDB to SL depreciation in Table 7.4.3 . Switching at the beginning of the 2nd year leaves 3 more years until the end of four years. The SL depreciation rate would then be 33.33% which is below the DDB depreciation rate. Switching at the beginning of the 3rd year leaves 2 more years until the end of four years. The SL depreciation rate would then be 50.00% which is the same as the DDB depreciation rate. Switching at the beginning of the 4th year leaves one more year until the end of four years. The SL depreciation rate would then be 100.00% which is the rate that should be used for zero salvage. The declining-balance method also uses a fixed depreciation rate of 150%/N which is often applied to used equipment and assets with long lives.

Section 7.4.4 - Sinking-Fund (SF) Depreciation Accounting

This method assumes that a sinking fund is established as a separate reserve in which funds will accumulate for a return of bonded indebtedness and for replacement purposes. The sinking-fund method is used by some government agencies to represent public projects that lose value slowly during the first years and more rapidly during later years. The annual sinking-fund depreciation charge (DC) is constant for N years of the lifespan, and it is calculated from the sinking-fund formula DC = (P-S)(A/F,i%,N) where $P = BV_0$, $S = BV_N$ and i% represents the annual interest rate earned in the sinking fund. The practice of keeping separate accounts of depreciation funds in an external sinking fund is seldom used in private industry. However, a firm that is having internal financial problems may be required by court settlements to set up a sinking fund for a depreciation reserve.

As an example of sinking-fund depreciation, let the annual interest rate be 7% and let BV_0 = \$5,000, BV_N = \$1,000 and N = 4 years as in the straight-line depreciation example of Table 7.4.1 . Then DC = (P-S)(A/F,7%,4) = \$4,000*0.22523 = \$900.92 per year as shown in Table 7.4.4 below. If 7% interest is included on the funds invested in the sinking fund, then the accumulation at the end of four years will be \$900.92*(F/A,7%,4) = \$900.92*4.43994 = \$4,000. The sinking-fund depreciation rate of 18.02% per year (100*\$900.92/\$5,000) is less than the 20% straight-line depreciation rate because of the interest earned on the funds allocated to the sinking fund.

Table 7.4.4 - 7% Sinking-Fund (SF) Depreciation Accounting for a 4-year Asset Life

(1)EOY	(2)Depr Rate (SF)	(3)BDC Book Value	(4)EOY Depr Charge	(5)ADC Book Value
0	18.02%	\$0.00	\$0.00	\$5,000.00
1	18.02%	\$5,000.00	-\$900.92	\$4,099.08
2	18.02%	\$4,099.08	-\$900.92	\$3,198.16
3	18.02%	\$3,198.16	-\$900.92	\$2,297.24
4	18.02%	\$2,297.24	-\$900.92	\$1,396.32

Section 7.4.5 - Cost and Percentage Depletion Accounting

Depletion may be defined as using up natural resources as a result of their removal. Capital investments in natural resources are being consumed as the natural resources are being removed and sold. In the case of mining natural resources, a company may sell itself out of business as it carries out its normal operations. The replacement of depleted natural resources is usually not possible except by discovering other mineral deposits through exploration. Therefore, a portion of the gross income derived from the production of natural resources is required for a return of the capital investment. Consequently, the need for depletion allowances to keep capital intact is essentially the same as the need for depreciation allowances to offset capital consumption.

There are two ways of computing depletion allowances: (1) the *cost method* and (2) the *percentage method*. For standing timber and most oil and gas wells excepting small domestic producing wells, the percentage method is not allowed and only the cost method is permissible. When the percentage method applies to a property, depletion allowances must be calculated by both the cost and percentage methods. The *larger* of the two allowances may be taken and used to reduce the adjusted cost basis of the property for the following tax year.

In the cost method, the cost of a *depletion unit* is determined by dividing the adjusted cost basis of a property by the number of units remaining to be mined or harvested. The number of recoverable depletion units can be estimated as tons of ore, cubic meters of gravel, barrels of oil, million cubic feet of natural gas, thousand board-feet of timber, etc. The cost depletion allowance for a given tax year is then calculated as the product of the number of depletion units *sold* during the year and the cost of a depletion unit.

The percentage method can be used to calculate the depletion allowances for mineral and geothermal deposits, and small domestic oil and gas wells. Unlike depreciation and cost depletion which allocate cost over useful life, percentage depletion is an annual allowance of a percentage of the gross income from the property as listed in Table 7.45 . Since percentage depletion is computed on the basis of gross income rather than the cost of the property, the total depletion allowances on a property may exceed the cost of the property. However, the allowable percentage depletion in any year is limited to not more than 50% of the taxable income from the property computed without the deduction for depletion.

Table 7.4.5 - Percentage Depletion Allowance for Selected Natural Deposits (1988)

Type of Deposit	Percent
Sulphur and uranium; domestic lead, zinc, nickel and asbestos	22%
Gold, silver, copper, iron ore and U.S. oil shale deposits	15%
U.S. geothermal wells and small domestic oil and gas producers	15%
Coal, lignite and sodium chloride	10%
Sand, gravel, stone, clam and oyster shells, brick and tile clay	5%
Most other minerals and metallic ores	14%

As an example of percentage depletion, consider a lignite mine which has a gross income of $2,000,000 and mining expenses of $1,680,000 for the year. From Table 7.4.5, the percentage depletion allowance of the lignite mine would be 10% of $2,000,000 which is $200,000 for the year. We must now compute the taxable income limitation of the mine:

Gross income	$2,000,000
Less expenses other than depletion	-$1,680,000
Taxable income (net income before taxes)	$320,000
Taxable income limitation (50%*$320,000)	$160,000

Since the taxable income limitation of $160,000 is less than the computed percentage depletion of $200,000, the allowable percentage depletion is the smaller deduction of $160,00.

Section 7.4.6 - ACRS Depreciation Accounting

Let us now consider the Accelerated Cost Recovery System (ACRS) of depreciation created by the Economic Recovery Tax Act of 1981 (ERTA). ACRS allows a business to write off the *cost basis* of depreciable property over a *recovery period*. Intangible property is not subject to ACRS depreciation. The unadjusted cost basis assumes no salvage value and is normally the cash purchase price of a property plus the cost of making the asset serviceable. The recovery period is obtained by assigning an asset to 3-, 5-, 10- or 15-year "property class lives" as discussed below.

The property class lives of ACRS depreciation were developed from a 1970 U. S. Treasury Department study of the actual useful lives in which assets were utilized. In 1971, they published Asset Depreciation Range (ADR) guidelines of lower, upper and midpoint limits of useful lives for about 100 classifications of depreciable assets. The ADR midpoint lives were somewhat shorter than actual average useful lives. The ADR midpoint-life guidelines have been incorporated into ACRS classifications so that most property classes are again shorter than ADR midpoint lives. Appendix 7A provides the ADR guideline periods for selected classifications of depreciable personal property from which it is generally possible to determine a property's ACRS class life of 3, 5, 10 or 15 years.

The percentage of the unadjusted cost basis that could be deducted from taxable income (i.e., net income before taxes) each year was stipulated as shown in Table 7.4.6 below. Subject to certain restrictions, taxpayers could elect to use straight-line depreciation over longer periods of time when they expected small taxable incomes in early years and large taxable incomes in later years.

Table 7.4.6 - ACRS Percentages for Property Placed in Service from 1/1/81 to 7/31/86

Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
3-Yr Property	25	38	37												
5-Yr Property	15	22	21	21	21										
10-Yr Property	8	14	12	10	10	10	9	9	9	9					
15-Yr Property	5	10	9	8	7	7	6	6	6	6	6	6	6	6	6

Section 7.4.7 - MACRS Depreciation Accounting

The modified ACRS (MACRS) rules created by the Tax Reform Act of 1986 (TRA) is now the principal means of writing off the cost basis of depreciable assets. The differences between ACRS and MACRS rules are mainly the class lives of assets and the methods of recovering their costs. Under MACRS, assets are assigned to one of six classes of depreciable personal property or to one of two classes of real property. The cost recovery rate in the more short-lived 3-, 5-, 7-, and 10-year property classes is based on the 200% declining-balance depreciation method. Switching to the straight-line depreciation method is permitted at the optimal time. The costs of assets in the 15- and 20-year classes are recovered using the 150% declining-balance depreciation method, again switching to the straight-line method at the optimal time. The straight-line method must be used for all real estate property. A summary of MACRS class lives and permissible depreciation methods is also provided in Appendix 7A.

MACRS uses several averaging conventions for depreciable property placed into service during the taxable year. Generally, a half-year convention applies to the first and last years of the class lives of depreciable personal property, and a mid-month convention applies to real estate property. A mid-quarter convention is used when more than 40% of asset additions are placed into service in the last quarter of the taxable year.

Section 7.4.8 - Multiple-Asset Depreciation Accounting

Until now, depreciation was considered for individual assets, but financial accounting often depreciates general equipment with different lifespans as a single group. Although the average life of assets in a group will be N years, some assets will last either less or more than N years. If an individual asset in a group was retired in less than N years, a loss of capital value would result. Conversely, if it was used more than N years and then sold for any amount, a capital gain would be realized. By using multiple-asset accounting, these losses and gains tend to cancel so that the capital-gain-and-loss adjustments needed for individual-asset accounting could be avoided.

However, the convenience of multiple-asset accounting is hardly worth losing the advantages of individual-asset accounting. Individual assets with short lives need to be replaced or retired much more often than assets with long lives. As compared to long-lived assets, the prolonged years of short-lived assets are usually filled with much more maintenance and far less productive capacity than could be justified on economic grounds. The average lifespan of assets in a group gives more weight to long-lived assets than the relative frequency with which they need to be replaced or retired. In general, the replacement or retirement of each asset should be evaluated on its own merits without complications from possible tax advantages of retaining other assets.

Section 7.5 - Federal and State Income Tax Rates

The present federal income tax dates back to 1913 with the passage of the Sixteenth Amendment to the Constitution which reads "The Congress shall have the power to lay and collect taxes on incomes, from whatever source derived, without apportionment among the several States, and without regard to any census or enumeration.". After the Sixteenth Amendment to the Constitution was ratified, the Revenue Act of 1913 was passed which permitted income taxes to be levied with a "reasonable allowance for depreciation by use, wear and tear of property, if any.". The regulations provided for depreciation "that arises from exhaustion, wear or tear, or obsolescence out of the uses to which the property is put."

Taxing authorities gradually realized that many industrial assets were used much longer than claims made for income tax purposes. Starting 1934, the United States Treasury Department placed the burden of proof on each taxpayer to justify the depreciation deductions claimed on income tax returns. At that time, the corporation tax rate was 13.75% which had only a minor effect on business decision-making. But tax rates were raised substantially by the Second Revenue Act of 1940 because of World War II. Federal income tax rates continued to be high for the next 40 years - about 46% on most corporate incomes and 70% on the highest bracket of personal incomes. In most states of the United States, there also were state taxes on corporate and individual incomes.

It became evident that the combination of high tax rates and low allowable depreciation rates were adversely affecting economic activity. Consequently, the first major changes in the tax laws were made with respect to depreciation accounting. In 1954, taxpayers were given the option of the declining-balance method providing the declining-balance rate was not more than 200% of the permissible straight-line rate. Another option was to use the sum-of-years-digits method. In 1962, "guideline lives" were authorized which were appreciably shorter than the ones previously required for most classes of depreciable assets. In 1971, the ADR (Asset Depreciation Range) system enabled taxpayers to use lives as much as 20% shorter or, at the taxpayers option, as much as 20% longer than guideline lives.

The concept of tying depreciation allowances to the "useful lives" of depreciable assets was completely changed by ACRS in the Economic Recovery Tax Act (ERTA) of 1981 (see Section 7.4). Under ACRS, eligible assets were divided into only four classes which had statutory cost recovery periods of 3, 5, 10 or 15 years. Subject to certain restrictions, taxpayers could elect to use straight-line depreciation over the same or certain stipulated longer lives. For all eligible assets, depreciation rates were computed as if salvage values were zero. If later on, it turned out that the actual net salvage value on disposal exceeded the remaining book value, then the "gain on disposal" would be taxable. Because taxpayers had been shifted into higher taxable income brackets by considerable price inflation since 1972, ERTA also changed the federal income tax rates from 70% to 50% on the highest bracket of personal incomes.

The Tax Reform Act (TRA) of 1986 greatly altered federal income tax rates in the United States and also modified some of the rules for determining taxable income. The changes in depreciation rates enacted by the Tax Reform Act of 1986 are described in Section 7.4 and Appendix 7A. The federal tax rate schedules on 1988 incomes of corporations and personal incomes in the United States as established by the U. S. Tax Reform Act of 1986 are listed below. The 1988 federal income tax rate schedules are not current, and they are not intended as a comprehensive treatment of federal income tax regulations. Instead, the purpose of the following presentation is to bring out some of the more important provisions of the Tax Reform Act of 1986 that can be understood in principle and applied without difficulty to many problems of economic decision-making in the private sector of the economy.

Taxable Income (TI) of Corporations	Tax Rate	$350,000 TI Example	
$0 < TI ≤ $50,000/yr	15%	0.15*($50,000 - $0)	= $7,500
$50,000 < TI ≤ $75,000/yr	25%	0.25*($75,000 - $50,000)	= $6,250
$75,000 < TI ≤ $100,000/yr	34%	0.34*($100,000- $75,000)	= $8,500
$100,000 < TI ≤ $335,000/yr	39%	0.39*($335,000-$100,000)	= $91,650
$335,000 < TI	34%	0.34*($350,000-$335,000)	= $5,100
			$119,000

The total income tax on TI = $350,000 is $119,000 which is 34% of $350,000. The 39% tax bracket is designed to phaseout the 34%-15% = 19% tax advantage of the 15% tax bracket and the 34%-25% = 9% tax advantage of the 25% tax bracket (i.e., 0.19*$50,000 = $9,500; 0.09*$25,000 = $2,250; and 0.05*$235,000 = $11,750 = $9,500+$2,250). Thus, corporate taxable incomes of $335,000 per year or greater would pay a flat rate of 34%.

Unlike corporations, individual taxpayers in the United States have special *exemptions* and *deductions* subtracted from their gross income to determine taxable income. In 1988, each taxpayer had a $1,950 exemption. Taxpayers who supported persons classified by law as *dependents* have an additional $1,950 exemption per dependent. Personal expenditures that could be itemized as deductions in 1988 included certain medical and dental expenses, taxes, interest paid, gifts to charity, casualty and theft losses, and other miscellaneous deductions. Individuals could take either a "standard deduction" ($3,000 for single or $5,000 for married taxpayers filing jointly), or itemize "nonbusiness deductions" in their federal income tax returns which could include a deduction for state income taxes.

Federal tax rates on 1988 incomes of single taxpayers and married taxpayers (filing joint returns) in the United States as established by U. S. Tax Act of 1986 is listed below.

Taxable Income of Single Taxpayers	Tax Rate	$125,000 TI Example (No Dependents)	
$0 < TI ≤ $17,850/yr	15%	0.15*($17,850 - $0)	= $2,677.50
$17,850 < TI ≤ $43,150/yr	28%	0.28*($43,150- $17,850)	= $7,084.00
$43,150 < TI ≤ $100,480/yr*	33%	0.33*($100,480- $43,150)	= $18,918.90
$100,480 < TI	28%	0.28*($125,000-$100,480)	= $6,865.60
			$35,546.00

Taxable Income of Married Taxpayers	Tax Rate	$225,000 TI Example (3 Dependents)	
$0 < TI ≤ $29,750/yr	15%	0.15*($29,750- $0)	= $4,462.50
$29,750 < TI ≤ $71,900/yr	28%	0.28*($71,900- $29,750)	= $11,802.00
$71,900 < TI ≤ $171,090/yr*	33%	0.33*($203,850- $71,900)	= $43,543.50
$171,090 < TI	28%	0.28*($225,000-$203,850)	= $5,922.00
			$65,730.00

*The top of the 33% tax bracket of both single and married taxpayers is increased by $10,920 for each dependent claimed. Thus, a married couple claiming three dependent children would increase the top of the 33% bracket from $171,090 to $171,090 + 3*$10,920 = $203,850.

The tax rate schedules of single and married taxpayers keep a flat 28% tax rate for wealthiest taxpayers while removing the tax savings from their individual tax exemptions. For example, a single taxpayer with no dependents has a single $1,950 tax exemption. If his or her taxable income was $125,000, the income tax would be $35,546. The tax benefits of the $1,950 exemption at the 28% tax rate is $546. Therefore, $546 was added to the 33% tax bracket so that the remaining $35,000 tax is exactly 28% of the $125,000 taxable income. Similarly, married taxpayers with three dependents receive a 5*$1,950 = $9,750 reduction in taxable income. If 28% of $9,750 or $2,730 is subtracted from the $65,730 tax payment, the remaining $63,000 is exactly 28% of the $225,000 taxable income.

Alternative Minimum Tax (AMT)

Besides lowering the maximum rate on corporate taxable income from 46% to 34%, the Tax reform Act of 1986 also created a new *alternative minimum tax* (AMT) system to ensure that corporations with large economic incomes would be subject to a minimum tax. Corporations must not only compute their income tax liabilty as described above, but also according to rather complex AMT rules that are beyond the scope of our discussion.

In addition to lowering the maximum rate on individual income taxes from 50% to 28%, the Tax reform Act of 1986 also created a new *alternative minimum tax* (AMT) system to ensure that individuals with large economic incomes would be subject to a minimum tax. As in the case of corporate AMT rules, the individual AMT rules are too complex for our discussion. The new AMT rules separate individual taxpayers into two distinct groups. One group consists of individuals with traditional sources of income such as salaries, wages, dividends and interest from savings accounts, money market funds, certificates of deposit and savings bonds.

Section 7.5.1 - Combined Federal and State Income Tax Rates

Many states (and some cities) in the United States impose income taxes on corporations and individuals in their jurisdictions. Although the income tax regulations of most states are patterned after the federal regulations, there are significant variations. State tax rates are generally much lower than federal rates and they often can be closely approximated as a constant percentage of federal rates. However, state income taxes are deductible expenses in computing taxable income for federal income taxes, but federal income taxes are not, in general, deductible expenses in computing taxable income for state income taxes. Therefore, state income taxes are applicable to larger taxable incomes than are federal income taxes.

If federal and state governments have the same rules for determining taxable income, then equation (7.5.1) would approximate the combined effect on income taxes of the federal and state income tax rates

$$t_E = t_S + (1-t_S)t_F \quad ....(7.5.1)$$

where t_E is the effective federal and state income tax rate, t_F is the federal income tax rate, and t_S is the state income tax rate. Since income tax rates vary with the level of taxable income, one must decide which income tax rates to use in each situation. In general, one should use the marginal tax rate that applies to the increment in taxable income projected in the economic analysis. For convenience in economic analysis, we shall only use the marginal tax rate that applies to increments of taxable incomes in the highest bracket of corporate or individual income taxes. The computation of taxable incomes to which the marginal tax rate applies is discussed in the next section.

Section 7.5.2 - Capital Gains (or Losses)

Whenever an asset disposal takes place at more or less than its book value, an increase or decrease occurs in the asset and owner's equity sections of the balance sheet. By accounting convention, this so-called capital gain or loss on disposal is assigned entirely to the year of the disposal regardless of when the asset was acquired. Before 1930, such capital gains or losses were included or deducted from taxable income in the year in which the asset disposal occurred. But in the 1930's, the law was changed by allowing long-term capital gains to be taxed at a lower rate than ordinary income and setting strict limits on the amount of capital losses that could be deducted from taxable income except to offset similar capital gains. The tax rules on capital gains or losses have changed many times and are full of technicalities.

For tax purposes, long-term capital gains or losses apply to assets held more than one year. For corporations, capital losses can be used to offset capital gains but not to reduce ordinary income. Losses can be carried over to later years. For individuals, capital losses must first be used to offset capital gains, and up to $3,000 of remaining losses can be used to reduce ordinary taxable income in a year. The tax rates on net long-term capital gains have been as high as 49% and as low as 20%. However, the Tax Reform Act of 1986 eliminated the differential between income tax rates on capital gains and ordinary income. Consequently, corporate capital gains are taxed at a maximum rate of 34%, and individual capital gains are taxed at a maximum rate of 28%.

Section 7.5.3 - Investment Tax Credit

In 1962, the U. S. Tax law enacted the Investment Tax Credit (ITC) to stimulate business activity. The ITC permits businesses to subtract up to 10% (and a maximum of $100,000) of the purchase price of certain specified types of depreciable property from their *income tax liability* for the year in which the investment was made. In effect, the government pays up to 10% of the invoiced cost of a qualifying investment and the investor pays the rest. If the credit could not be used that year, it could be carried forward or backward to offset tax liabilities in other years. The credit could be partially recaptured by the government if the assets are disposed of too soon.

The credit was suspended several months in 1966-67, repealed in 1969, and reinstated in 1971. The Tax Reform Act of 1986 revoked the ITC in conjunction with broadening the tax base and lowering maximum corporate tax rates from 46% to 34% and maximum individual tax rates from 50% to 28%. Judging from the "on again-off again" status of ITC in the past, it is likely that Congress will restore some form of the credit when the national economy is weak.

Section 7.6 - Managerial Accounting for Project Evaluation

The object of managerial accounting, also called cost accounting, is to estimate and analyze differences in cash flow that stem from the possible replacement of ongoing with engineering and financial alternatives. Because income taxes are major cash outflows, managerial accounting is concerned with results both before and after income taxes. Since financial accounting provides the basic quantitative information system in virtually every enterprise, managerial accounting derives much of its data from the work of financial accountants. However, there is cost information that accounting records will not give, and some financial accounting results need to be modified for the purposes of managerial accounting.

Although accounting records are a good source of historical data needed in describing the cost and revenue structure of ongoing alternatives, they have important limitations when estimating incremental costs and revenues of engineering and financial alternatives. Where the economy of some new equipment or process is involved, estimates of operating and material costs must be obtained by experimental studies rather than from accounting records. And where the quantity and quality of the output is altered, marketing studies are needed to estimate the change in revenues. Besides accounting records, there are usually persons and data within an organization as well as outside consultants and publications that may be used as sources for cost and revenue estimation. Otherwise, the only alternative may be research and development activities to generate the necessary information.

The common practice of multiple-asset depreciation should be changed to individual-asset depreciation for the purposes of managerial accounting. From a financial accounting viewpoint, multiple-asset depreciation is easier and costs somewhat less than individual-asset depreciation. But situations could arise where replacing or refurbishing old machinery with efficient modern equipment and controls is substantially delayed because the methods of multiple-asset depreciation do not take advantage of the accelerated depreciation allowances that individual-asset depreciation would permit under the MACRS rules of 1986. As a consequence, the firm is saddled with the continued use of inefficient machinery and the government continues to have a stagnant tax base with which less taxes are collected.

Because of the periodic need for financial accounting information, revenues and expenses are classified and aggregated into the same accounting period even though they are not causally related. Also, time equivalences of a dollar at the beginning and end of an accounting period are not taken into account. Moreover, the financial accounting system is rigidly categorized into various types of assets, liabilities, net worth, income and expenses which may be suitable for financial summaries, but which is rarely appropriate for the needs of economic decision-making involving engineering design, project alternatives and long-term considerations.

Managerial accounting uses the accrual basis of accounting by recognizing the positive economic effect of business revenues at the time they are earned, and recognizing the negative economic effect of business expenses at the time the goods and services are consumed. Alternatively, under the *cash basis of accounting*, revenues are recorded only on receipt of cash, and expenses are recorded only when cash payments are made. Most individuals and professionals such as physicians and lawyers maintain accounting records on a cash basis because it delays the incidence of taxable income. The accrual basis of accounting is used for economic analysis to avoid the timing distortions of uncollected revenues which have been earned and expenses which have been incurred but not paid,.

The following is a general system of managerial accounting for cash flow estimates of project alternatives on a dated basis both before and after income taxes.

(1)	(2)	(2')	(3)	(4)	(5)	(6)	(6')	(7)	(8)
EOY	NOI	CapTr	Int	BTCF	Depr	IncTax	CapTax	LoanCF	ATCF

(1)EOY = End-of-year.

(2)NOI = Net Operating Income: Obtained by deducting from *gross revenues* all ordinary and necessary expenses required to carry out the project, *except for capital transactions* involved in acquiring depreciable assets whose limited lifespans are greater than one year, and *interest* payable on borrowed capital. Capital transactions and interest are posted in columns (2') and (3) respectively.

(2')CapTr = Capital Transactions: The capital transaction component of BTCF in column (2') consists of capital expenditures (i.e., both equity and borrowed capital, but not leased capital) for all depreciable assets whose limited lifespans are greater than one year. Net salvage values realized upon disposing depreciated assets are posted in column (2'). Capital transactions are a component of before-tax-cash-flows, BTCF in column (4), but they are not part of taxable incomes until they are recovered as depreciation in column (5).

(3)Int = Interest: Interest expenses are (-) for capital borrowed by the organization as a whole in order to acquire business assets. In managerial accounting, interest is treated as a separate expense rather than be included with other operating expenses. Interest incomes are (+) when they are received from project assets.

(4)BTCF = Before-tax cash flow = (2) + (2') + (3) which defines BTCF every year. Therefore, ΔNPV{BTCF} = ΔNPV{(2)+(2')+(3)}.

(5)Depr = Depreciation allowances are treated as expenses (-) for the recovery of unallocated capital costs (Section 7.4) in the computation of taxable income. Depreciation allowances for income tax purposes are based strictly on historical costs rather than the basis of either market valuation or physical depreciation in the form of wear and tear, deterioration, obsolescence, or depletion. When facilities or productive equipment with useful lives in excess of one year are acquired by business enterprises, the total cost of those assets may not be treated as an expense in the year that they are acquired. Instead, the capital cost of such long-term assets are first capitalized in column (2') and later expensed in column (5) through depreciation allowances over the useful life of those assets. Special assets such as cash and inventories are not considered depreciable. Land and precious metals with unlimited useful lives are also nondepreciable, and costs of acquiring them in column (2') are recoverable only through sales and salvage. All other facilities and productive equipment with useful lives of one year or less are expensed in column (2).

(6)IncTax = Income taxes are levied on taxable incomes which is the sum of net operating incomes in column (2), interest expense in column (3) and depreciation allowances in column (5). The taxable income of a project is usually positive. If projects have negative taxable incomes, they could offset positive taxable incomes from other projects within the same enterprise. If the enterprise as a whole has negative taxable incomes, they may be carried forward to later years. Assuming the taxable incomes in (2)+(3)+(5) are positive, the effective tax rate, t_E, is defined in equation (7.5.1) for both federal and state income taxes and the income taxes in column (6) = $- t_E[(2)+(3)+(5)]$. Since income tax rates vary with the level of taxable income, one must decide which income tax rates to use in each project. In general, one should use the marginal tax rate that applies to the increment in taxable income estimated for the organization as a whole. For convenience in economic analysis, we will use only the marginal tax rate that applies to increments of taxable incomes in the highest bracket of corporate or individual income taxes.

(6')CapTax = Capital Taxes & Credits: Special tax rates may apply to net capital gains in column (2'). In 1989, investment tax credits to reduce income-tax liabilities were no longer applicable. Capital taxes and credits were not considered in the following examples.

(7)LoanCF = Loan-Capital-Flow portion of Capital Transactions which are not subject to income taxes. Inflow (+) and Outflow (-).

(8)ATCF = After-tax cash flow: (8) = (4)+(6)+(6')+(7). This defines ATCF every year. Therefore, ΔNPV equation (7.6.1) below is satisfied when discounted with any interest rate.

ΔNPV{BTCF}+ΔNPV{IncTax}+ΔNPV{CapTax}+ΔNPV{LoanCF} = ΔNPV{ATCF} ....(7.6.1)

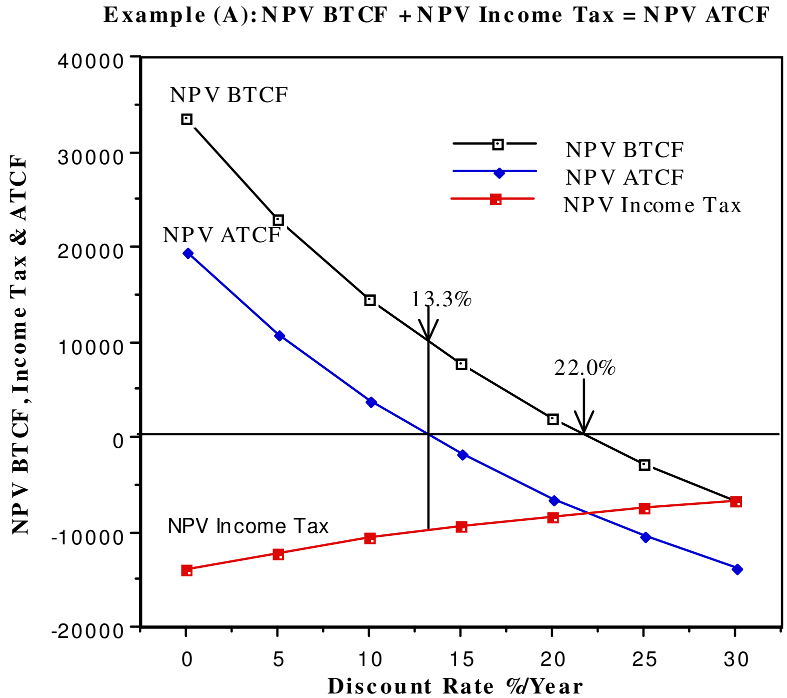
Example: A firm plans to buy a $45,000 refurbished testing machine which would have no salvage value at the end of its 5-year useful life. The testing machine is estimated to save $23,000 per year in maintenance costs with annual operating costs of $7,300. The firm would use straight-line depreciation and it has an effective income tax rate of 42%. The firm's cost of borrowing money before taxes (COBM(BT)) is 9%/year, and the firm's cost of borrowing money after taxes (COBM(AT)) is 9*(1-0.42) = 5.22%/year. The firm's minimum attractive rate of return before taxes (MARR(BT)) is 15%/year, and the firm's minimum attractive rate of return after taxes (MARR(AT)) is 15*(1-0.42) = 8.7%/year. The firm has three mutually exclusive alternatives: (A) 100% equity funding; (B) $23,000 equity funding and $22,000 debt financing with the loan repaid at the end of 5 years; and (C) $23,000 equity funding and $22,000 debt financing with $4,400 repayments of principal each year. Find the best way of financing the purchase of the testing machine?

(A) 100% Equity Funding.

(1) EOY	(2) NOI	(2') CapTr	(3) Int	(4) BTCF	(5) Depr	(6) IncTax	(6') CapTax	(7) LoanCF	(8) ATCF
0	0	-45,000	0	-45,000	0	0	0	0	-45,000
1	15,700	0	0	15,700	-9,000	-2,814	0	0	12,886
2	15,700	0	0	15,700	-9,000	-2,814	0	0	12,886
3	15,700	0	0	15,700	-9,000	-2,814	0	0	12,886
4	15,700	0	0	15,700	-9,000	-2,814	0	0	12,886
5	15,700	0	0	15,700	-9,000	-2,814	0	0	12,886
∑	78,500	-45,000	0	33,500	-45,000	-14,070	0	0	19,430

IRR(BTCF) = 21.958%; IRR(ATCF) = 13.294%; IRR(BTCF)·0.58 = 12.736%

Interest Rate %	ΔNPV BTCF	+ ΔNPV IncTax	+ ΔNPV LoanCF	= ΔNPV ATCF
"As is" 0	33500.00	-14070.00	0	19430.00
COBM(AT) **5.22**	22561.39	-12109.41	0	**10451.98**
MARR(AT) 8.7	16545.85	-11031.21	0	5514.64
COBM(BT) 9	16067.53	-10945.48	0	5122.05
IRR(ATCF)13.294	9826.70	-9826.70	0	0.00
MARR(BT)15	7628.83	-9432.96	0	-1804.13
20	1952.61	-8415.58	0	-6462.97
IRR(BTCF)21.958	0.00	-8065.53	0	-8065.53
30	-6761.55	-6853.69	0	-13615.25



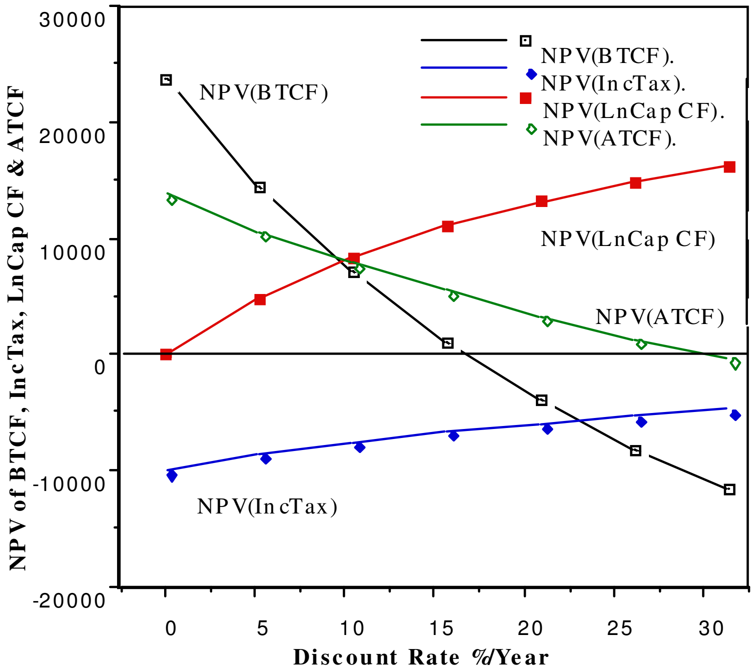
(B) $23,000 equity funding and $22,000 debt financing with loan repaid at end of 5 years.

(1) EOY	(2) NOI	(2') CapTr	(3) Int	(4) BTCF	(5) Depr	(6) IncTax	(6') CapTax	(7) LoanCF	(8) ATCF
0	0	-45,000	0	-45,000	0	0.0	0	22,000	-23,000.0
1	15,700	0	-1,980	13,720	-9,000	-1,982.4	0	0	11,737.6
2	15,700	0	-1,980	13,720	-9,000	-1,982.4	0	0	11,737.6
3	15,700	0	-1,980	13,720	-9,000	-1,982.4	0	0	11,737.6
4	15,700	0	-1,980	13,720	-9,000	-1,982.4	0	0	11,737.6
5	15,700	0	-1,980	13,720	-9,000	-1,982.4	0	-22,000	-10,262.4
∑	78,500	-45,000	-9,900	23,600	-45,000	-9,912.0	0	0	13,688.0

IRR(BTCF) = 15.927% IRR(ATCF) = 28.957% IRR(BTCF)·0.58 = 9.238%

Interest Rate	ΔNPV BTCF	+ ΔNPV IncTax	+ ΔNPV LoanCF	= ΔNPV ATCF
"As is" 0	23600.00	-9912.00	0.00	13688.00
COBM(AT) **5.22**	14040.91	-8530.81	4941.88	**10451.98**
MARR(AT) 8.7	8784.02	-7771.24	7503.11	8515.88
COBM(BT) 9	8366.02	-7710.84	7701.51	8356.69
MARR(BT)15	991.57	-6645.31	11062.11	5408.37
IRR(BTCF)15.927	0.00	-6502.02	11492.49	4990.47
20	-3968.80	-5928.59	13158.69	3261.30
IRR(ATCF)28.957	-10904.84	-4926.40	15831.24	0.00
30	-11583.98	-4828.27	16074.76	-337.49

Example (B) NPV(BTCF)+NPV(IncTax)+NPV(LnCap CF)=NPV(ATCF)



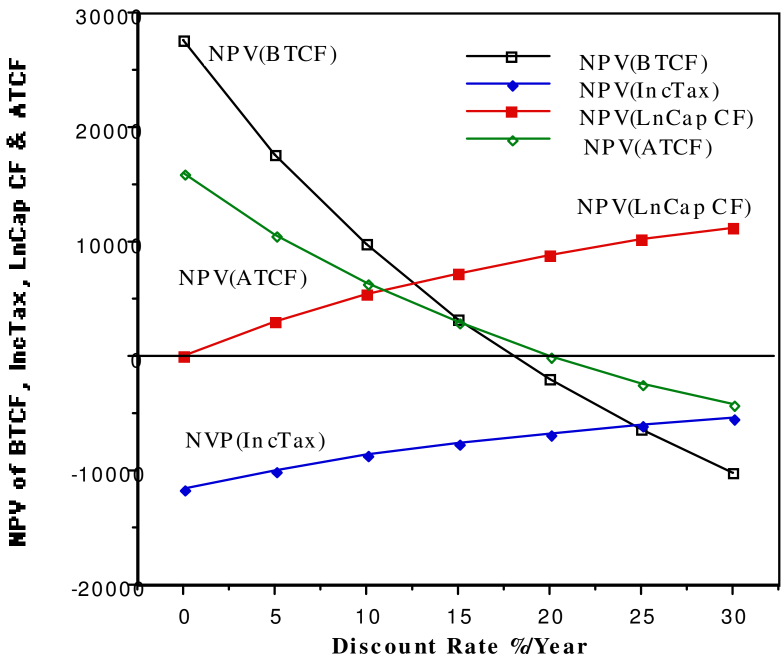
(C) $23,000 equity funding; $22,000 debt financing with $4,400 principal repayments/yr.

(1) EOY	(2) NOI	(2') CapTr	(3) Int	(4) BTCF	(5) Depr	(6) IncTax	(6') CapTax	(7) LoanCF	(8) ATCF
0	0	-45,000	0	-45,000	0	0.00	0	22,000	-23,000.00
1	15,700	0	-1,980	13,720	-9,000	-1,982.40	0	-4,400	7,338.60
2	15,700	0	-1,584	14,116	-9,000	-2,149.72	0	-4,400	7,567.28
3	15,700	0	-1,188	14,512	-9,000	-2,315.04	0	-4,400	7,797.96
4	15,700	0	-792	14,908	-9,000	-2,481.36	0	-4,400	8,027.64
5	15,700	0	-396	15,304	-9,000	-2,648.68	0	-4,400	8,256.32
∑	78,500	-45,000	-5,940	27,560	-45,000	-11,575.20	0	0	15,984.80

IRR(BTCF) = 17.977% IRR(ATCF) = 20.140% 17.977% ·0.58 = 9.713%

Interest Rate	ΔNPV BTCF	+ ΔNPV IncTax	+ ΔNPV LoanCF	= ΔNPV ATCF
"As is" 0	27560.00	-11575.20	0.00	15984.80
COBM(AT) **5.22**	17275.87	-9889.49	3065.60	10451.98
MARR(AT) 8.7	11630.53	-8968.28	4751.48	7413.73
COBM(BT) 9	11181.99	-8893.55	4885.53	7173.97
MARR(BT)15	3278.52	-7605.83	7250.52	2923.21
IRR(BTCF)17.977	0.00	-7074.40	8233.28	1158.88
20	-2025.98	-6744.58	8841.31	70.75
IRR(ATCF)20.140	-2160.66	-6723.84	8884.50	0.00
30	-10146.60	-5431.97	11283.49	-4295.08

Example (C): NPV BTCF+NVP IncTax + NPV LnCap CF = NPV ATCF



The best alternative can be considered from three viewpoints: **(1)** the government collecting income taxes, **(2)** the bank lending money, and **(3)** the firm using the equipment. We will first show the firm's choice is immaterial to the government. For this purpose, a summary of the undiscounted cash flows of the three alternatives is given in Table 7.6.1 .

Table 7.6.1 - Summary of results for undiscounted cash flows of alternatives A, B and C.

Alternative	ΔNPV(BTCF) +	ΔNPV(Inc Tax) +	ΔNPV(LoanCF) =	ΔNPV(ATCF)
(A)	33,500	-14,070.00	0	19,430.00
(B)	23,600	-9,912.00	0	13,688.00
(C)	27,560	-11,575.20	0	15,984.80

(1) The $14,070 income tax the government would collect from the firm in (A) is $4,158 more than the $9,912 it would collect from the firm in (B). However, the bank has $9,900 more taxable income in (B) than in (A). If the $9,900 greater taxable income of the bank is also in the 42% income tax bracket, the government would collect 0.42*$9,900 = $4,158 more from the bank in (B) than in (A) which makes the firm's choice of (A) or (B) a matter of indifference to the government insofar as income tax collections are concerned.

Similarly, the $14,070 income tax the government would collect from the firm in (A) is $2,494.80 more than the $11,575.20 it would collect from the firm in (C). However, the bank has $5,940 more taxable income in (C) than in (A). If the bank is also subject to the 42% income tax rate, the government would collect 0.42*$5,940 = $2,494.80 more income taxes from the bank in (C) than in (A). Therefore, the government will collect $14,070 in taxes from the firm and the bank regardless whether the firm chooses (A), (B) or (C).

(2) The bank earns 9%/year interest before taxes and (1-0.42)*9% = 5.22%/year after taxes whether the firm chooses (B) or (C) as shown in Table 7.6.2 below. The bank's BTCF and ATCF obey a general IRR relationship as given in equation (7.6.2) regardless of how much money the firm borrows or how soon the loan is paid back.

IRR{ATCF} = (1 - effective tax rate)*IRR{BTCF} ....(7.6.2)

Table 7.6.2 - Summary of results for the bank lending money to the firm.

(B) Cash flow analysis of bank's $22,000 loan @ 9%/yr interest repaid at the end of 5 years.

(1)	(2)	(2')	(3)	(4)	(5)	(6)	(6')	(7)	(8)
EOY	NOI	CapTr	Int	BTCF	Depr	IncTax	CapTax	LoanCF	ATCF
0	0	-22,000	0	-22,000	0	0.0	0	0	-22,000.0
1	1,980	0	0	1,980	0	-831.6	0	0	1,148.4
2	1,980	0	0	1,980	0	-831.6	0	0	1,148.4
3	1,980	0	0	1,980	0	-831.6	0	0	1,148.4
4	1,980	0	0	1,980	0	-831.6	0	0	1,148.4
5	1,980	22,000	0	23,980	0	-831.6	0	0	23,148.4
∑	9,900	0	0	9,900	0	-4,158.0	0	0	5,742.0

IRR{BTCF} = 9.00%/year IRR{ATCF} = 5.22%/year

(C) Cash flow analysis of bank's $22,000 loan @ 9%/yr interest repaid principal $4,400/yr.

(1)	(2)	(2')	(3)	(4)	(5)	(6)	(6')	(7)	(8)
EOY	NOI	CapTr	Int	BTCF	Depr	IncTax	CapTax	LoanCF	ATCF
0	0	-22,000	0	-22,000	0	0.00	0	0	-22,000.00
1	1,980	4,400	0	6,380	0	-831.60	0	0	5,548.40
2	1,584	4,400	0	5,984	0	-665.28	0	0	5,318.72
3	1,188	4,400	0	5,588	0	-498.96	0	0	5,089.04
4	792	4,400	0	5,192	0	-332.64	0	0	4,859.36
5	396	4,400	0	4,796	0	-166.32	0	0	4,629.68
∑	5,940	0	0	5,940	0	-2,494.80	0	0	3,445.20

IRR{BTCF} = 9.00%/year IRR{ATCF} = 5.22%/year

The before-tax (BT) and after-tax (AT) rate of return analysis of the bank's cash flows in (B) and (C) exhibit a general relationship as given in equation (7.6.3) regardless of how much money the firm borrows or how soon it pays back the loan.

AT(cost of borrowing money) = (1 - tax rate)*BT(cost of borrowing money) ....(7.6.3)

(3) In order to remove the influence which the use of equity funding or debt financing could have on economic decision-making of the firm, the proper discount rate that should be used is the firm's after-tax cost of borrowing money as shown in Table 7.6.3 below.

Table 7.6.3 - ΔNPV(ATCF) results of the 5.22%/year after-tax discount rate: (A)=(B)=(C).

Alternative	ΔNPV(BTCF) +	ΔNPV(Inc Tax) +	ΔNPV(LoanCF) =	ΔNPV(ATCF)
(A)	22,561.39	-12,109.41	0.00	10,451.98
(B)	14,040.91	-8,530.81	4,941.88	10,451.98
(C)	17,275.87	-9,889.49	3,065.60	10,451.98

The $10,451.98 net present value of ATCF discounted at 5.22% per year is constant regardless of using equity funding or debt financing at 9% per year. Consequently, (A), (B) and (C) are equally good from the firm's viewpoint. If the $45,000 testing machine ranks sufficiently high to be undertaken as an independent alternative, the firm should first use as much of its own money as it can spare and borrow the rest from the bank at 9%/yr interest.

The ΔNPV(ATCF) results in Table 7.6.3 should be compared with those in Table 7.6.4 based on the 8.7%/year after-tax MARR discount rate which would be obtained by applying equation (7.6.3) to the before-tax MARR discount rate of 15%/year.

Table 7.6.4 - ΔNPV(ATCF) results of the 8.7%/year after-tax discount rate: (B)>(C)>(A).

Alternative	ΔNPV(BTCF)	+ ΔNPV(Inc Tax)	+ ΔNPV(LoanCF)	= ΔNPV(ATCF)
(A)	16,067.53	-10,945.48	0.00	5,122.05
(B)	8,366.02	-7,710.84	7,701.51	8,356.69
(C)	11,181.99	-8,893.55	4,885.53	7,173.97

The ΔNPV(ATCF) = $5,122.05 of (A) indicates it is the worst alternative if the firm uses its own money. On the other hand, borrowing $22,000 from the bank at 9% per year in (B) or (C) is more profitable than using the firm's own money. Moreover, (B) is more profitable than (C) because the $22,000 loan was outstanding longer in (B) than in (C). In reality, a borrowed dollar is indistinguishable from the firm's own dollar. But when discounting at a rate which is greater than the after-tax cost of borrowing money, then ΔNPV measurements make it appear as if borrowing is more profitable than using the firm's own money.

Section 7.7 - Rate of Return Before and After Taxes

Income from bonds issued by municipal and state agencies are usually exempt from federal income taxes. Investors may compare higher taxable rates of return from corporate bonds to lower tax-exempt rates of return from municipal and state bonds as follows:

$$\text{Tax-exempt rate of return} = (1 - t_E)*\text{Taxable rate of return} \qquad ....(7.7.1)$$

For example, an investor lives in a state with a 10% income tax rate and is in a 28% federal income tax bracket. According to equation (7.51), the effective tax rate is $t_E = t_S + (1-t_S)t_F = 0.10 + (1-0.10)*0.28 = 0.352$ or 35.2%. Suppose the investor could choose between an 11%/year taxable bond and a 7%/year tax-exempt bond. According to equation (7.7.1), the yield of the 11% bond is 11*(1-0.352) = 7.128% per year *after* taxes and the 7% bond yields 7/(1-0.352) = 10.802% per year *before* taxes.

Since income taxes are a major expense in the private sector of the economy, the after-tax rate of return is needed for analyzing after-tax cash flows in economic decision-making. In this connection, it is convenient to derive the after-tax cash flow (ATCF) of each alternative from its before-tax cash flow (BTCF), and then calculate the internal rate of return of ATCF. Therefore, let us derive the after-tax rate of return IRR(ATCF) from the before-tax rate of return IRR(BTCF) by equation (7.7.2) below, just like the before-and-after-tax cost of the bank in equations (7.6.2) and (7.6.3), and the tax-exempt and taxable rates of return in equation (7.7.1).

$$\text{IRR(ATCF)} = (1 - \text{effective tax rate})*\text{IRR(BTCF)} \qquad ....(7.7.2)$$

However, equations (7.6.2), (7.6.3) and (7.7.1) apply exactly to financial alternatives where capital transactions are nondepreciable. Equation (7.7.2) approximates IRR(ATCF) from IRR(BTCF) for engineering alternatives involving only equity funding as shown by alternative (A) in Section 7.6. Alternatives (B) and (C) involve debt financing which makes equation (7.7.2) misleading for estimating IRR(ATCF) from IRR(BTCF). In order to satisfy equation (7.7.2) exactly, the cash flow description of alternative (A) in Section 7.6 which is repeated in Table 7.7.1 below needs to be reinterpreted as if it was a financial loan.

Table 7.7.1 - Cash flow description of engineering alternative (A) in Section 7.6 .

(1) EOY	(2) NOI	(2') CapTr	(3) Int	(4) BTCF	(5) Depr	(6) IncTax	(6') CapTax	(7) LoanCF	(8) ATCF
0	0	-45,000	0	-45,000	0	0	0	0	-45,000
1	15,700	0	0	15,700	-9,000	-2,814	0	0	12,886
2	15,700	0	0	15,700	-9,000	-2,814	0	0	12,886
3	15,700	0	0	15,700	-9,000	-2,814	0	0	12,886
4	15,700	0	0	15,700	-9,000	-2,814	0	0	12,886
5	15,700	0	0	15,700	-9,000	-2,814	0	0	12,886
∑	78,500	-45,000	0	33,500	-45,000	-14,070	0	0	19,430

IRR(BTCF) = 21.9575%/yr; 0.58*IRR(BTCF) = 12.7354%/yr; IRR(ATCF) = 13.2938%/yr.

In order for IRR(ATCF)= 13.2938%/year to equal 0.58*IRR(BTCF) = 12.7354%/year according to equation (7.7.2), the BTCF of engineering alternative (A) must be reinterpreted as if it was a financial loan as shown in Table 7.7.2 .

Table 7.7.2 - Reinterpreting the BTCF of engineering alternative (A) as a financial loan.

(1) EOY	(2) Int Rate	(3) ACF Bal	(4) BTCF	(5) BCF Bal	(6) IntBTCF	(6') CapBTCF	(7) IntAT	(8) ATCF
0	21.9575%	0	-45,000	45,000	0	-45,000	0	-45,000
1	21.9575%	54,881	15,700	39,181	9,881	5,819	5,731	11,550
2	21.9575%	47,784	15,700	32,084	8,603	7,097	4,990	12,087
3	21.9575%	39,129	15,700	23,429	7,045	8,655	4,086	12,741
4	21.9575%	28,573	15,700	12,873	5,144	10,556	2,983	13,539
5	21.9575%	15,700	15,700	0	2,827	12,873	1,639	14,512
∑		186,067	33,500	152,567	33,500	0	19,429	19,429

IRR(BTCF) = 21.9575%/yr; 0.58*IRR(BTCF) = 12.7354%/yr; IRR(ATCF) = 12,7354%/yr.

Columns 1 to 5 of Table 7.7.2 are the A' B' C' rules for present time equivalences of BTCF in column (4) discounted at the IRR(BTCF) = 21.9575%/year in column (2). The BTCF is then subdivided into interest and capital components (i.e., (4) = (6) + (6')) by multiplying BCF Bal in column (5) by the 21.9575% interest rate in the following year. Thus, the IntBTCF = \$9,881 at the end of year one is the product of BCF Bal = \$45,000 at the end of year zero and the 21.9575% interest rate during the first year. The CapBTCF = \$5,819 in column (6') is then obtained by subtracting IntBTCF = \$9,881 in column (6) from BTCF = \$15,700 in column (4) at the end of year one. The *interest-after-taxes* IntAT = \$5,731 in column (7) equals IntBTCF = \$9,881 multiplied by the 58% complement of the 42% tax rate. The ATCF = \$11,550 in column (8) equals the sum of columns (6') and (7). Columns (6) and (6') of Table 7.7.2 reinterprets the BTCF in column (4) as a financial loan as shown in columns (2) and (2') respectively of Table 7.7.3 below.

Table 7.7.3 - Reinterpreting the BTCF of engineering alternative (A) as a financial loan.

(1) EOY	(2) NOI	(2') CapTr	(3) Int	(4) BTCF	(5) Depr	(6) IncTax	(6') CapTax	(7) LoanCF	(8) ATCF
0	0	-45,000	0	-45,000	0	0	0	0	-45,000
1	9,881	5,819	0	15,700	0	-4,150	0	0	11,550
2	8,603	7,097	0	15,700	0	-3,613	0	0	12,087
3	7,045	8,655	0	15,700	0	-2,959	0	0	12,741
4	5,144	10,556	0	15,700	0	-2,160	0	0	13,539
5	2,827	12,873	0	15,700	0	-1,188	0	0	14,512
∑	33,500	0	0	33,500	0	-14,070	0	0	19,429

IRR(BTCF) = 21.9575%/yr; 0.58*IRR(BTCF) = 12.7354%/yr; IRR(ATCF) = 12,7354%/yr.

Tables 7.7.2 and 7.7.3 indicate that engineering alternatives involving equity financing only could be reinterpreted as financial loans where IRR(ATCF) could be estimated exactly from IRR(BTCF) with equation (7.7.2). But ranking engineering alternatives in descending order of their IRR(ATCF)'s could give distorted estimates of MARR after taxes.

Equation (7.7.2) is always satisfied by the cost of borrowing money, 9%/year before taxes and (1-0.42)*9% = 5.22%/year after taxes, independently of the amounts of debt financing in alternatives (A), (B) and (C). Also, ΔNPV(ATCF) = $10,451.98 for each of these three alternatives. Since leasing is an alternative to owning the $45,000 machine, let us determine an annual lease, **x**, for 5 years which also has ΔNPV(ATCF) = $10,451.98 . The annual after-tax cash flow from the lease would then be (1-0.42)*(15,700-**x**). Solving for the net-operating-income (15,700-**x**) after lease expenses from equation (7.7.3) below, we get 15,700-**x** = $4,187.66/year as shown in Table 7.7.4 below. If the lessor pays $45,000 for the machine in order to get $11,512.34/year, the IRR of the lessor would be 8.8105%/year before taxes excluding any salvage value of the machine after 5 years.

$$\$10,451.98 = (1-0.42)\cdot(15,700-\mathbf{x})\cdot(P/A,5.22\%,5) \quad ....(7.7.3)$$

Table 7.7.4 - Cash-flow of an equivalent lease of the $45,000 machine in Section 7.6 .

(1) EOY	(2) NOI	(2') CapTr	(3) Int	(4) BTCF	(5) Depr	(6) IncTax	(6') CapTax	(7) LoanCF	(8) ATCF
0	0	0	0	0	0	0.00	0	0	0.00
1	4,187.66	0	0	4,187.66	0	-1,758.82	0	0	2,428.84
2	4,187.66	0	0	4,187.66	0	-1,758.82	0	0	2,428.84
3	4,187.66	0	0	4,187.66	0	-1,758.82	0	0	2,428.84
4	4,187.66	0	0	4,187.66	0	-1,758.82	0	0	2,428.84
5	4,187.66	0	0	4,187.66	0	-1,758.82	0	0	2,428.84
∑	20,938.30	0	0	20,938.30	0	-8,794.10	0	0	12,144.20

ΔNPV(BTCF)@9% = 4,187.66 ·(P/A,9%,5) = $16,288.53 IRR(BTCF) not defined
ΔNPV(ATCF)@5.22% = 2,428.84 ·(P/A,5.22%,5) = $10,451.96 IRR(ATCF) not defined

The reason why the IRR(ATCF) of (B) and (C) are so much greater than that of (A) is that the 9% loan cash flow before taxes (or 5.22% after taxes) is subtracted in (B) and (C) from the after-tax cash flow of (A), thereby making their IRR(ATCF) greater than that of (A). Conventional economic decision-making does not permit borrowed money in time equivalence accounting because it makes investment alternatives appear better than they would have if they were evaluated on their own merits without any borrowing. It is argued that the opportunity cost of capital (MARR) is, by definition, the rate of return foregone by investing in the project under consideration. Therefore, the idea of borrowing money for a project at a lower rate of interest circumvents the concept underlying investment opportunity costs.

Section 7.8 - Summary of Chapter Seven

Accounting is the language of business and government enterprises. It is essential to understand the bookkeeping for project alternatives in regard to paying income taxes and making economic decisions. A basic document of financial accounting is the Balance Sheet (Section 7.1) which is a statement of the financial position of a firm at a given point of time. It balances the firm's assets with claims against those assets according to the equation

Assets = Liabilities + Owners' Equity

where "assets" are things of monetary value that the system possesses, and "liabilities" and "owners' equity" tell us who supplied those resources to the business and how much was supplied by each group of creditors or owners. The assets of the Balance Sheet always equals the total of liabilities and owners' equity or net worth because everything that a business owns has been supplied either by creditors or owners.

The assets and liabilities sections of the balance sheet are essentially the same for corporations as for single proprietorships and partnerships of two or more persons. However, significant differences appear in the owners' equity section of the balance sheet for corporations The net worth of corporate balance sheets is called Stockholders' Equity, which consists of Capital Stock and Retained Earnings. Capital originally paid in by stockholders is permanent capital not subject to withdrawal. The unit of corporate ownership is a share of stock which represents the share of net income to be distributed to stockholders as dividends. Retained Earnings are the accumulation of net incomes that have not been distributed to the stockholders as dividends since the inception of the business.

A corporation's profit is taxed before dividends are paid to stockholders. The part of the remaining profit that stockholders receive in the form of dividends must be declared as income on their individual income tax returns. Thus, corporate profit is really taxed twice. Corporate income taxes could be postponed by accumulating profits in the Retained Earnings account, but additional taxes are imposed on corporations "improperly accumulating surplus". Under certain conditions, corporations with 35 or fewer stockholders can elect to be taxed as partnerships. Such corporations are called Subchapter S Corporations. Ordinarily, all business incomes of partnerships and single proprietorships are subject to individual income taxes in the year that they are earned whether or not any business profits were withdrawn by the owners.

Financial activities during the time interval between successive balance sheets are documented periodically in the form of Income Statements (Section 7.2) which list revenues and expenses during each accounting period according to the equation

Revenues - Expenses = Net Income, or Net Profits

Income Statements use single-entry bookkeeping for revenues and expenses in each accounting period. Balance Sheets use double-entry bookkeeping at the beginning and end of each accounting period which is synchronized with Income Statements by the equation:

Assets = Liabilities + {Beginning Ownership + Revenues - Expenses}

At the top of the income statement is *gross revenues* which may include both *operating* and *nonoperating* income. Operating income refers to the sales of goods and services of the firm. Nonoperating income refers to revenues from cash deposits, marketable securities, and gains or losses on the sale of equipment. From gross revenues, we subtract various expenses which may also be classified as *operating* and *nonoperating*. Operating expenses consist of the *cost of goods sold* which includes material and energy costs, wages of workers, and certain overhead expenses, as well as the *cost of selling* which covers the cost of advertising and the marketing costs. Nonoperating expenses cover salaries of managers not directly engaged in production, depreciation allowances and interest on borrowed capital. The *net income before taxes* (NIBT) or *taxable income* is determined by subtracting all operating and nonoperating expenses from the gross revenues during the accounting period. After taxable income is computed, the applicable tax rate is used to compute the income taxes which are deducted from net income before taxes to arrive at the *net income after taxes* (NIAT). *Cash dividends* declared for various classes of stock are then deducted from NIAT and the remainder is added to the retained earnings of the balance sheet.

Accounting audits require revenues and expenses to be recorded on an "as is" basis without adjusting for the time value of money. When an *accrual basis of accounting* is used, revenues and expenses are recorded in the accounting period in which they are earned or incurred rather than the period they are collected or paid for in cash. Alternatively, under the *cash basis of accounting*, revenues and expenses are recorded only on receipt of cash or when cash expenditures are made. Most individual businesses maintain their accounting records on a cash basis, and corporations generally use an accrual basis. Because the cash basis of accounting ignores earned revenues which have not yet been received and incurred expenses which have not yet been paid, economic analysts prefer the accrual basis of accounting.

Because of the periodic need for financial accounting information, revenues and expenses are classified into the same accounting period even though they are not causally related. Time equivalences of a dollar at the beginning and end of an accounting period are not taken into account. Moreover, the financial accounting system is rigidly categorized into various types of operating and nonoperating revenues and expenses which may be suitable for income tax purposes, but which is rarely appropriate for the needs of economic decision-making. In particular, different methods of allocating operating expenses such as the *cost of goods sold*, and nonoperating expenses such as *depreciation allowances* and *interest on debt* can strongly affect income statements and balance sheets for many accounting periods.

Different methods of allocating the cost of goods sold are discussed in Section 7.3 under the heading of Inventory Accounting. Four methods of inventory pricing were considered, namely, (1) the FIFO method, (2) the LIFO method, (3) the specific identification method and (4) the average cost method. Different methods of allocating depreciation are discussed in Section 7.4 under the heading of Depreciation Accounting. Prior to the Economic Recovery Tax Act (ERTA) of 1981, the methods of depreciation allowances placed a strong emphasis on estimates of useful lives and salvage values of depreciable assets. The methods of depreciation used prior to 1981 are (1) Straight-Line (SL) depreciation, (2) Sum-of-Years-Digits (SOYD) depreciation, (3) Declining-Balance depreciation, (4) Sinking-Fund (SF) depreciation, (5) Cost and Percentage depletion, and (6) Multiple-Asset depreciation.

The Accelerated Cost Recovery System (ACRS) of depreciation was created by the Economic Recovery Tax Act of 1981 (ERTA). ACRS allows a business to write off the *cost basis* of depreciable property over a statutory *recovery period* which is usually much shorter than actual useful lives Intangible property is not subject to ACRS depreciation. The unadjusted cost basis assumes no salvage value and is normally the cash purchase price of a property plus the cost of making the asset serviceable. If actual salvage values were greater than the book value of the asset, a capital gains tax would be imposed on the disposed asset.

The property class lives of ACRS depreciation were developed from a 1970 U. S. Treasury Department study of the actual useful lives in which assets were utilized. In 1971, they published Asset Depreciation Range (ADR) guidelines of lower, upper and midpoint limits of useful lives for about 100 classifications of depreciable assets. The ADR midpoint lives were somewhat shorter than actual average useful lives. The ADR midpoint-life guidelines have been incorporated into ACRS classifications so that most property classes are again shorter than ADR midpoint lives. Appendix 7A provides the ADR guideline periods for selected classifications of depreciable personal property from which it is generally possible to determine a property's ACRS class life of 3, 5, 10 or 15 years.

For each ACRS class life, the percentage of the unadjusted cost basis of an asset that could be deducted from taxable income each year is shown in Table 7.46 below. Since depreciation allowances cannot b e accumulated, taxpayers could elect to use straight-line depreciation over longer periods of time when they expected small taxable incomes in early years and large taxable incomes in later years.

The modified ACRS (MACRS) rules created by the Tax Reform Act of 1986 (TRA) is now the principal means of writing off the cost basis of depreciable assets. The differences between ACRS and MACRS rules are mainly the class lives of assets and the methods of recovering their costs. Under MACRS, assets are assigned to one of six classes of depreciable personal property (3-, 5-, 7-, 10-, 15-, and 20-year classes) or to one of two classes of real property. The cost recovery rate on personal property is based on declining-balance depreciation switching to straight-line depreciation at the optimal time. The straight-line method must be used for all real estate property. A summary of MACRS class lives and permissible depreciation methods is provided in Appendix 7B.

The Tax Reform Act (TRA) of 1986 greatly altered federal income tax rates and also modified some of the rules for determining taxable income. The federal tax rate schedule on 1988 incomes of U. S. corporations and individuals as established by TRA is listed in Section 7.5 which also describes the Alternative Minimum Tax (AMT), tax rates on capital gains (or losses),the Investment Tax Credit (ITC), and combined Federal and State income tax rates.

Sections 7.6 and 7.7 of Chapter Seven are concerned with cash flow and managerial accounting and how they differ from financial accounting. Long-term capital expenditures may not all be expensed for income tax purposes in the year of acquiring the assets, but such equity and borrowed capital expenditures are nevertheless cash flows. Since depreciation is not a cash flow, it cannot be added or subtracted from cash flows. However, depreciation is used to determine taxable income from which we calculate income taxes that are cash flows. When borrowed capital is involved, the interest charges are cash flows which are deductible from taxable income, and the loan receipts and amortizations are cash flows that do not affect taxable income. It is shown that before-tax cash flows (BTCF) should be discounted with the cost of borrowing money before taxes, and after-tax cash flows (ATCF) should be discounted with the cost of borrowing money after taxes. The discount rate after taxes equals the discount rate before taxes multiplied by the complement of the tax rate. The net present value (NPV) of ATCF is then independent of the relative amounts of equity funding and debt financing.

It is shown in Section 7.7 that the internal rate of return of before-tax cash flows (IRR{BTCF}) from engineering alternatives could be evaluated as ordinary financial alternatives using IRR{BTCF} as the discount rate. The restructured cash flows of engineering alternatives would then have an after-tax internal rate of return (IRR{ATCF}) equal to the complement of the tax rate multiplied by the before-tax internal rate of return (IRR{BTCF}) whether equity funding or debt financing was used for capital acquisition. This suggests there is considerable room for questioning the applicability of internal-rate-of-return measurements for the evaluation of engineering alternatives.

Appendix 7A - Asset Depreciation Range (ADR) Guideline Periods

The Economic Recovery Tax Act (ERTA) of 1981 established an Accelerated Cost Recovery System (ACRS) of depreciation that could be used for the purposes of U. S. income taxes. ACRS depreciation uses (1) "property class lives" that are less than "actual useful lives", and (2) zero salvage values. The Tax Reform Act (TRA) of 1986 continued to use ACRS depreciation with a number of modifications. MACRS uses six personal property classes (3-, 5-, 7-, 10-, 15, and 20-years) instead of four (3-, 5-, 10-, and 15-years), and two real-estate property classes (27.5 years for residential rental property and 31.5 years for nonresidential real property) instead of one (15-years for all buildings).

ACRS and MACRS classes of depreciable property are based on a 1970 U. S. Treasury Department study of asset utilization. In 1971, they published guidelines for about 100 broad classifications of depreciable assets. Each classification has a midpoint, and lower and upper limits of useful life, called the Asset Depreciation Range, or ADR. The ADR midpoint lives are shorter than actual average useful lives, and they are used in the definition of ACRS and MACRS property classes. Table 7A.1lists the ADR guideline lives of selected assets.

Table 7A.1 - Selected Asset Depreciation Range (ADR) Guideline Lives

Description of Depreciable Assets	Guideline Life (Years)
Transportation	
Automobiles, taxis	3
Buses	9
Light-duty trucks	4
Heavy-duty trucks	6
Commercial air transport	12
Petroleum	
Exploration and drilling assets	14
Refining and marketing assets	16
Manufacturing	
Sugar and sugar products	18
Tobacco and tobacco products	15
Carpets and apparel	9
Lumber, wood products, and furniture	10
Chemicals and allied products	9.5
Cement	20
Fabricated metal products	12
Electronic components	6
Rubber products	14
Communication	
Central-office telephone buildings	45
Telephone poles, cables, etc.	34
Radio and television broadcasting	6
Electric utility	
Hydraulic plant	50
Nuclear plant	20
Transmission and distribution	30
Services	
Office furniture and equipment	10
Computers and peripheral equipment	6
Recreation - bowling alleys, theaters, etc.	10

Appendix 7B - MACRS Class Lives and Depreciation Methods

The first step in MACRS depreciation is to determine the class life and the depreciation method for the asset being depreciated from the ADR guideline periods outlined in Table 7B.1 .

Table 7B.1 - MACRS Class Lives and Depreciation Methods

3-year class life, 200% declining-balance depreciation - includes property with an ADR class life of 4 years or less. Excludes cars and light trucks.

5-year class life, 200% declining-balance depreciation - includes property with an ADR class life of more than 4 years to less than 10 years. Includes cars and light trucks, semiconductor manufacturing equipment, qualified technological equipment, computer-based cetral office switching equipment, some renewable and biomass power facilities, and research and development property.

7-year class life, 200% declining-balance depreciation - includes property with an ADR class life of more than 10 years to less than 16 years. Includes single-purpose agricultural agricultural and horticultural structures, railroad track, and property with no ADR midpoint.

10-year class life, 200% declining-balance depreciation - includes property with an ADR class life of more than 16 years to less than 20 years.

15-year class life, 150% declining-balance depreciation - includes property with an ADR class life of more than 20 years to less than 25 years. Includes sewage treatment plants, telephone distribution plants, and equipment for two-way voice and data communication.

20-year class life, 150% declining-balance depreciation - includes property with an ADR class life of 25 years or more. Excludes real property with ADR midpoint of 27.5 years or more. Includes municipal sewers.

27.5-year class life, straight-line depreciation - includes all residential rental property, regardless of the ADR class life.

31.5-year class life, straight-line depreciation - includes all nonresidential real-estate property, regardless of the ADR class life.

After assigning assets to one of six MACRS classes of depreciable personal property or to one of two MACRS classes of real-estate property described in Table 7B.1, the next step is to apply the corresponding depreciation method. The declining-balance depreciation methods that apply to the six classes of depreciable personal property all permit switching to straight-line depreciation when it is needed to optimize deductions. Under MACRS rules established in the Tax Reform Act of 1986, the half-year convention treats all depreciable personal property placed in service, or disposed of, during a tax year as having been placed in service, or disposed of, at the midpoint of that tax year. A half-year of depreciation is then allowed during the tax year the property is placed in service, regardless of when the property was actually placed in service. For each of the remaining years, a full year of depreciation is taken. If property is held for the entire recovery period, a half-year of depreciation is allowed for the tax year following the end of the recovery period. If property is disposed of before the end of the recovery period, only a half-year of depreciation is allowed for the year of disposal. The depreciation rates that result from the MACRS depreciation methods for depreciable personal property under the half-year convention is listed in Table 7B.2 .

Table 7B.2 - MACRS depreciation rates for personal property (Tax Reform Act of 1986)

Year	Class Life	3-years	5-years	7-years	10-years	15-years	20-years
1		33.33%	20.00%	14.29%	10.00%	5.00%	3.75%
2		44.45%	32.00%	24.49%	18.00%	9.50%	7.22%
3		14.81%	19.20%	17.49%	14.40%	8.55%	6.68%
4		7.41%	11.52%	12.49%	11.51%	7.70%	6.18%
5			11.52%	8.93%	9.22%	6.93%	5.71%
6			5.76%	8.92%	7.37%	6.23%	5.28%
7				8.93%	6.55%	5.90%	4.89%
8				4.46%	6.55%	5.90%	4.52%
9					6.56%	5.91%	4.47%
10					6.55%	5.90%	4.47%
11					3.28%	5.91%	4.46%
12						5.90%	4.46%
13						5.91%	4.46%
14						5.90%	4.46%
15						5.91%	4.46%
16						2.95%	4.46%
17							4.46%
18							4.46%
19							4.46%
20							4.46%
21							2.23%

The MACRS depreciation rates listed in Table 7B.2 are percentages of the original cost of the asset. Percentages of the original cost can be derived from the declining-balance percent depreciation rate of the current book value as shown in Table 7.43 in the following manner. Let us assume the original cost is $1,000.00, and we will use 200% declining-balance depreciation for an asset with a 5-year class life with the half-year convention. The 200% declining-balance depreciation rate for a 5-year class life is 200%/5 = 40% per year, but the rate is only 20% in the first year because of the half-year convention. At the beginning of the fourth year, the book value is $288.00 with 2.5 years remaining for cost recovery by straight-line depreciation. When EOY depreciation charges in column (4) of Table 7B.3 are expressed in percentage form, the results are the same as in Table 7B.2 .

Table 7B.3 - MACRS 200% Declining-Balance Depreciation for 5-year property

(1)MOY	(2)Depr Rate (DDB)	(3)BDC Book Value	(4)EOY Depr Charge	(5)ADC Book Value
1	20.00%	$1,000.00	-$200.00	$800.00
2	40.00%	$800.00	-$320.00	$480.00
3	40.00%	$480.00	-$192.00	$288.00
4	40.00%SL	$288.00	-$115.20	$172.80
5	40.00%SL	$172.80	-$115.20	$57.60
6	20.00%SL	$57.60	-$57.60	$0.00

If the aggregate cost bases of all personal property placed in service during the last 3 months of a tax year exceed 40% of all personal property placed in service during that tax year, a mid-quarter convention is used to calculate all depreciation deductions. Under a mid-quarter convention, all property placed in service, or disposed of, during any quarter of a tax year is treated as if it was placed in service, or disposed of, at the midpoint of the quarter. The depreciation rate for the full tax year is multiplied by 87.5%, 62.5%, 37.5% and 12.5% for the 1st, 2nd, 3rd and 4th quarters of the tax year the property is placed in service. Table 7B.4 lists MACRS depreciation rates for 3-, 5-, and 7-year property under the half-year and mid-quarter conventions. For most personal property, straight-line is applied with half-year convention, and a mid-month convention must be used for all real estate.

Table 7B.4 - 3-,5-,7-yr MACRS depreciation rates under half-year & mid-quarter conventions

Year	Half-Year	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
1	33.33%	58.33%	41.67%	25.00%	8.33%
2	44.45%	27.78%	38.89%	50.00%	61.11%
3	14.81%	12.35%	14.14%	16.67%	20.37%
4	7.41%	1.54%	5.30%	8.33%	10.19%
Year	Half-Year	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
1	20.00%	35.00%	25.00%	15.00%	5.00%
2	32.00%	26.00%	30.00%	34.00%	38.00%
3	19.20%	15.60%	18.00%	20.40%	22.80%
4	11.52%	11.01%	11.37%	12.24%	13.68%
5	11.52%	11.01%	11.37%	11.30%	10.94%
6	5.76%	1.38%	4.26%	7.06%	9.58%
Year	Half-Year	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
1	14.29%	25.00%	17.85%	10.71%	3.57%
2	24.49%	21.43%	23.47%	25.51%	27.55%
3	17.49%	15.31%	16.76%	18.22%	19.68%
4	12.49%	10.93%	11.97%	13.02%	14.06%
5	8.93%	8.75%	8.87%	9.30%	10.04%
6	8.92%	8.74%	8.87%	8.85%	8.73%
7	8.93%	8.75%	8.87%	8.86%	8.73%
8	4.46%	1.09%	3.33%	5.53%	7.64%

Although the faster write-offs of MACRS depreciation methods are generally beneficial for ongoing businesses, several situations arise in which slower write-offs are preferable. For example, suppose a new business is starting up in which low income or losses are expected in the beginning. However, depreciation is deducted annually. Even though the deduction may give no benefit because other deductions already exceed income in early years, the depreciation deduction may not be postponed for accumulation in high income years. Therefore, the faster write-offs of MACRS depreciation methods may waste depreciation deductions that could be used in later years when income increases.

Sometimes the faster write-offs of MACRS depreciation methods are not used because they adversely affect the net income reported in current financial statements. This could create difficulties in public offerings of stock, in raising more capital through loans, in images of management, etc. Also, when the 200% declining-balance rate is used for regular tax purposes, taxpayers may become subject to the alternative minimum tax (AMT). For AMT purposes, the 150% rate must then be used to adjust taxable income. To avoid this double accounting, the 150% rate may be elected for regular tax purposes.

If it is desired to write off depreciation at an even pace, one may irrevocably elect to use the straight-line method over the the regular MACRS recovery period, *or* to use the alternative straight-line method (without regard to salvage value) over the designated recovery period for the class life. For the purposes of the alternative straight-line rule, the designated recovery period for cars, light trucks, and computers is 5 years. For business furniture and fixtures, the alternative straight-line recovery period is 10 years, but instead an election may be made to use straight-line recovery over the regular 7-year MACRS recovery period. The recovery period for personal property with no class life is 12 years. For nonresidential real property and residential rental property, one may elect straight-line recovery over 40 years.

Appendix 7C - Accounting for Inflation

Inflation is defined as a general rise in prices brought about by increases in currency and credit relative to the amount of goods and services available. More succinctly, inflation is described as "too much money chasing too few goods". The general rise in prices during an inflationary period does not mean that all prices have the same rate of inflation. The prices of different items change at different rates, and the prices of some items may even decrease during an inflationary period. When prices generally fall relative to the amount of goods and services available, the situation is called *deflation*. History shows inflation is much more common than deflation, and prolonged periods of mild inflation are often called *stagflation*.

During inflation, the purchasing power of money develops more slowly than the earning power of money. Economists blame inflation on a wide spectrum of causes. *Cost-push* inflation is caused by the escalation of consumer prices at a faster rate than producers' costs are increasing. *Demand-pull* inflation is caused by excessive spending and borrowing of consumers who "buy ahead" in belief that prices will inflate and loans can be repaid in cheaper dollars. People on fixed incomes suffer the greatest loss of purchasing power, and wage earners experience "tax creep" when their inflated incomes push them into higher tax brackets. Depreciation allowances based on recovery of historical costs are usually much smaller than replacement costs. *War* inflation caused by military buildup reduces the amount of goods and services available for civilian consumption. International trade imbalances and governmental budget deficits are other major factors which have an appreciable impact on inflation.

In order to take inflation and deflation into account in economic decision-making, the cash flows of engineering and financial alternatives should first be described in terms of *actual* or *current* dollar units, $\$_N$, as of the time N that those cash flows occur. Actual dollar units incorporate allowances for expected price inflation or deflation. Cash flow descriptions in terms of actual dollar units must then be discounted to *real* or *present* dollar units, $\$_0$, in terms of which measurements of constant purchasing power are expressed.

Inflation makes a dollar $\$_N$ earned or spent N years in the future less valuable than a dollar $\$_0$ earned or spent at the present time. Historical data on price level changes for individual commodities, or composite groups of goods and services, are gathered and compiled into price indices by various agencies of the federal government as well as private organizations. The most familiar composite index, is the Consumer Price Index (CPI) which represents relative changes in retail prices for a "market basket" of consumer purchases of clothing, food, housing, transportation and utilities. The CPI index is designed to measure relative changes in retail prices required to maintain a fixed standard of living for the "average" consumer. Other well-known composite price indices gathered by federal agencies are the Producer Price Index (PPI) and the Implicit Price Index for the Gross National Product (IPI-GNP) which measure historical price-level changes within the United States Economy.

When we are attempting to reflect price changes for particular sets of engineering and financial alternatives, price indices should be selected to measure price changes that are relevant to the cash flow descriptions of those alternatives. However, the method of using a price index to measure how much a group of prices is changing is essentially the same for all price indices. Therefore, instead of dealing with a multitude of different price indices, only the CPI will be used as a typical example of the underlying methodology. As shown in Table 7C.1 below, the CPI has increased from a value of 100.0 in the base year of 1967 to a value of 340.4 in 1987. We will now explain the relationship between actual dollars, $\$_N$, and real dollars, $\$_0$, with respect to CPI measurements.

Table 7C.1 - Consumer Price Index (CPI) and Derived Percent Inflation Rate (g and g_c).

Year	CPI	g%	g_c%	Year	CPI	g%	g_c%
1967	100.0	2.900	2.859	1978	195.4	7.658	7.379
1968	104.2	4.200	4.114	1979	217.4	11.259	10.669
1969	109.8	5.374	5.235	1980	246.8	13.523	12.684
1970	116.3	5.920	5.751	1981	272.4	10.373	9.869
1971	121.3	4.299	4.209	1982	289.1	6.131	5.950
1972	125.3	3.298	3.244	1983	298.4	3.217	3.166
1973	133.1	6.225	6.039	1984	311.1	4.256	4.168
1974	147.7	10.969	10.408	1985	322.2	3.568	3.506
1975	161.2	9.140	8.746	1986	328.4	1.924	1.906
1976	170.5	5.769	5.609	1987	340.4	3.654	3.589
1977	181.5	6.452	6.252				

Price indices utilize an annual percentage rate to represent the relative increase or decrease in prices over a one-year time span. Since the annual percentage rate is based on the previous year's prices, the percentage rate has an annual compounding effect. Assuming discrete end-of-year compounding, the annual percent inflation rate, g, can be derived from the formula F = P(F/P,g,1) = P(1+g) of equation (3.5.1). For example, g = 3.654% in 1987 is obtained from 1+g = F/P = CPI(1987)/CPI(1986) 340.4/328.4 = 1.03654 . If we assume a constant inflation rate during the year, then the continuously compounded inflation rate, g_c, can be derived from the formula $F = P(F/P,g_c,1) = Pe^{g_c}$ of equation (3.5.2). In the previous example where F = 340.4 and P = 328.4, we get $g_c = \ln\{1.03654\} = 0.03589$ or g_c = 3.589%.

In accounting for inflation, it is convenient to use continuously rather than discretely compounded inflation rates. For example, suppose we are given CPI(1983)= 298.4 and the set of four annual percent inflation rates (both g and g_c) from 1984 to 1987. The problem is to determine CPI(1987) and the average annual inflation rate from 1984 to 1987.

<u>Calculations based on percent inflation rates with discrete annual compounding</u>:
CPI(1987)=298.4(1.04256)(1.03568)(1.01924)(1.03654) = 298.4*1.14075 = 340.4
Average g (1984-1987) = $(1.14075)^{1/4}$ = 1.03347 or Average g = 3.347%/yr

<u>Calculations based on percent inflation rates with continuous annual compounding</u>:
CPI(1987) = $298.4e^{(0.04168+0.03506+0.01906+0.03589)} = 298.4e^{0.13169}$= 340.4
Average g (1984-1987) = $e^{(0.13169/4)}$ = 1.03347 or Average g = 3.347%/yr

The CPI is the most widely used measure of general inflation. Its importance is emphasized by its use for adjusting income payments to wage earners, social-security beneficiaries, food-stamp recipients, and retirees. It also forms the eligibility basis in many welfare programs.

Therefore, it is important for both borrowers and lenders to adjust real interest rates for their expectations of inflation.

Inflation helps long-term borrowers of money by enabling them to repay their debt in dollars with reduced purchasing power. Conversely, inflation is unfavorable to long-term lenders of money. Deflation affects borrowers and lenders in an opposite way than inflation. Therefore, it is important for both borrowers and lenders to adjust real interest rates for their expectations of either inflation or deflation.

In the case of discrete annual compounding, the annual interest rate i* adjusted for inflation is related to real interest rate i and annual inflation rate g by the formula of Section 3.8 .

$$1+i^* \equiv (1+i)/(1+g) \quad \text{or} \quad i^* \equiv (i-g)/(1+g)$$

In the case of continuous compounding, the interest rate adjustment for inflation is $r-g_c$.

Chapter Seven - Exercises

7-1a A plant manager wants equipment that costs $22,000 and has a 5-year life with no salvage value. The equipment would provide an EOY savings of $7,600/yr before taxes. Use straight-line depreciation and an effective income tax rate of 34% per year. If the equipment was acquired by 100% equity funding, use the table below to determine the cash flows, NPV and IRR before and after taxes, using a cost of borrowing money = 9%(BT) and 5.94%(AT) per yr, and MARR = 12%(BT) and 7.92%(AT) per year. (see Table 7.7.1)

(1)EOY (2)NOI (2')CapTr (3)Int (4)BTCF (5)Depr (6)IncTax (6')CapTax (7)LoanCF (8)ATCF

0____

_

1____

_

2____

_

3____

_

4____

_

5____

_

NPV(BTCF)@9.00%:	NPV(BTCF)@12.00%:	IRR(BTCF):
NPV(ATCF)@ 5.94%:	NPV(ATCF)@ 7.92%:	IRR(ATCF):
IRR(BTCF)	IRR(ATCF)	0.66*IRR(BTCF)

7-1b Suppose the equipment was acquired using $8,000 equity funds and borrowing $14,000 at 9%/yr interest to be paid back with interest and equal EOY principal reductions of $2,800 for 5 years. Use the table below to determine the cash flows, NPV and IRR before and after taxes under the same financial conditions as in problem **7-1a**.

(1)EOY (2)NOI (2')CapTr (3)Int (4)BTCF (5)Depr (6)IncTax (6')CapTax (7)LoanCF (8)ATCF

0____

_

1____

_

2____

_

3____

_

4____

_

5____

_

NPV(BTCF)@9.00%:	NPV(BTCF)@12.00%:	IRR(BTCF):
NPV(ATCF)@ 5.94%:	NPV(ATCF)@ 7.92%:	IRR(ATCF):
IRR(BTCF)	IRR(ATCF)	0.66*IRR(BTCF)

7-1c Restructure the BTCF = (2) + (2') + (3) of the engineering alternative in problem **7-1a** as a financial loan at the IRR(BTCF) discount rate. (See Table 7.7.2)

EOY	Int Rate	ACF Bal	BTCF	BCF Bal	Int BTCF	Cap BTCF	Int AT	ATCF
0								
1								
2								
3								
4								
5								

IRR(BTCF)	IRR(ATCF)	0.66*IRR(BTCF)

7-1d Use the tables of problem **7-1a** to show that the restructured BTCF of problem **7-1c** satisfies equation (7.7.2). (See Table 7.7.4)

(1)	(2)	(2')	(3)	(4)	(5)	(6)	(6')	(7)	(8)
EOY	NOI	CapTr	Int	BTCF	Depr	IncTax	CapTax	Loan CF	ATCF
0									
1									
2									
3									
4									
5									

IRR(BTCF)	IRR(ATCF)	0.66*IRR(BTCF)

7-1e Determine an equivalent ΔNPV(ATCF) 5-year EOY uniform lease payments for the $22,000 machine and construct a cash flow analysis similar to Table 7.7.4 on page 185. If the lessor pays $22,000 for the machine, estimate the IRR of the lessor before taxes excluding any salvage value of the machine after 5 years.

Chapter Eight - Single Variable Optimization

Section 8.1 - Breakeven Analysis

In this chapter, we are dealing with economic problems that have a single variable over which we exercise some control by our decision-making. We will study the effects of varying the control variable upon the magnitude of an objective quantity of the problem. An optimal decision can then be reached by finding the value of the control variable which will either maximize or minimize the objective quantity of the problem.

Breakeven analysis is a study of cost-volume-profit relationships. The control variable is the output volume per unit of time. The breakeven point is the output volume per unit of time in which the firm has zero profits (i.e., the output rate where total revenues equal total costs, resulting in zero profits). If the firm has a higher (or lower) output rate than the breakeven point, the firm realizes a profit (or incurs a loss). Most breakeven analyses use linear approximations which apply over a limited period of time and a limited range of output rates in the neighborhood of the breakeven point.

The total output volume per unit of time may be expressed either as a rate of physical output, Q, or as a rate of dollar output, $Q, or as a percent of production capacity, %Q. The choice of Q, $Q, or %Q as the control variable depends upon the nature of available information. Managerial accountants often apply breakeven analysis when information is available from balance sheets and income statements. In such cases, the control variable would be $Q and possibly %Q. When breakeven analysis is applied to physical production problems, the control variable would be Q with which we will start our description of breakeven analysis.

In a unit period of time, the sale of Q units of output at a price of 'p' per unit yields total revenues of TR (i.e., TR = p•Q). The total cost TC of producing Q units of output may be divided into variable costs VC and fixed costs FC (i.e., TC = VC + FC). The variable cost per unit of output v = VC/Q tends to remain constant as the volume of production Q varies in the vicinity of the breakeven point. Thus, the variable costs VC = v•Q would be proportional to Q. Fixed costs do not change in total, but the fixed cost per unit of output FC/Q becomes smaller as the output per unit period of time, Q, increases.

The profit Z may now be expressed as an equation Z = TR - TC = p•Q - v•Q - FC = (p - v)•Q - FC. Solving for Q, we get

$$Q = (Z + FC)/(p - v) \quad \text{....(8.1.1)}$$

$$Q_B = FC/(p - v) \quad \text{....(8.1.2)}$$

which is depicted on the left-hand side of Figure 8.1.1 below. The denominator (p - v) of (8.1.1) and (8.1.2) represents the marginal contribution to profit of a unit increase in output. Profit Z is maximized by setting its derivative with respect to Q equal to zero. Equation (8.1.3) shows profit Z is maximized when *marginal revenues* dTR/dQ = p equals *marginal costs* dVC/dQ = v. Substituting p-v = 0 in (8.1.1) shows Q = ∞ when Z is maximized.

$$\frac{dZ}{dQ} = \frac{dTR}{dQ} - \frac{dVC}{dQ} - \frac{dFC}{dQ} = 0 \quad \text{or} \quad \frac{dTR}{dQ} = \frac{dVC}{dQ} \quad \text{....(8.1.3)}$$

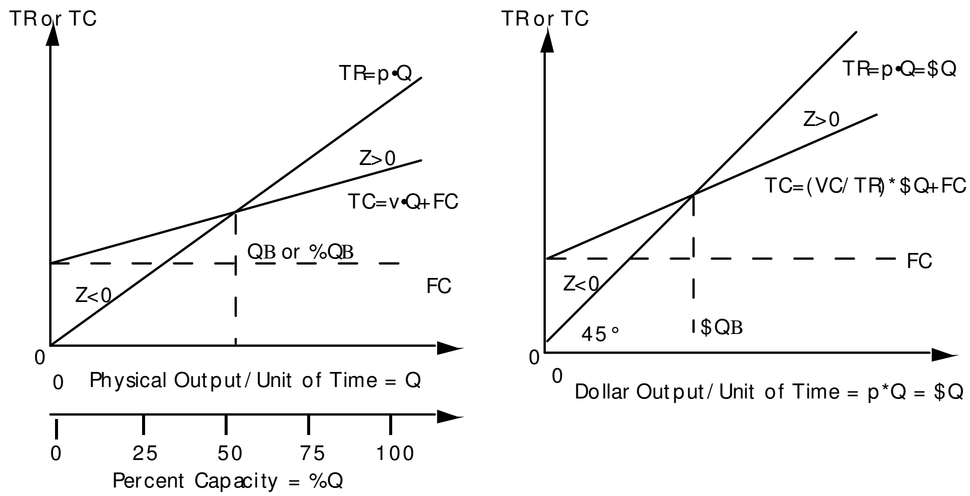


Figure 8.1.1 - Graphs of Cost-volume-profit Relationships.

The firm's income statement often provides data on TR, VC, FC, and Z, but not on Q, p, and v. Equations (8.1.1) and (8.1.2) can be adapted to such information by multiplying both sides with p, and multiplying both numerator and denominator of v/p by Q. Since $p \bullet Q = \$Q$ and $p \bullet Q_B = \$Q_B$ by definition, we get equations (8.1.4) and (8.1.5) below.

$$\$Q = (Z+FC)/\{1-(v/p)\} = (Z+FC)/\{1-(v \bullet Q)/(p \bullet Q)\} = (Z+FC)/\{1-(VC/TR)\} \quad ....(8.1.4)$$

$$\$Q_B = FC/\{1-(v/p)\} = FC/\{1-(VC/TR)\} \quad ....(8.1.5)$$

The denominator {1-(VC/TR)} is the marginal contribution to profits of a dollar increase of output. The breakeven analysis for \$Q and $\$Q_B$ is shown on the right of Figure 8.1.1.

<u>Problem</u>: When the 1988 Income Statement of ABC Inc. below was compiled, the manager suggested an increase of annual fixed costs by increasing advertising outlays \$100,000 per year, as a result of which sales could be expected to increase substantially. Assuming the suggestion is adopted and the ratio of variable costs to sales revenues remains constant, find the annual dollar sales that are needed (a) to break even, and (b) to make an annual profit of \$100,000.

Revenues:	Sales	\$1,200,000
Expenses:	Variable Costs	\$750,000
	Fixed Costs	\$500,000
Profit (or Loss):		(\$50,000)

Solution: (a) FC = \$500,000 + \$100,000 = \$600,000
{1-(VC/TR)} = {1-(\$750,000/\$1,200,000)} = 0.375
$\$Q_B$ = FC/{1-(VC/TR)} = \$600,000/0.375 = \$1,600,000

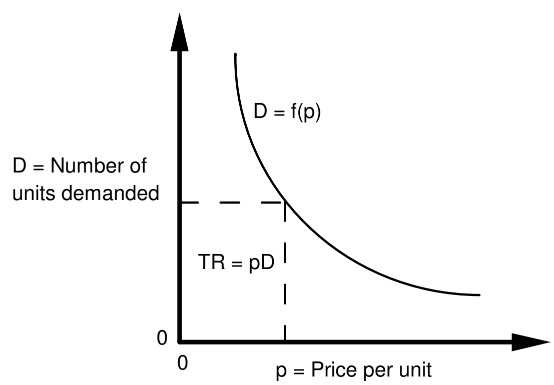
(b) Z+FC = \$100,000 + \$600,000 = \$700,000
{1-(VC/TR)} = {1-(\$750,000/\$1,200,000)} = 0.375
\$Q = (Z+FC)/{1-(VC/TR)} = \$700,000/0.375 = \$1,866,667

Section 8.2 - Demand Analysis

In the study of breakeven analysis, the volume of production, represented by output rate Q, was treated as a single variable over which was controlled by our decision-making. However, the volume of production supplied to consumers tends to be in a price equilibrium with the amount that consumers demand as expressed in the laws of supply and demand. If the price per unit of product is high, producers would be willing to supply more but consumers would be willing to demand less. Conversely, if the price is low, consumers would be willing to demand more but producers would be willing to supply less. Because of these price mechanisms, consumer demand in a given period of time can be treated as a single variable over which we exercise some control by the price we charge per unit of product.

The demand function, D = f(p), shown as a curve in Figure 8.2.1 below, expresses the relationship of a firm to its market by giving the demand for a product which can be sold, denoted by D, as a function of the price charged per unit of product, denoted by p. Empirical demand curves indicate how much product consumers would be willing to buy at a given price. Reasonably good estimates of a firm's demand curves are obtainable from market surveys, sales data, and other techniques of demand analysis. The information derived from such a demand curve indicates how much product could be sold at a given price and the total revenue which would be obtained from that volume of sales and production.

Figure 8.2.1 - Hypothetical demand curve.



The characteristics of demand curves are described in the economic literature by means of *elasticities*. Therefore, it is necessary to understand the meaning of elasticity in order to read about demand analysis in the literature, quite aside from the conceptual advantages derived from its use. Elasticity μ of demand with respect to price is defined as

$$\mu = \frac{\text{relative change in demand}}{\text{relative change in price}} = \frac{-dD / D}{dp / p} \quad \text{....(8.2.1)}$$

where D is the number of units demanded and p is the price per unit. Since relative changes of demand and price are both dimensionless, the elasticity ratio μ is also dimensionless.

When demand for a product is sensitive to price, the demand for the product is said to be *elastic* which is indicated by $\mu > 1$. With elastic demand, a 1% decrease in price would increase demand by <u>more</u> than 1%, so that the total revenue TR = pD would <u>increase</u>. When prices have little or no effect on total revenues from the sale of a product, the demand for the product is said to have *unit elasticity* (i.e., $\mu = 1$). When demand for a product is insensitive to price, the demand for the product is said to be *inelastic* (i.e., $\mu < 1$). With inelastic demand, a 1% decrease in price would increase demand by less than 1%, so that the total revenue TR = pD would decrease. Demand elasticity properties are conveniently summarized by differentiating TR = pD with respect to D as follows:

$$\frac{dTR}{dD} = p + D\frac{dp}{dD} = p + p\frac{dp/p}{dD/D} = p[1-(1/\mu)] \qquad ....(8.2.2)$$

Thus, when demand is elastic and $\mu > 1$, then $dTR/dD > 0$ and TR increases with demand. When demand has unit elasticity, then $dTR/dD = 0$ and TR is constant with increasing demand. Lastly, when $\mu < 1$ and demand is inelastic, then $dTR/dD < 0$ and TR decreases with increasing demand.

The profit Z may now be expressed in equation form as total revenue TR(D) minus total costs TC(D). The total costs TC are again divided into two classes, variable costs VC(D) and fixed costs FC. As a result, we have Z = TR(D) - VC(D) - FC. The profit Z is maximized by setting its derivative with respect to D equal to zero as follows:

$$\frac{dZ}{dD} = \frac{dTR}{dD} - \frac{dVC}{dD} - \frac{dFC}{dD} = 0 \quad \text{or} \quad \frac{dTR}{dD} = \frac{dVC}{dD} \qquad ....(8.2.3)$$

Equations (8.1.3) and (8.2.3) show that profit Z is maximized when *marginal revenues* $dTR/dD = p[1-(1/\mu)]$ equal *marginal costs* $dVC/dD = v$. However, the linear approximations of both TR and TC in equation (8.1.3) require the output rate Q to be infinite in order to maximize profit Z. More realistically, total revenues TR do not increase linearly with output rate Q because the price p declines when the demand becomes very large. Consequently, profit is maximized when the output volume is finite. The rest of this chapter similarly makes more realistic extensions to the linear breakeven analysis presented in Section 8.1 .

Section 8.3 - Marginal and Average Revenues and Costs

In order to maximize profits in breakeven analysis, marginal revenues should equal marginal costs. However, breakeven analyses are only applicable over a limited period of time. In the long run, various alternatives arise in connection with different possible scales of a project. Scale alternatives are usually coupled to time alternatives because the scale of a project can be increased only gradually over a period of time.

Opportunities to increase the scale of a project form a set of *mutually inclusive alternatives* whose input rates X(t) and output rates Y(t) have an S-shape growth curve in the course of time t as depicted in Figure 8.3.1 which is plotted from the data of Table 8.3.1. The Y-axis of Figure 8.3.1 represents the output rate Y(t) from an input rate X(t) at time t of a fixed proportion of capital-and-labor as a single variable. The startup of a project does not begin at the origin because substantial input is usually required before obtaining any output. At the lower part of the S-curve, the output first increases slowly, but then increases more rapidly relative to increasing input. As the project matures in the upper part of the S-curve, increments of output begin to diminish relative to increments of input. Thus, the S-curve spans a wide range of economically significant stages in the development of a project.

Figure 8.3.1 - Set of Mutually Inclusive Scale Alternatives

Total (Y), Average (Y/X), and Marginal (dY/dX) Output Rates from Input Rate X(t)

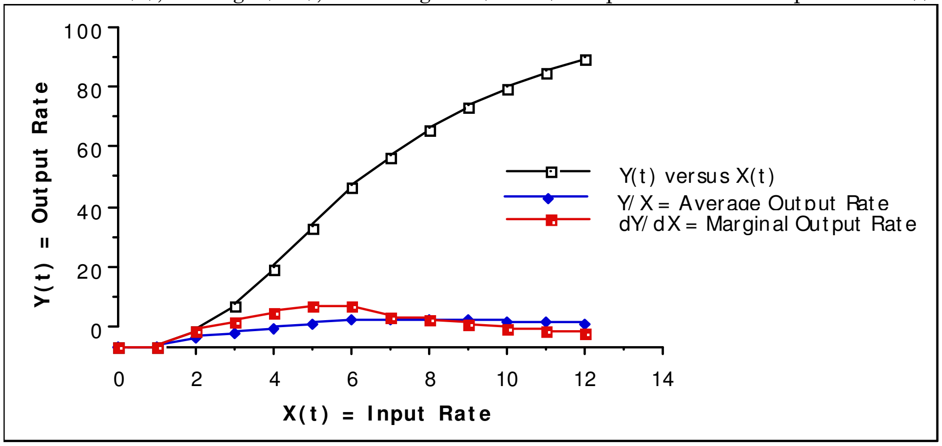


Table 8.3.1 - Set of Discrete Mutually Inclusive Scale Alternatives
Rates of Total Output Y, Average Output Y/X, Marginal Output dY/dX, and Change of Marginal Output d^2Y/dX^2 from Input Rate X(t).

X	Y	Y/X	dY/dX	d^2Y/dX^2
1	0.0	0.00	0.0	5.0
2	5.0	2.50	5.0	3.0
3	13.0	4.33	8.0	3.0
4	24.0	6.00	11.0	2.0
5	37.0	7.40	13.0	0.0
6	50.0	8.33	13.0	-3.5
7	59.5	8.50	9.5	-2.0
8	68.0	8.50	8.5	-1.0
9	75.0	8.33	7.0	-1.0
10	81.0	8.10	6.0	-1.0
11	86.0	7.82	5.0	-1.0
12	90.0	7.50	4.0	- -

Although Table 8.3.1 lists discrete scale alternatives, it is simpler to view Figure 8.3.1 as a continuous S-curve which can be subdivided into three regions. The lower part of the S-curve is concave upward where the slope of the curve is increasing. The middle part of the S-curve has little or no curvature where the slope of the curve is relatively constant. The upper part of the S-curve is concave downwards where the slope of the curve is decreasing.

The *average output rate* Y/X of input rate X is defined as the output rate Y divided by the input rate X. Equivalently, the average output rate Y/X is given by the slope of the position vector from the origin to the point (X,Y) on the curve. The average output curve of Figure 8.31 was plotted from data in the first and third columns of Table 8.3.1 .

The *marginal output rate* dY/dX of input rate X is defined as the change of output rate Y along the S-curve divided by the corresponding change of input rate X. Equivalently, the marginal output dY/dX of input rate X is given by the slope of the tangent to the curve at the point (X,Y). The marginal output curve of Figure 8.3.1 was plotted from data in the first and fourth columns of Table 8.3.1 .

Proposition 1: If the marginal output rate dY/dX is greater than the average output rate Y/X, then the average output rate Y/X is increasing as input rate X(t) increases. Proof:

We have to show that d(Y/X)/dX > 0 when dY/dX > Y/X Differentiation yields d(Y/X)/dX = [(dY/dX) - (Y/X)]/X which is greater than zero when dY/dX > Y/X as stated in Proposition 1. Thus, the marginal output dY/dX is greater than the average output Y/X from X = 2 to X = 7 in Figure and Table 8.3.1 .

Proposition 2: If the marginal output rate dY/dX equals the average output rate Y/X, then the average output rate Y/X is a maximum. When the tangent to the curve coincides with the position vector from the origin in Figure 8.3.1, the average output rate Y/X is a maximum. This occurs at X = 8 where the marginal equals the average output curve.

Proposition 3: If the marginal output rate dY/dX is less than the average output rate Y/X, then the average output rate Y/X is decreasing as input rate X(t) increases. Proof:

We have to show that d(Y/X)/dX < 0 when dY/dX > Y/X Differentiation yields d(Y/X)/dX = [(dY/dX) - (Y/X)]/X which is less than zero when dY/dX < Y/X as stated in Proposition 3. Thus, the marginal output dY/dX is smaller than the average output Y/X from X = 9 to X = 12 in Figure and Table 8.3.1 .

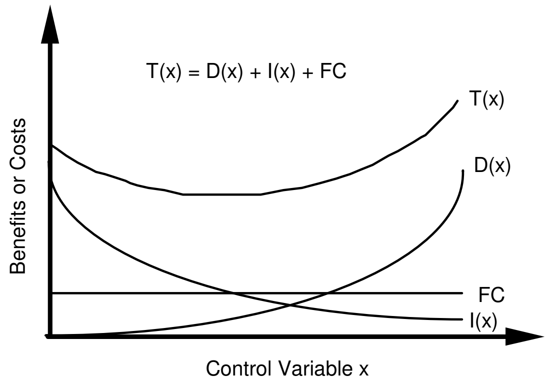
The inflection point of the S-curve occurs in the middle part of the curve where the slope of the curve is relatively constant. From a mathematical viewpoint, the inflection point occurs where the S-curve has no change of slope as indicated by the second derivative d^2Y/dX^2 being equal to zero. The marginal output curve has a maximum at X = 5 in Figure and Table 8.3.1 . The inflection point has little economic significance in economic decision-making as compared to the marginal and average output relationships discussed here. Other important applications of average and marginal relationships are presented in Chapter Nine where the input factors to the production process have variable rather than fixed proportions.

Section 8.4 - Kelvin Law Problems

Many problems in engineering economics, inventory control, production scheduling, and operations analysis concern a single control variable x. As x increases, directly varying costs D(x) tend to increase, inversely varying costs I(x) tend to decrease, and fixed costs FC tend to remain constant as depicted in Figure 8.4.1 below. The problem is to determine x which either minimizes input for a given output or maximizes output for a given input.

The earliest attempts to solve such problems are attributable to British physicist and mathematician Sir William Thomson (Lord Kelvin). In particular, Kelvin tried to find the optimal cross-sectional area x of a wire conductor to be used for supplying a distant location with a given amount of electrical energy per year for a number of years. If x was too large, then the capital recovery costs D(x) would also be too large. If x was too small, then the cost of energy transmission losses I(x) would be too large. Hence, the optimal cross-sectional area x which would minimize the cost must lie somewhere in between.

Figure 8.4.1 - Kelvin Law Optimization Model



Let us first formulate the general problem before solving any specific Kelvin Law problem. Let $T(x) = D(x) + I(x) + FC$ denote the total annual cost or benefit. The <u>necessary condition</u> for T(x) to be either a minimum or maximum is $T'(x) = 0$.

$$T'(x) = D'(x) + I'(x) = 0, \quad \text{or} \quad D'(x) = -I'(x) \qquad ....(8.4.1)$$

Thus, the optimum value of T(x) occurs at the value of x where the slopes of the D(x) and I(x) curves are equal in magnitude and opposite in sign. In the special case where $D(x) = Ax$ and $I(x) = B/x$, we get $D'(x) = A$ and $I'(x) = -B/x^2$. Equation (8.4.1) then implies

$$A = -(-B/x^2), \quad \text{or} \quad x_{opt} = \sqrt{B/A} \qquad ....(8.4.2)$$

Since $D(x_{opt}) = A\sqrt{B/A} = \sqrt{AB}$ and $I(x_{opt}) = B/\sqrt{B/A} = \sqrt{AB}$, it follows that the optimum value of T(x) occurs not only where D(x) and I(x) curves have equal slopes of opposite signs, but also where D(x) and I(x) curves intersect (i.e., $D(x_{opt}) = I(x_{opt}) = \sqrt{AB}$).

In order to distinguish between a minimum and a maximum when $T'(x_{opt}) = 0$, the slopes at $T(x_{opt})$ need to be investigated. The T(x) curve at x_{opt} would be a <u>minimum</u> if its slopes changed from negative to positive, and it would be a <u>maximum</u> if its slopes changed from positive to negative. Accordingly, the <u>sufficient conditions</u> for $T(x_{opt})$ to be either a <u>minimum</u> or a <u>maximum</u> is given by the sign of the second derivative $T''(x_{opt})$.

$$T''(x_{opt}) = D''(x_{opt}) + I''(x_{opt}) > 0 \quad \text{for a minimum} \qquad ....(8.4.3)$$

$$T''(x_{opt}) = D''(x_{opt}) + I''(x_{opt}) < 0 \quad \text{for a maximum} \qquad ....(8.4.4)$$

In the special case where $D(x) = Ax$ and $I(x) = B/x$, $D''(x) = 0$ and $I''(x) = 2B/x^3$ so that

$$T''(x) = D''(x) + I''(x) = 0 + 2B/x^3 \quad \text{and} \quad T''(x_{opt}) = 2B(B/A)^{-3/2} > 0$$

which guarantees the solution is a minimum when $A, B > 0$. If the necessary condition for a minimum or maximum of T(x) is not satisfied, then boundary values of x must be investigated for an optimal solution.

Example 8.4.1 - Economical Size of an Electrical Conductor

Let us first review the units of electrical measurement. Ohm's law states that the current I flowing through an electrical conductor is proportional to the electromotive driving force E and inversely proportional to the resistance R (i.e., $I \approx E/R$). When I is measured in amperes and E in volts, then the resulting measurement of R is in ohms.

The power loss or rate of energy loss of electrical current flowing in a conductor is measured by the product of driving force E and current flow I. The product of E volts and I amperes equals W watts which are the units for measuring power loss. Upon substituting Ohm's Law in the product EI, we find that power loss can also be measured by I^2R and E^2/R. The choice of EI, I^2R, or E^2/R for measuring power loss depends upon the information given in a problem. Since a watt is a small rate of energy loss, it is customary to use kilowatts KW which are 1,000 watts. Multiplying the rate of energy loss KW by the time in hours H of current flow gives the amount of energy loss KWH expressed in kilowatt-hours.

The resistance of an electrical conductor is usually given in ohms per 1,000 feet of wire 0.001 inches in diameter which is defined as a cross-sectional area of one circular mil. If a wire is x circular mils in cross-section, then its 1,000-foot circular-mil (CM) resistance needs to be divided by x in order to determine its actual 1,000-foot resistance.

Problem - An electrical conductor is needed to carry 25 amperes for 5,400 hours per year. Electrical energy costs $0.048 per KWH. Copper wire costs $0.75 cents per pound, and 1,000 feet of one circular-mil wire weighs 0.00302 pounds and has 10,580 ohms resistance at 20 degrees Centigrade. The life of the project is 25 years with zero salvage value. Capital recovery costs are 12% per year. Find the most economical cross-sectional area x_{opt} for the electrical conductor. (Adapted from E. L. Grant and W. G. Ireson, Principles of Engineering Economy, 5th Ed., Ronald Press, 1970).

Solution: x = cross-sectional area of wire in circular mils (CM).
D(x) = annual capital recovery cost per 1,000 feet of wire ($/yr/Kft).
I(x) = annual operating cost per 1,000 feet of wire ($/yr/Kft).
FC = fixed cost ($/yr/Kft)

(lbs/CM/Kft) * CM * ($/lb) * (($/yr)/$) = ($/yr/Kft)

$$D(x) = (0.00302) * (x) * (0.75) * (A/P,12,25) = 2.88787*10^{-4}\, x$$

(amps)² * (ohms/Kft) * (KW/watt) * (Hrs/yr) * ($/KWH) = ($/yr/Kft)

$$I(x) = (25)^2 * (10{,}580/x) * (1/1{,}000) * (5{,}400) * (0.048) = 1{,}713{,}960/x$$

Since D(x) = Ax and I(x) = B/x where $A = 2.88787*10^{-4}$ and B = 1,713,960 as described in the special case above, it follows that $x_{opt} = (B/A)^{1/2} = (1{,}713{,}960/2.88787*10^{-4})^{1/2} =$ 77,039 circular mils. Although the cross-sectional area x was treated here as a continuous variable, conductors are only available in discrete sizes. By examining a table of American Wire Gage (AWG) sizes, the closest available conductors are AWG Nos. 2 and 1. The table below shows that AWG No. 1 is the most economical, commercially available, wire size.

x	AWG #2 = 66,370CM	x_{opt} = 77,039CM	AWG #1 = 83,690CM
D(x)	$19.17	$22.25	$24.17
I(x)	$25.82	$22.25	$20.48
T(x)	$44.99	$44.50	$44.65

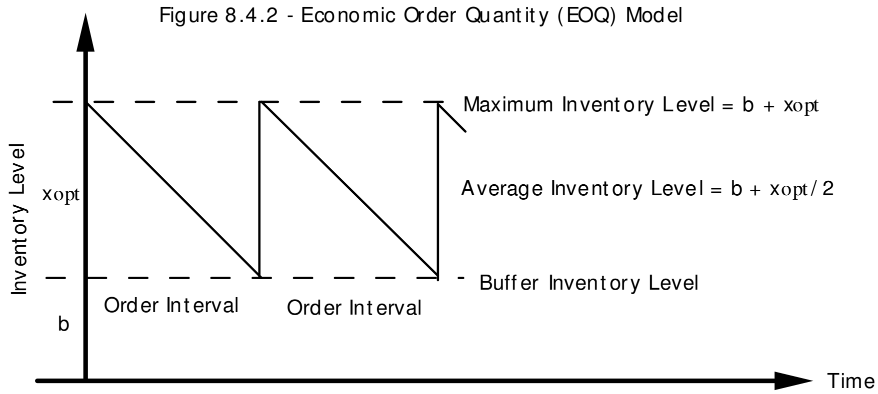
Example 8.4.2 - Economic Order Quantity (EOQ)

Inventory control concerns the process of ordering and storing quantities of stock in order to meet demand requirements. Usually a certain buffer level of stock is maintained so that there is enough stock on hand to meet unexpected demand requirements or supply delays. While demand must be met, inventory maintenance costs should be minimized.

In inventory control, the relevant costs include the cost of *ordering*, the cost of *handling* orders by placing them in inventory from which they are sold, and the cost of *holding* inventory for a given period of time. As the lot size x (i.e., order quantity) increases, the costs of handling and holding more materials in inventory increases, but the costs of ordering decreases because fewer orders are needed to meet the demand in a given period of time. The economic order quantity x_{opt} minimizes the sum of directly varying handling and holding costs and inversely varying ordering costs in a given time period.

Problem - A distributor sells 8,000 gallons of paint per year at an average price of \$6.30 per gallon. The costs of ordering are \$300 per order. Handling costs are prorated at \$0.42 per gallon per year based on the maximum inventory level in the warehouse. A buffer inventory of 400 gallons is maintained at all times. Holding costs (insurance and interest charges) are 12% per year based on the value of the average inventory level in the warehouse. The problem is to find the economic order quantity. (Adapted from J. L. Riggs, Engineering Economics, 2nd Ed., McGraw-Hill Book Company, 1982).

Figure 8.4.2 - Economic Order Quantity (EOQ) Model



Solution: x = order quantity (gals/order); x = gallons of inventory.
D(x) = annual costs of handling and holding gallons of inventory (\$/yr).
I(x) = annual costs of ordering (\$/yr).

	(gals)	*	(\$/gal/yr)	+	(gals)	*	(\$/gal/yr)	=	(\$/yr)
D(x) =	(x+400)	*	(0.42)	+	(x/2)	*	(6.30*0.12)	=	(0.798x+168)

	(\$/order)	*	(gals/yr)	/	(gals/order)	=	(\$/yr)
I(x) =	(300)	*	(8,000)	/	(x)	=	(2,400,000/x)

$x_{opt} = \sqrt{B/A} = \sqrt{2,400,000/0.798} = 1,734$ gallons per order.
Ordering frequency = 8,000(gals/yr)/1,734(gals/order) = 4.56(orders/yr).
Ordering interval = 4.56^{-1} (yrs/order) * 12(months/yr) = 2.63(months/order)

Example 8.4.3 - Economic Order Quantity (EOQ) with Price Discounts

An important variation of EOQ problems is that price discounts may be offerred by suppliers to encourage large orders which make possible large-scale economies of production and transportation. With price discounts, the effects of discounting on the purchase cost of the materials must be included in the fixed cost of meeting demand. Although larger price breaks lower unit costs, the customer incurs greater risks of keeping larger inventories.

Problem - A firm sells 400,000 electronic part units per year. Ordering costs are $25 per order. Handling and holding costs of storing each item in inventory per year are $0.06 + 0.2*P where P is the price per unit based on average inventory. Price discounts are:

P1 = $0.50 per unit for $x < 10{,}000$ units
P2 = $0.47 per unit for $10{,}000 \le x < 25{,}000$ units
P3 = $0.45 per unit for $x \ge 25{,}000$ units

The problem is to determine the *economic order quantity* x_{opt} and the *lowest-total-cost order quantity* of meeting the demand of 400,000 units per year. Solution:

(\$/yr) = (units) * (\$/unit/yr)	(\$/yr) = (units/yr) * (\$/unit)
D{x(P)} = (x/2) * (0.06+0.2P) = (0.03+0.1P)x:	FC(P) = 400,000 * P
D{x(P1)} = 0.080x;	FC(P1) = $200,000/yr;
D{x(P2)} = 0.077x;	FC(P2) = $188,000/yr;
D{x(P3)} = 0.075x.	FC(P3) = $180,000/yr.

(\$/yr) = (\$/order) * (units/yr) / (units/order)
I(x) = (25) * (400,000) / (x) = $10^7/x$.

$x_{opt}(P1) = (10^7/0.080)^{1/2} = 11{,}180$;
$x_{opt}(P2) = (10^7/0.077)^{1/2} = 11{,}396$;
$x_{opt}(P3) = (10^7/0.075)^{1/2} = 11{,}547$.

x = 10,000: T{x(P2)} = D{x(P2)} + I(x) + FC(P2) = 770 + 1,000 + 188,000 = $189,770/yr.
x = 11,396: T{x(P2)} = D{x(P2)} + I(x) + FC(P2) = 877 + 877 + 188,000 = $189,754/yr.
x = 25,000: T{x(P3)} = D{x(P3)} + I(x) + FC(P3) = 1,875 + 400 + 180,000 = $182,275/yr.

The *economic order quantity* is 11,396 units per order. The *lowest-total-cost order quantity* is 25,000 units per order at P3 = $0.45 per unit, resulting in a total cost of $182,275/year.

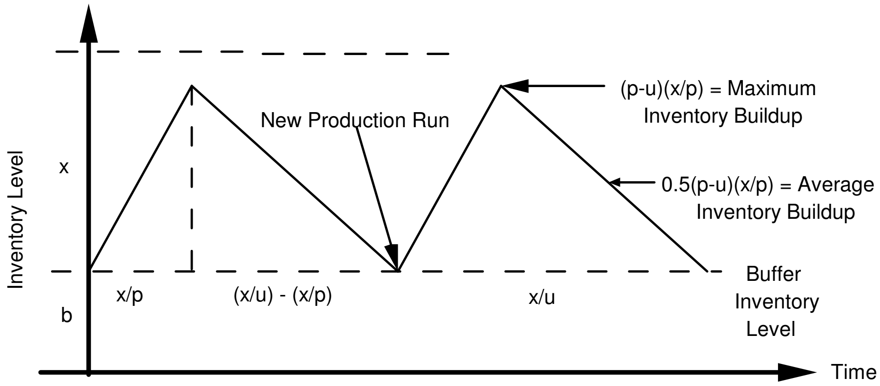
Example 8.4.4 - Economic Production Quantity (EPQ)

Suppose an item is being produced at the rate of 'p' units per period and used simultaneously at the rate of 'u' units per period. Then inventory builds up at the rate (p-u) when the item is being produced and used, and depletes at the rate 'u' when the item is being used but not produced. If x denotes the number of items produced in a production run, then the time it would take to get to the point of maximum inventory buildup would be x/p, and the amount of maximum inventory buildup would be (p-u)(x/p) as shown in Figure 8.4.4 below. The time it would take to deplete the maximum inventory buildup to the buffer inventory level would be (p-u)(x/pu). The problem is to find the economic production quantity x_{opt} which will minimize the total cost of handling and holding inventory and of changing manufacturing setups.

Figure 8.4.4 - Economic Production Quantity (EPQ) Model.

x_{opt} = Economic Production Quantity (EPQ)

b = Buffer Inventory Level



Problem - A company manufactures an electrical component which is used in a product that is sold at the rate of 400 units per week. After a setup cost of $125, the electrical component can be manufactured at the rate of 2,200 units/week with direct labor and material costs of $0.60 per unit. Interest and insurance charges are 15% per year based on the average inventory level. Assume the manufacturing plant operates 50 weeks per year. The problem is to determine the economic production quantity x_{opt} in units per setup that would minimize the total annual cost of producing the electrical component, and to determine the time intervals of inventory buildup and depletion between production runs of the electrical component.

Solution: x = production quantity (units/setup); x_{opt} = economic production quantity.
D(x) = annual costs of handling and holding units in inventory ($/yr).
I(x) = annual costs of changing manufacturing setups ($/yr).

(units/week) * (weeks) = (units)
Maximum inventory buildup: (p - u) * (x/p) = (p-u)(x/p)
Average inventory buildup: 0.5*(p-u)(x/p) (units)
Time interval of maximum inventory buildup: x/p (weeks)
Time interval of dissipating maximum inventory buildup: (x/u) - (x/p) (weeks)

(units of avg inventory) * ($/unit/yr) = ($/yr)
D(x) = 0.5*(2,200 - 400)*(x/2,200) * (0.60 * 0.15) = (0.0368x)

($/setup) * (units/wk) * (wks/yr) / (units/setup) = ($/yr)
I(x) = (125) * (400) * (50) / (x) = (2,500,000/x)

$x_{opt} = \sqrt{B/A} = \sqrt{2500000/0.0368}$ = 8,240 units/production-run.

Time interval of maximum inventory buildup: 8,240/2,200 = 3.75 weeks
Time interval to deplete maximum inventory buildup: (8,240/400) - 3.75 = 16.85 weeks
Time interval between production runs: 3.75 + 16.85 = 20.6 weeks.

Section 8.5 - Economic Life of a Growing Forest or Aging Wine

A growing forest or an aging wine are examples of products whose values appreciate in the course of time with little maintenance or operating costs. Other examples of such products are precious metals and jewelry, art and antiques, and land. The economic life of a forest or wine is defined as the optimal time for cutting down the forest or selling the wine. If the product was sold too soon, most of the value appreciation would be lost. But delaying such decisions excessively causes the holding costs to offset much of the value appreciation.

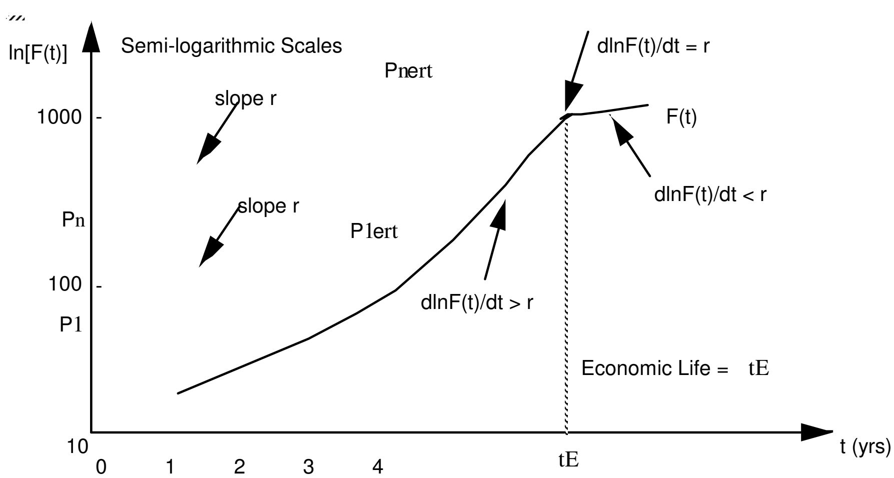
Let F(t) denote the future selling price of the product t years from now, and let us assume that F(t) is a known function of time. The present value of the future selling price, denoted by P(t), is equal to $F(t)e^{-rt}$, where r is the nominal annual interest rate. The necessary and sufficient conditions for maximizing P(t) are:

Necessary condition: $P'(t) = e^{-rt}[F'(t)-rF(t)] = 0; \ r = F'(t)/F(t) = d\ln F(t)/dt$ ....(8.5.1)

Sufficient condition: $P''(t) = e^{-rt}[F''(t)-rF'(t)] - r e^{-rt}[F'(t)-rF(t)] < 0$
Since $[F'(t)-rF(t)] = 0$ when $P'(t) = 0$, we get $e^{-rt}[F''(t)-rF'(t)] < 0$;
$F''(t) - rF'(t) < 0$; or $r > F''(t)/F'(t) = d\ln[F'(t)]/dt$ ....(8.5.2)

Equations (8.5.1) and (8.5.2) can be readily applied as follows. Plot the selling price F(t) on a logarithmic scale versus the time t on a linear scale which is depicted as an S-shaped curve in Figure 8.5.1 below. On the same semi-logarithmic graph, plot a sequence of parallel lines with slope r representing the future values P_1e^{rt}, ..., P_ne^{rt} of increasing present-value amounts $P_1 < ..... < P_n$ invested at the market rate of interest. The least present-value amount P_1 whose line P_1e^{rt} is tangent to the F(t) curve indicates the economic life t_E and the present value of the optimal selling price. In order to optimize P_1 and t_E, equation (8.5.2) requires the relative growth rate dlnF(t)/dt to be larger than r before t_E, and smaller than r after t_E.

Figure 8.5.1 - Economic Life of a Growing Forest or Aging Wine



Section 8.6 - Maintenance, Replacement and Retirement Policy

As the service life of equipment increases, decisions must be made between keeping and maintaining equipment in service and the mutually exclusive alternatives of either replacing or retiring the equipment. In order to determine which decision is best, it is necessary to evaluate the *marginal annual costs* of keeping and maintaining the existing equipment in service for an additional year and the *average annual costs* of replacing or retiring the equipment, assuming all other revenues and costs are unchanged.

Economic service life problems differ from those of the economic life of appreciating capital assets described in Section 8.5 where costs of operation and maintenance are negligible and the value of the product appreciates with time. In the case of economic service life problems, the value of the equipment tends to decrease with time due to wear and tear as well as obsolescence. Moreover, revenues derived from using the equipment may either increase due to market demand or decrease because of substitution effects or increased competition. Therefore, the problem of determining how long equipment should be kept and maintained before it is replaced or retired is quite different than the problem of determining when a growing forest should be harvested or an aging wine should be sold.

To fix ideas, let us consider the following problem. A chemical company needs an insulated stainless steel tank to store corrosive chemicals. It is estimated a new tank costs $10,000 and its salvage value at end-of-year N is given by $5,500 - $500N for N = 1, ... , 9. Annual end-of-year costs (-) of operation and maintenance are:

End-of-year N:	1	2	3	4	5	6	7	8	9
Annual Costs:	$1,000	$1,200	$1,600	$2,000	$2,200	$2,600	$3,000	$3,400	$3,800
Salvage Value:	$5,000	$4,500	$4,000	$3,500	$3,000	$2,500	$2,000	$1,500	$1,000

(a) Determine the *marginal annual costs* of keeping and maintaining the existing tank in service during the 1st, 2nd, ... and 9th years using an interest rate of 10% per year.
(b) Determine the *average annual costs* of the tank for a period of one, two, ... and nine years. Assume the average annual costs are measured by the equivalent uniform annual costs at an interest rate of 10% per year that would result from identical tank replacements.
(c) Determine the *economic service life* of replacing the tank when its average annual cost equals the marginal annual cost of keeping and maintaining the existing tank in service.
(Adapted from W. T. Morris, Engineering Economy, Richard D. Irwin, Inc., 1960)

Part (a) Solution - The marginal annual cost of keeping and maintaining the tank in service during a year is due to annual costs of (1) operation and maintenance, (2) the change in salvage value during the year, and (3) the interest cost of holding last-year's salvage value for one year. Consequently, the answer to part (a) is given in Table 8.6.1 below.

Table 8.6.1 - Marginal Annual Costs (-) of Keeping and Maintaining the Tank in Service.

End of Year j:	1	2	3	4	5	6	7	8	9
Op&Maint AC:	$1,000	$1,200	$1,600	$2,000	$2,200	$2,600	$3,000	$3,400	$3,800
Δ Salvage AC:	$5,000	$500	$500	$500	$500	$500	$500	$500	$500
Holding AC:	$1,000	$500	$450	$400	$350	$300	$250	$200	$150
Marginal AC:	$7,000	$2,200	$2,550	$2,900	$3,050	$3,400	$3,750	$4,100	$4,450

The data on the last line of Table 8.6.1 are defined as the *marginal annual costs* of keeping and maintaining the existing tank in service for one year. These marginal costs must be weighed against marginal revenues for decisions to continue keeping the equipment.

Part (b) Solution - *Average annual costs* are defined as the equivalent annual worth of the future value of its marginal annual costs. This definition of average annual costs provides a basis of averaging the marginal annual costs of the tank over the years before its replacement or retirement. Therefore, the marginal annual costs on the bottom line of Table 8.6.1 can be treated as end-of-year cash flows as shown in column (4) of Table 8.6.2 . Using a 10% per year discount rate in column (2), the ABC accounting rules enable us to determine the future value of the marginal annual costs in column (5) of Table 8.6.2 . Column (6) lists the factor (A/F,10%,n) of equation (3.6.2) with which to multiply the future values in column (5) in order to get the average annual costs of the tank in column (7) of Table 8.6.2 .

Table 8.6.2 - Average Annual Costs (AC) of the Tank Replacement

(1)EOY	(2)Int/Yr	(3)BCF Bal	(4)Marginal AC	(5)ACF Bal	(6)(A/F,10,n)	(7)Average AC
0	10.0%	$0	$0	$0	-.------	$0
1	10.0%	$0	-$7,000	-$7,000	1.00000	-$7,000
2	10.0%	-$7,700	-$2,200	-$9,900	0.47619	-$4,714
3	10.0%	-$10,890	-$2,550	-$13,440	0.30212	-$4,060
4	10.0%	-$14,784	-$2,900	-$17,684	0.21547	-$3,811
5	10.0%	-$19,452	-$3,050	-$22,502	0.16380	-$3,686
6	10.0%	-$24,753	-$3,400	-$28,153	0.12961	-$3,649
7	10.0%	-$30,968	-$3,750	-$34,718	0.10541	-$3,659
8	10.0%	-$38,190	-$4,100	-$42,290	0.08745	-$3,698
9	10.0%	-$46,519	-$4,450	-$50,969	0.07364	-$3,753

Treating marginal annual costs as cash flows in column (4) of Table 8.6.2 is not fully justified because holding costs and changes in salvage values are not cash flows. However, the same results as column (5) of Table 8.6.2 are obtained in column (7) of Table 8.6.3 by the ABC rules using the cash flows of annual operation and maintenance costs in column (4) of Table 8.6.3, and adding the cash flows of terminal salvage values in column (6) to ACF Balances in column (5) in order to get the future-value marginal annual costs in column (7).

Table 8.6.3 - Future Value of Marginal Annual Costs from Cash Flow Analysis.

(1)EOY	(2)Int/Yr	(3)BCF Bal	(4)EOY Csh Flw	(5)ACF Bal	(6)EOY Salv	(7)FV Marg AC
0	10.0%	$0	-$10,000	-$10,000	$10,000	$0
1	10.0%	-$11,000	-$1,000	-$12,000	$5,000	-$7,000
2	10.0%	-$13,200	-$1,200	-$14,400	$4,500	-$9,900
3	10.0%	-$15,840	-$1,600	-$17,440	$4,000	-$13,440
4	10.0%	-$19,184	-$2,000	-$21,184	$3,500	-$17,684
5	10.0%	-$23,302	-$2,200	-$25,502	$3,000	-$22,502
6	10.0%	-$28,053	-$2,600	-$30,653	$2,500	-$28,153
7	10.0%	-$33,718	-$3,000	-$36,718	$2,000	-$34,718
8	10.0%	-$40,390	-$3,400	-$43,790	$1,500	-$42,290
9	10.0%	-$48,169	-$3,800	-$51,969	$1,000	-$50,969

Part (c) Solution - The *economic service life* of the tank is defined as the time in years when its marginal annual costs equals its average annual cost. As the service life of the tank increases, its operation and maintenance costs tend to increase. But as the service life of the existing tank increases, its initial cost is spread over a longer period of time during which salvage values are usually decreasing. After a period of time, the operation and maintenance annual costs increase faster than the decrease in the time average of the initial equipment cost adjusted for terminal salvage values. Consequently, marginal annual costs increase faster than average annual costs after reaching the economic service life. (see Section 8.3 for relations between marginal and average revenues and costs).

Table 8.6.2 shows the 6th-year marginal annual cost of the tank is -$3,400 which is *less* than the 6-year average annual cost of -$3,649. However, the 7th-year marginal annual cost of the tank is -$3,750 which is *greater* than the 7-year average annual cost of -$3,659. Consequently, the tank reaches its economic service life during the seventh year.

Afterwards, marginal annual costs are increasing faster than average annual costs, and the replacement or retirement of equipment is recommended in a steady-state situation.

Economic service life may be defined as the time when average annual costs are minimized without looking at trends of increasing marginal annual costs. This could change the economic service life as defined above. For example, column (7) of Table 8.6.2 shows the average annual cost is a minimum in the 6th year, but marginal equals average annual costs in the 7th year. Although these results differ by only one year, their conceptual differences are economically significant. Measurements of equivalent annual costs assume identical replacements which are rarely made in practice. New equipment may last longer, and they frequently provide important changes in quality, reliability, efficiency and capacity of output as well as operation and maintenance costs, all of which factor into the initial cost of the equipment. Consequently, it is important that average annual costs be compared to marginal annual costs and revenues in any dynamic equipment policy.

Moreover, average annual costs are slowly varying when they are spread over a number of years. Table 8.6.2 shows the minimum average annual cost is -\$3,649 in the 6th year, whereas the average annual cost is only \$10 more in the 7th year. It is not reasonable to keep the tank in service only 6 years just because the average annual cost increases \$10 from the 6th to the 7th year. When making such a decision, the marginal annual costs of keeping the tank in service should be determined (i.e., operating and maintenance costs, changes in salvage value, and holding costs) because they lie at the razor's edge of economic decision-making.

The definition of economic service life in terms of marginal or average costs ignores the revenues derived from using the equipment. If replacing equipment could generate a better or faster output, then such increased revenues should shorten economic service life and accelerate the replacement of existing equipment. As the service life of equipment increases, net revenues derived from using the equipment usually decline. The choice may then be between keeping and maintaining equipment in service or retiring the equipment. The equipment should not be retired as long as marginal revenues exceed marginal costs.

It is convenient to describe marginal costs in a continuous form. For this purpose, let C(t) denote the capital value of the equipment at any time t, where $0 \leq t \leq T$. Then C(0) is the initial equipment cost, C(T) is its salvage value at time T, C'(t) is the rate of depreciation at time t and rC(t) is the rate of holding costs at time t where r is the market rate of interest. Also, let m(t) denote the rate of maintenance and operating costs at time t. Letting F(T) denote the <u>f</u>uture-<u>v</u>alue-<u>m</u>arginal-<u>c</u>osts (FVMC) at time T, then F(T) is given by equation (8.6.1).

$$F(T) = \int_0^T [m(t) + C'(t) - rC(t)]\, e^{r(T-t)}\, dt \qquad \text{....(8.6.1)}$$

Because $[C'(t) - rC(t)]e^{-rt} = \frac{d}{dt}[C(t)e^{-rt}]$, equation (8.6.1) simplifies to

$$F(T) = \int_0^T m(t)\, e^{r(T-t)}\, dt + C(T) - C(0)e^{rT} \qquad \text{....(8.6.2)}$$

Hence, F(T) = FVMC depends only on the initial cost C(0) and the terminal salvage value C(T), and not on intermediate salvage values C(t) for $0 < t < T$. The present value of F(T), denoted by P(T) = PVMC, is obtained by multiplying equation (8.62) with e^{-rT} as shown in equation (8.6.3).

$$P(T) = \int_0^T m(t)\, e^{-rt}\, dt + C(T)e^{-rT} - C(0) \qquad ....(8.6.3)$$

Equations (8.6.2) and (8.6.3) can be used for most problems in replacement analysis. For example, let us consider the problem of determining the economic life of a growing forest or aging wine described in Section 8.5 where it was assumed that operating and maintenance costs were negligible. Since $m(t) = 0$, equation (8.6.3) reduces to

$$P(T) = C(T)e^{-rT} - C(0) \qquad ....(8.6.4)$$

The meanings of C(T) and C(0) in equations (8.6.3) and (8.6.4) are not the same. In equation (8.6.3), C(0) represents a high initial equipment cost which depreciates to a terminal salvage value at time T. In equation (8.6.4), C(0) represents goods with a relatively small initial cost whose maintenance and operating costs are negligible and which appreciates in value to C(T) at time T when it is sold. The meaning of C(T) in equation (8.6.4) is analogous to F(t) in equations (8.5.1) and (8.5.2) which show that the present value of the future sales prices would be maximized when the relative growth rate in the sales price equals the market rate of interest.

The economic service life problem presented at the beginning of Section 8.6 can also be analyzed by means of equation (8.6.2). In this connection, let m(t) represent the cash flow rate of operating and maintenance costs, C(0) represent the initial cost, C(T) represent the salvage value at time T, and let F(T) represent FVMC in either column (5) of Table 8.6.2 or column (7) of Table 8.6.3 . The average annual cost of F(T), denoted by $A_E(T)$, is given by

$$A_E(T) = F(T)(A_c/F,r,T) = rF(T)[e^{rT}-1]^{-1} \qquad ....(8.6.5)$$

The average annual cost $A_E(T)$ is a minimum when its derivative equals zero.

$$dA_E(T)/dT = rF'(T)[e^{rT}-1]^{-1} - r^2e^{rT}F(t)[e^{rT}-1]^{-2} = 0 \text{ or}$$

$$F'(T)/F(T) = re^{rT}[e^{rT}-1]^{-1} = r/[1- e^{-rT}] = (A_c/P,r,T) \qquad ....(8.6.6)$$

Equation (8.6.6) may be applied to the marginal annual costs of Table 8.6.1 by comparing discrete approximations of F'(T)/F(T) and $(A_c/P,r,T)$ in columns (6) and (7) of Table 8.6.4.

Table 8.6.4 - Minimum average annual cost calculations for equation (8.6.6)

(1)EOY	(2)Int/Yr	(3)BCF Bal	(4)Marg AC	(5)F(t)=FVMC	(6)F'(T)/F(T)	(7)(A/P,10,n)
0	10.0%	$0	$0	$0	-.------	-.------
1	10.0%	$0	-$7,000	-$7,000	0.41429	1.10000
2	10.0%	-$7,700	-$2,200	-$9,900	0.35758	0.57619
3	10.0%	-$10,890	-$2,550	-$13,440	0.31577	0.40212
4	10.0%	-$14,784	-$2,900	-$17,684	0.27245	0.31547
5	10.0%	-$19,452	-$3,050	-$22,502	0.25113	0.26380
6	10.0%	-$24,753	-$3,400	-$28,153	0.23319	0.22961
7	10.0%	-$30,968	-$3,750	-$34,718	0.21810	0.20541
8	10.0%	-$38,190	-$4,100	-$42,290	0.20523	0.18745
9	10.0%	-$46,519	-$4,450	-$50,969	-.------	0.17364

The entries in column (7) of Table 8.6.4 are greater than those of column (6) up until the sixth year when the entries of column (7) become smaller than those of column (6). Therefore, the minimum average annual cost occurs in the sixth year just as we have previously determined in Table 8.6.2 . This example verifies the application of equation

(8.6.6) for determining economic service life whenever it is defined as occurring at the time of minimum average annual costs.

Equation (8.6.2) can be used to approximate the future value of marginal annual costs when m(t) and C(t) are estimated as continuous functions. For this purpose, let us approximate the annual rate of operation and maintenance costs in Table 8.6.1 by the function $m(t) = -910 - 310*t$ for $0 \le t \le 9$. The total operation and maintenance cost over 9 years would then be $20,745 as compared to $20,800 in Table 8.6.1. The initial cost of the tank is C(0) = $10,000, and the terminal salvage values of the tank are C(t) = $5,500 - $500*t for $1 \le t \le 9$. The interest rate of 10% per year compounded annually is equivalent to the nominal annual interest rate r = ln(1.10) = 0.0953102 or 9.53102%/year compounded continuously. Therefore, the future value of marginal annual costs F(T) is given by

$$F(T) = \int_0^T -[910+310*t]e^{r(T-t)}dt + 5{,}500 - 500*T - 10{,}000e^{rT} \qquad ....(8.6.7)$$

The integral above can be evaluated by the formula $F(T) = [A_c+(G_c/r)](F/A_c,r,T)-(TG_c/r)$, where $A_c = -910$, $G_c = -310$ and r = 0.0953102 . Table 8.6.5 below compares the results of evaluating F(T) by continuous functions in equation (8.6.7) to the results in Tables 8.6.2 and 8.6.3 of evaluating F(T) in discrete steps.

Table 8.6.5 - Continuous and discrete evaluations of future-value marginal annual costs.

t years	Continuous evaluation of F(T)	Discrete evaluation of F(T)
1	-$7,115	-$7,000
2	-$10,266	-$9,900
3	-$14,008	-$13,440
4	-$18,400	-$17,684
5	-$23,506	-$22,502
6	-$29,397	-$28,153
7	-$36,153	-$34,718
8	-$43,860	-$42,290
9	-$52,613	-$50,969

Table 8.6.5 illustrates that the cash flows of marginal annual costs can be estimated in good approximation as continuous functions, and that the future values of those marginal annual costs can be evaluated directly by equation (8.6.2) which avoids the need for ABC spreadsheet calculations. It should be noted that equation (8.6.2) is a continuous analog of the annual-cost definitions given in Section 8.6 . Consequently, equation (8.6.2) may be used when marginal revenues are included with marginal costs, as well as before-and-after-tax cash flow analyses.

Section 8.7 - Matching Impedances

A recurring problem of electrical power supply concerns a load with a continuously variable resistance of x ohms. The battery supplies the load with fixed voltage E and it has an internal resistance of R ohms. What should the load resistance x be adjusted to in order to maximize the power dissipation of the load? If x was made twice as large as R, what fraction of the total energy dissipation of the system would be dissipated at the load?

Solution: The current I must pass through resistances R and X which are in series. Hence, by Ohm's law, $I = E/(R+x)$ and $I^2 = E^2/(R+x)^2$. Therefore, the energy dissipation of the load and of the total system would be given by equations (8.7.1) and (8.7.2) respectively.

$$I^2x = E^2x/(R+x)^2 \quad \text{....(8.7.1)}$$

$$I^2(R + x) = E^2(R+x)/(R+x)^2 = E^2/(R+x) \quad \text{....(8.7.2)}$$

The necessary and sufficient conditions for maximizing I^2x are:

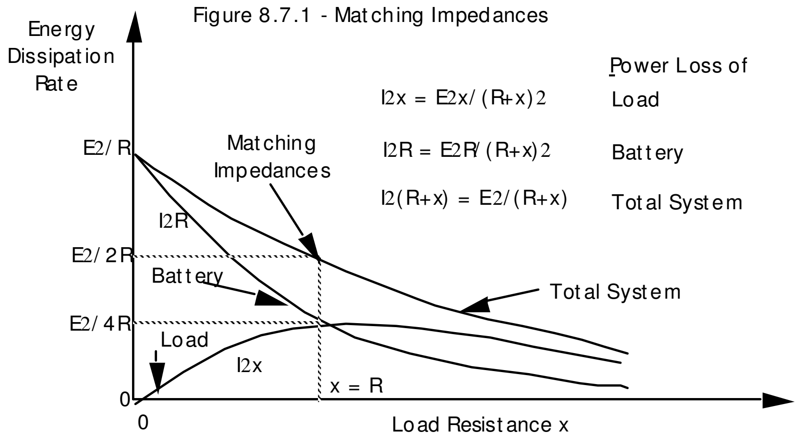
$$d(I^2x)/dx = E^2(R + x)^{-2}[1 - 2x(R+x)^{-1}] = 0$$
$$1 = 2x(R+x)^{-1}; \quad R+x = 2x; \quad \text{or} \quad x = R \quad \text{....(8.7.3)}$$

$$d^2(I^2x)/dx^2|_{x=R} = E^2(R+x)^{-2}[-2(R+x)^{-1} + 2x(R+x)^{-2}]|_{x=R}$$
$$= [E^2/4R^{-2}]\{(-2/2R)+(2R/4R^2)\} = -E^2/8R^3 < 0 \quad \text{....(8.7.4)}$$

Equations (7.5.3) and (7.5.4) show that useful power at the load is maximized when the load and battery resistances match. The load would then have 50% of the power of the system. If x = 2R, the load would have two-thirds of the power of the system.

Figure 8.7.1 below depicts the energy dissipation rates I^2x and $I^2(R + x)$ of the load and total system as functions of load resistance x. The power losses in the battery and load equals the power loss of the system. However, the issue is not to maximize the power of the system. Such an optimum would occur when the load is short-circuited (i.e., x = 0), and all the power would be dissipated at the battery. Instead, the problem is either to maximize the useful power available at the load, or to get more than half of the power of the total system at the load.

Figure 8.7.1 - Matching Impedances

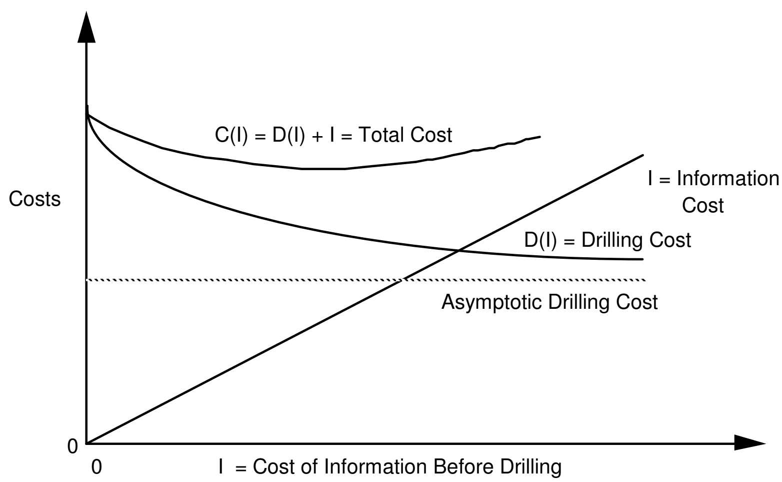


Section 8.8 - Economic Purchase of Information

The outcome of many projects can be enhanced if more information is available prior to undertaking the project. However, there are costs associated with purchasing useful information. If too much information is obtained, then the cost of the additional information could exceed its benefits. Deciding when to drill in oil and gas exploration is essentially a problem of determining how much information should be bought prior to drilling.

Consider the exploration costs D(I) of drilling for oil as a function of the cost I for geological and geophysical information which is bought before drilling. If little or no information is bought before drilling (i.e., I = 0), then D(I) is apt to be very large because wildcat drilling usually results in a large number of dry holes. On the other hand, with perfect information on the location of an oil pool, D(I) would decrease and approach an asymptotic drilling cost as shown in Figure 8.8.1 below. The problem is to decide how much information should be bought in order to minimize the total exploration cost C(I) consisting of the drilling cost D(I) and the cost I of buying prior information as expressed in the equation C(I) = D(I) + I.

Figure 8.8.1 - Minimizing Exploration Costs by Buying Information Before Drilling



The curve D(I) has its maximum value when I = 0, and it declines to the asymptote where information is sufficient to avoid drilling dry holes. The necessary condition for minimizing the total exploration cost C(I) is that its first derivative should equal zero.

$$C'(I) = D'(I) + 1 = 0; \quad \text{or} \quad D'(I) = -1 \qquad ....(8.8.1)$$

The sufficient condition for a minimum is C"(I) > 0, or C"(I) = D"(I) > 0. Both the necessary and sufficient conditions for a minimum need to be satisfied in order to minimize total costs. If the sufficient condition D"(I) > 0 is satisfied, but the necessary condition D'(I) = -1 is nowhere satisfied, then buying information should not be a prerequisite to drilling. The possibility that the optimum may occur at the boundary (i.e., I = 0) marks the drilling problem as a special Kelvin Law problem.

Although most Kelvin Law problems involve minimizing input for a given output, the process of bidding for a contract is an important problem of maximization. If the bid is too high, there is a good chance the bidder will be awarded the contract, but there is a small chance of making a profit. If the bid is too low, there is a good chance the bidder will make a profit, but there is a small chance of being awarded the contract. Since the chances of realizing a profit depends upon the product of the chances of being awarded the contract and of making a profit if the bidder is awarded the contract, the maximum realizable profit is obtained somewhere in between the high and low bid prices. Admittedly, the directly and inversely varying curves of drilling and bidding problems are difficult to quantify, but Kelvin's Law provides at least a conceptual framework for analyzing these problems.

Section 8.9 - Summary of Chapter Eight

In this chapter we are concerned with optimization problems having a single control variable. By varying the magnitude of the control variable, an objective quantity of the problem can be either maximized or minimized. Therefore, our aim is to define different types of single variable optimization problems in each section of the chapter, and to develop methods of finding the magnitude of the control variable which would either maximize or minimize the objective quantity of the problem.

Breakeven analysis (Section 8.1) is the study of cost-volume-profit relationships. The control variable is the volume of output per unit of time expressed in either physical units Q, sales volume \$Q, or percent capacity %Q. The rate of volume output is controlled in the vicinity of the *breakeven point* at which sales revenues equals the sum of fixed and variable costs, resulting in zero profits. Most breakeven analyses use linear approximations of revenues and costs in the neighborhood of the breakeven point. In general, the profit Z is maximized when marginal revenues equal marginal costs. Owing to the indefinite range of linear approximations in conventional breakeven analysis, profit would be maximized when the output per unit of time is infinite.

Conventional breakeven analyses assume the price per unit of output does not depend upon the volume of demand. But demand analysis (Section 8.2) assumes the demand D varies inversely with the price p charged per unit of product as expressed by the demand function $D = f(p)$. The characteristics of demand curves are usually described in terms of *elasticities*. The elasticity μ of demand with respect to price is defined as the ratio of the relative increase in demand to the relative decrease in price. When $\mu < 1$, the demand is said to be inelastic because the relative increase in demand is less than the relative decrease in price so that total revenues decrease. If demand is inelastic, sales revenues will decrease as the price is decreased because the demand does not increase proportionately. Consequently, when demand is taken into account in conventional breakeven analysis, profit would be maximized when the output per unit of time is finite.

Over a period of time, the rates of input $X(t)$ and output $Y(t)$ of a project usually have an S-shaped growth curve. Opportunities to increase the scale of a project form a set of *mutually inclusive alternatives* over time. When all input factors are used in fixed proportions, the input can be treated as a single control variable. Section 8.3 analyzes relationships between the marginal and average output per unit of input during different stages of growth. In early stages of development, the marginal output dY/dX is greater than the average output Y/X which is increasing as $X(t)$ increases. In later stages of development, the marginal output dY/dX is smaller than the average output Y/X which is decreasing as $X(t)$ increases. When the marginal and average outputs are equal, the average output Y/X is a maximum.

Kelvin law problems (Section 8.4) involve a single control variable x which, if increased, will increase directly varying costs D(x) and decrease inversely varying costs I(x). The optimal value of the total cost T(x) occurs at the value of x where the slopes of the D(x) and I(x) curves are equal in magnitude and opposite in sign. It is necessary to evaluate the second derivative of T(x) to determine whether the optimal value of the total cost T(x) is a minimum or maximum.

Kelvin's law was first applied to finding the optimal cross-sectional area x of a wire conductor (Example 8.5.1) which minimizes the sum of directly varying capital recovery costs of the wire conductor and inversely varying costs of power transmission loss. Kelvin's law is also used in inventory control to find economic order quantities (EOQ) (Example 8.5.2). The economic order quantity x minimizes the sum of directly varying handling and holding costs and inversely varying ordering costs in a given time period. Suppliers usually offer price discounts to encourage large orders, and this problem is solved as a special case of the EOQ problem (Example 8.5.3). Another application of Kelvin's law is to find the economic production quantity (EPQ) (Example 8.5.4) of an item being produced. The economic production quantity x minimizes the total cost of handling and holding inventory and of changing manufacturing setups.

The problem of finding the optimal time to harvest a growing forest or to sell an aging wine (Section 8.6) is a special type of problem where the value of the product appreciates with time while operation and maintenance costs are negligible. Other examples of such products are precious metals and jewelry, art and antiques, and land. The selling price F(t) is assumed to be a known function of time and is plotted on a logarithmic scale versus the time t on a linear scale. The optimal time for selling the product occurs at the point of time where the relative rate of growth of the selling price $d\ln F(t)/dt = F'(t)/F(t)$ is equal to the nominal annual interest rate r. This optimal economic life solution for capital appreciating assets assumes that the relative growth rate of the sales price is first *greater* than, and later *smaller* than, the market rate of interest.

Maintenance, replacement and retirement policy (Section 8.7) deals with deciding between keeping and maintaining equipment in service and the mutually exclusive alternatives of replacing or retiring the equipment. To make such decisions, the *marginal annual costs* for keeping and maintaining the equipment in service during each year must be determined by summing the annual costs of (1) operation and maintenance, (2) change in salvage value during the year and (3) interest cost of holding last-year's salvage value for one year. These marginal annual costs can be formulated in either discrete or continuous models. The future value of the marginal annual costs of each replacement candidate is then determined as a function of its years of service, and the equivalent annual worth of those future values are defined as the *average annual costs* of the equipment.

The *economic service life* of equipment is defined as the number of years it would need to be in service before its marginal annual cost would be equal to its average annual cost. As the service life of equipment increases, its initial cost is spread over a longer period of time which tends to reduce its average annual cost. As the service life of the existing equipment increases, its marginal annual costs tend to increase due to increased operation and maintenance costs. After the marginal and average costs are equal, the decrease in average annual costs due to the decreasing time average of the initial equipment cost tends to be smaller than the increase in average annual cost due to increasing marginal annual costs. The economic service life then indicates a point in time when both marginal and average annual costs are increasing (see Proposition 1 of Section 8.4).

Economic service life is often defined as the number of years where average annual costs are minimized. Average annual costs are largely evaluated to plan how long equipment should be kept in service rather than to decide whether or not it should be replaced or retired from service. Measurements of the average annual costs of equipment in service are insensitive because they are spread over a number of years. Consequently, at the time of deciding to replace equipment in service, it is necessary to determine the marginal annual costs of the equipment in service because it lies at the razor's edge of economic decision-making. There is a widespread belief that keeping an asset is a passive decision once the asset is owned. However, keeping an asset is an active decision which needs to be investigated periodically by comparing marginal costs and revenues of equipment in service.

The definition of economic service life as a minimum of average annual costs largely ignores the revenues derived from using the equipment. If replacing equipment could generate output faster or of a better quality, then such revenue enhancements should affect the measurement of economic service life so that the existing equipment would be replaced sooner. As the service life of equipment increases, the revenues derived from using the equipment may be declining. The choice may then be between keeping and maintaining the equipment in service or retiring the equipment. The equipment should not be retired as long as marginal revenues exceed marginal costs.

Matching impedances (Section 8.7) deals with problems of electrical power supply concerning a load which has a continuously variable resistance. The power losses in the battery power supply and the load equals the power loss of the system. The issue is to determine the load resistance at which the power dissipation at the load is maximized. The useful power at the load is maximized when the load and battery resistances match, wherin the load would have one half of the power of the system. The power of the system would be maximized if the load were short-circuited, with all the energy being dissipated at the battery. However, the important concern about matching impedances is either to maximize the useful power available to the load, or to get more than half of the power dissipation of the total system at the load.

Economic purchase of information (Section 8.8) concerns how much information should be bought before taking decisive action. Deciding when to drill in oil and gas exploration is essentially a problem of determining how much information should be bought prior to drilling. As more information is bought, the cost of drilling is reduced until it reaches a minimum drilling cost that would be incurred if the exact location of the prospect were known. Since the cost of perfect information is apt to be very large, we must decide how much information should be bought prior to drilling in order to minimize the total cost of drilling and information. The minimum total cost occurs either where the drilling and information cost curves have slopes that are equal in magnitude but opposite in sign, or where drilling costs diminish very slowly as more information is acquired.

The process of bidding for a contract is an important problem of profit maximization. If the bid is too high, there is a good chance the bidder will be awarded the contract, but there is a small chance of making a profit. If the bid is too low, there is a good chance the bidder will make a profit, but there is a small chance of being awarded the contract. Since the chances of realizing a profit depends upon the product of the chances of being awarded the contract and of making a profit if the contract is awarded to the successful bidder, the maximum realizable profit is obtained somewhere in between the high and low bid prices. Admittedly, the directly and inversely varying curves of drilling and bidding problems are very difficult to quantify, but Kelvin's Law provides at least a conceptual framework for analyzing these problems.

Chapter Eight - Exercises

8-1a A company imports lawn mowers at $30 per unit. There is a fixed annual cost of $45,000 for importing the lawn mowers. The sales personnel of the company receive a commission of 25% of the selling price for each lawn mower that they sell. If 5000 lawn mowers are sold each year, what should be the selling price of a lawn mower in order for the company to breakeven?

8-1b If $8000 worth of advertisement is added to the fixed annual cost of importing the lawn mowers, the company expects that it could reduce sales commissions from 25% to 15% and maintain the same sales volume and selling price of the lawn mowers. How much would the company's net income change as a result of such advertisement?

8-2a A machine is 5 years old and its present market value is $5,000. An analysis is being made of its marginal and average annual costs in order to evaluate its possible replacement. Salvage values and annual operating and maintenance costs are estimated for the next 11 years in the table below. Determine the *marginal annual costs* of keeping and maintaining the existing machine in service during the 1st, 2nd, ... and 11th years using an interest rate of 10% per year. (see Section 8.6 and Table 8.6.1 on page 211)

8-2b Determine the *average annual costs* of the machine for a period of one, two, ... and eleven years. Assume the average annual costs are measured by the equivalent uniform annual costs at an interest rate of 10% per year that would result from identical machine replacements. (see Table 8.6.2)

8-2c Determine the *economic service life* of replacing the machine when its marginal annual cost is equal to the average annual cost of operating and maintaining the existing machine in service. In what year is the average annual cost a minimum?

8-2d Suppose there are $400 per year tax savings from the remaining depreciation allowances of the machine at the ends of the first and second years. Determine the *economic service life* of replacing the machine when its marginal annual cost is equal to the average annual cost of operating and maintaining the existing machine in service when these tax savings are included with the operating and maintenance costs at the ends of the first and second years. In what years are the average annual costs a minimum?

End-of-year	EOY Salvage Value	Operation & Maintenance AC
0	$5,000	$0
1	$4,000	$0
2	$3,500	$100
3	$3,000	$200
4	$2,500	$300
5	$2,000	$400
6	$2,000	$500
7	$2,000	$600
8	$2,000	$700
9	$2,000	$800
10	$2,000	$900
11	$2,000	$1,000

Chapter Nine - Multiple Variable Optimization

Section 9.1 - Engineering Objective in Linear Programming (LP)

During World War II, British scientists pioneered in the development of *operations research* (OR) techniques which enabled them to model and solve military logistics problems that required the optimal allocation of limited resources. A major thrust of OR is a mathematical technique called *linear programming* (LP) which involves translating problems into linear relationships between variables that are subject to either equality or inequality constraints in order to make some objective quantity as large or small as possible. When LP is applied to problems of optimizing selections from production alternatives, questions are raised about objectives to be optimized, effectiveness of input and output constraints, and linear and nonlinear relationships of variables. Chapter Nine explains the origins of these questions and how they could be resolved by engineering production functions.

Let us apply the LP method of optimization to a product-mix problem involving two input and two output variables subject to input constraints. Consider an owner of a coffee plantation and factory which processes two kinds of coffee beans, *arabica and robusta beans*, into two brands of instant coffee, *Taster's and Connoisseur's Brews*. The plantation's weekly inputs to the factory are 3 tons of *arabica beans* and 4 tons of *robusta beans*. The factory's weekly outputs are one ton of *Taster's* (x_T= 1) and one ton of *Connoisseur's Brews* (x_C= 1) which sell for p_T= $4,000 and p_C= $5,000 and which yield $x_T p_T + x_C p_C$ = $9,000 sales revenues per week. Table 9.1.1 lists coffee-bean input requirements per ton of coffee-brew outputs.

Table 9.1.1 - Coffee-Bean Inputs Needed for Instant Coffee-Brew Outputs

	Tons of Coffee-Bean Input Required per Ton of Instant Coffee-Brew Output		Coffee-Bean Inputs
	Taster's Brew	Connoisseur's Brew	(tons/week)
Arabica Beans	a_T=1.0	a_C=2.0	$a = a_T+a_C$= 3
Robusta Beans	r_T=2.5	r_C=1.5	$r = r_T+r_C$= 4
Output Prices/ton	p_T=$4,000	p_C=$5,000	

The *engineering objective* for producing instant-coffee brew outputs from the coffee-bean inputs is to minimize leftover inputs. The *financial objective* for producing instant-coffee brews from the inputs is to maximize sales revenues from the outputs.

Suppose (a_C,r_C) = (2,1.5) was planned to make x_C= 1 worth $x_C p_C$ = $5,000. Instead, Taster's Brew was made with r_T = 1.5 and a_T = 1.5*(a_T/r_T) = 0.6 so that x_T= 0.6/a_T= 0.6 worth $x_T p_T$ = 0.6*$4,000 = $2,400 which is $2,600 less than $x_C p_C$ = $5,000. Producing Taster's Brew from (a_C,r_C) = (2,1.5) leaves over a_C-$a_T x_T$ = 2-1.0*0.6 = 1.4 tons of arabica beans whose *implicit costs* are y_a = $2,600/1.4 = $1,857.14 per ton.

Suppose (a_T,r_T) = (1,2.5) was planned to make x_T= 1 worth $x_T p_T$ = $4,000. Instead, Connoisseur's Brew was made with a_C = 1 and r_C = 1*(r_C/a_C) = 0.75 so that x_C= 0.75/r_C= 0.5 worth $x_C p_C$ = 0.5*$5,000 = $2,500 which is $1,500 less than $x_T p_T$ = $4,000. Producing Connoisseur's Brew from (a_T,r_T) = (1,2.5) leaves over r_T-$r_C x_C$ = 2.5-1.5*0.5 = 1.75 tons of robusta beans whose *implicit costs* are y_r = $1,500/1.75 = $857.14 per ton.

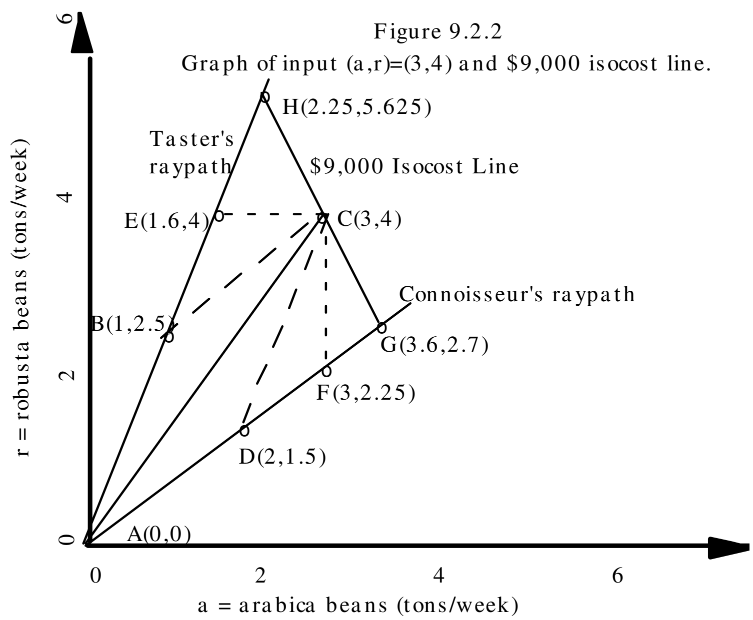
Implicit costs of arabica and robusta beans serve as transfer prices between the plantation and coffee factory. More specifically, the factory could realize x_Cp_C = \$5,000 sales revenues from the plantation's output of (a_C,r_C) = (2,1.5), the implicit costs of which are

$$a_Cy_a+r_Cy_r = 2*\$1,857.14+1.5*\$857.14 = \$3,714.29+\$1,285.71 = \$5,000$$

The LP solution of the product-mix problem is depicted graphically in Figure 9.1.1. Point C(a,r) = C(3,4) represents the weekly arabica and robusta bean input constraints. The C(3,4) inputs would be fully utilized for making instant-coffee brews without any left overs if $a = a_T+a_C= 3$ and $r = r_T+r_C= 4$. These conditions can be satisfied in general by constructing Taster's and Connoisseur's Brew raypaths whose slopes are $m_T = r_T/a_T = 2.50$ and $m_C = r_C/a_C =$ 0.75 respectively. Parallelogram ABCD is constructed by drawing a line parallel to Connoisseur's raypath from C(3,4) until it intersects Taster's raypath at B(1,2.5). Another line is drawn parallel to Taster's raypath from C(3,4) until it intersects Connoisseur's raypath at D(2,1.5).

If Taster's Brew only is made, its raypath intersects r = 4 at $E(a_E,r)=E(a_E,4)$ where $a_E= 4/m_T=$ 1.6 and $x_T= 1.6/a_T= 1.6$, leaving over $a-a_E= 3-1.6 = 1.4$ tons of arabica beans each week. If Connoisseur's Brew only is made, its raypath intersects a = 4 at $F(a_F,r_F)=F(4,r_F)$ where $r_F=$ $a_F*m_C= 3$ and $x_C= r_F/r_C= a_F/a_C= 2.0$, leaving over $r-r_F = 6-3 = 3$ tons of robusta beans each week.

Figure 9.2.2
Graph of input (a,r)=(3,4) and \$9,000 isocost line.



$a-a_Tx_T = 3-1.0*1.6 = 1.4$ tons of arabica beans would be leftover each week which effectively costs \$2,600/1.6 = \$1,857.14 per ton. To determine production rates x_T and x_C without any leftovers, each input constraint is set equal to its use in Taster's and Connoisseur's Brew outputs as shown in material-balance equations (9.1.1) and (9.1.2).

$a_Tx_T + a_Cx_C = a$ or $1.0x_T+2.0x_C = 4$ tons/week of *arabica beans* ....(9.1.1)
$r_Tx_T + r_Cx_C = r$ or $2.5x_T+1.5x_C = 6$ tons/week of *robusta beans* ....(9.1.2)

To solve equations (9.1.1) and (9.1.2) for $x_T = x_T(a,r)$, multiply (9.1.1) by r_C and (9.1.2) by a_C and eliminate $a_C r_C x_C$ by subtraction to get equation (9.1.3). Because $x_T(a,r)$ cannot be negative, numerator $ar_C - a_C r$ in equation (9.1.3) must be negative from which it follows that $(r/a) \geq (r_C/a_C = 3/4)$. Partial derivatives $\partial x_T/\partial a = -3/7$ and $\partial x_T/\partial r = 4/7$ in equations (9.1.4) and (9.1.5) are called *marginal productivities* which measure the change in Taster's output rate due to changing only the arabica or robusta input rates respectively. Linear systems of equations exhibit *constant returns to scale* as shown by constant marginal productivities which are independent of the scale of the (a,r) input constraints.

$$x_T(a,r) = \frac{ar_C - a_C r}{a_T r_C - a_C r_T} = \frac{4\times 1.5 - 2.0\times 6}{1.0\times 1.5 - 2.0\times 2.5} = \frac{-6}{-3.5} = \frac{12}{7} \qquad ....(9.1.3)$$

$$\frac{\partial x_T}{\partial a} = \frac{r_C}{a_T r_C - a_C r_T} = \frac{1.5}{1.0\times 1.5 - 2.0\times 2.5} = \frac{1.5}{-3.5} = \frac{-3}{7} \qquad ....(9.1.4)$$

$$\frac{\partial x_T}{\partial r} = \frac{-a_C}{a_T r_C - a_C r_T} = \frac{-2.0}{1.0\times 1.5 - 2.0\times 2.5} = \frac{-2.0}{-3.5} = \frac{4}{7} \qquad ....(9.1.5)$$

Similarly, we can solve for $x_C = x_C(a,r)$ by multiplying (9.1.1) by r_T and (9.1.2) by a_T which enables us to eliminate $a_T r_T x_T$ by subtraction and obtain equation (9.1.6).

$$x_C(a,r) = \frac{a_T r - ar_T}{a_T r_C - a_C r_T} = \frac{1.0\times 6 - 4\times 2.5}{1.0\times 1.5 - 2.0\times 2.5} = \frac{-4.0}{-3.5} = \frac{8}{7} \qquad ....(9.1.6)$$

$$\frac{\partial x_C}{\partial a} = \frac{-r_T}{a_T r_C - a_C r_T} = \frac{-2.5}{1.0\times 1.5 - 2.0\times 2.5} = \frac{-2.5}{-3.5} = \frac{5}{7} \qquad ....(9.1.7)$$

$$\frac{\partial x_C}{\partial r} = \frac{a_T}{a_T r_C - a_C r_T} = \frac{1.0}{1.0\times 1.5 - 2.0\times 2.5} = \frac{1.0}{-3.5} = \frac{-2}{7} \qquad ....(9.1.8)$$

Because $x_C(a,r)$ cannot be negative, numerator $a_T r - ar_T$ in equation (9.1.6) must be negative from which it follows that $(r/a) \leq (r_T/a_T = 5/2)$. Partial differentiation of equation (9.1.6) results in equations (9.1.7) and (9.1.8). *Marginal productivities* $\partial x_C/\partial a = 5/7$ and $\partial x_C/\partial r = -2/7$ measure changes in Connoisseur's output rate due to changing only arabica or robusta input rates respectively. Partial derivatives $\partial x_C/\partial a$ and $\partial x_C/\partial r$ depend only on input/output coefficients a_T, a_C, r_T and r_C regardless of the scale of the (a,r) input constraints.

Engineering resource planning requires a map of (a,r) input and (x_T, x_C) output rates. For this purpose, it is convenient to use total differential equations (9.1.9) and (9.1.10) which are expressed in terms of the determinant $|A| = a_T r_C - a_C r_T = -7/2$.

$$dx_T(a,r) = \frac{\partial x_T}{\partial a}da + \frac{\partial x_T}{\partial r}dr = \frac{r_C}{|A|}da + \frac{-a_C}{|A|}dr = \frac{-3}{7}da + \frac{4}{7}dr \qquad ....(9.1.9)$$

$$dx_C(a,r) = \frac{\partial x_C}{\partial a}da + \frac{\partial x_C}{\partial r}dr = \frac{-r_T}{|A|}da + \frac{a_T}{|A|}dr = \frac{5}{7}da - \frac{2}{7}dr \qquad ....(9.1.10)$$

The map between (a,r) input and (x_T,x_C) output rates in Table 9.1.2 represent static equilibrium points that satisfy the engineering objective. The initial point at the upper left of Table 9.1.2 is (a,r) = (2,6) from which (x_T,x_C) = (18/7,-2/7) is calculated by equations (9.1.3) and (9.1.6). Since da and dr in equations (9.1.9) and (9.1.10) are independent variables, let da=1 and dr=0 to calculate (x_T,x_C) for column changes of (a,r), and let da=0 and dr=-1 to calculate (x_T,x_C) for row changes of (a,r). Thus, in the last column of Table 9.1.2, (a,r) goes from (2,3) to (5,3) and (x_T,x_C) goes from (6/7,4/7) to (-3/7,19/7). The input constraint ratio r/a must lie inside the interval $(r_T/a_T=5/2) \geq (r/a) \geq (r_C/a_C=3/4)$ in order for both (x_T,x_C) to be nonnegative. Hence, at the upper left of Table 9.1.2, x_C=-2/7 is negative when (a,r) = (2,6) because (r/a)=(6/2) is greater than $(r_T/a_T=5/2)$. Also, at the lower right of Table 9.1.2, x_T=-3/7 is negative when (a,r) = (5,3) because (r/a)=(3/5) is less than $(r_C/a_C=3/4)$.

Table 9.1.2 - Full-utilization Map Between (a,r) Input and (x_T,x_C) Output Rates.

(a,r)	(2,6)	(2,5)	(2,4)	(2,3)
(x_T,x_C)	(18/7,-2/7)	(14/7,0/7)	(10/7,2/7)	(6/7,4/7)
(a,r)	(3,6)	(3,5)	(3,4)	(3,3)
(x_T,x_C)	(15/7,3/7)	(11/7,5/7)	(7/7,7/7)	(3/7,9/7)
(a,r)	(4,6)	(4,5)	(4,4)	(4,3)
(x_T,x_C)	(12/7,8/7)	(8/7,10/7)	(4/7,12/7)	(0/7,14/7)
(a,r)	(5,6)	(5,5)	(5,4)	(5,3)
(x_T,x_C)	(9/7,13/7)	(5/7,15/7)	(1/7,17/7)	(-3/7,19/7)

Because $x_T(a,r)$ and $x_C(a,r)$ in equations (9.1.3) and (9.1.6) are both functions of (a,r), they need to be tested for functional independence. Jacobian determinant $|J|_X$ in (9.1.11) consists of all first-order partial derivatives of $x_T(a,r)$ and $x_C(a,r)$. The linear or nonlinear functional independence of equations (9.1.3) and (9.1.6) exist if and only if $|J|_X \neq 0$. Since $|J|_X$ = -2/7 in equation (9.1.11), $x_T(a,r)$ and $x_C(a,r)$ are functionally independent except if the determinant $|A| = a_T r_C - a_C r_T = 0$.

$$|J|_X = \left|\frac{\partial x_T, \partial x_C}{\partial a, \partial r}\right| = \begin{vmatrix} \frac{\partial x_T}{\partial a} & \frac{\partial x_T}{\partial r} \\ \frac{\partial x_C}{\partial a} & \frac{\partial x_C}{\partial r} \end{vmatrix} = \begin{vmatrix} \frac{r_C}{|A|} & \frac{-a_C}{|A|} \\ \frac{-r_T}{|A|} & \frac{a_T}{|A|} \end{vmatrix} = \frac{|A|}{|A|^2} = -\frac{2}{7} \qquad \text{....(9.1.11)}$$

The conditions under which determinant $|A|$ = 0 depend upon the input/output coefficients a_T, a_C, r_T and r_C. For example, if a_C and r_C were a multiple k of a_T and r_T, then determinant $|A| = a_T r_C - a_C r_T = k a_T r_T - k a_T r_T = 0$. This would mean that arabica and robusta beans are used in the same proportions for both products in which case Taster's and Connoisseur's Brews would merely be different names of the same product.

The LP solution of the product-mix problem is depicted graphically in Figure 9.1.1. Point C(a,r) = C(4,6) represents the weekly arabica and robusta bean input constraints. Taster's and Connoisseur's Brew raypath slopes are $m_T = r_T/a_T = 2.50$ and $m_C = r_C/a_C = 0.75$ respectively. If Taster's Brew only is made, its raypath intersects r = 6 at $E(a_E,r_E)=E(a_E,6)$ where $a_E= r_E/m_T= 2.4$ and $x_T= a_E/a_T= r_E/r_T= 2.4$, leaving over $a-a_E$ = 4-2.4 = 1.6 tons of arabica beans each week. If Connoisseur's Brew only is made, its raypath intersects a = 4 at $F(a_F,r_F)=F(4,r_F)$ where $r_F= a_F * m_C= 3$ and $x_C= r_F/r_C= a_F/a_C= 2.0$, leaving over $r-r_F$ = 6-3 = 3 tons of robusta beans each week.

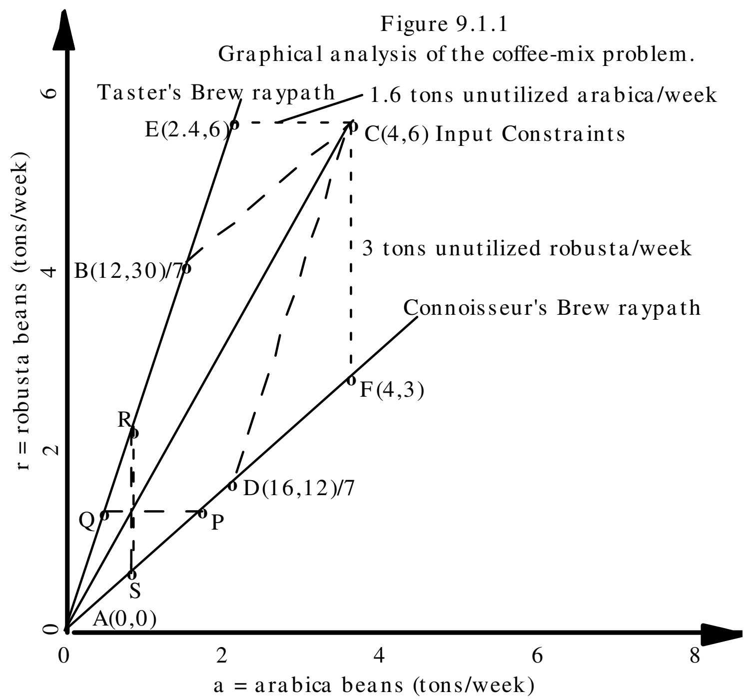


Figure 9.1.1
Graphical analysis of the coffee-mix problem.

Let us translate Connoisseur's raypath parallel to itself to points C(a,r) = C(4,6) and $B(a_B,r_B)$ on Taster's raypath. Since $r_B = m_T a_B$ and $r-r_B = m_C(a-a_B)$, the $B(a_B,r_B)$ coordinates are $a_B = (r-m_C a)/(m_T-m_C) = 12/7$ and $r_B = m_T a_B = 30/7$ as shown in Figure 9.1.1. Similarly, let us translate Taster's raypath parallel to itself to points C(a,r) = C(4,6) and $D(a_D,r_D)$ on Connoisseur's raypath. Since $r_D = m_C a_D$ and $r-r_D = m_T(a-a_D)$, the $D(a_D,r_D)$ coordinates are $a_D = (r-m_T a)/(m_C-m_T) = 16/7$ and $r_D = m_C a_B = 12/7$ as shown in Figure 9.1.1. Parallelogram ABCD satisfies input constraints $a = a_B+a_D = (12+16)/7 = 4$ and $r = r_B+r_D = (30+12)/7 = 6$ which map into $x_T = a_B/a_T = r_B/r_T = 12/7$ and $x_C = a_D/a_C = r_D/r_C = 8/7$ as shown in Table 9.1.2.

If C(a,r) lies between Taster's and Connoisseur's raypaths in Figure 9.1.1, then $B(a_B,r_B)$ and $D(a_D,r_D)$ are in the first quadrant which make (x_T,x_C) positive. But if C(a,r) lies above Taster's raypath (i.e., $(r/a) > (r_T/a_T) = m_T$), then D would be in the third quadrant of Connoisseur's raypath making $x_C < 0$ and $x_T > a/a_T$ in excess of the arabica input constraint. Thus, (a,r) = (2,6) maps into $(x_T,x_C) = (18/7,-2/7)$. Also, if C(a,r) lies below Connoisseur's raypath (i.e., $(r/a) < (r_C/a_C)$), B would be in the third quadrant making $x_T < 0$ and $x_C > r/r_C$ in excess of the robusta input constraint. Thus, (a,r) = (5,3) maps into $(x_T,x_C) = (-3/7,19/7)$. Therefore, $(r_T/a_T) \geq (r/a) \geq (r_C/a_C)$ is necessary for satisfying the engineering objective

Equations (9.1.4) and (9.1.7) can be derived from Figure 9.1.1 by locating $P(a_C,r_C)$ on Connoisseur's raypath where $x_P = 1$. Draw PQ parallel to CE making triangles APQ and BCE similar and determining $Q(r_C/m_T,r_C)$ on Taster's raypath where $x_Q = r_C/r_T$. By definition, $\partial x_T/\partial a$ is the change from A(0,0) to $x_Q = r_C/r_T$ divided by the change from a_P to a_Q with r_C fixed.

$$\frac{\partial x_T}{\partial a} = \frac{x_{AQ}}{a_{PQ}} = \frac{(r_C / r_T) - 0}{(r_C / m_T) - a_C} = \frac{r_C}{a_T r_C - a_C r_T} = \frac{r_C}{|A|} = \frac{-3}{7} \qquad \text{....(9.1.4)}$$

Similarly, $\partial x_C/\partial a$ is defined as the change from A(0,0) to x_P= 1 divided by the change from a_Q to a_P with r_C fixed.

$$\frac{\partial x_C}{\partial a} = \frac{x_{AP}}{a_{QP}} = \frac{1-0}{a_C - (r_C / m_T)} = \frac{r_T}{a_C r_T - a_T r_C} = \frac{-r_T}{|A|} = \frac{5}{7} \qquad \text{....(9.1.7)}$$

Equations (9.1.5) and (9.1.8) can be derived from Figure 9.1.1 by locating $R(a_T,r_T)$ on Taster's raypath where x_R= 1. Draw RS parallel to CF making triangles ARS and DCF similar and intersecting $S(a_T,a_T m_C)$ on Connoisseur's raypath where $x_S= a_T/a_C$. By definition, $\partial x_T/\partial r$ is the change from A(0,0) to $x_R= r_C/r_T$ divided by the change from r_R to r_S with a_T fixed.

$$\frac{\partial x_T}{\partial r} = \frac{x_{AR}}{r_{SR}} = \frac{1-0}{r_T - a_T m_C} = \frac{a_C}{a_C r_T - a_T r_C} = \frac{-a_C}{|A|} = \frac{4}{7} \qquad \text{....(9.1.5)}$$

Similarly, $\partial x_C/\partial r$ is defined as the change from A(0,0) to $x_S= a_T/a_C$ divided by the change from r_S to r_R with a_T fixed.

$$\frac{\partial x_C}{\partial r} = \frac{x_{AS}}{a_{RS}} = \frac{a_T / a_C}{a_T m_C - r_T} = \frac{a_T}{a_T r_C - a_C r_T} = \frac{a_T}{|A|} = \frac{-2}{7} \qquad \text{....(9.1.8)}$$

The graphical analysis of Figure 9.1.1 is difficult to display in problems with more than two input and two output variables. However, the analysis can be generalized to more variables by matrix algebra which can be carried out with microcomputer spreadsheets. For this reason, equations (9.1.1) and (9.1.2) are repeated in matrix form in equation (9.1.12).

$a_T x_T + a_C x_C = a$ or $1.0x_T+2.0x_C = 4$ (arabica material balance) ....(9.1.1)
$r_T x_T + r_C x_C = r$ or $2.5x_T+1.5x_C = 6$ (robusta material balance) ....(9.1.2)

$$AX = \begin{bmatrix} a_T & a_C \\ r_T & r_C \end{bmatrix} \begin{bmatrix} x_T \\ x_C \end{bmatrix} = \begin{bmatrix} a_T x_T + a_C x_C \\ r_T x_T + r_C x_C \end{bmatrix} = \begin{bmatrix} a \\ r \end{bmatrix} = B \qquad \text{....(9.1.12)}$$

Production-coefficient matrix A in equation (9.1.12) is postmultiplied by output column vector $X \equiv [x_T, x_C]^t$ to equal input column vector $B \equiv [a,r]^t$. When AX is expanded by matrix multiplication, the intermediate result is exactly the same as equations (9.1.1) and (9.1.2).

We can determine X from matrix equation $A \cdot X = B$ by premultiplying both sides with inverse matrix A^{-1} to get $A^{-1} \cdot A \cdot X = A^{-1} \cdot B$. Inverse matrix A^{-1} shown below was derived algebraically from production coefficients a_T, r_T, a_C, r_C and determinant $|A|$ of matrix A. Matrix product $A^{-1} \cdot A \cdot X$ reduces to IX = X because $A^{-1} \cdot A$ equals the identity matrix I as shown below. Consequently, we get $X = A^{-1} \cdot B$ in equation (9.1.13) which is expanded by matrix multiplication into equations (9.1.14) and (9.1.15) which are the same as equations (9.1.3) and (9.1.6) respectively.

$$A^{-1}A = \begin{bmatrix} \frac{r_C}{|A|} & \frac{-a_C}{|A|} \\ \frac{-r_T}{|A|} & \frac{a_T}{|A|} \end{bmatrix} \begin{bmatrix} a_T & a_C \\ r_T & r_C \end{bmatrix} = \begin{bmatrix} \frac{r_C a_T - a_C r_T}{|A|} & \frac{r_C a_C - a_C r_C}{|A|} \\ \frac{-r_T a_T + a_T r_T}{|A|} & \frac{-r_T a_C + a_T r_C}{|A|} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$X = \begin{bmatrix} x_T \\ x_C \end{bmatrix} = \begin{bmatrix} \frac{r_C}{|A|} & \frac{-a_C}{|A|} \\ \frac{-r_T}{|A|} & \frac{a_T}{|A|} \end{bmatrix} \begin{bmatrix} a \\ r \end{bmatrix} = \begin{bmatrix} \frac{\partial x_T}{\partial a} & \frac{\partial x_T}{\partial r} \\ \frac{\partial x_C}{\partial a} & \frac{\partial x_C}{\partial r} \end{bmatrix} \begin{bmatrix} a \\ r \end{bmatrix} = A^{-1}B \qquad \text{....(9.1.13)}$$

$$x_T(a, r) = \frac{r_C}{|A|} a - \frac{a_C}{|A|} r = \frac{-3}{7} a + \frac{4}{7} r; ..x_T(4,6) = \frac{-3}{7} 4 + \frac{4}{7} 6 = \frac{12}{7} \text{ tons / week} \qquad \text{....(9.1.14)}$$

$$x_C(a, r) = \frac{-r_T}{|A|} a + \frac{a_T}{|A|} r = \frac{5}{7} a - \frac{2}{7} r; .....x_C(4,6) = \frac{5}{7} 4 - \frac{2}{7} 6 = \frac{8}{7} \text{ tons / week} \qquad \text{....(9.1.15)}$$

Constant coefficients of equations (9.1.14) and (9.1.15) are the marginal-productivity elements of inverse matrix A^{-1} previously called first-order partial derivatives of the Jacobian determinant. Linear systems of equations have *constant returns to scale* where marginal productivities $\partial x_T/\partial a = r_C/|A|$, $\partial x_T/\partial r = -a_C/|A|$, $\partial x_C/\partial a = -r_T/|A|$ and $\partial x_C/\partial r = a_T/|A|$ depend only on coefficients a_T, r_T, a_C and a_C independently of the scale of (a,r) input constraints.

Matrix equations (9.1.12) to (9.1.15) may be calculated with spreadsheet programs as shown in Table 9.1.3. Cells A1 to D5 contain Table 9.1.1 data with matrix A = {B3:C4}. To evaluate determinant |A|, select B7 and enter cell-C7 command "=MDETERM(B3:C4)". Select cells {A9:B10} and enter cell-C9 command "=7*MINVERSE(B3:C4)" by pressing keys command-shift-return. Inverse matrix $7{*}A^{-1}$ is then returned in cells {A9:B10}. Similarly, matrix products $A^{-1}{\cdot}A$ and $A^{-1}{\cdot}B$ are executed and returned in cells {A12:B13} and {B15:B16}.

Table 9.1.3 - Microsoft Excel 5.0 Matrix Analysis of Coffee Product-Mix Problem
Error! Not a valid link.

Section 9.2 - Financial Objective in Linear Programming (LP)

In Section 9.1, the coffeet-mix problem was treated without any input or output prices. This procedure is admissable if we only need to satisfy the engineering objective of fully utilizing the (a,r) input constraints without any leftovers. Aside from the engineering objective, the financial objective is to maximize weekly sales revenues for given weekly input costs. More specifically, the financial objective function Z is defined in equation (9.2.1) as weekly sales revenues from Taster's and Connoisseur's Brew outputs. When the given weekly inputs are (a,r)=(4,6), then the outputs $(x_T,x_C)=(12/7,8/7)$, $(p_T,p_C)=(\$4000,\$5000)$ and Z=\$12,571 satisfy *both* the engineering and financial objectives.

$$Z(p_T,p_C,x_T,x_C) \equiv p_Tx_T+p_Cx_C = \$4,000(12/7)+\$5,000(8/7) = \$12,571 \qquad \text{....(9.2.1)}$$

If either $(p_T/p_C) > (r_T/r_C= 5/3)$ or $(p_T/p_C) < (a_T/a_C=1/2)$, the financial objective would be satisfied by producing only Taster's or Connoisseur's Brew respectively, even though the engineering objective would not be satisfied due to leftover constraints. When (a,r)=(4,6),

$(x_T,x_C)=(12/7,8/7)$ and $(p_T/p_C)=(\$5{,}500/\$2{,}750)>(5/3)$, then the engineering objective is satisfied with $Z = p_Tx_T+p_Cx_C= \$5{,}500(12/7)+\$2{,}750(8/7) = \$12{,}571$. But producing Taster's Brew only yields $(x_T,x_C)= (2.4,0)$ worth $Z = p_Tx_T= \$5{,}500*2.4 = \$13{,}200$ with 1.6 tons/week of arabica beans leftover. Also, when $(p_T/p_C)=(\$2{,}750/\$6{,}875)<(1/2)$, then $Z = \$2{,}750(12/7) + \$6{,}875(8/7) = \$12{,}571$. But producing Connoisseur's Brew only yields $(x_T,x_C)=(0,2)$ worth $Z = p_Cx_C= \$6{,}875*2 = \$13{,}750$ with 3 tons/week of robusta beans leftover.

Herein lies the distinction between engineering and financial objectives. The engineering objective uses equality input constraints in a production function which assumes output is maximized when inputs are efficiently combined without any leftovers. The financial objective uses input constraints to maximize outputs in a monetary sense. But if p_T/p_C is either greater than 5/3 or less than 1/2 as shown in Figure 9.2.1, then weekly sales from outputs of $(x_T,x_C) = (2.4,0)$ or $(x_T,x_C) = (0,2)$ with leftover input constraints would be greater than those from outputs of $(x_T,x_C) = (12/7,8/7)$ that satisfy the engineering objective.

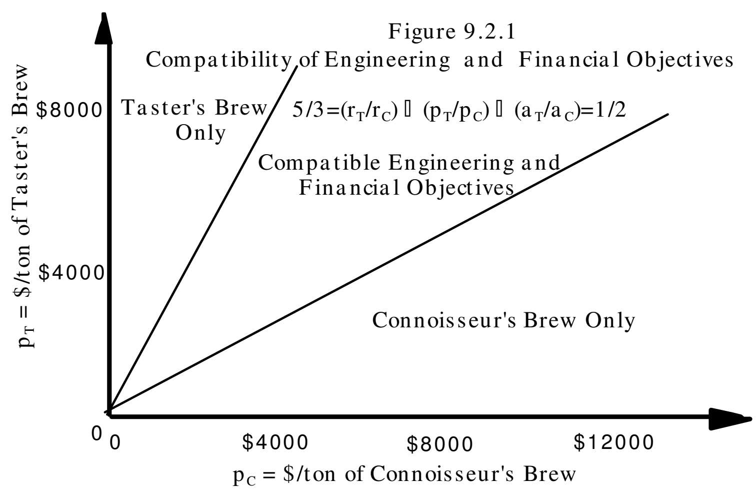


Figure 9.2.1
Compatibility of Engineering and Financial Objectives

The problem of maximizing $Z(p_T,p_C,x_T,x_C)$ subject to equality input constraints (9.1.1) and (9.1.2) is conveniently solved by *Lagrange undetermined multipliers*. The Lagrangian technique combines the financial objective function and the engineering input constraints into one equation, called the *Lagrangian function*. This insures that **(a)** the financial objective function will be kept intact and **(b)** the engineering objective will be satisfied. For these purposes, the input constraints will be written as $a\text{-}a_Tx_T\text{-}a_Cx_C = 0$ and $r\text{-}r_Tx_T+r_Cx_C = 0$. The Lagrangian function $L(x_T,x_C,y_a,y_r)$ is then defined as follows:

$$L(x_T,x_C,y_a,y_r) \equiv p_Tx_T+p_Cx_C + y_a(a\text{-}a_Tx_T\text{-}a_Cx_C) + y_r(r\text{-}r_Tx_T+r_Cx_C) \qquad \text{....(9.2.2)}$$

where y_a and y_r are undetermined multipliers, called the *implicit costs* of arabica and robusta beans. $L(x_T,x_C,y_a,y_r)$ is maximized by setting each partial derivative equal to zero.

$$\partial L/\partial x_T = p_T - a_Ty_a - r_Ty_r = 0; \qquad a_Ty_a + r_Ty_r = p_T \qquad \text{....(9.2.3)}$$

$$\partial L/\partial x_C = p_C - a_Cy_a - r_Cy_r = 0; \qquad a_Cy_a + r_Cy_r = p_C \qquad \text{....(9.2.4)}$$

$\partial L/\partial y_a = a - a_T x_T - a_C x_C = 0;$ $\quad a_T x_T + a_C x_C = a$ ....(9.2.5)

$\partial L/\partial y_r = r - r_T x_T - r_C x_C = 0;$ $\quad r_T x_T + r_C x_C = r$ ....(9.2.6)

Equations (9.2.5) and (9.2.6) are the same as material-balance rate equations (9.1.1) and (9.1.2) which are called the *primal* LP problem for satisfying the engineering objective as discussed in Section 9.1 . Cash-balance equations (9.2.3) and (9.2.4) are called the *dual* LP problem for satisfying the financial objective. Let us first analyze equations (9.2.3) and (9.2.4) of the dual LP problem in order to determine the mathematical conditions for satisfying *both* the engineering and financial objectives.

To solve equations (9.2.3) and (9.2.4) for $y_a=y_a(p_T,p_C)$, multiply (9.2.3) by r_C and (9.2.4) by r_T and eliminate $r_T r_C y_r$ by subtraction to get equation (9.2.7) for the implicit cost of arabica beans. Since $y_a(p_T,p_C)$ is presumed to be nonnegative and denominator $a_T r_C - a_C r_T$ of equation (9.2.7) is negative, numerator $p_T r_C - p_C r_T$ cannot be positive which means $(p_T/p_C) \le (r_T/r_C)=5/3$. *Marginal costs* $\partial y_a/\partial p_T = -3/7$ and $\partial y_a/\partial p_C = 5/7$ in equations (9.2.8) and (9.2.9) measure changes in arabica bean costs due to price changes of Taster's or Connoisseurs' Brews only. It is worth noting that $\partial y_a/\partial p_T = -3/7$ and $\partial y_a/\partial p_C = 5/7$ in equations (9.2.8) and (9.2.9) are the same as $\partial x_T/\partial a = -3/7$ and $\partial x_C/\partial a = 5/7$ in equations (9.1.4) and (9.1.7).

$$y_a(p_T,p_C) = \frac{p_T r_C - p_C r_T}{a_T r_C - a_C r_T} = \frac{4000 \times 1.5 - 5000 \times 2.5}{1.0 \times 1.5 - 2.0 \times 2.5} = \frac{-6500}{-3.5} = \frac{13000}{7} \quad ....(9.2.7)$$

$$\frac{\partial y_a}{\partial p_T} = \frac{r_C}{a_T r_C - a_C r_T} = \frac{\partial x_T}{\partial a} = \frac{1.5}{1.0 \times 1.5 - 2.0 \times 2.5} = \frac{1.5}{-3.5} = \frac{-3}{7} \quad ....(9.2.8)$$

$$\frac{\partial y_a}{\partial p_C} = \frac{-r_T}{a_T r_C - a_C r_T} = \frac{\partial x_C}{\partial a} = \frac{-2.5}{1.0 \times 1.5 - 2.0 \times 2.5} = \frac{-2.5}{-3.5} = \frac{5}{7} \quad ....(9.2.9)$$

Similarly, to solve equations (9.2.3) and (9.2.4) for $y_r=y_r(p_T,p_C)$, multiply (9.2.3) by a_C and (9.2.4) by a_T and eliminate $a_T a_C y_a$ by subtraction to get equation (9.2.10) for the implicit cost of robusta beans. Because $y_r(p_T,p_C)$ is presumed to be nonnegative, the $p_C a_T - p_T a_C$ numerator of (9.2.10) cannot be positive from which it follows that $(p_T/p_C) \ge (a_T/a_C)=1/2$.

$$y_r(p_T,p_C) = \frac{p_C a_T - p_T a_C}{a_T r_C - a_C r_T} = \frac{5000 \times 1.0 - 4000 \times 2.0}{1.0 \times 1.5 - 2.0 \times 2.5} = \frac{-3000}{-3.5} = \frac{6000}{7} \quad ....(9.2.10)$$

$$\frac{\partial y_r}{\partial p_T} = \frac{-a_C}{a_T r_C - a_C r_T} = \frac{\partial x_T}{\partial r} = \frac{-2.0}{1.0 \times 1.5 - 2.0 \times 2.5} = \frac{-2.0}{-3.5} = \frac{4}{7} \quad ....(9.2.11)$$

$$\frac{\partial y_r}{\partial p_C} = \frac{a_T}{a_T r_C - a_C r_T} = \frac{\partial x_C}{\partial r} = \frac{1.0}{1.0 \times 1.5 - 2.0 \times 2.5} = \frac{1.0}{-3.5} = \frac{-2}{7} \quad ....(9.2.12)$$

Partial differentials $\partial y_r/\partial p_T = 4/7$ and $\partial y_r/\partial p_C = -2/7$ in (9.2.11) and (9.2.12) are *marginal implicit costs* which measure change in robusta bean costs due to price changes of Taster's or Connoisseurs' Brews only. It is worth noting that $\partial y_r/\partial p_T = \partial x_T/\partial r = 4/7$ and $\partial y_r/\partial p_C = \partial x_C/\partial r = -2/7$ in equations (9.2.11) and (9.2.12) are the same as equations (9.1.5) and (9.1.8).

Financial planners for factory and plantation resources could map (p_T,p_C) output prices and (y_a,y_r) implicit costs from total differential equations (9.2.13) and (9.2.14).

$$dy_a(p_T,p_C) = \frac{\partial y_a}{\partial p_T}dp_T + \frac{\partial y_a}{\partial p_C}dp_C = \frac{r_C}{|A^t|}dp_T + \frac{-r_T}{|A^t|}dp_C = \frac{-3}{7}dp_T + \frac{5}{7}dp_C \quad \text{.... (9.2.13)}$$

$$dy_r(p_T,p_C) = \frac{\partial y_r}{\partial p_T}dp_T + \frac{\partial y_r}{\partial p_C}dp_C = \frac{-a_C}{|A^t|}dp_T + \frac{a_T}{|A^t|}dp_C = \frac{4}{7}dp_T - \frac{2}{7}dp_C \quad \text{.... (9.2.14)}$$

The map between (p_T,p_C) output prices and (y_a,y_r) implicit costs (in \$1,000's/ton) in Table 9.2.1 represent static equilibrium points which satisfy the engineering and financial objectives providing $2.5=(r_T/a_T) \geq (r/a) \geq (r_C/a_C)=(3/4)$ and $(5/3)=(r_T/r_C) \geq (p_T/p_C) \geq (a_T/a_C)=(1/2)$. The initial point at the upper left of Table 9.2.1 is $(p_T,p_C) = (3,7)$ from which we calculate $(y_a,y_r) = (26/7,-2/7)$ by equations (9.2.7) and (9.2.10). Since da and dr are independent variables in equations (9.2.13) and (9.2.14), let $dp_T=0$ and $dp_C=-1$ to calculate (y_a,y_r) for column changes of (p_T,p_C), and let $dp_T=1$ and $dp_C=-1$ to calculate (y_a,y_r) for row changes of (p_T,p_C). Thus, in the first column of Table 9.2.1, (p_T,p_C) goes from (3,7) to (3,4) while (y_a,y_r) goes from (26/7,-2/7) to (11/7,4/7). It is worth noting when $(p_T/p_C) < (1/2)$ in the upper left of Table 9.2.1, then $y_r<0$; and when $(p_T/p_C) > (5/3)$ in the lower right of Table 9.2.1, then $y_a<0$.

Table 9.2.1 - Financial Map Between (p_T,p_C) Output Prices and (y_a,y_r) Implicit Costs.

(p_T,p_C)	(3,7)	(4,6)	(5,5)	(6,4)
(y_a,y_r)	(26/7,-2/7)	(18/7,4/7)	(10/7,10/7)	(2/7,16/7)
(p_T,p_C)	(3,6)	(4,5)	(5,4)	(6,3)
(y_a,y_r)	(21/7,0/7)	(13/7,6/7)	(5/7,12/7)	(-3/7,18/7)
(p_T,p_C)	(3,5)	(4,4)	(5,3)	(6,2)
(y_a,y_r)	(16/7,2/7)	(8/7,8/7)	(0/7,14/7)	(-8/7,20/7)
(p_T,p_C)	(3,4)	(4,3)	(5,2)	(6,1)
(y_a,y_r)	(11/7,4/7)	(3/7,10/7)	(-5/7,16/7)	(-13/7,22/7)

Because $y_a(p_T,p_C)$ and $y_r(p_T,p_C)$ implicit costs in equations (9.2.7) and (9.2.10) are both functions of (p_T,p_C) output prices, they need to be tested for functional independence. The Jacobian determinant $|J|_Y$ shown in (9.2.15) consists of all first-order partial derivatives of $y_a(p_T,p_C)$ and $y_r(p_T,p_C)$. The functional independence of equations (9.2.7) and (9.2.10) exist if and only if $|J|_Y \neq 0$. Since transpose determinant $|A^t| = |A| = a_T r_C - a_C r_T$, it follows that $|J|_Y = -2/7$ in equation (9.2.15) just as $|J|_X = -2/7$ in equation (9.1.11). Therefore, $y_a(p_T,p_C)$ and $y_r(p_T,p_C)$ are functionally independent except if $|A^t| = |A| = a_T r_C - a_C r_T = 0$ which could occur if a_C and r_C were a common multiple of a_T and r_T. This would mean Taster's and Connoisseur's Brews use arabica and robusta beans in the same proportion.

$$|J|_Y = \left|\frac{\partial y_a, \partial y_r}{\partial p_T, \partial p_C}\right| = \begin{vmatrix} \frac{\partial y_a}{\partial p_T} & \frac{\partial y_a}{\partial p_C} \\ \frac{\partial y_r}{\partial p_T} & \frac{\partial y_r}{\partial p_C} \end{vmatrix} = \begin{vmatrix} \frac{r_C}{|A^t|} & \frac{-r_T}{|A^t|} \\ \frac{-a_C}{|A^t|} & \frac{a_T}{|A^t|} \end{vmatrix} = \frac{|A^t|}{|A^t|^2} = -\frac{2}{7} \quad \text{....(9.2.15)}$$

Although the first and last two equations from (9.2.3) to (9.2.6) are functionally independent, all four equations are linearly dependent as shown by multiplying equations (9.2.3) with x_T, (9.2.4) with x_C, (9.2.5) with y_a and (9.2.6) with y_r.

$$a_T x_T y_a + r_T x_T y_r = p_T x_T \quad ....(9.2.16)$$
$$a_C x_C y_a + r_C x_C y_r = p_C x_C \quad ....(9.2.17)$$
$$a_T x_T y_a + a_C x_C y_a = a y_a \quad ....(9.2.18)$$
$$r_T x_T y_r + r_C x_C y_r = r y_r \quad ....(9.2.19)$$

When equations (9.2.16) and (9.2.17) are added together, we get equation (9.2.20) below. Also, adding equations (9.2.18) and (9.2.19) together gives equation (9.2.21) below.

$$a_T x_T y_a + r_T x_T y_r + a_C x_C y_a + r_C x_C y_r = p_T x_T + p_C x_C \quad ....(9.2.20)$$
$$a_T x_T y_a + a_C x_C y_a + r_T x_T y_r + r_C x_C y_r = a y_a + r y_r \quad ....(9.2.21)$$

Since the lefthand sides of equations (9.2.20) and (9.2.21) are equal, it follows that the righthand sides of these two equations must also be equal as shown in equation (9.2.22).

$$Z(p_T,p_C,x_T,x_C) \equiv p_T x_T + p_C x_C = a y_a + r y_r \quad ....(9.2.22)$$

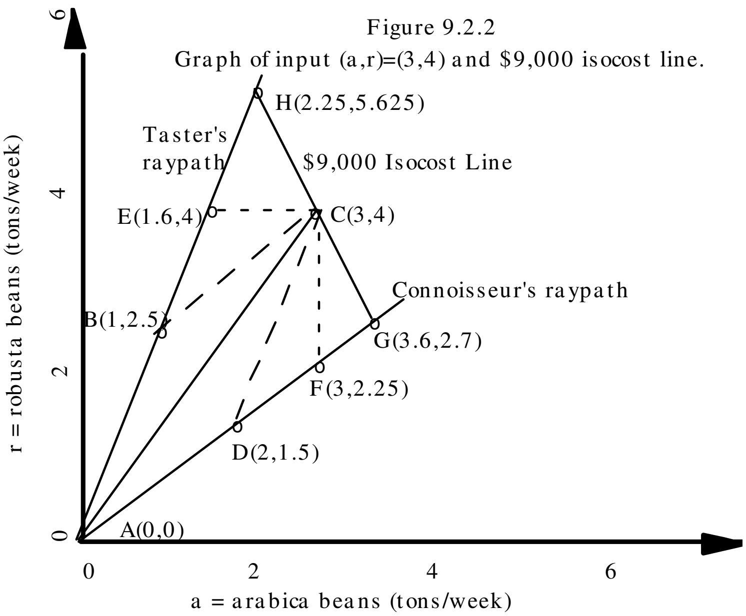
Equation (9.2.22) expresses a basic characteristic of LP problems, namely, that Lagrange undetermined multipliers (i.e., implicit costs) allocate total sales revenues exactly among the input constraints. This property of Lagrange undetermined multipliers enables the dual LP problem to be formulated as an alternative to the primal LP problem for the purposes of computing solutions and checking results.

Interpretations of equation (9.2.22) often refer to Lagrange undetermined multipliers y_a and y_r under such synonyms as "implicit values", "shadow prices", "opportunity costs", "internal transfer prices", "incremental prices", "breakeven prices", "dual variables" and "efficiency prices" which have different connotations than actual market prices of the input constraints. However, interpreting equation (9.2.22) by these synonyms is misleading because the coffee producer would not be willing to pay “opportunity costs” for relaxing input constraints if the resulting sales revenues would only breakeven with the change of input costs. This would not leave any room for other operating costs or any margin for profit.

In simple terms, Lagrange undetermined multipliers y_a and y_r are defined as the change of total weekly sales revenues divided by the corresponding change of an input while other inputs are fixed. Since $y_a(p_T,p_C)$ and $y_r(p_T,p_C)$ are independent of the input constraints (see equations (9.2.7) and (9.2.10)), we can graph an isocost line through an arbitrary (a,r) = (3,4) input constraint as shown in Figure 9.2.2 without any loss of generality.

Figure 9.2.2

Graph of input (a,r)=(3,4) and $9,000 isocost line.



By definition, y_a equals the change of weekly sales divided by the change of arabica bean input keeping the robusta bean input fixed. Let us determine the decrease of weekly sales from $3y_a+4y_r$ at C(3,4) to $E(a_E,4) = E(4/m_T,4) = E(1.6,4)$ where weekly sales of Taster's Brew are $x_T= 1.6/a_T= 4/r_T= 1.6$ tons worth $p_T x_T$=$4,000*1.6=$6,400. The corresponding decrease in arabica bean input from C(3,4) to E(1.6,4) is in the denominator of equation (9.2.23) which is simplified on the right to the coordinates of E(1.6,4).

$$y_a = \frac{3y_a + 4y_r - \$6{,}400}{3 - 1.6} \quad \text{or} \quad 1.6y_a + 4y_r = \$6{,}400 \qquad ....(9.2.23)$$

Likewise, y_r equals the change of weekly sales divided by the change of robusta bean input keeping the arabica bean input fixed. Let us determine the decrease of weekly sales from $3y_a+4y_r$ at C(3,4) to $F(3,r_F) = F(3,3m_C) = F(3,2.25)$ where weekly sales of Connoisseur's Brew are $x_C= 3/a_C= 2.25/r_C= 1.5$ tons worth $p_C x_C$=$5,000*1.5=$7,500. The decrease in robusta bean input from C(3,4) to F(3,2.25) is in the denominator equation (9.2.24) which is simplified on the right to the coordinates of F(3,2.25).

$$y_r = \frac{3y_a + 4y_r - \$7{,}500}{4 - 2.25} \quad \text{or} \quad 3y_a + 2.25y_r = \$7{,}500 \qquad ....(9.2.24)$$

To solve for y_a, let us multiply (9.2.23) by 2.25 and (9.2.24) by 4 and eliminate $9y_r$ by subtraction to get equation (9.2.25).

$$y_a = \frac{\$30{,}000 - \$14{,}400}{12 - 3.6} = \frac{\$15{,}600}{8.4} = \frac{\$13{,}000}{7} = \$1{,}857\ /\ \text{ton} \quad \text{arabica beans} \qquad ....(9.2.25)$$

Similarly, to solve for y_r, let us multiply (9.2.23) by 3 and (9.2.24) by 1.6 and eliminate $4.8y_a$ by subtraction to get equation (9.2.26).

$$y_r = \frac{\$19{,}200 - \$12{,}000}{12 - 3.6} = \frac{\$7{,}200}{8.4} = \frac{\$6{,}000}{7} = \$857 \text{ / ton robusta beans} \qquad ....(9.2.26)$$

The equivalence between equations (9.2.7) and (9.2.25), and between (9.2.10) and (9.2.26) illustrates y_a and y_r are independent of the (a,r) input constraints. Moreover, the weekly sales revenues $p_Tx_T+p_Cx_C = ay_a+ry_r$ of equation (9.2.22) may now be verified for the data (p_T,p_C) = (\$4000,\$5000) and (y_a,y_r) = (\$13000/7,\$6000/7) in Table 9.2.1, and (a,r) = (3,4) and (x_T,x_C) = (7/7,7/7) = (1,1) in Table 9.1.2 as follows:

\$4,000*7/7 + \$5,000*7/7 = [3*\$13,000+4*\$6,000]/7 = \$63,000/7 = \$9,000/week

Basic properties of LP problems are conveniently described by *isoquants*. The prefix "iso" means equal. An isoquant curve represents all the combinations of inputs that have a specified quantity of costs or revenues. The \$9,000-isocost line in Figure 9.2.2 through C(3,4) also intersects points G(3.6,2.7) and H(2.25,5.625). At G(3.6,2.7) where $(x_T,x_C) = (0,3.6/a_C)$, we have p_Cx_C= \$5000*3.6/a_C = ay_a+ry_r= [3.6*\$13000+2.7*\$6000]/7 = \$9,000/week. Similarly, at H(2.25,5.625) where $(x_T,x_C) = (2.25/a_T,0)$, we have p_Tx_T = \$4000*2.25/a_T = ay_a+ry_r= [2.25*\$13,000+5.625*\$6,000]/7 = \$9,000/week.

The \$9,000-isocost line intercepts on the (a,r) axes are \$9,000/y_a and \$9,000/y_r respectively. Therefore, the equation of the \$9,000-isocost line is

$$r = -[(\$9{,}000/y_a)/(\$9{,}000/y_r)]a + \$9{,}000/y_r = -(y_a/y_r)a + \$9{,}000/y_r \qquad ....(9.2.27)$$

All isocost lines are parallel with slopes $-(y_a/y_r)$. When (p_T/p_C)=(\$3,000/\$6,000)=(a_T/a_C)=1/2, the \$9,000-isocost line is vertical because the numerator of equation (9.2.10) equals zero making y_r= 0. The intercept of the vertical \$9,000-isocost line on the a-axis is \$9,000/y_a= 3 or y_a= \$3,000/ton of arabica beans. When (p_T/p_C)=(\$5,625/\$3,375)=(r_T/r_C)=5/3, the \$9,000-isocost line is horizontal because the numerator of equation (9.2.7) equals zero making y_a= 0. The intercept of the horizontal \$9,000-isocost line on the r-axis is \$9,000/y_r= 4 or y_r= \$2,250/ton of robusta beans.

Input constraint C(a,r) divides the length of the isocost line between Connoisseur's and Taster's raypaths in the inverse ratio of the weekly sales revenues of the outputs. Thus, the ratio of distances GC/CH in Figure 9.2.2 equals p_Tx_T/p_Cx_C as shown in equation (9.2.28).

$$\frac{GC}{CH} = \frac{\sqrt{(3.6-3)^2+(2.7-4)^2}}{\sqrt{(3-2.25)^2+(4-5.625)^2}} = \frac{4}{5} = \frac{p_T x_T}{p_C x_C} = \frac{\$4{,}000 \times 1}{\$5{,}000 \times 1} \qquad ....(9.2.28)$$

The graphical analysis of Figure 9.2.2 is difficult to display in problems with more than two input and two output variables. However, the analysis can be generalized to more variables by matrix algebra which can be carried out with microcomputer spreadsheets. For this reason, equations (9.2.3) and (9.2.4) are expressed in matrix form in equation (9.2.29). In the matrix notation $A^tY=P$ of equation (9.2.29), A^t is the transpose of matrix A. The solution of $A^tY=P$ is $Y= [A^t]^{-1}P = [A^{-1}]^tP$ which is shown in equation (9.2.30) below.

$$A^t Y = \begin{bmatrix} a_T & r_T \\ a_C & r_C \end{bmatrix}\begin{bmatrix} y_a \\ y_r \end{bmatrix} = \begin{bmatrix} a_T y_a + r_T y_r \\ a_C y_a + r_C y_r \end{bmatrix} = \begin{bmatrix} p_T \\ p_C \end{bmatrix} = P \qquad ....(9.2.29)$$

$$Y = \begin{bmatrix} y_a \\ y_r \end{bmatrix} = \begin{bmatrix} \frac{r_C}{|A|} & \frac{-r_T}{|A|} \\ \frac{-a_C}{|A|} & \frac{a_T}{|A|} \end{bmatrix}\begin{bmatrix} p_T \\ p_C \end{bmatrix} = \begin{bmatrix} \frac{\partial y_a}{\partial p_T} & \frac{\partial y_a}{\partial p_C} \\ \frac{\partial y_r}{\partial p_T} & \frac{\partial y_r}{\partial p_C} \end{bmatrix}\begin{bmatrix} p_T \\ p_C \end{bmatrix} = \left[A^t\right]^{1} P \qquad ....(9.2.30)$$

Matrix equation (9.2.30) is equivalent to equations (9.2.7) and (9.2.10) which are evaluated from the data a_T= 1.0, r_T=2.5, a_C= 2.0 and r_C= 1.5 of Table 9.1.1. The value of determinant $|A|=|A^t|$ is $a_T r_C - a_C r_T$ = (1.0)(1.5) - (2.0)(2.5) = -3.5 = -7/2. The Microsoft Excel 5.0 solution of equation (9.2.30) is analogous to the solution of equation (9.1.13) in Table 9.1.3 .

$$y_a(p_T, p_C) = \frac{r_C}{|A|} p_T - \frac{r_T}{|A|} p_C = \frac{-3}{7}\$4000 + \frac{5}{7}\$5000 = \frac{\$13000}{7} = \$1857\ /\ \text{ton} \qquad ....(9.2.7)$$

$$y_r(p_T, p_C) = \frac{-a_C}{|A|} p_T + \frac{a_T}{|A|} p_C = \frac{4}{7}\$4000 - \frac{2}{7}\$5000 = \frac{\$6000}{7} = \$857\ /\ \text{ton} \qquad ....(9.2.10)$$

The mappings between $X \equiv (x_T, x_C)^t$ and $B \equiv (a,r)^t$ in tons/week and between $Y \equiv (y_a, y_r)^t$ and $P \equiv (p_T, p_C)^t$ in prices/ton are summarized by matrix equations (9.1.13) and (9.2.30).

$$X = \begin{bmatrix} x_T \\ x_C \end{bmatrix} = \begin{bmatrix} \frac{r_C}{|A|} & \frac{-a_C}{|A|} \\ \frac{-r_T}{|A|} & \frac{a_T}{|A|} \end{bmatrix}\begin{bmatrix} a \\ r \end{bmatrix} = \begin{bmatrix} \frac{\partial x_T}{\partial a} & \frac{\partial x_T}{\partial r} \\ \frac{\partial x_C}{\partial a} & \frac{\partial x_C}{\partial r} \end{bmatrix}\begin{bmatrix} a \\ r \end{bmatrix} = \begin{bmatrix} \frac{-3}{7} & \frac{4}{7} \\ \frac{5}{7} & \frac{-2}{7} \end{bmatrix}\begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{12}{7} \\ \frac{8}{7} \end{bmatrix} = A^{-1}B \qquad ....(9.1.13)$$

$$Y = \begin{bmatrix} y_a \\ y_r \end{bmatrix} = \begin{bmatrix} \frac{r_C}{|A|} & \frac{-r_T}{|A|} \\ \frac{-a_C}{|A|} & \frac{a_T}{|A|} \end{bmatrix}\begin{bmatrix} p_T \\ p_C \end{bmatrix} = \begin{bmatrix} \frac{\partial y_a}{\partial p_T} & \frac{\partial y_a}{\partial p_C} \\ \frac{\partial y_r}{\partial p_T} & \frac{\partial y_r}{\partial p_C} \end{bmatrix}\begin{bmatrix} p_T \\ p_C \end{bmatrix} = \begin{bmatrix} \frac{-3}{7} & \frac{5}{7} \\ \frac{4}{7} & \frac{-2}{7} \end{bmatrix}\begin{bmatrix} \$4000 \\ \$5000 \end{bmatrix} = \begin{bmatrix} \$1857 \\ \$857 \end{bmatrix} = \left[A^t\right]^{1} P \qquad ....(9.2.30)$$

Comparisons of A^{-1} and $[A^t]^{-1}$ indicate $\partial x_T/\partial a = \partial y_a/\partial p_T = r_C/|A| = -3/7$; $\partial x_T/\partial r = \partial y_r/\partial p_T = -a_C/|A| = 4/7$; $\partial x_C/\partial a = \partial y_a/\partial p_C = -r_T/|A| = 5/7$; and $\partial x_C/\partial r = \partial y_r/\partial p_C = a_T/|A| = -2/7$. Equations (9.1.13) and (9.2.30) must both obey production constraints $(5/2)=(r_T/a_T) > (r/a) > (r_T/a_T)=(3/4)$. If price constraints $(5/3)=(r_T/r_C) > (p_T/p_C) > (a_T/a_C)=(1/2)$ are satisfied, isocost lines have negative slopes and engineering and financial objectives are compatible as shown in Figure 9.2.1.

Section 9.3 - Inequality Constraints in Linear Programming (LP)

Economic decisions faced by engineering and financial managers often involve large and unequal numbers of input and output variables subject mostly to inequality constraints. Such problems are much more complex than the coffee-mix problem which had only two input and two output variables subject to equality constraints. However, our purpose in presenting the coffee-mix problem was not so much in the mechanics of its solution as it was to illustrate the similarities and differences between engineering and financial objectives, to evaluate the effectiveness of constraints, and to examine the economic significance of primal and dual solutions of LP formulations. Moreover, LP problems that have unique optimal solutions end up with equal numbers of input and output variables subject to equality constraints just like the coffee-mix problem.

The need for solving LP problems arose from concerns about theories of input-output measurements during the Great Depression of the 1930s. This lead Professor Wassily Leontieff of Harvard University to develop a linear programming model representing the entire U.S. economy. In 1947, George B. Dantzig who was on assignment with the U.S. Air Force after World War II, developed the simplex method of solving the LP problem formulated by Leontieff. The simplex method is a computational routine which works in a repeated pattern to reach the best solution of a LP model formulation.

Each iteration of the simplex method yields a value of a linear objective function that is at least as good as (and usually better than) the previous solution. Using the simplex method to solve even small-scale LP problems by hand is a tedious process. For this reason, the simplex algorithm has been coded into programs which are available for use on mainframe computers (i.e., MPSX and SAS/OR), on free-standing microcomputers (i.e., LINDO and LPSBA), and on programs which interface with microcomputer spreadsheets such as LOTUS and EXCEL. Recently, Narendra Karmarkar of AT&T Bell Laboratories developed a new method of solving very large scale LP problems which would be impractical to solve by the simplex method.

Because LP models usually describe business, industrial, and governmental problems in terms of inequality rather than equality constraints, it is necessary to examine how inequality constraints affect the solutions. The simplex algorithm is generally required for solving LP models with inequality constraints. Many excellent textbooks describe how to solve LP problems by the simplex method, the references for which can be found at the end of this chapter. Therefore, there is no need for a detailed description of the simplex algorithm in this section. Instead, we can use algebraic and graphical methods of formulating and solving relatively simple LP problems with inequality constraints. The methodology is illustrated with the coffee-mix problem given in Table 9.3.1 below which was reproduced from previous Table 9.1.1 for convenience.

Table 9.3.1 - Arabica and Robusta Beans Required for Taster's and Connoisseur's Brews

	Tons of Coffee-Bean Input Required per Ton of Instant Coffee Output		Coffee-Bean Input Constraints (tons/week)
	Taster's Brew	Connoisseur's Brew	
Arabica Beans	a_T=1.0	a_C=2.0	a=4 tons/week
Robusta Beans	r_T=2.5	r_C=1.5	r=6 tons/week
Output Prices/ton	p_T=\$4,000	p_C=\$5,000	

The ***primal*** LP for the coffee-mix problem specifies that the (a,r) weekly input constraints are less than, or at most equal to, the arabica and robusta inputs in the x_T and x_C weekly output decision variables. The algebraic expressions for these inequalities are

$a_Tx_T+a_Cx_C = 1.0x_T+2.0x_C \leq a=4$ tons/week (arabica inequality constraint) ....(9.3.1)

$r_Tx_T+r_Cx_C = 2.5x_T+1.5x_C \leq r=6$ tons/week (robusta inequality constraint) ....(9.3.2)

It would be meaningless for output decision variables x_T and x_C to be negative. Therefore, *nonnegativity constraints* are imposed on x_T and x_C as follows:

$x_T \geq 0$ (Taster's Brew nonnegativity constraint) ....(9.3.3)

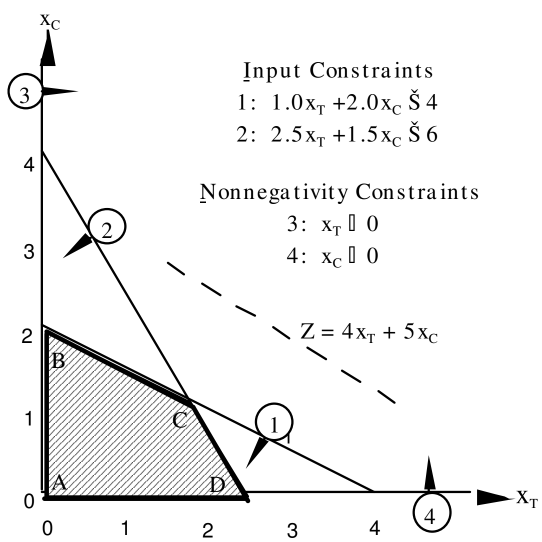
$x_C \geq 0$ (Connoisseur's Brew nonnegativity constraint) ....(9.3.4)

The objective function of the primal LP problem is the weekly sales revenues $Z(x_T,x_C) \equiv x_Tp_T+x_Cp_C$ at given output prices of p_T=4 and p_C=5 (in \$1,000's/ton). The object of the primal LP problem is to maximize objective function (9.3.5) by determining output decision variables x_T and x_C subject to inequality constraints (9.3.1) to (9.3.4).

$Z(x_T,x_C) \equiv p_Tx_T+p_Cx_C = 4x_T+5x_C$ (in \$1,000's/week) ....(9.3.5)

A primal LP problem involving two output variables can readily be displayed and solved by *graphical optimization*. The x_T and x_C axes of Figure 9.3.1 represent the *output decision variables*. Constraints (9.3.1) to (9.3.4) are represented by straight-line equations and the arrows show the directions in which their inequalities are not violated. Intersections A, B, C and D of two constraints are called *extreme* or *corner* points, the coordinates of which may be optimal values of x_T and x_C output decision variables.

Figure 9.3.1 - Graphical Analysis of the Primal Linear Programming Problem



The shaded *feasible region* on the graph is formed by connecting extreme points A, B, C, and D with straight lines. The feasible region is a convex polygon which constitutes the *solution space* of the problem because it consists of all points (x_T,x_C) that satisfy all constraints. Figure 9.3.1 shows objective function $Z(x_T,x_C)$ as a dashed line for an arbitrarily large value. As the dashed line is translated parallel to itself away from the origin towards B, C and D, its value becomes as large or larger than previously, until maximum occurs at a corner point or edge of the feasible region as summarized in Table 9.3.2 below.

Point D is where $2.5x_T+1.5x_C= 6$ intersects the x_T-axis. Solving $2.5x_T+1.5x_C= 6$ and $x_C=0$ gives $x_T = 2.4$. Thus, $D(x_T,x_C)=D(2.4,0)$ and $Z(2.4,0) = 4\cdot 2.4+5\cdot 0$ = \$9,600.

Point B is where $x_T+2x_C = 4$ intersects the x_C-axis. Solving $x_T+2x_C= 4$ and $x_T= 0$ gives $x_C= 2$. Thus, $B(x_T,x_C) = B(0,2)$ and $Z(0,2) = 4\cdot 0+5\cdot 2$ = \$10,000.

Point C is where lines $x_T+2x_C= 4$ and $2.5x_T+1.5x_C= 6$ intersect. Solving $x_T+2x_C= 4$ and $2.5x_T+1.5x_C= 6$ simultaneously gives $x_T= 12/7$ and $x_C = 8/7$. Thus, $C(x_T,x_C) = C(12/7,8/7)$ and $Z(12/7,8/7) = 4\cdot(12/7)+5\cdot(8/7) = 88/7$ = \$12,571.

Table 9.3.2 - Primal Objective Function and Constraint Utilization at Extreme Points

Extreme Points (x_T,x_C)	$(4x_T,5x_C)$	$Z=4x_T+5x_C$	$1.0x_T+2.0x_C$	$2.5x_T+1.5x_C$
$D(x_T,x_C)= D(2.4,0)$	(9.600, 0.000)	9.600	2.4	6.0
$B(x_T,x_C)= B(0,2)$	(0.000, 10.000)	10.000	4.0	3.0
$C(x_T,x_C)= C(12/7,8/7)$	(6.857, 5.714)	<u>12.571</u>	4.0	6.0

As $Z=4x_T+5x_C$ moves parallel to itself away from the origin, it first intersects D(2.4,0) where Z(2.4,0)=\$9,600/week with 2.4 tons of arabica beans unused. Point B(0,2) is next where Z(0,2) = \$10,000/week with 3.0 tons of robusta beans unused. Point C(12/7,8/7) is last where Z(12/7,8/7) = \$6,857.14+\$5,714.29 = \$12,571.43/week which uses $1.0x_T+2.0x_C= 4$ tons of arabica and $2.5x_T+1.5x_C= 6$ tons of robusta beans to satisfy both the engineering and financial objectives. The slope, $-p_T/p_C$, of objective function $Z(x_T,x_C) \equiv x_Tp_T+x_Cp_C$ may equal either the -1/2 slope of line BC or the -5/3 slope of line CD. Then all points on BC or CD would satisfy the financial objective but B and D would not satisfy the engineering objective.

The ***dual*** LP for the coffee-mix problem requires the p_T and p_C prices of coffee-brew outputs to be either smaller than or equal to the *implicit costs* of their coffee-bean inputs. This is expressed algebraically by inequalities (9.3.6) and (9.3.7) which are similar to equations (9.2.3) and (9.2.4). Assuming implicit costs y_a and y_r could not be negative, inequalities (9.3.8) and (9.3.9) are imposed as *nonnegativity constraints*. The objective function of the dual LP problem are the implicit costs $W(y_a,y_r) \equiv ay_a+ry_r$ of the coffee-bean inputs per week as shown in (9.3.10).

$a_Ty_a + r_Ty_r = 1.0y_a+2.5y_r \geq p_T$ (Taster's price constraint) ....(9.3.6)

$a_Cy_a + r_Cy_r = 2.0y_a+1.5y_r \geq p_C$ (Connoisseur's price constraint) ....(9.3.7)

$y_a \geq 0$ (arabica implicit-cost nonnegativity constraint) ....(9.3.8)

$y_r \geq 0$ (robusta implicit-cost nonnegativity constraint) ....(9.3.9)

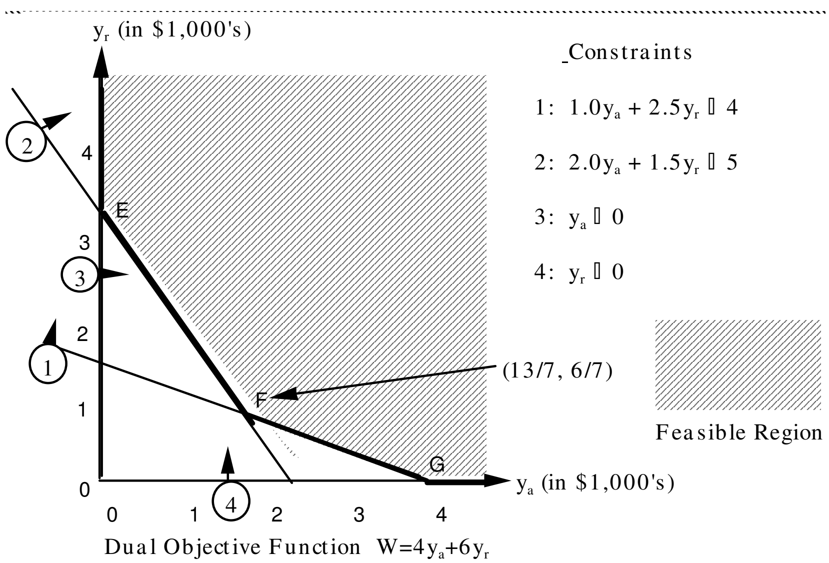
$W(y_a,y_r) \equiv ay_a+ry_r = 4y_a+6y_r$ (in \$1,000's/week) ....(9.3.10)

The object of the dual LP problem is to determine *implicit-cost decision variables* y_a and y_r which minimize $W(y_a,y_r)$ subject to constraints (9.3.6) to (9.3.9). Greater-than-or-equal ("≥") p_T and p_C constraints of (9.3.6) and (9.3.7) are coefficients of x_T and x_C in primal objective $Z(x_T,x_C)$, and less-than-or-equal ("≤") 'a' and 'r' constraints of (9.3.1) and (9.3.2) are coefficients of y_a and y_r in dual objective $W(y_a,y_r)$. Maximizing $Z(x_T,x_C)$ and minimizing $W(y_a,y_r)$ then become interchangeable. This is a major characteristic of the *simplex algorithm* developed by G. B. Dantzig in the 1940s (see "Linear Programming and Extensions" by George B. Dantzig, Princeton University Press, Princeton, N. J., 1963).

The simplex algorithm requires inequalities to be converted into strict equalities by adding *slack variables* to less-than-or-equal constraints and subtracting *surplus variables* from greater-than-or-equal constraints. The iterations of the simplex algorithm then sort out which variables are nonbinding or redundant. In well-defined LP problems, unique solutions contain equal numbers of remaining input and output variables. Readers interested in details of the sinplex method may find references at the end of the chapter.

The dual LP for the coffee-mix problem can also be solved graphically in a manner similar to the method previously described in Figure 9.3.1 for the primal problem. However, the graph of the dual problem uses (y_a,y_r) rather than (x_T,x_C) coordinates, and the feasible region of the dual is a convex polygon in the upper right-hand corner of Figure 9.3.2 rather than the lower lefthand corner of Figure 9.3.1. The Taster's and Connoisseur's Brews price-constraint inequalities (9.3.6) to (9.3.9) were plotted in Figure 9.3.2 with arrows showing directions in which the constraints are not violated.

Figure 9.3.2 - Graphical Analysis of the Dual LP Problem



In order to determine the optimal solution, objective function $W(y_a,y_r)$ is drawn as a dashed line in Figure 9.3.2 . At the origin where $W(y_a,y_r) = 0$, the dashed line is translated parallel to itself towards the feasible region. The minimum $W(y_a,y_r)$ value occurs where the dashed line first touches either a corner point or edge of the feasible region.

Point F lies where $2y_a+1.5y_r=5$ and $y_a+2.5y_r=4$ intersect. Solving these equations gives $y_a=13/7$ and $y_r=6/7$. Thus, $F(y_a,y_r)=F(13/7,6/7)$ and $W(y_a,y_r)=W(13/7,6/7)= 4\cdot(13/7)+6\cdot(6/7)=88/7$ or \$12,571/week. Point F corresponds to $C(x_T,x_C)=C(12/7,8/7)$ of Figure 9.3.1 and $Z(x_T,x_C)= Z(12/7,8/7)=4\cdot(12/7)+5\cdot(8/7)=88/7$ or \$12,571/week.

Point G lies where $y_a+2.5y_r= 4$ and the y_a-axis intersect. Solving these equations gives $y_a=4$ and $y_r=0$. Therefore, $G(y_a,y_r)=G(4,0)$ and $W(y_a,y_r)=W(4,0)=4\cdot 4+6\cdot 0=16$ or

\$16,000/week. Point G of Figure 9.3.2 corresponds to the point of Figure 9.3.1 where $x_T+2x_C=4$ and the x_T-axis intersect. Therefore, $Z(x_T,x_C)=Z(4,0)= 4{\cdot}4+5{\cdot}0=16$ or \$16,000/week.

Point E lies where $2y_a+1.5y_r= 5$ intersects the y_r-axis. Solving these equations gives $y_a=0$ and $y_r= 10/3$. Thus, $E(y_a,y_r)=E(0,10/3)$ and $W(y_a,y_r)=W(0,10/3)=4{\cdot}0+6{\cdot}(10/3)=20$ or \$20,000/week. Point E of Figure 9.3.2 corresponds to the point of Figure 9.3.1 where $2.5x_T+1.5x_C= 6$ and the x_C-axis intersect. Thus, $Z(x_T,x_C)=Z(0,4)=4{\cdot}0+5{\cdot}4=20$ or \$20,000/week.

Table 9.3.3 - Implicit-cost Minimization at Dual Extreme Points (in \$1,000's)

Extreme Points	$(4y_a,6y_r)$	$W=4y_a+6y_r$	$1.0y_a+2.5y_r$	$2.0y_a+1.5y_r$
$F(y_a,y_r)=F(13/7,6/7)$	(52/7,36/7)	<u>88/7=12.571</u>	28/7=4.000	35/7=5.000
$G(y_a,y_r)=G(4,0)$	(16,0)	16.000	4.000	8.000
$E(y_a,y_r)=E(0,10/3)$	(0,20)	20.000	25/3=8.333	15/3=5.000

Frequently, LP problems have unequal rather than equal numbers of input and output variables. Let us consider the solution of an LP problem with <u>two</u> input and <u>four</u> output variables. Suppose "ABC Brew" and "XYZ Brew" are two new instant coffee brands which sell for \$2,000 and \$2,500 per ton respectively. The reason for introducing ABC and XYZ Brews is to take advantage of the greater availability and lower price of robusta beans. The raw-bean requirements per ton of the four instant-coffee brews are shown in Table 9.3.4 below. The data of Table 9.3.4 are restated (in bold notation) in normalized lots of an arbitrary market value, say \$10,000 per week, as shown in Table 9.3.5 below.

Table 9.3.4 - Arabica and Robusta Inputs Required for Four Coffee Brew Outputs.

	Tons of Coffee-Bean Input Required per ton of Instant Coffee Brew Output				Coffee-Bean Input Constraints
	Taster's	Connoisseur's	ABC	XYZ	(tons/week)
Arabica Beans	$a_T=1.0$	$a_C=2.0$	$a_A=0.35$	$a_X=0.5$	a=4
Robusta Beans	$r_T=2.5$	$r_C=1.5$	$r_A=2.8$	$r_X=2.4$	r=6
Brew Price/ton	$p_T=\$4,000$	$p_C=\$5,000$	$p_A=\$2,500$	$p_X=\$2,500$	

Table 9.3.5 - Raw Bean Inputs Required for Weekly Sales Revenues of \$10,000.

	Tons of Coffee-Bean Input Required per \$10,000 of Instant Coffee Brew Output				Coffee-Bean Input Constraints
	Taster's	Connoisseur's	ABC	XYZ	(tons/week)
Arabica Beans	$\mathbf{a_T}=2.50$	$\mathbf{a_C}=4.0$	$\mathbf{a_A}=1.4$	$\mathbf{a_X}=2.0$	a=4
Robusta Beans	$\mathbf{r_T}=6.25$	$\mathbf{r_C}=3.0$	$\mathbf{r_A}=11.2$	$\mathbf{r_X}=9.6$	r=6
Market Value	\$10,000	\$10,000	\$10,000	\$10,000	

The input and nonnegativity constraints of the primal LP problem can now be stated in inequalities (9.3.11), (9.3.12) and (9.3.13) below, where output decision variables $\mathbf{x_T}$, $\mathbf{x_C}$, $\mathbf{x_A}$ and $\mathbf{x_X}$ are all expressed in \$10,000 lot sizes. The *objective function* of the primal LP is the weekly sales revenues $\mathbf{Z(x_T,x_C,x_A,x_X)}$ defined in (9.3.14). The primal LP objective is to determine output decision variables $\mathbf{x_T}$, $\mathbf{x_C}$, $\mathbf{x_A}$ and $\mathbf{x_X}$ which maximize objective function (9.3.14) without violating inequality constraints (9.3.11), (9.3.12) and (9.3.13).

$$2.50\mathbf{x_T}+ 4.0\mathbf{x_C} + 1.4\mathbf{x_A} + 2.0\mathbf{x_X} \le 4 \quad \text{(arabica inequality constraint)} \qquad ....(9.3.11)$$

$6.25\mathbf{x_T} + 3.0\mathbf{x_C} + 11.2\mathbf{x_A} + 9.6\mathbf{x_X} \leq 6$ (robusta inequality constraint) ....(9.3.12)

$\mathbf{x_T}, \mathbf{x_C}, \mathbf{x_A}, \mathbf{x_X} \geq 0$ (output nonnegativity constraints) ....(9.3.13)

$\mathbf{Z}(\mathbf{x_T},\mathbf{x_C},\mathbf{x_A},\mathbf{x_X}) \equiv \$10{,}000(\mathbf{x_T}+\mathbf{x_C}+\mathbf{x_A}+\mathbf{x_X})$ (objective function) ....(9.3.14)

Figure 9.3.3 presents the data of Table 9.3.5 as a two-dimensional (a,r) graph with two input and four output variables. The feasible region for decision variables $\mathbf{x_T}$, $\mathbf{x_C}$, $\mathbf{x_A}$ and $\mathbf{x_X}$ is the shaded rectangle with input constraints a=4 and r=6 intersecting at R(4,6).

Figure 9.3.3 - Graphical Analysis of the Coffee-Mix Problem in Table 9.3.5.

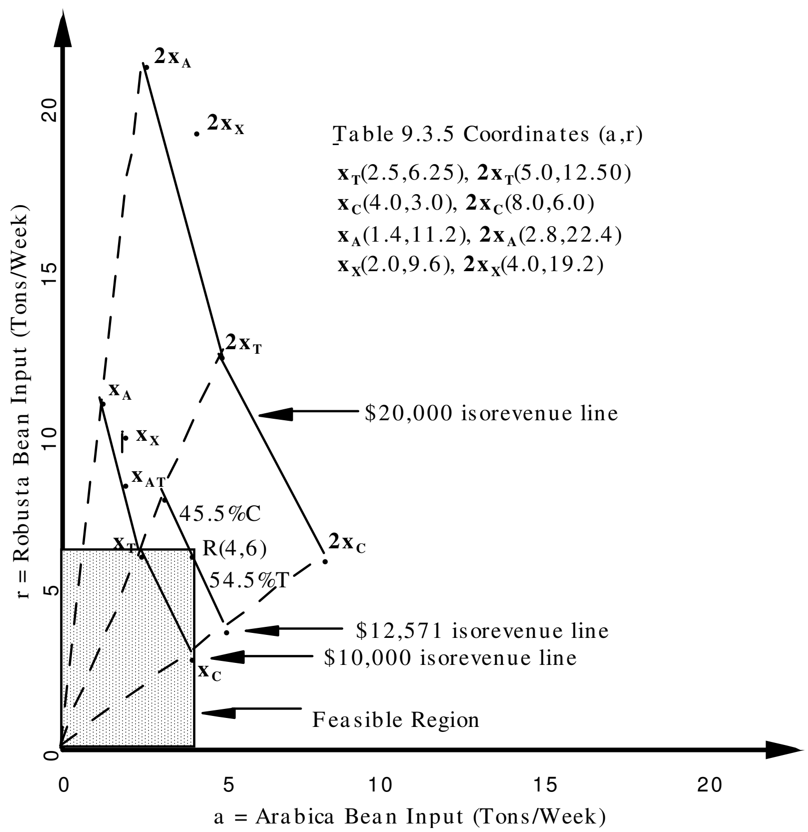


Figure 9.3.3 is a plot of $\mathbf{x_T}(2.5,6.25)$, $\mathbf{x_C}(4,3)$, $\mathbf{x_A}(1.4,11.2)$ and $\mathbf{x_X}(2,9.6)$ about which a convex envelope, called the \$10,000 isorevenue line, is drawn from $\mathbf{x_A}(1.4,11.2)$ to $\mathbf{x_T}(2.5,6.25)$ to $\mathbf{x_C}(4,3)$. To show $\mathbf{x_X}(2,9.6)$ is on the concave side of the envelope, a vertical line is drawn from $\mathbf{x_X}(2,9.6)$ to $\mathbf{x_{AT}}$ which divides the line from $\mathbf{x_A}(1.4,11.2)$ to $\mathbf{x_T}(2.5,6.25)$ in a ratio of (1.4-2):(2-2.5) or 6:5. Combining $(6/11)*\mathbf{x_T}(2.5,6.25)$ and $(5/11)*\mathbf{x_A}(0,12.5)$ equals $\mathbf{x_{AT}}(2,8.5)$ worth \$10,000/week just like $\mathbf{x_X}(2,9.6)$ which uses 1.1 tons more robusta beans per week. Hence, XYZ Brew is not on the convex envelope of the \$10,000 isorevenue line.

The next step of the graphical analysis is to draw raypaths from the origin to points $\mathbf{x_A}$(1.4,11.2), $\mathbf{x_T}$(2.5,6.25) and $\mathbf{x_C}$(4,3) of the \$10,000 isorevenue line. The slopes of the raypaths are $\mathbf{r_T}/\mathbf{a_T}$=2.5, $\mathbf{r_C}/\mathbf{a_C}$=0.75 and $\mathbf{r_A}/\mathbf{a_A}$=8 as compared to the slope r/a=6/4=1.5 of the input constraints. Since $\mathbf{r_T}/\mathbf{a_T}$=2.5 > r/a=1.5 > $\mathbf{r_C}/\mathbf{a_C}$=0.75, the (a,r)=(4,6) input constraints lie between Taster's and Connoisseur's raypaths, while $\mathbf{x_A}$(1.4,11.2) lies outside these two raypaths. Therefore, ABC Brew is excluded from the optimal solution.

Equation (9.2.27) indicates that the equation of the \$10,000 isorevenue line from $\mathbf{x_T}$(2.5,6.25) to $\mathbf{x_C}$(4,3) is r = -(y_a/y_r)a + \$10,000/y_r. Slope -y_a/y_r of the line from $\mathbf{x_T}$(2.5,6.25) to $\mathbf{x_C}$(4,3) is (6.25-3)/(2.5-4) = -13/6. Substituting a=4 and r=3 from $\mathbf{x_C}$(4,3) on the \$10,000 isorevenue line then gives 3 = -(13/6)*4 + \$10,000/y_r or y_r = \$60,000/70 = \$857/ton of robusta beans. It follows that y_a = (13/6) *y_r = \$130,000/70 = \$1,857/ton of arabica beans. The implicit costs of $\mathbf{x_T}$(2.5,6.25), $\mathbf{x_C}$(4,3), $\mathbf{x_A}$(1.4,11.2) and $\mathbf{x_X}$(2,9.6) are evaluated below.

2.5·\$1,857.14 + 6.25·\$857.14 = \$10,000 (Taster's Brew implicit cost)
4.0·\$1,857.14 + 3.0·\$857.14 = \$10,000 (Connoisseur's Brew implicit cost)
1.4·\$1,857.14 + 11.2·\$857.14 = \$10,714.25 (ABC Brew implicit cost)
2.0·\$1,857.14 + 9.6·\$857.14 = \$11,942.82 (XYZ Brew implicit cost)

The slope r/a=6/4=1.5 of input constraint R(4,6) intersects the \$10,000 isorevenue line from $\mathbf{x_T}$(2.5,6.25) to $\mathbf{x_C}$(4,3) at $\mathbf{x_{TC}}$($\mathbf{a_{TC}}$,$\mathbf{r_{TC}}$). Solving the equations $\mathbf{r_{TC}}$= 1.5$\mathbf{a_{TC}}$ and $\mathbf{r_{TC}}$= -(13/6)$\mathbf{a_{TC}}$ + \$10,000/(\$857.14), we get $\mathbf{a_{TC}}$= 70/22 and $\mathbf{r_{TC}}$ = 105/22. Therefore, the line segment from $\mathbf{x_T}$(2.5,6.25) to $\mathbf{x_C}$(4,3) is divided by $\mathbf{x_{TC}}$(70/22,105/22) in a ratio of (2.5-(70/22)):((70/22)-4) or 5:6. Hence, 5/11 or 45.45% of the \$10,000 revenues comes from sales of Connoisseur's Brew and 6/11 or 54.54% comes from sales of Taster's Brew.

Isorevenue lines between Taster's and Connoisseur's raypaths are all parallel and increasing in value as their distance from the origin increases. Let Z denote the unknown value of the isorevenue line through input constraint R(4,6) whose equation is 6 = -(13/6)*4 + Z/\$857.14 which can be solved to give Z = \$12,571.43 per week. Connoisseur's sales would then be 0.4545*\$12,571.43 = \$5,714.29 from sales of \$5,714.29/\$5,000 = 8/7 tons/week. Taster's sales would be 0.5454*\$12,571.43 = \$6,857.14 from sales of \$6,857.14/\$4,000 = 12/7 tons/week.

The simplex trial-and-error procedure for the dual LP model of the coffee-mix problem minimizes the implicit costs of the inputs by picking a pair of outputs, calculating their implicit costs, and shifting, if necessary, to another pair of outputs. The process eventually eliminates ABC and XYZ Brews and determines the minimum from the implicit costs of Taster's and Connoisseur's Brews. Similarly, the simplex procedure for the primal LP model eliminates ABC and XYZ Brews and determines the maximum weekly sales revenues from those of Taster's and Connoisseur's Brews.

Section 9.4 - Engineering Production Functions

Linear programming (LP) is only one of a number of linear and nonlinear analytical techniques, known collectively as *mathematical programming* (MP), used to solve constrained optimization problems. Linear programming can be used with readily available computer algorithms to solve large-scale problems of many different economic organizations. However, owing to losses of accuracy resulting from assumptions of linearity and constraints, LP models are limited to short-run problems where linear approximations will not seriously degrade the analysis.

An engineering production function is a mathematical description of the technological relationship between the magnitudes of input and output variables in a production process. The inputs are assumed to be combined in the most economical and technologically efficient manner for producing the outputs. In seeking a mathematical relationship between inputs and outputs that have different units of measurement, dimensional analysis requires that the ratio of any two concrete examples of output to be independent of the size of the units in which the inputs are measured. This 'ratio requirement' of dimensional analysis, also known as the product theorem, restricts the form of the production function to a product of the powers of the input and output quantities (see the proof of the product theorem in Dimensional Analysis by P.W. Bridgman, Yale University Press, 1931).

The ratio requirement of dimensional analysis enables different physical quantities to be expressed in terms of the same dimension in order to make sensible comparisons. The accuracy of productivity measurements depends primarily on dimensional analysis which is a mathematical tool of great utility in science and engineering. It enables equations of physical hypotheses to be checked for logical consistency without solving the equations. The chicken-and-egg production function presented in Appendix 9B is a simple example which illustrates how the ratio requirement of dimensional analysis helps to formulate a production function and understand the physical significance of its parameters.

To fix ideas, let us formulate a production function to describe the coffee-mix problem from data in the first two columns of Table 9.3.5. Let P represent the production of weekly sales revenues, and let 'a' and 'r' represent the inputs of arabica and robusta beans in tons per week respectively. The production function which relates P, a and r takes the form

$$P(a,r) = A(a)^{\alpha}(r)^{\beta} \qquad ....(9.4.1)$$

where parameters α and β are dimensionless exponents and A is a dimensional constant. Equation (9.4.1) is called a homogeneous production function of degree $\alpha+\beta \equiv n$ which measures how much the output changes relative to equiproportional changes of all inputs. The ratio β/α measures the substitutability of arabica for robusta beans when P is constant.

In productivity experiments where all input variables are changed by the same proportion (i.e., equiproportionally), the resulting changes in output are called *scale effects* which fall into three broad categories. *Constant returns to scale* means the output changes in the same proportion as changes in all input variables (i.e., $\alpha + \beta \equiv n = 1$). All LP models assume constant returns to scale which implies that technological efficiency does not vary with the scale of operations. *Increasing returns to scale* means the output changes in a greater proportion than changes in all input variables. *Decreasing returns to scale* means the output changes in a smaller proportion than changes in all input variables.

Equiproportional changes of all input variables in scale-effect experiments implies fixed proportions between all inputs whose changes can be treated as a single variable. However, owing to the *indivisibilities* of diverse input variables, it may be difficult to design experiments where all inputs can change in the same proportion. The range of inputs in productivity experiments is largely determined by budget constraints and output demand which dominate the input constraints in production processses.

The main focus of productivity experiments in the formulation of production functions is on measuring the *substitution effects* of changing the proportion of only one input variable relative to all other input variables. Empirical evidence in measurements of substitution effects led to a pillar of economic analysis in the past two centuries known as the law of variable proportions, more commonly called the law of diminishing returns, which may be stated as follows: *If the quantity of one input variable is increased by equal increments while the quantities of all other input variables remain fixed, then the resulting increments of output will decrease after a certain point*, called the *point of diminishing returns*.

It is commonly thought that constant returns to scale implies that the point of diminishing returns has not yet been reached. However, constant returns to scale states that if A and B yields P, then 2A and 2B yields 2P; and the point of diminishing returns states that if A and B yields P, then 2A and B yields less than 2P. Therefore, constant returns to scale may include the point of diminishing returns. Under conditions of constant returns to scale, it is quite possible for the law of variable proportions to rule from the very beginning. Consequently, there is no sense in the common assertion that an input variable should never be incremented beyond the point of diminishing returns.

The scale and substitution effects of productivity experiments can be represented by common parameters of a single production function where all outputs are measured in dollar values per week rather than units of individual outputs. Thus, the outputs of the coffee-mix problem could be expressed in dollars of sales revenues instead of tons of Taster's and Connoisseur's Brews which have different prices per ton. The inputs to the production function are still described in physical terms which are assumed to be combined in the most economical and technologically efficient manner.

Table 9.4.1 - Arabica and Robusta Inputs Required for Two Coffee Brew Outputs.

	Tons of Coffee-Bean Input Required per ton of Instant Coffee Brew Output		Coffee-Bean Input Constraints
	Taster's	Connoisseur's	(tons/week)
Arabica Beans	a_T=1.0	a_C=2.0	a=3
Robusta Beans	r_T=2.5	r_C=1.5	r=4
Brew Price/ton	p_T=\$4,000	p_C=\$5,000	

Table 9.4.2 - Raw Bean Inputs Required for Common Weekly Revenues of $4,000.

	Tons of Coffee-Bean Input Required per $5,000 of Instant Coffee Brew Output		Coffee-Bean Input Constraints
	Taster's	Connoisseur's	(tons/week)
Arabica Beans	a_T=1.0	0.8a_C=1.6	a=3
Robusta Beans	r_T=2.5	0.8r_C=1.2	r=4
Weekly Output Value	$4,000	$4,000	

Production function (9.4.1) has three unknown parameters A, α and β which can be determined from three independent productivity experiments. Table 9.4.2 has two different inputs with a common output. Thus, one ton of arabica beans and 2.5 tons of robusta beans produce $4,000/week of sales from Taster's Brew. Similarly, 1.6 tons of arabica beans and 1.2 tons of robusta beans produce $4,000/week of sales from Connoisseur's Brew. These two data sets form equations (9.4.2) and (9.4.3) which represent points on the production function surface. The ratio of equations (9.4.3) to (9.4.2) gives equation (9.4.4).

$$P(a_T,r_T) = \$4,000 = A(1.0)^{\alpha}(2.5)^{\beta} \qquad ....(9.4.2)$$

$$P(0.8a_C,0.8r_C) = \$4,000 = A(1.6)^{\alpha}(1.2)^{\beta} \quad ....(9.4.3)$$

$$1 = (0.8a_C/a_T)^{\alpha}(0.8r_C/r_T)^{\beta} = (1.6/1.0)^{\alpha}(1.2/2.5)^{\beta} = (1.6)^{\alpha}(0.48)^{\beta} \quad ....(9.4.4)$$

Taking logarithms of both sides of (9.4.4) gives us $0 = \alpha\ \ln(1.60) + \beta\ \ln(0.48)$. This enables us to determine the ratio $\beta/\alpha = -\ln(1.60)/\ln(0.48) = 0.6403588$ which measures the substitutability of arabica for robusta beans in producing \$4,000/week.

Besides the substitutability-effect experiment obtained from Table 9.4.2, we need a scale-effect experiment to determine $\alpha+\beta \equiv n$ which measures the returns to scale of increasing all inputs equiproportionally (see the definition of a homogeneous function of degree n in Appendix 9D). Increasing the scale of the inputs may increase output by a greater proportion which could be offset by lower sales prices needed to stimulate demand for more product. As a guess, let us assume <u>constant returns to scale</u> (i.e., $n = 1$) just like the LP example. The relations $\alpha+\beta = 1$ and $\beta/\alpha = 0.6403588$ enable us to calculate $\alpha(1+0.6403588) = 1$, or $\alpha = 1/(1.6403588) = 0.6096227$ and $\beta = 1-\alpha = 0.3903773$.

Having determined the α and β dimensions of the (a,r) inputs, the dimensional constant A needs to be determine from any set of data on either Taster's or Connoisseur's raypath in either Table 9.4.1 or 9.4.2 as indicated in equations (9.4.5) and (9.4.6) below.

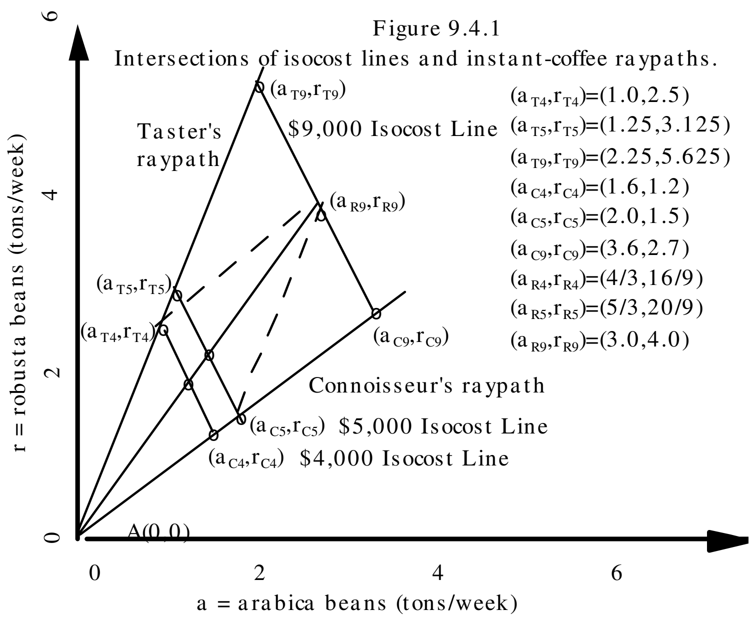
$$P(a_T,r_T) = \$4,000 = A(1.0)^{0.6096227}(2.5)^{0.3903773}\ ; \quad A = \$2797.1338/\text{week} \quad ....(9.4.5)$$
$$P(a_C,r_C) = \$5,000 = A(2.0)^{0.6096227}(1.5)^{0.3903773}\ ; \quad A = \$2797.1338/\text{week} \quad ....(9.4.6)$$

When $P(a_T,r_T)$ and $P(a_C,r_C)$ are evaluated with their common A, α and β parameters, they give the same numerical values as the p_T and p_C prices of the LP model because their (a,r) coordinates which lie on Taster's and Connoisseur's raypaths were used to formulate the production function. For example, Table 9.4.1 data give $P(a_T,r_T) = P(1.0,2.5) = \$4,000$/week and $P(a_C,r_C) = P(2.0,1.5) = \$5,000$/week which are the same as the prices per ton of Taster's and Connoisseur's Brews as shown in equations (9.4.7) and (9.4.8). But when (a,r) coordinates lie between Taster's and Connoisseur's raypaths, say $(a_T+a_C,r_T+r_C) = (1.0+2.0,2.5+1.5) = (3.0,4.0)$ as in equation (9.4.9), then $P(3.0,4.0)$ is greater than the sum of the p_T and p_C prices per ton of Taster's and Connoisseur's Brews.

$$P(1.0,2.5) = \$2797.1338*(1.0)^{0.6096227}(2.5)^{0.3903773} = \$4,000/\text{week} = p_T \quad ....(9.4.7)$$
$$P(2.0,1.5) = \$2797.1338*(2.0)^{0.6096227}(1.5)^{0.3903773} = \$5,000/\text{week} = p_C \quad ....(9.4.8)$$
$$P(3.0,4.0) = \$2797.1338*(3.0)^{0.6096227}(4.0)^{0.3903773} = \$9,388.75/\text{week} > p_T+p_C \quad ....(9.4.9)$$

Equations (9.4.7), (9.4.8) and (9.4.9) bring out basic similarities and differences between production function and LP models. LP models measure (a,r) input costs by their implicit prices y_a and y_r which may differ from market prices used in production functions. Moreover, LP models specify individual (a,r) input constraints and (p_T,p_C) price constraints whereas production functions only specify budget constraints to determine (a,r) inputs at market prices which maximize the marginal productivities of arabica and robusta beans.

Isocost lines and isorevenue curves coincide in LP models because equation (9.2.22) states implicit input costs, ay_a+ry_r, equal individual output revenues, $p_Tx_T+p_Cx_C$. But production processes rarely, if ever, have inputs and outputs of equal value. In general, output revenues are greater than input costs in production function models. Figure 9.4.1 based on data from Table 9.4.1 will now be used to explore the similarities and differences between production function and LP models in greater detail.



The input costs of $(a_{R9},r_{R9})=(3,4)$ in Figure 9.4.1 are $ay_a+ry_r = 3*\$1,857.14+4*\$857.14 = \$9,000$/week assuming y_a=\$1,857.14 and y_r=\$857.14 are both implicit and market prices per ton of arabica and robusta beans. Point $(a_{T9},r_{T9})=(2.25,5.625)$ on the \$9,000-isocost line at Taster's raypath satisfies the equations $a_{T9}*\$1,857.14+r_{T9}*\$857.14 = \$9,000$ and $r_{T9}/a_{T9} = r_T/a_T = 2.5/1.0$. Also, point $(a_{C9},r_{C9}) = (3.6,2.7)$ on the \$9,000-isocost line at Connoisseur's raypath satisfies the equations $a_{C9}*\$1,857.14+r_{C9}*\$857.14 = \$9,000$ and $r_{C9}/a_{C9} = 1.5/2.0$.

According to the LP model, the three points on the \$9,000-isocost line also have \$9,000 output revenues. Thus, input $(a_{T9},r_{T9})=(2.25,5.625)$ on the \$9,000-isocost line at Taster's raypath has output revenues of $p_Tx_T+p_Cx_C = \$4,000*(2.25/a_T)+\$5,000*(0) = \$9,000$. The input $(a_{C9},r_{C9})=(3.6,2.7)$ on the \$9,000-isocost line at Connoisseur's raypath also has output revenues of $p_Tx_T+p_Cx_C = \$4,000*(0)+\$5,000*(3.6/a_C) = \$9,000$. Lastly, the input $(a_{R9},r_{R9})=(3,4)$ on the \$9,000-isocost line between Taster's and Connoisseur's raypaths has output revenues of $p_Tx_T+p_Cx_C = \$4,000*(1)+\$5,000*(1) = \$9,000$ (see Table 9.1.2).

A common property of production function and LP models is that the (a,r) input point divides the length of the isocost line between Taster's and Connoisseur's raypaths in inverse proportion to the input costs of Taster's and Connoisseur's Brews that are produced. Thus, the input $(a_{R9},r_{R9}) = (3,4)$ in Figure 9.4.1 divides the length of the \$9,000-isocost line from $(a_{T9},r_{T9}) = (2.25,5.625)$ to $(a_{C9},r_{C9}) = (3.6,2.7)$ in inverse proportion to the \$4,000 and \$5,000 cost of producing one ton of Taster's and Connoisseur's Brews as verified below.

$$\frac{\$4,000}{\$9,000} = \frac{\sqrt{(3.6-3)^2+(2.7-4)^2}}{\sqrt{(3.6-2.25)^2+(2.7-5.625)^2}}\ ;\quad \frac{\$5,000}{\$9,000} = \frac{\sqrt{(3-2.25)^2+(4-5.625)^2}}{\sqrt{(3.6-2.25)^2+(2.7-5.625)^2}}$$

Therefore, any point on the \$9,000-isocost line from Taster's to Connoisseur's raypaths may be selected to determine complementary fractions of \$9,000/week that should be allocated to producing Taster's and Connoisseur's Brews. Moreover, whatever input fractions of \$9,000 are selected, their output revenues will be the same in LP models because $p_T x_T + p_C x_C = a y_a + r y_r$ of equation (9.2.22) makes isorevenue curves indistinguishable from isocost lines. As a result, the choice of points on the \$9,000-isocost line would be arbitrary if individual (a,r) input constraints were not specified in LP models.

Production function models use budget constraints whose (a,r) inputs are evaluated at (y_a, y_r) market prices. LP models use individual (a,r) input constraints which are evaluated by (y_a, y_r) implicit input prices that are determined from the dual LP model by specifying (p_T, p_C) output price constraints. In production function models, (y_a, y_r) market prices are not constrained by (p_T, p_C) output prices. Both production function and LP models have parallel isocost lines whose slopes $-(y_a/y_r)$ are constant everywhere between raypaths of Taster's and Connoisseur's Brew (see equation (9.2.27)). On the other hand, production functions have parallel isorevenue curves which are convex to the origin with constant slopes along each raypath (see Appendix 9A, Isoquant Properties I to V).

Production functions have one optimal raypath between Taster's and Connoisseur's raypaths where isocost lines and isorevenue curves are tangent. In order to determine the slope of isorevenues curves along any raypath, it is convenient to use the total differential of $P(a,r) = A(a)^{\alpha}(r)^{\beta}$ given in equation (9.4.10). The evaluations of *marginal productivities* $[\partial P(a,r)/\partial a]_r$ and $[\partial P(a,r)/\partial r]_a$ given in (9.4.11) and (9.4.12) show that they are proportional to their *average productivities* P(a,r)/a and P(a,r)/r. The marginal productivity ratio determines slope dr/da of isorevenues curves along any raypath as given in (9.4.13).

$$dP(a,r) = [\partial P(a,r)/\partial a]_r da + [\partial P(a,r)/\partial r]_a dr = 0 \quad ....(9.4.10)$$

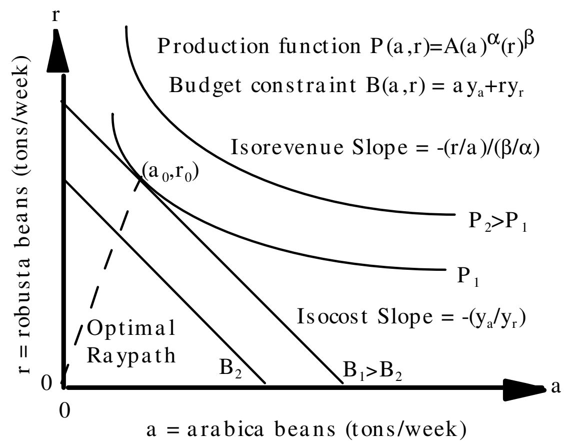
$$[\partial P(a,r)/\partial a]_r = A\alpha(a)^{\alpha-1}(r)^{\beta} = \alpha P(a,r)/a \quad ....(9.4.11)$$

$$[\partial P(a,r)/\partial r]_a = A\beta(a)^{\alpha}(r)^{\beta-1} = \beta P(a,r)/r \quad ....(9.4.12)$$

$$dr/da = -(\partial P/\partial a)_r/(\partial P/\partial r)_a = -\alpha r/\beta a \quad ....(9.4.13)$$

At Taster's raypath, isorevenue-curve slope $-\alpha r_T/\beta a_T = -2.5\alpha/\beta = -2.5*1.5616244 = -3.9040609$ is steeper than isocost-line slope $-y_a/y_r = -13/6 = -2.166...$ At Connoisseur's raypath, isorevenue-curve slope $-\alpha r_C/\beta a_C = -0.75 \times 1.5616244 = -1.1712183$ is shallower than the slope of isocost lines. Since slopes of isorevenue curves change continuously from Taster's to Connoisseur's raypaths, Figure 9.4.2 indicates an optimal raypath slope r_O/a_O occurs where isorevenue curves are tangent to isocost lines.

Figure 9.4.2 - Tangency of isorevenue curves and isocost lines.



Optimal raypath slope r_O/a_O satisfies the equation $-\alpha r_o / \beta a_o = -13/6$ where isocost lines and isorevenue curves are tangent. Hence, $r_O/a_O = 1.3874442$. At the intersection with the \$9,000-isocost line, $a_O y_a + r_O y_r$ = \$9,000 or \$1,857.14a_O + \$857.14*2.5a_O = \$9,000 from which we determine a_O = 2.954332 tons/week of arabica beans and r_O = 1.3874442a_O = 4.0989708 tons/week of robusta beans as compared to (a_{R9},r_{R9})=(3,4) in Figure 9.4.1.

Point (a_O,r_O) = (2.954332,4.0989708) divides the \$9,000-isocost line from (a_{T9},r_{T9}) = (2.25,5.625) to (a_{C9},r_{C9}) = (3.6,2.7) in inverse proportion to the production costs of Taster's and Connoisseur's Brews. Thus, the production costs of Taster's and Connoisseur's Brews may be calculated as shown in equations (9.4.14) and (9.4.15).

$$\frac{\$9{,}000\sqrt{(3.6-a_o)^2+(2.7-r_o)^2}}{\sqrt{(3.6-2.25)^2+(2.7-5.625)^2}} = \$4{,}304.51 \text{ / week} \quad \text{(Taster's cost)} \qquad ....(9.4.14)$$

$$\frac{\$9{,}000\sqrt{(a_o-2.25)^2+(r_o-5.625)^2}}{\sqrt{(3.6-2.25)^2+(2.7-5.625)^2}} = \$4{,}695.49 \text{ / week} \quad \text{(Connoisseur's cost)}....(9.4.15)$$

Let (a_{T4O},r_{T4O}) denote the tons/week of arabica and robusta beans which cost \$4,304.51/week to make Taster's Brew. By similar triangles, we have $(a_{T4O}/a_{T9}) = (r_{T4O}/r_{T9})$ = (\$4,304.51/\$9,000) from which we determine a_{T4O} = 1.0761282 and r_{T4O} = 2.6903206 tons per week. Taster's output would then be $(a_{T4O}/a_T) = (r_{T4O}/r_T)$ = 1.0761282 tons/week. Also, Taster's \$4,304.51-isocost line intersects the optimal raypath at (a_{R4O},r_{R4O}) = (1.4129946,1.9604512) and Connoisseur's raypath at (a_{C4O},r_{C4O}) = (1.721804,1.291353) where $(a_{R4O}/a_0) = (r_{R4O}/r_0) = (a_{C4O}/a_{C9}) = (r_{C4O}/r_{C9})$ = (\$4,304.51/\$9,000).

Let (a_{C5O},r_{C5O}) denote the tons per week of arabica and robusta beans which cost \$4,695.49/week to make Connoisseur's Brew. By similar triangles, we have $(a_{C5O}/a_{C9}) = (r_{C5O}/r_{C9})$ = (\$4,695.49/\$9,000) from which we get a_{C5O} = 1.878196 and r_{C5O} = 1.408647 tons per week. Connoisseur's output would be $(a_{C5O}/a_C) = (r_{C5O}/r_C)$ = 0.9390974 tons per week. Also, Connoisseur's \$4,695.49-isocost line intersects the optimal raypath at (a_{R5O},r_{R5O}) = (1.5413374, 2.1385196) and Taster's raypath at (a_{T5O},r_{T5O}) = (1.1738725,2.9346812) where $(a_{R5O}/a_0) = (r_{R5O}/r_0) = (a_{T5O}/a_{T9}) = (r_{T5O}/r_{T9})$ = = (\$4,695.49/\$9,000).

Let us evaluate the production function along Taster's, Connoisseur's and optimal raypaths as follows: (A=\$2797.1338/week, α=0.6096227, β=0.3903773)

$P(a_{T40},r_{T40}) = A(1.0761282)^{\alpha}(2.6903206)^{\beta}$ = \$4,304.51/week ....(9.4.16)
$P(a_{T50},r_{T50}) = A(1.1738725)^{\alpha}(2.9346812)^{\beta}$ = \$4,695.49/week ....(9.4.17)
$P(a_{T9},r_{T9}) = A(2.25)^{\alpha}(5.625)^{\beta}$ = \$9,000/week Taster's output only ....(9.4.18)
$P(a_{T40},r_{T40})/(a_{T40}/a_T)$ = \$4,304.51/1.0761282 = \$4,000/ton Taster's ....(9.4.19)

$P(a_{C40},r_{C40}) = A(1.721804)^{\alpha}(1.291353)^{\beta}$ = \$4,304.51/week ....(9.4.20)
$P(a_{C50},r_{C50}) = A(1.878196)^{\alpha}(1.408647)^{\beta}$ = \$4,695.49/week ....(9.4.21)
$P(a_{C9},r_{C9}) = A(3.6)^{\alpha}(2.7)^{\beta}$ = \$9,000/week Connoisseur's output only ....(9.4.22)
$P(a_{C50},r_{C50})/(a_{C50}/a_C)$=\$4,695.49/0.9390974= \$5,000/ton Connoisseur's ....(9.4.23)

$P(a_{R40},r_{R40}) = A(1.4129946)^{\alpha}(1.9604512)^{\beta}$ = \$4,491.29/week ....(9.4.24)
$P(a_{R50},r_{R50}) = A(1.5413374)^{\alpha}(2.1385196)^{\beta}$ = \$4,899.24/week ....(9.4.25)
$P(a_O,r_O)=A(2.954332)^{\alpha}(4.0989708)^{\beta}$= \$9,390.53/week optimum output....(9.4.26)

Equations (9.4.16) to (9.4.23) represent outputs of either Taster's or Connoisseur's Brew only with a budget constraint of \$9,000/week and input costs equal to output revenues. The optimum outputs of equations (9.4.24) to (9.4.26) are 4.339% greater than their input costs.

Optimum outputs can also be determined by *Lagrange undetermined multipliers*. The Lagrangian technique combines production function P(a,r) and budget constraint B(a,r), written as \$9,000-ay_a-ry_r, into one equation, called the *Lagrangian function*, L(a,r,λ), to insure **(a)** the production function remains intact and **(b)** the budget constraint is satisfied.

$$L(a,r,\lambda) \equiv P(a,r) + \lambda B(a,r) = A(a)^{\alpha}(r)^{\beta} + \lambda(\$9{,}000 - ay_a - ry_r) \quad ....(9.4.27)$$

Parameter λ is the undetermined multiplier. Inputs (a,r) which maximize P(a,r) subject to budget constraint B(a,r) can be found by differentiating L(a,r,λ) partially with respect to unknown parameters a, r and λ and equating the derivatives to zero. Equations (9.4.28) and (9.4.29) show marginal productivities $\alpha P/a$ and $\beta P/r$ of arabica and robusta beans equal λy_a and λy_r respectively. Equations (9.4.28) can each be solved for λ as shown in equation (9.4.31) which indicates $\alpha P = \lambda ay_a$ and $\beta P = \lambda ry_r$ so that $(\alpha+\beta)P = \lambda(ay_a + ry_r)$. This gives another solution for λ as shown in equation (9.4.32).

$$\partial L/\partial a = A\alpha(a)^{\alpha-1}(r)^{\beta} - \lambda y_a = (\alpha P/a) - \lambda y_a = 0 \quad ....(9.4.28)$$
$$\partial L/\partial r = A\beta(a)^{\alpha}(r)^{\beta-1} - \lambda y_r = (\beta P/r) - \lambda y_r = 0 \quad ....(9.4.29)$$
$$\partial L/\partial \lambda = \$9{,}000 - ay_a - ry_r = 0 \quad ....(9.4.30)$$
$$\lambda = \alpha P / ay_a = \beta P / ry_r \quad ....(9.4.31)$$
$$\lambda = (\alpha+\beta)P/(ay_a + ry_r) = nP(a,r)/B(a,r) \quad ....(9.4.32)$$

The optimal value of λ_O = 1.0433922 is obtained in equations (9.4.24) to (9.4.26) with (a_0,r_0), (a_{R50},r_{R50}) and (a_{R40},r_{R40}) on the optimal raypath. Specifying other (a,r) inputs on either side of the optimal raypath for a fixed B(a,r) = ay_a+ry_r budget constraint evaluated at (y_a,y_r) market prices modifies the ratio of Taster's to Connoisseur's output and reduces λ and output revenues nP(a,r) based on the marginal productivities of arabica and robusta beans.

Section 9.5 - Cobb-Douglas Production Function

Engineering production functions are mathematical descriptions of input/output rates in production processes. Aggregate production functions describe input/output relationships in an industry or a nation. The earliest aggregate production function satisfying the 'ratio requirement' of dimensional analysis is attributable to Charles W. Cobb, Professor of Mathematics at Amherst College, and Paul H. Douglas, Professor of Economics at the University of Chicago and later a United States Senator from Illinois (see *A Theory of Production*, American Economic Review, 18 (Supplement), 1928, pages 139-165). The Cobb-Douglas production function states that

$$P = A * L^{\alpha} * C^{\beta} \qquad ....(9.5.1)$$

where P, L, and C are indices representing industrial output, and labor and capital input per year respectively, and A, α and β are unknown parameters.

The Cobb-Douglas production function was fitted by log-linear regression analysis to annual input and output data of United States industry for 24 years from 1899 to 1922. The indices of P, L, and C were all unitized at the base year of 1899. It was assumed United States had constant returns to scale (i.e., $\alpha+\beta = 1$) during the 24-year study period, thereby reducing the unknown parameters to A, α and $1-\alpha$. The results were as follows:

$$P = 1.01 * L^{0.75} * C^{0.25} \qquad ....(9.5.2)$$

Exponents $\alpha = 0.75$ and $\beta = 1-\alpha = 0.25$ were interpreted as imputed shares of labor (L) and capital (C) from industrial output (P), and dimensional constant A = 1.01 was interpreted as a least-squares residual representing technological change.

These interpretations of parameters A, α and β differ significantly from those of engineering production functions because the regression analysis assumed constant returns to scale in order to have fewer unknown parameters and improve statistical reliability. When constant returns to scale is assumed in an aggregate production function, it implies that every size is just as efficient as any other size and that two United States are twice as productive as one United States. However, when testing the laws of return on a smaller scale of a factory or an industrial process, then measurements of increasing, decreasing or constant returns to scale need to be based on more specific empirical observations.

The assumption of constant returns to scale in fitting the Cobb-Douglas production function leads to two other problems, namely, **(a)** interpreting dimensional constant A as a measure of technological change and **(b)** interpreting dimensionless exponents α and β as measures of the imputed shares of labor and capital in the value of the total output.

(a) In the engineering production function of Section 9.4 and the chicken-and-egg production function of Appendix 9B, parameter A was estimated for dimensional validity rather than the need to improve statistical reliability. In particular, dimensional constant A = \$2,797.1338/week could be determined from equations (9.4.5) and (9.4.6) even if $\alpha+\beta$ did not equal one. In the chicken-and-egg production function of Appendix 9B, dimensional constant $A = E_0/C_0^{\alpha}D_0^{\beta} = 2/3$ eggs per chicken per day with $\alpha=\beta=1$ and $\alpha+\beta=2$ which implies increasing returns to scale. Therefore, least-square residuals used to determine A = 1.01 may not be a proper measure of technological change.

The parameters of the Cobb-Douglas production function can be interpreted by means of the total differential dP given in equation (9.5.3). The partial derivatives of equation (9.5.3) are evaluated in equations (9.5.4) and (9.5.5).

$$dP = (\partial P/\partial L)_C dL + (\partial P/\partial C)_L dC \quad ....(9.5.3)$$

$$(\partial P/\partial L)_C = \alpha A L^{\alpha-1} C^{\beta} = \alpha P/L \quad ....(9.5.4)$$

$$(\partial P/\partial C)_L = \beta A L^{\alpha} C^{\beta-1} = \beta P/C \quad ....(9.5.5)$$

Upon substituting equations (9.5.4) and (9.5.5) into equation (9.5.3), we get

$$dP = \alpha P\frac{dL}{L} + \beta P\frac{dC}{C} \quad \text{or} \quad \frac{dP}{P} = \alpha\frac{dL}{L} + \beta\frac{dC}{C} \quad ....(9.5.6)$$

In scale-effect experiments, the relative change μ of labor L is equal to the relative change of capital C (i.e., $\mu = dL/L = dC/C$). Consequently, equation (9.5.6) shows the relative change dP/P of output P in scale-effect experiments equals $(\alpha+\beta)\mu$. If $(\alpha+\beta)$ was either greater than, equal to, or less than one, then United States industry would have had increasing, constant, or decreasing returns to scale from 1899 to 1922. Instead of assuming $\beta = 1-\alpha$ when fitting the Cobb-Douglas production function to the data, Appendix 9C shows that α and β could be estimated independently from parametric equations of the annual growth rates m_P of Production, m_L of Labor and m_C of Capital.

In scale- and substitution-effect experiments, let $P+\Delta P = P(1+\Delta_P)$, $L+\Delta L = L(1+\Delta_L)$ and $C+\Delta C = C(1+\Delta_C)$, where $\Delta_P=\Delta P/P$, $\Delta_L=\Delta L/L$ and $\Delta_C=\Delta C/C$ denote small relative changes of P, L and C. Using a first order approximation of the binomial expansion for small relative changes of P, L and C, equation (9.5.1) could be rewritten as follows:

$$P[1+\Delta_P] = A[L(1+\Delta_L)]^{\alpha}[C(1+\Delta_C)]^{\beta} = AL^{\alpha}(1+\alpha\Delta_L)C^{\beta}(1+\beta\Delta_C)$$

$$P[1+\Delta_P] = AL^{\alpha}C^{\beta}(1+\alpha\Delta_L+\beta\Delta_C+\alpha\beta\Delta_L\Delta_C) \quad ....(9.5.7)$$

Cancelling $P = AL^{\alpha}C^{\beta}$ on both sides of equation (9.5.8) and assuming the product of small relative changes $\alpha\beta\Delta_L\Delta_C$ is negligible, we get

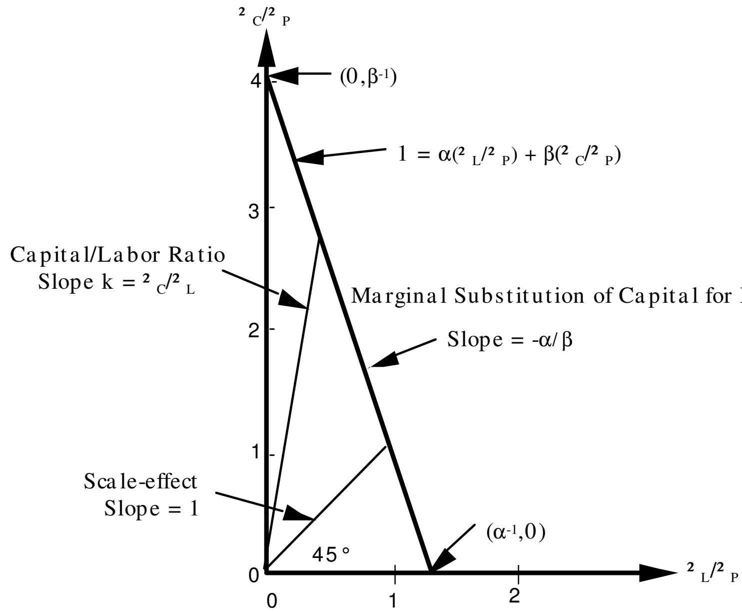
$$\Delta_P = \alpha\Delta_L + \beta\Delta_C \quad \text{or} \quad 1 = \alpha\Delta_L/\Delta_P + \beta\Delta_C/\Delta_P \quad ....(9.5.8)$$

Equation (9.5.8) is plotted as a straight-line function of variables Δ_L/Δ_P and Δ_C/Δ_P in Figure 9.5.1. The intercepts on the Δ_L/Δ_P and Δ_C/Δ_P axes are $(1/\alpha,0)$ and $(0,1/\beta)$ respectively, and its slope is $-\alpha/\beta$. Scale-effect experiments require $\Delta_L/\Delta_P = \Delta_C/\Delta_P$ which occurs on a 45° line from the origin whose intersection with the production-function line is $(1/(\alpha+\beta),1/(\alpha+\beta))$. The definition of *average productivity* is based on scale-effect experiments where $\Delta_L/\Delta_P = \Delta_C/\Delta_P$ and $\Delta_P/\Delta_L = \Delta_P/\Delta_C = \alpha+\beta$. The *capital/labor ratio* $(\Delta_C/\Delta_P)/(\Delta_L/\Delta_P) = \Delta_C/\Delta_L = k > 0$ occurs on the line $(\Delta_C/\Delta_P) = k(\Delta_L/\Delta_P)$ which intersects the production-function line at $(1/(\alpha+k\beta),k/(\alpha+k\beta))$. When k is greater or less than one, it is called capital or labor intensive respectively. The *marginal substitution of capital for labor* is defined as the slope $-\alpha/\beta$ of the production-function line between the average productivity and capital/labor ratio lines.

Let us apply equation (9.5.8) with $\alpha=0.75$ and $\beta=0.25$ to annual growth trends in the United States of $\Delta_P = 3.25\%$, $\Delta_L = 1.50\%$ and $\Delta_C = 8.50\%$ during the early 1900's. Assuming scale-effect experiments with constant returns to scale, equation (9.5.8) indicates the $\Delta_P =$

3.25% annual growth rate in industrial production would occur with 3.25% annual growth rates in labor and capital (i.e., $\Delta_P = \Delta_L = \Delta_C = 3.25\%$).

Figure 9.5.1 - Scale and Substitution Effects of the Cobb-Douglas Production Function



However, instead of scale-effect experiments, the Δ_P = 3.25% annual growth rate of industrial production occurred with Δ_L = 1.50% annual growth rate of labor and Δ_C = 8.50% annual growth rate of capital. This indicates the capital/labor ratio was $k = \Delta_C/\Delta_L$ = 8.50%/1.50% = 5.67 which is shown in Figure 9.5.1 by the line drawn from the origin with a 5.67 slope. Upon substituting $\Delta_C = 5.67\Delta_L$ into equation (9.5.8), we find again Δ_P = 3.25%.

$$\Delta_P = \alpha\Delta_L + \beta\Delta_C = 0.75\Delta_L + 0.25*(5.67\Delta_L) = 1.125\% + 2.125\% = 3.25\%$$

The *marginal substitution of capital for labor* is defined as the slope $-\alpha/\beta = -3$ of the production-function line between the average productivity and capital/labor ratio lines. It is important to note a common confusion that the definition of marginal substitution of capital for labor is independent of the definition of capital/labor ratio as shown in equation (9.5.9). This suggests that the sum $\alpha+\beta$ and ratio $-\alpha/\beta$ rather than dimensional constant A should be used as measures of technological change.

$$\frac{\text{Marginal substitution}}{\text{of capital for labor}} \equiv \frac{k(\alpha+k\beta)^{-1}-(\alpha+\beta)^{-1}}{(\alpha+k\beta)^{-1}-(\alpha+\beta)^{-1}} = \frac{\alpha(k-1)}{\beta(1-k)} = -\frac{\alpha}{\beta} = -3 \quad ....(9.5.9)$$

(b) The interpretation of exponents α and β as imputed shares of labor and capital in the value of the output is based on the integration of equations (9.5.3), (9.5.4) and (9.5.5) which are summarized in equation (9.5.10) below for convenience.

$$dP = (\partial P/\partial L)_C dL + (\partial P/\partial C)_L dC = \alpha P(dL/L) + \beta P(dC/C) \quad ....(9.5.10)$$

Because the Cobb-Douglas production function is homogeneous of degree $\alpha+\beta$, Euler's theorem indicates the total output $(\alpha+\beta)P$ is determinable by summing the inputs of labor and capital multiplied by their marginal productivities as shown in equation (9.5.11) (see Euler's Theorem on Homogeneous Functions in Appendix 9D).

$$L(\partial P/\partial L)_C + C(\partial P/\partial C)_L = (\alpha P/L)*L + (\beta P/C)*C = (\alpha+\beta)P \qquad ....(9.5.11)$$

The Ricardian theory of rent described in Section 1.2 indicates labor L and capital C would be paid according to their marginal productivities $(\partial P/\partial L)_C$ and $(\partial P/\partial C)_L$ in a competitive equilibrium. If so, equation (9.5.11) shows the imputed shares of labor and capital in the total output would be αP and βP so that their imputed fractions of the total output would be $\alpha/(\alpha+\beta)$ and $\beta/(\alpha+\beta)$ rather than just α and β. This permits the imputed shares of labor and capital to be 100% of the total output in a competitive equilibrium even though the sum $(\alpha+\beta)$ may be greater than, equal to or less than one.

However, under imperfect competition, the wage rate, W, per unit of labor and the interest rate, I, per unit of capital could deviate from the values of their marginal productivities. If labor and capital were the only two factors of production, then the total output value, $(\alpha+\beta)PV$, would still be completely distributed to labor and capital as shown in equation (9.5.12), but underpayments relative to labor's marginal productivity would result in overpayments relative to capital's marginal productivity and vice versa.

$$(\alpha+\beta)PV = LW + CI \qquad ....(9.5.12)$$

When the wage rate is less than the value of the marginal productivity of labor, the role of economic decision-making would be to hire more labor and slow down the rate of capital investment. When the wage rate is less than the value of the marginal productivity of labor, then less labor would be hired and the rate of capital investment would be increased. In any case, the effects of such economic decision-making would be to move in a direction of competitive equilibrium.

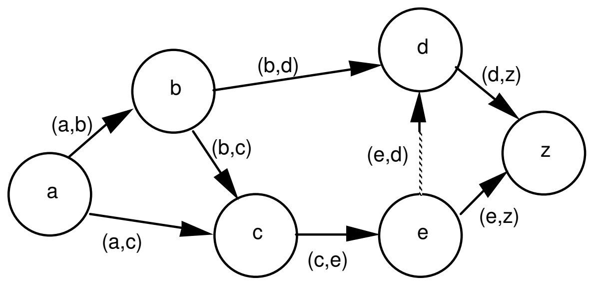
In this connection, it is interesting to compare theories of production and distribution outlined in Section 1.2 of Chapter One. Labor and capital receive wage rates and interest rates which are later exchanged for shares of the total value of the output. Adam Smith assumed the increased productivity due to the division of labor and industrial specialization would be distributed accordingly to labor and capital by the medium of exchange. After allowing for capital consumption during production, Karl Marx assumed capital was remunerated even though its marginal productivity was zero. Consequently, labor could receive its marginal productivity only if all means of production was owned by the state.

Although theories of Adam Smith and Karl Marx may be partially true, their validity need to be tested by well-defined and accurate measurements. Applying dimensional analysis to formulating engineering and aggregate productions functions promises to be the best guide for comparing input factor payments to their marginal productivities.

Section 9.6 - Project Management and the Critical Path Method (CPM)

Engineers often have the managerial responsibility of planning a project and proposing methods of carrying it out. The most well-known and useful method of project management is the "critical path method, or CPM". Various tasks (or activities) that make up a project can be represented by a network diagram as shown in Figure 9.6.1. Some activities in a network cannot start unless others are completed. For example, activities (b,c) and (b,d) cannot start before (a,b) is completed, and activity (c,e) cannot start before both (a,c) and (b,c) are completed.

Figure 9.6.1 - Logic of a Network Diagram



Network logic does not permit activities to cycle or backtrack. For example, let (c,a) denote the activity in the reverse direction of activity (a,c). Then the start of (c,a) must wait for (b,c) to finish, the start of (b,c) must wait for (a,b) to finish and the start of (a,b) must wait for (c,a) to finish. Thus, (c,a) cannot start until it finishes which is impossible. To prevent cycling or backtracking, the network has only one starting node 'a' from which activities leave but cannot enter, and only one finishing node 'z' where activities enter but cannot leave. Network logic also does not permit dangling activities whose start and finish are not both part of a complete string of activities from start to finish of the complete project.

CPM networks use either arrow diagramming (Figure 9.6.4) or precedence diagramming (Figure 9.6.5) to plan, schedule and monitor all activities of a project. The arrows represent activities which obey network logic from start to finish of a project. Every activity is part of a string of arrows from starting node 'a' to finishing node 'z' of the project. Each arrow is designated by a unique pair of starting and finishing nodes (i,j) that represent points of time between which the activity could be executed. If two or more distinct activities in parallel had the same pair of starting and finishing nodes, their designations would not be unique. To circumvent this problem, a *node designation dummy activity* is inserted in series with parallel activities to make their node designations unique. Node designation dummy activities are represented by dashed lines in arrow diagrams, and they do not require any time for their execution.

Logical dummy activities may also be needed to specify the logic of a network with parallel strings of activities that start and end at different nodes. For example, activities (b,d) and (c,e) in Figure 9.6.1 have starting nodes 'b' and 'c', and ending nodes 'd' and 'e'. But activity (d,z) cannot start until <u>both</u> activities (b,d) and (c,e) are completed. In this case, a logical dummy activity (e,d) must be inserted from the ending node of (c,e) to the ending node of (b,d), thereby forcing both (b,d) and (c,e) to be completed before (d,z) can start.

Arrows express the logic of the network, but the lengths of the arrows do not have to be related to the durations and delays of the activities. However, it is useful to have real-time arrow diagrams in which the lengths of arrows are scaled to the durations of the activities. The slack at the start and finish of each activity (i.e., the backward and forward-floats of each activity) also needs to be drawn to scale. We will now describe forward and backward-float calculations which are needed for drawing real-time arrow diagrams.

<u>Forward-Float Calculations</u> begin at the <u>earliest starting time ES(a)</u> of starting node 'a' and end at the <u>earliest starting time ES(z)</u> of finishing node 'z'. All activities starting from node 'i' have the same earliest starting time ES(i). Upon adding duration time $T(i,j) \geq 0$ of the ij-th activity to the earliest starting time ES(i), we obtain the early finish time EF(i,j) of the ij-th activity as defined in equation (9.6.1).

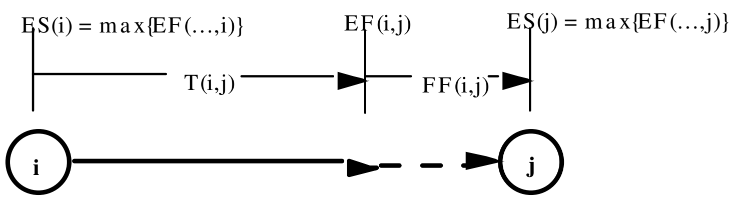
$$ES(i) + T(i,j) = EF(i,j) \qquad ....(9.6.1)$$

The earliest starting time ES(j) of all activities leaving node 'j' is defined as the maximum of the early finish times of all activities ending at node 'j'. Thus, ES(j) = max{EF(...,j)}. Assuming the ij-th activity starts at the earliest starting time ES(i), the early finish time EF(i,j) is either equal to or less than ES(j). The forward float FF(i,j) measures the slack time, if any, from early finish time EF(i,j) of the ij-th activity to the earliest starting time ES(j) from node 'j' as defined in equation (9.12) and displayed in Figure 9.6.2 .

$$EF(i,j) + FF(i,j) = ES(j) = \max\{EF(...,j)\}; \quad FF(i,j) \geq 0 \qquad ....(9.6.2)$$

Arrow diagrams without a time scale could be converted into real-time forward-float diagrams by adding T(i,j) to ES(i) in order to obtain EF(i,j), and then adding FF(i,j) to obtain ES(j) as shown in Figure 9.6.2. Since forward float FF(i,j) occurs *after* the early finish time EF(i,j) of the ij-th activity, forward float FF(i,j) is an "after-the-fact" measure of slack time which is not as useful as measuring the slack *before* the late start of the ij-th activity.

Figure 9.6.2 - Forward-float, earliest-start, real-time arrow diagram.



Forward-float calculations finish at ending node 'z' where earliest starting time ES(z) = max{EF(...,z)} is the maximum early finish time of all activities ending at node 'z' which is the earliest completion date of the project. ES(z) is set equal to the latest finishing time LF(z) of all activities ending at node 'z' in order to begin the backward-float calculations.

Backward-Float Calculations begin at latest finishing time LF(z) of node 'z' and end at latest finishing time LF(a) of node 'a'. All activities ending at node 'j' have the same latest finishing time LF(j). Upon subtracting the duration time T(i,j) of the ij-th activity from LF(j), we obtain the late start time LS(i,j) of the ij-th activity as defined in (9.6.3).

$$LS(i,j) = - T(i,j) + LF(j) \qquad ....(9.6.3)$$

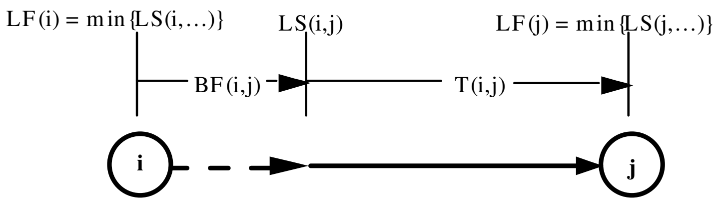
The latest finishing time LF(i) of all activities ending at node 'i' is defined as the minimum of the late start times of all activities that leave node 'i'. Thus, LF(i) = min{LS(i,...)}. Assuming the ij-th activity finishes at latest finishing time LF(j), the late start time LS(i,j) is either equal to or greater than LF(i). The backward-float BF(i,j) measures the slack time from the latest finishing time LF(i) of all activities finishing at the i-th node to the late start time LS(i,j) of the ij-th activity as defined in equation (9.14) and displayed in Figure 9.6.3 .

$$\min\{LS(i,...)\} = LF(i) = - BF(i,j) + LS(i,j); \quad BF(i,j) \geq 0 \qquad ....(9.6.4)$$

Arrow diagrams without a time scale could be converted into real-time backward-float diagrams by subtracting T(i,j) from LF(j) in order to obtain LS(i,j), and then subtracting

BF(i,j) from LS(i,j) in order to obtain LF(i) as shown in Figure 9.6.3 below. Since the backward float BF(i,j) measures slack time "before-the-fact" of the late start of the ij-th activity, it is more useful than FF(i,j) for scheduling activities without delaying the project completion date.

Figure 9.6.3 - Backward-float, latest-finish, real-time arrow diagram.



At node 'a', backward-float calculations determine the latest finishing time LF(a) which is the latest time a project could start without delaying the project completion date. If all forward- and backward-float calculations were carried out correctly, the earliest-starting and latest-finishing times at node 'a' should be equal (i.e., ES(a) = LF(a)).

Total float TF(i,j) measures the slack time of the ij-th activity between latest finishing time LF(j) at node 'j' and earliest starting time ES(i) at node 'i' as shown in equation (9.6.5).

$$TF(i,j) = LF(j) - T(i,j) - ES(i) = \min\{LS(j,...)\} - T(i,j) - \max\{EF(...,i)\} \geq 0 \quad ....(9.6.5)$$

Backward-float calculations are needed to determine LF(j) and forward-float calculations are needed to determine ES(i). Therefore, unlike BF(i,j) and FF(i,j), the measurement of the total float TF(i,j) of the ij-th activity requires both backward- and forward-float calculations. It will now be shown that total float TF(i,j) is greater than or equal to either FF(i,j) or BF(i,j). For this purpose, substitute -EF(i,j) from equation (9.6.1) for -T(i,j)-ES(i) in equation (9.6.5), and then substitute FF(i,j)-ES(j) for -EF(i,j) from equation (9.6.2) to get

$$TF(i,j) = LF(j) + FF(i,j) - ES(j) \geq 0 \quad ....(9.6.6)$$

Since LF(j) - ES(j) $\geq$ 0, it follows that TF(i,j) $\geq$ FF(i,j). Similarly, substituting equations (9.6.3) and (9.6.4) in equation (9.6.5) gives equation (9.6.7) from which it follows that TF(i,j) $\geq$ BF(i,j).

$$TF(i,j) = LF(i) + BF(i,j) - ES(i) \geq 0 \quad ....(9.6.7)$$

The critical path(s) from start to finish of a project consist of activities whose total floats are all zero. The project duration would be shorter if activities along the critical path(s) could be arranged in parallel. The project completion date could also be realized sooner if critical activities were shortened by *crashing*. *Crashing* consists of extra workers, equipment, overtime or subcontracting in order to shorten activity duration time. Crashing is often expensive, but the benefits of a reduced time schedule may be worth the extra cost.

Manpower leveling concerns stabilizing the number of workers involved during a project, or in scheduling highly skilled or short-supply workers. Real-time arrow diagrams show where a project has slack time which can be used to spread out the total number of workers needed at any time, and even out the demand for more critical or highly skilled workers. Manpower leveling may require many economic comparisons for optimal tradeoffs between time savings and increased costs under existing manpower and financial constraints. Arrow diagrams often reveal intermediate outputs which could generate income before completing the project. The critical path method is illustrated by the following example shown as an arrow diagram in Figure 9.14 and a precedence diagram in Figure 9.15 . Tables 9.11 to 9.13 give the boundary timetable and real-time forward- and backward-float diagrams.

Figure 9.6.4 - Arrow Diagramming of Project Activities

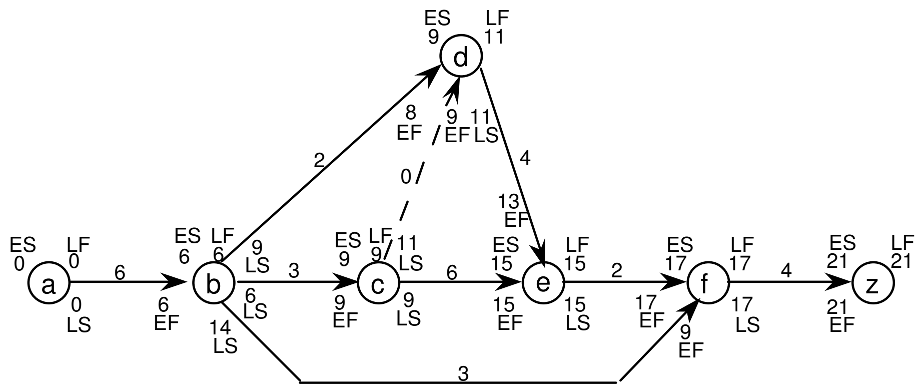


Table 9.6.1 - Boundary timetable of project activities

ES(i) + T(i,j) = EF(i,j) LF(j) - T(i,j) = LS(i,j)
EF(i,j) + FF(i,j) = ES(j) LS(i,j) - BF(i,j) = LF(i)

		Forward Float Calculations			Backward Float Calculations			
Activity (i,j)	Duration T(i,j)	Earliest Starting ES(i)	Early Finish EF(i,j)	Forward Float FF(i,j)	Latest Finishing LF(j)	Late Start LS(i,j)	Backward Float BF(i,j)	Total Float TF(i,j)
a,b	6	0	6	0	6	0	0	0
b,d	2	6	8	1	11	9	3	3
b,c	3	6	9	0	9	6	0	0
b,f	3	6	9	8	17	14	8	8
c,d	0	9	9	0	11	11	2	2
c,e	6	9	15	0	15	9	0	0
d,e	4	9	13	2	15	11	0	2
e,f	2	15	17	0	17	15	0	0
f,z	4	17	21	0	21	17	0	0

Table 9.6.2 - Real-time forward-float diagram.

Real-Time Forward-Float Diagram: ES(i) + T(i,j) ‘ EF(i,j) + FF(i,j) ‘ ES(j) Float between the <u>early finish time</u> EF(i,j) and <u>earliest starting time</u> ES(j).

```
Time Scale:  0 . . . . 5 . . . .10 . . . .15 . . . .20 . . . .25
             a          b     c          e   f       z
(abcefz)     . . . . . . . . . . . . . . . . . . . . . . . . .
             a          b     cd       e + e   f       z
(abcdefz)    . . . . . . . . . . . . . . . . . . . . . . . . .
             a          b   d+d        e + e   f       z
(abdefz)     . . . . . . . . . . . . . . . . . . . . . . . . .
             a          b     f + + + + + + + f       z
(abfz)       . . . . . . . . . . . . . . . . . . . . . . . . .
```

Table 9.6.3 - Real-time backward-float diagram.

Real-Time Backward-Float Diagram: LF(i) ' -BF(i,j) + LS(i,j) ' -T(i,j) + LF(j)
Float between the <u>late start time</u> LS(i,j) and the <u>latest finishing time</u> LF(i).

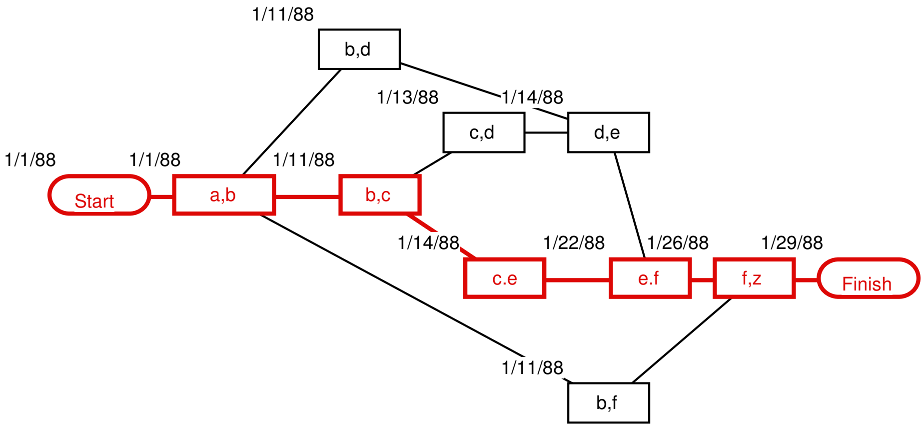
```
Time Scale:  0 . . . . 5 . . . .10 . . . .15 . . . .20 . . . .25
             a           b     c           e   f       z
(abcefz)     . . . . . . . . . . . . . . . . . . . . . . . . . .
             a           b     c – cd      e   f       z
(abcdefz)    . . . . . . . . . . . . . . . . . . . . . . . . . .
             a           b – – b   d       e   f       z
(abdefz)     . . . . . . . . . . . . . . . . . . . . . . . . . .
             a           b – – – – – – – b     f       z
(abfz)       . . . . . . . . . . . . . . . . . . . . . . . . . .
```

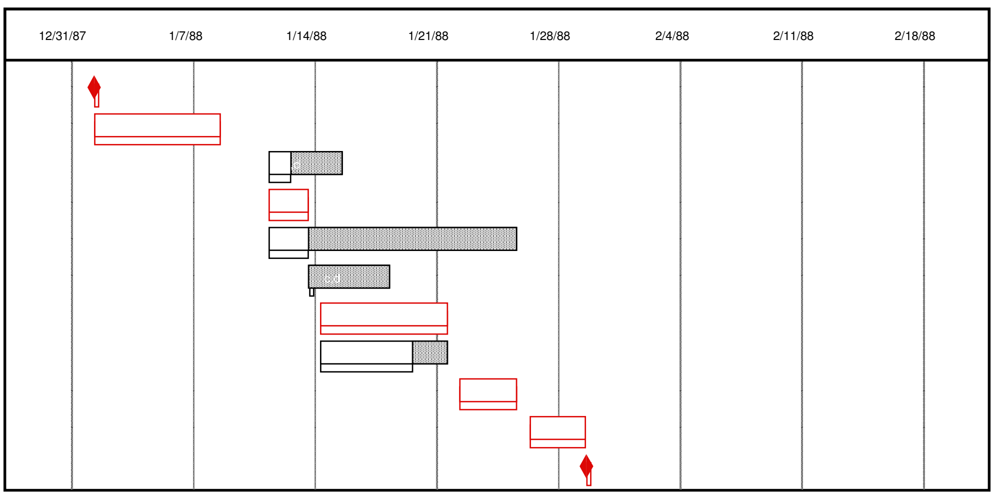
Various project management programs are available for IBM-compatible and Macintosh personal computers. The output from personal computer programs are usually based on precedence rather than arrow diagrams, and float is usually depicted as total float rather than distinguishing between forward and backward float. Precedence diagrams use boxes in which each activity is described, and arrows are used in a left-to-right direction to describe the precedence relationships between activities. Weekends and holidays are taken into account for timing activities and total floats. For this reason, the following calendar is useful in reading the output from Figure 9.15 which illustrates the precedence and total float diagramming from the MacProject II program of the Claris Corporation for Macintosh personal computers.

Table 9.6.4 - January 1988 Calendar

	January			1988		
Sun	Mon	Tue	Wed	Thur	Fri	Sat
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

Figure 9.6.5 - Precedence and Total Float Diagramming of Project Activities.


1/11/88
b,d
1/13/88
1/14/88
c,d
d,e
1/1/88
1/1/88
1/11/88
Start
a,b
b,c
1/14/88
1/22/88
1/26/88
1/29/88
c.e
e.f
f,z
Finish
1/11/88
b,f


12/31/87
1/7/88
1/14/88
1/21/88
1/28/88
2/4/88
2/11/88
2/18/88

Section 9.7 - Summary of Chapter Nine

In this chapter we deal with problems of economic decision-making for the optimal allocation of limited resources involving two or more variables. The problems of optimizing selections from production alternatives raise important questions about the objectives to be optimized, the effectiveness of input and output constraints, and the economic significance of linear and nonlinear relationships of the variables.

Linear programming (LP) is a mathematical technique of formulating problems with linear functions and finding optimal solutions for a given set of limiting constraints. Section 9.1 concerns engineering objectives in linear programming subject to equality constraints. A product-mix problem is given with two output variables (instant coffees called Taster's and Connoisseur's Brews) and a production matrix with two input variables (called arabica and robusta beans) subject to equality constraints. If only one output is selected, it would use all of one input constraint but the other constraint would have some leftovers. The engineering objective is to maximize the outputs without any leftover input constraints.

The *primal* problem of linear programming in Section 9.1 is formulated and solved both algebraically and graphically. The arabica and robusta inputs (a,r) are shown by the Jacobian matrix to be functionally independent of the Taster and Connoisseur outputs (x_T,x_C). A map is then constructed between the (a,r) input constraints and the (x_T,x_C) outputs which satisfy the engineering objective. The solution of this product-mix problem by matrix algebra functions of microcomputer spreadsheet programs is briefly explained.

Section 9.2 introduces prices (p_T,p_C) of the Taster and Connoisseur outputs. The financial objective in linear programming is defined as maximizing the sales revenues of the (x_T,x_C) outputs subject to the (a,r) input constraints. Herein lies the distinction between engineering and financial objectives. The engineering objective assumes output is maximized when inputs are combined without any leftovers. The financial objective maximizes outputs in a monetary sense regardless of leftover input constraints. If p_T/p_C is greater than 5/3 or less than 1/2, sales revenues with leftover input constraints would be greater than those that satisfy the engineering objective. The remainder of Section 9.2 is limited to $5/3 > p_T/p_C > 1/2$ where engineering and financial objectives are compatible.

The Lagrangian technique combines the financial objective and the engineering input constraints into one *Lagrangian function* $L(x_T,x_C,y_a,y_r)$. Undetermined multipliers y_a and y_r of $L(x_T,x_C,y_a,y_r)$ are called the *implicit costs* of arabica and robusta beans. $L(x_T,x_C,y_a,y_r)$ is maximized by setting the partial derivatives of each of its four arguments equal to zero. Partial derivatives of y_a and y_r set equal to zero result in two material-balance equations called the *primal* problem of linear programming. Partial derivatives of x_T and x_C set equal to zero result in two cash-balance equations called the *dual* problem of linear programming.

The objective of the primal problem is to maximize output decision variables (x_T,x_C) subject to the (a,r) input constraints. The objective of the dual problem is to minimize input decision variables (y_a,y_r) subject to the (p_T,p_C) output price constraints. The Jacobian matrix shows that not only (a,r) and (x_T,x_C) are functionally independent, but also (y_a,y_r) and (p_T,p_C). However, (a,r) and (x_T,x_C) are functionally dependent on (y_a,y_r) and (p_T,p_C). Consequently, the dual LP problem can be formulated as an alternative to the primal LP problem for the purposes of computing solutions and checking results. But it also raises the more important question of whether minimizing input costs is equivalent to maximizing output revenues.

To examine this question further, Lagrange implicit costs y_a and y_r are defined as the change of total sales revenues divided by the change of either arabica or robusta bean input while the other input is fixed. This enables us to plot isocost line ay_a+ry_r through an arbitrary (a,r) input constraint which would be equivalent to the isorevenue line $p_Tx_T+p_Cx_C$. It is then shown that the (a,r) input constraint divides the length of the isocost line between Taster's and Connoisseur's raypaths in the inverse ratio p_Cx_C/p_Tx_T (see Figure 9.2.2).

Section 9.3 uses the same coffee-mix problem as before but in connection with inequality constraints that may involve unequal numbers of input and output variables. The primal and dual problems of two input and two output variables are formally defined in linear programming language and both problems are solved graphically in (x_T,x_C) and (y_a,y_r) coordinates respectively. The coffee-mix problem is then expanded to two input and four output variables and the problem is solved graphically in (a,r) coordinates using convex envelopes of isocost lines to determine two surplus variables from the four output variables. The result is the same as two input and two output variables subject to equality constraints. Because of the large literature and off-the-shelf computer programs which are available, we only mention the simplex method of G. B. Dantzig which systematically determines slack and surplus variables in large-scale linear programming problems.

An engineering production function (Section 9.4) is a mathematical description of the technological relationship between input and output variables in a production process. A single budget constraint replaces the arabica and robusta input constraints. Taster's and Connoisseur's market prices are used as sales revenues rather than the price constraints used for determining arabica and robusta implicit costs in linear programming. Engineering production functions use market prices of arabica and robusta beans to determine their input costs. Consequently, isocost lines differ from isorevenue curves and the differences between input costs and output revenues permit the realization of profit.

In seeking a mathematical relationship between inputs and outputs that have different units of measurement, dimensional analysis requires the ratio of any two concrete examples of output to be independent of the size of the units in which inputs are measured. This 'ratio requirement' of dimensional analysis restricts the form of the production function to a product of the powers of the input and output quantities (see the product theorem in *Dimensional Analysis* by P.W. Bridgman, Yale University Press, 1931). Appendix 9A explains properties of isocost lines and isorevenue curves from homogeneous production functions which satisfy the product theorem of dimensional analysis.

The engineering production function describing the previous coffee-mix problem is

$$P(a,r) = A(a)^{\alpha}(r)^{\beta} \qquad ....(9.4.1)$$

where P represents weekly revenues from sales of Taster's and Connoisseur's Brews produced from inputs of 'a' and 'r' tons/week of arabica and robusta beans respectively. Parameters α and β are dimensionless exponents and A is a dimensional constant. Equation (9.4.1) is called a homogeneous production function of degree $\alpha+\beta \equiv n$ which measures how much the output changes relative to equiproportional changes of both inputs. The ratio β/α measures the substitutability of arabica for robusta beans when P is constant. From the production coefficient matrix of the coffee-mix problem, we determined that $\beta/\alpha=0.6403588$. For comparisons with linear programming, it was assumed that $\alpha+\beta = 1$. This assumption enabled us to determine exponents $\alpha = 0.6096227$ and $\beta = 0.3903773$.

From production coefficients required to produce one ton of Taster's or Connoisseur's Brew and their market prices p_T=\$4,000 and p_C=\$5,000 per ton, equation (9.4.1) enabled us to determine A = \$2797.1338/week. Assuming y_a = \$1,857.14 and y_r = \$857.14 are both implicit and market prices per ton of arabica and robusta beans, an input of (a,r) = (3,4) tons/week would cost and earn \$9,000/week in the LP model. Although an input of (a,r) = (3,4) tons/week in the production function model also costs \$9,000/week, it earns \$9,388.75/week. If a budget constraint of \$9,000/week is used in the production function, the optimum input is (a,r) = (2.954332,4.0989708) tons/week which earns \$9,390.53/week.

Section 9.5 concerns aggregate production functions which describe input/output relationships in an industry or a nation. An aggregate production function satisfying the 'ratio requirement' of dimensional analysis is the Cobb-Douglas production function

$$P = A * L^{\alpha} * C^{\beta} \qquad ....(9.5.1)$$

where P, L, and C are annual indices representing industrial output, and labor and capital input, and A, α and β are unknown parameters. The Cobb-Douglas production function was fitted by log-linear regression analysis to United States industrial data from 1899 to 1922. It was assumed United States had constant returns to scale (i.e., $\alpha+\beta = 1$) during the 24-year study period. The results were A = 1.01, $\alpha = 0.75$ and $\beta = 1-\alpha = 0.25$. Exponents $\alpha = 0.75$ and $\beta = 1-\alpha = 0.25$ were interpreted as imputed shares of labor (L) and capital (C), and A = 1.01 was interpreted as a least-squares residual representing technological change.

Parameters A, α and β in the Cobb-Douglas production function were interpreted very differently than in engineering production functions where parameter A is estimated only for dimensional validity rather than improving statistical reliability by assuming constant returns to scale. It is explained in Section 9.5 as well as the chicken-and-egg production function of Appendix 9B that technological change is defined and measured by the sum and ratio of parameters α and β instead of parameter A. Appendix 9C suggests a statistical method of estimating the parameters of the Cobb-Douglas production function that does not assume constant returns to scale.

The Ricardian theory of rent described in Section 1.2 implies labor and capital would be paid according to their marginal productivities in a competitive equilibrium. If so, Euler's theorem on homogeneous functions presented in Appendix 9D shows the imputed shares of labor and capital in the total output $(\alpha+\beta)P$ would be $\alpha/(\alpha+\beta)$ and $\beta/(\alpha+\beta)$ instead of α and β. Under imperfect competition, underpayments relative to labor's marginal productivity would result in overpayments relative to capital's marginal productivity and vice versa.

Section 9.6 deals with the critical path method (CPM) of solving project management problems.. The time-dependent activities activities which comprise a project are represented in an arrow network diagram which shows the relationships between activities in terms of the order in which they are performed and the duration of each activity. The arrows in the diagram specify the logic of the network, but their lengths are not necessarily proportional to the duration of the activities they represent. However, in order to schedule and moniter the project it is useful to have real-time arrow diagrams in which the arrow lengths are proportional to the duration of the activities. For this, it is necessary to incorporate the slack times at the start and finish of each activity by calculating the backward and forward-float of each activity. The critical path (or longest path) is the string of activities whose total floats are all zero. Projects may be completed earlier if critical activities can be shortened either by crashing or manpower leveling.

Appendix 9A - Properties of Homogeneous Production Functions

Linear programming is widely used for multiple-variable optimization and it is often applied to large-scale problems of different economic organizations. Equality and inequality constraints are specified for input and output variables which are supplemented with slack and surplus variables to satisfy the constraints. The LP algorithm maximizes or minimizes a linear objective function of the variables iteratively until an optimum is reached.

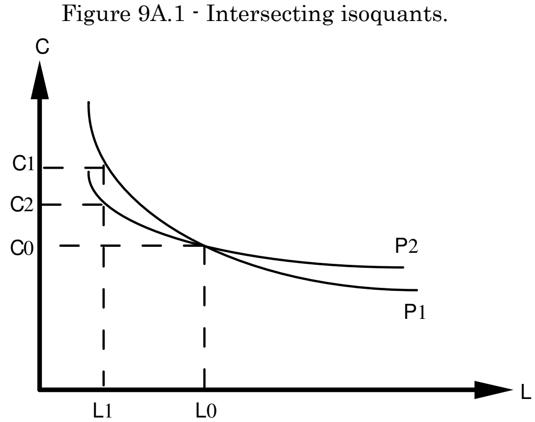
Section 9.4 explains how multiple-variable optimization could be carried out in small-scale problems with homogeneous production functions which satisfy the ratio requirement and product theorem of dimensional analysis. A production function is a mathematical description of the technological relationship between input and output variables of a production system per unit of time. The basic properties of production functions will be presented here in the context of the Cobb-Douglas production function as stated in equation (9.5.1) which is reproduced below as equation (9A.1). When equation (9A.1) is fitted to empirical data by log-linear regression analysis, multicollinearity problems often arise which may be resolved by the technique suggested in Appendix 9C.

$$P = A * L^{\alpha} * C^{\beta} = f(L,C) \qquad ....(9A.1)$$

The properties of production functions can be examined graphically using three-dimensional surfaces which depict the outputs for every combination of two inputs. However, it is easier to visualize production function surfaces in two dimensions by their intersections with parallel planes of constant-output which are commonly called *surface contours* or *isoquant curves* Each isoquant curve specifies all combinations of inputs that produce an equal quantity of output. Each *isocost* line specifies all combinations of inputs that have an equal cost. From five isoquant properties explained below, we can show that output is maximized for a given input and input is minimized for a given output at points of tangency between isoquant curves and isocost lines.

Isoquant Property I - Isoquants of P = f(L,C) do not intersect because f(L,C) is a single-valued function. To prove isoquants of single-valued production functions do not intersect, let us consider two intersecting isoquants as shown in Figure 9A.1.

Figure 9A.1 - Intersecting isoquants.



Suppose $P_2 > P_1$ and their isoquants intersect at (L_0,C_0). Since $P_2 > P_1$ by hypothesis, the combination (L_0,C_0) would be used to produce P_2 in order to maximize output for a given input. Let us now consider input $L_1 < L_0$. The combination (L_1,C_2) produces P_2 and the combination (L_1,C_1) produces P_1. In other words, fixed input L_1 and input $C_2 < C_1$ would produce $P_2 > P_1$ (i.e., less input is used to produce more output). This contradicts the basic property of production functions that every combination of inputs is used in a technologically efficient manner to produce the most output. This proves that production functions do not have intersecting isoquants.

Isoquant Property II - Isoquants of P = f(L,C) have negative slopes, dC/dL < 0. Proof: Set the total differential dP of equation (9A.1) equal to zero and solve for dC/dL.

$$dP = (\partial P/\partial L)_C dL + (\partial P/\partial C)_L dC = 0\,; \quad dC/dL = -(\partial P/\partial L)_C/(\partial P/\partial C)_L \qquad ....(9A.2))$$

Substituting $(\partial P/\partial L)_C = \alpha A L^{\alpha-1}C^{\beta} = \alpha P/L$ and $(\partial P/\partial C)_L = \beta A L^{\alpha}C^{\beta-1} = \beta P/C$ from equations (9.5.4) and (9.5.5) in equation (9A.2), we get equation (9A.3). Since α, β, C and L are positive, it follows from equation (9A.3) that dC/dL is negative.

$$dC/dL = -\alpha C/\beta L < 0 \qquad ....(9A.3)$$

Isoquant Property III - The isoquants of P = f(L,C) are convex to the origin, (i.e., $d^2C/dL^2 > 0$). Proof: Determine the total differential of dC/dL and divide the result by dL to get d^2C/dL^2. Substitute equation (9A.3) in the differential equation to evaluate d^2C/dL^2.

$$d\left\{\frac{dC}{dL}\right\} = \frac{\partial}{\partial L}\left\{\frac{dC}{dL}\right\}dL + \frac{\partial}{\partial C}\left\{\frac{dC}{dL}\right\}dC$$

$$\frac{d}{dL}\left\{\frac{dC}{dL}\right\} = \frac{d^2C}{dL^2} = \frac{\partial}{\partial L}\left\{\frac{dC}{dL}\right\} + \frac{\partial}{\partial C}\left\{\frac{dC}{dL}\right\}\left\{\frac{dC}{dL}\right\}$$

$$\frac{d^2C}{dL^2} = \frac{\partial}{\partial L}\left(-\frac{\alpha C}{\beta L}\right) + \frac{\partial}{\partial C}\left(-\frac{\alpha C}{\beta L}\right)\left(-\frac{\alpha C}{\beta L}\right) = \frac{\alpha C}{\beta L^2} + \left(-\frac{\alpha}{\beta L}\right)\left(-\frac{\alpha C}{\beta L}\right)$$

$$\frac{d^2C}{dL^2} = \frac{\alpha C}{\beta L^2} + \left(\frac{\alpha}{\beta L}\right)\left(\frac{\alpha C}{\beta L}\right) = \frac{\alpha\beta C}{\beta^2 L^2} + \frac{\alpha^2 C}{\beta^2 L^2} = \frac{\alpha C}{\beta^2 L^2}(\beta + \alpha) > 0 \qquad ....(9A.4)$$

Since α, β, C and L are positive, it follows from equation (9A.4) that dC^2/dL^2 is positive.

Isoquant Property IV - A ray path with a fixed C/L input ratio intersects every isoquant of P = f(L,C) at the same angle (i.e., the isoquants are parallel at the points of intersection with the same ray path), and small equiproportionate changes of L and C inputs (i.e., scale changes) result in n times that proportionate change in output P, where n = α + β is the degree of homogeneity of the production function.

By definition, a production function P = f(L,C) is homogeneous of degree n, if for all real values of $\lambda > 0$, $f(\lambda L,\lambda C) = \lambda^n f(L,C)$, or $f(L,C) = \lambda^{-n} f(\lambda L,\lambda C)$. Hence, when each input factor is multiplied by a real positive constant λ, the constant λ^n can be completely factored out of the functional expression. For example, in the Cobb-Douglas production function, we have $f(L,C) = A L^{\alpha} C^{\beta}$ and $\lambda^{-n} f(\lambda L,\lambda C) = \lambda^{-n} A (\lambda L)^{\alpha} (\lambda C)^{\beta} = \lambda^{-n}\lambda^{\alpha+\beta} A L^{\alpha} C^{\beta} = A L^{\alpha} C^{\beta} = f(L,C)$.

In particular, let $\lambda = L^{-1}$ which results in $P = f(L,C) = L^n f(1,C/L) = AL^n (C/L)^{\beta}$. Since $L = \rho\cos\theta$, where $\rho \equiv \sqrt{L^2 + C^2}$ is the length of the ray path from the origin to the isoquant of output P, and θ is the angle the ray path makes with the L-axis. We may now rewrite the production function as follows:

$$P = A L^{\alpha} C^{\beta} = A L^n (C/L)^{\beta} = A \rho^n (\cos\theta)^n (C/L)^{\beta} \qquad ....(9A.5)$$

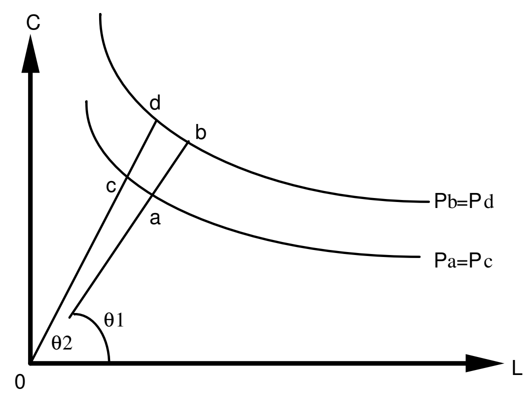
In order to show that a ray path intersects every isoquant at the same angle, and small equiproportional changes of all inputs result in n times that proportional change in output, let us consider two isoquants which differ slightly in output and their intersections with two ray paths which differ slightly in angle as shown in Figure 9A.2 below. We then have

$$P_a = A (\rho_a)^n (\cos\theta_1)^n (C_a/L_a)^{\beta} = P_c = A (\rho_c)^n (\cos\theta_2)^n (C_c/L_c)^{\beta}$$

$$P_b = A (\rho_b)^n (\cos\theta_1)^n (C_b/L_b)^{\beta} = P_d = A (\rho_d)^n (\cos\theta_2)^n (C_d/L_d)^{\beta}$$

Dividing the second set of equations by the first set, we get $\frac{P_b}{P_a} = \frac{(\rho_b)^n}{(\rho_a)^n} = \frac{(\rho_d)^n}{(\rho_c)^n}$ or $\frac{\rho_b}{\rho_a} = \frac{\rho_d}{\rho_c}$. Hence, triangles 0bd and 0ac are similar, and bd is parallel to ac.

Figure 9A.2 - Two adjacent isoquants intersecting with two adjacent ray paths.



Also, let $P_b = P_a + \Delta P$ and $\rho_b = \rho_a + \Delta\rho$. Then

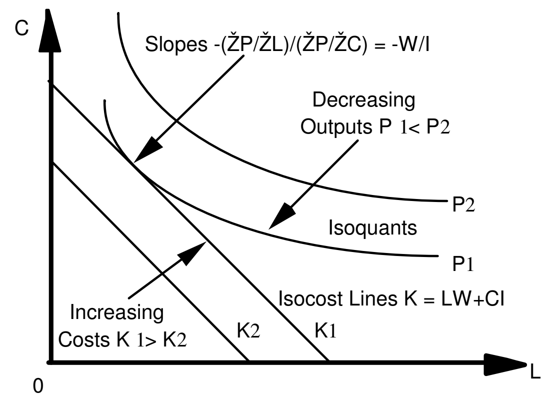
$$\frac{P_b}{P_a} = 1 + \frac{\Delta P}{P_a} = \frac{(\rho_b)^n}{(\rho_a)^n} = \frac{(\rho_a + \Delta\rho)^n}{(\rho_a)^n} = (1 + \frac{\Delta\rho}{\rho_a})^n \cong 1 + \frac{n\Delta\rho}{\rho_a} \quad \text{(binomial approximation)}$$

Therefore, small equiproportional changes of L and C inputs result in n times that change in output P, where $n = \alpha + \beta$ is the degree of homogeneity of the production function.

Isoquant Property V - The least-cost combination of inputs L and C for a specified output of P requires their marginal physical productivities $\partial P/\partial L$ and $\partial P/\partial C$ to be in the same proportion as their prices. This means that an additional dollar's worth of either input adds as much to the total output as does a dollar's worth of the other input.

Graphically, the minimum cost of L and C occurs at the point of tangency between decreasing isoquant curves and increasing isocost lines LW+CI, where W and I are the prices per unit of L and C respectively as shown in Figure 9A.3. Consequently, isoquant slope $-(\partial P/\partial L)/(\partial P/\partial C)$ equals the isocost slope -W/I. Adding the price V per unit of P to the analysis converts isoquant into isorevenue curves and makes the total value of the output equal to $(\alpha+\beta)*PV$ of which $LW = [\alpha/(\alpha+\beta)]*PV$ is the share of labor and $CI = [\beta/(\alpha+\beta)]*PV$ is the share of capital under a competitive equilibrium (see Section 9.5).

Figure 9A.3 - Point of tangency of isoquant curves and isocost lines.



Appendix 9B - Chicken-and-Egg Production Function

The measurement of marginal productivity is one of the most difficult problems in the study of economics. When the proportion of input factors is fixed, the measurement of average productivity is straightforward. But when the proportion of input factors is variable, it is not only important to measure average productivity but also how much output is attributable to changes of each input factor while other factors are fixed. Such measurements of marginal productivity are not simple economic problems as shown by the following example (see G. J. Stigler, The Theory of Price, Macmillan Co., 1947).

Suppose a crew of 10 men with 10 shovels could dig a ditch 50 yards long per day. The average productivity of a man-and-shovel is 5 yards of ditches per day. If an eleventh man was added keeping the number of shovels fixed, how could we measure the marginal productivity of an eleventh man? Eleven men could rotate 10% resting periods in order to measure how many more yards of ditch would be dug per day. If an eleventh shovel was added keeping the number of men fixed, measuring the marginal productivity of an eleventh shovel would be even more problematical then that eleven men. We could use eleven shovels that were 10% smaller than before in order to measure the scale effect of adding an eleventh man-and-shovel, but that also is of little use for carrying out substitution-effect experiments of varying men keeping shovels fixed, or of varying shovels keeping men fixed.

All scale and substitution effects experiments can be incorporated in a production function to measure average and marginal productivities. The measurement of productivity uses the 'ratio requirement' of dimensional analysis which is a tool of great utility in science and engineering. The underlying premise is that physical quantities cannot be compared in mathematical equations unless they are expressed in terms of the same dimensions.

In this connection, let us consider a chicken-and-egg production problem where the role of dimensional analysis and the physical significance of production function parameters are more easily explained. An experiment shows 1.5 chickens (C_0=1.5) produce 1.5 eggs (E_0=1.5) in 1.5 days (D_0=1.5). The question is, " How many eggs would a dozen chickens produce in a week?". It would seem the data $C_0 = E_0 = D_0 = 1.5$ could be scaled down so that one chicken would produce one egg in one day. It would then follow 12 chickens would take 7 days to produce 84 eggs as shown by equation (9B.1) which is correct numerically.

$$84 \text{ eggs} = (12 \text{ chickens})*(7 \text{ days}) \qquad ....(9B.1)$$

Equation (9B.1) assumes the production function is $E = C*D$, where E denotes the output of Eggs, and C and D denote the inputs of Chickens and Days. But substituting $C_0 = E_0 = D_0 = 1.5$ in $E = C*D$ results in the contradiction 1.5 equals 1.5^2. Moreover, $E = C*D$ implies Eggs equals Chickens times Days which are obviously incomparable.

Dimensional analysis requires the ratio of any two concrete examples of a secondary quantity (such as output E) to be independent of the size of the units in which the primary quantities (such as inputs C and D) are measured. This 'ratio requirement' restricts the form of the production function to be independent of the size of the units in which E, C or D are measured. If the 'ratio requirement' is satisfied, then the production function must be expressible as a product of the powers of the primary quantities. (See P. W. Bridgman, Dimensional Analysis, Yale University Press, 1931). Thus,

$$E/E_0 = (C/C_0)^{\alpha}(D/D_0)^{\beta} \qquad ....(9B.2)$$

where both sides of equation (9B.2) are dimensionless ratios. Exponents α and β are dimensionless, and they can be either positive, negative, integral or fractional.

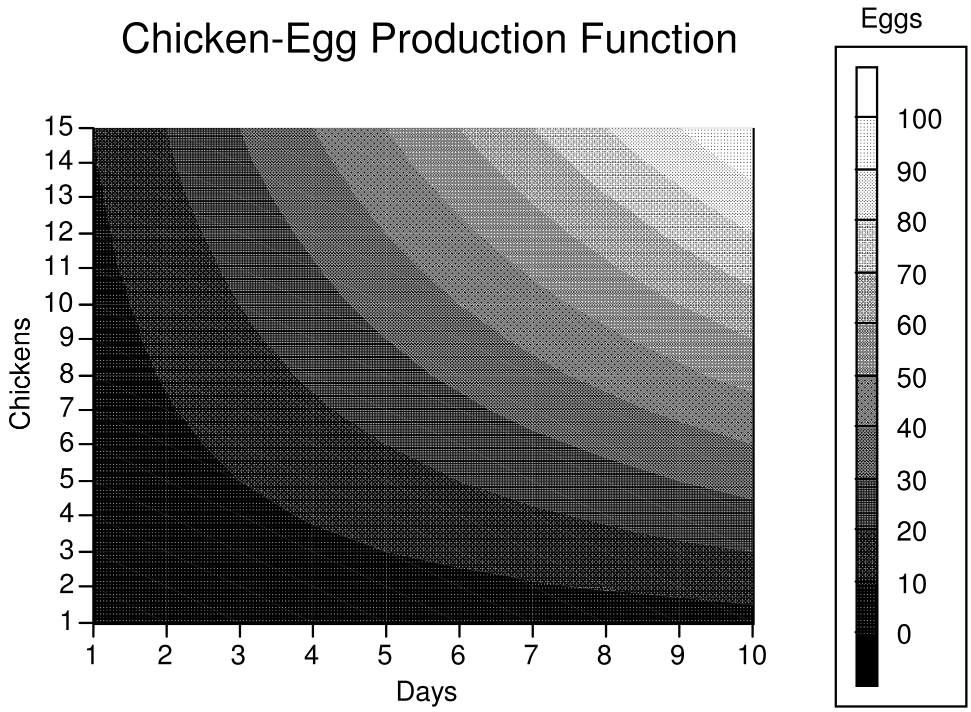
If the chickens are doubled keeping the days fixed, egg production should double. Substituting $E/E_0 = C/C_0 = 2$ and $D/D_0 = 1$ in equation (9B.2), we get $2 = 2^\alpha \times 1$ which indicates $\alpha = 1$. If the days are doubled keeping the chickens fixed, egg production should also double. Substituting $E/E_0 = D/D_0 = 2$ and $C/C_0 = 1$ in equation (9B.2), we get $2 = 1 \times 2^\beta$ which indicates $\beta = 1$. Hence, the chicken-and-egg production function becomes

$$E = \{E_0 / C_0^\alpha D_0^\beta\} C^\alpha D^\beta = A*C^\alpha*D^\beta \qquad ....(9B.3)$$

where constant $A \equiv \{E_0 / C_0^\alpha D_0^\beta\} = 1.5/1.5^2 = 2/3$ has dimensions of eggs per chicken per day. Upon substituting A = 2/3, C = 12, D = 7 and $\alpha = \beta = 1$ into (9B.3), we find 56 instead of 84 eggs would be produced by 12 chickens in a week. Both sides of equation (9B.3) have egg dimensions because dimensional constant A combines with the dimensions of C*D to give dimensions of eggs. The chicken-egg production function surface consists of egg outputs obtained from various input combinations of chickens and days as listed and plotted below.

Table 9B.1 - Chicken-and-Egg Production Function Surface

• • Chicken-and-Egg Production Function• • • • • • • • • • • • • • Eggs =(2/3)*(Chickens)*(Days)• 1• 2• 3• 4• 5• 6• 7• 8• 9• 10• **Days**• • 1• 0.67• 1.33• 2.00• 2.67• 3.33• 4.00• 4.67• 5.33• 6.00• 6.67• • • 2• 1.33• 2.67• 4.00• 5.33• 6.67• 8.00• 9.33• 10.67• 12.00• 13.33• • • 3• 2.00• 4.00• 6.00• 8.00• 10.00• 12.00• 14.00 • 16.00• 18.00• 20.00• • • 4• 2.67• 5.33• 8.00• 10.67• 13.33• 16.00• 18.67• 21.33• 24.00• 26.67 • • • 5• 3.33• 6.67• 10.00• 13.33• 16.67• 20.00• 23.33• 26.67• 30.00• 33.33• • • 6• 4.00• 8.00• 12.00• 16.00• 20.00• 24.00• 28.00• 32.00• 36.00• 40.00• • • 7• 4.67• 9.33• 14.00• 18.67• 23.33 • 28.00• 32.67• 37.33• 42.00• 46.67• • • 8• 5.33• 10.67• 16.00• 21.33• 26.67• 32.00• 37.33• 42. 67• 48.00• 53.33• • • 9• 6.00• 12.00• 18.00• 24.00• 30.00• 36.00• 42.00• 48.00• 54.00• 60.00• • • 10• 6.67• 13.33• 20.00• 26.67• 33.33• 40.00• 46.67• 53.33• 60.00• 66.67• • • 11• 7.33• 14.6 7• 22.00• 29.33• 36.67• 44.00• 51.33• 58.67• 66.00• 73.33• • • 12• 8.00• 16.00• 24.00• 32.00• 40.00• 48.00• **56.00**• 64.00• 72.00• 80.00• • • 13• 8.67• 17.33• 26.00• 34.67• 43.33• 52.00• 60. 67• 69.33• 78.00• 86.67• • • 14• 9.33• 18.67• 28.00• 37.33• 46.67• 56.00• 65.33• 74.67• 84.00 • 93.33• • • 15• 10.00• 20.00• 30.00• 40.00• 50.00• 60.00• 70.00• 80.00• 90.00• 100.00• • • **Ch ickens**• • • • • • • • • • • **Eggs**• •



Let us analyze how scale and substitution-effect experiments are stored in the production function. Let Δ_E denote the relative change of E (i.e., the change ΔE divided by its magnitude E), and let ΔC and ΔD denote the relative changes of C and D respectively. Then for small relative changes of E, C and D, equation (9B.3) could be written as

E(1+ΔE) = A*C□□(1+ΔC)□ *D□(1+ΔD)□□ = A*C□*D□(1+□□ΔC+□□ΔD+□□□ΔCΔD) ....(9B.4)

Since $E = A{*}C^{\alpha}{*}D^{\beta}$ in equation (9B.3), we may cancel these terms on both sides of equation (9B.4). Assuming relative changes $\alpha\Delta_C$ and $\beta\Delta_D$ are much less than one, their product $\alpha\beta\Delta_C\Delta_D$ would be negligible. Consequently, equation (9B.4) reduces to

$$\Delta_E = \alpha\Delta_C + \beta\Delta_D \quad \text{or} \quad 1 = \alpha\Delta_C/\Delta_E + \beta\Delta_D/\Delta_E \qquad ....(9B.5)$$

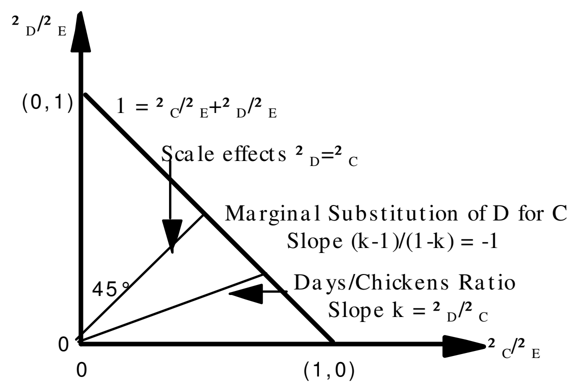
Applying equation (9B.5) to the previous example, if increases Δ_C and Δ_D were each 10% and $\alpha=\beta=1$, then the increase in output Δ_E would be 20% instead of 21% evaluated from equation (9B.4). This discrepancy can be explained by the approximation that $\Delta_C\Delta_D$ is negligible.

Equation (9B.5) can be plotted as a straight-line function of variables Δ_C/Δ_E and Δ_D/Δ_E as shown in Figure 9B.1. The intercepts of the production function line on the Δ_C/Δ_E and Δ_D/Δ_E axes are $(1/\alpha,0) = (1,0)$ and $(0,1/\beta) = (0,1)$ and the slope of the production function line is $-\alpha/\beta = -1$ respectively. Scale-effect experiments require $\Delta_C/\Delta_E = \Delta_D/\Delta_E$ which occurs on a 45° line that intersects the production-function line at $(1/(\alpha+\beta),1/(\alpha+\beta)) = (1/2,1/2)$. The *average productivity of chickens or days* is defined in scale-effect experiments as $\Delta_E/\Delta_C = \Delta_E/\Delta_D = \alpha+\beta = 2$ eggs per chicken or day.

The *days/chickens ratio* is defined on the line $(\Delta_D/\Delta_E) = k(\Delta_C/\Delta_E)$ with a slope $\Delta_D/\Delta_C = k > 0$ which intersects the production-function line at $(1/(\alpha+k\beta),k/(\alpha+k\beta)) = (1/(1+k),k/(1+k))$. The *marginal substitution of days for chickens* is defined as the slope $-\alpha/\beta = -1$ of the production-function line between the lines of average productivity of chickens or days and days/chickens ratio line as verified in equation (9B.6) below.

$$\frac{\text{Marginal substitution}}{\text{of days for chickens}} \equiv \frac{k(\alpha+k\beta)^{-1}-(\alpha+\beta)^{-1}}{(\alpha+k\beta)^{-1}-(\alpha+\beta)^{-1}} = \frac{\alpha(k-1)}{\beta(1-k)} = -\frac{\alpha}{\beta} = -1 \quad ....(9B.6)$$

Figure 9B.1 - Scale and Substitution Effects of the Chicken-and-Egg Production Function



To fix ideas on applying equation (9B.5), let us consider how a 20% increase in egg output could be obtained from increased inputs of chickens and days. As discussed under scale-effect experiments, a 20% increase in egg output could be achieved along the 45° line where $\Delta_C = \Delta_D$. Upon substituting $\Delta_C = \Delta_D$, $\alpha=\beta=1$ and $\Delta_E = 20\%$ in equation (9B.5), we get $\Delta_C = \Delta_D = 10\%$. Similarly, if the days/chickens ratio $k = \Delta_D/\Delta_C = 3$ and $\alpha=\beta=1$, then $\Delta_E = 20\%$ would be obtained with $\Delta_C = 5\%$ and $\Delta_D = 15\%$.

The chicken-and-egg production function is a purely technological description which provides rational measurements of cause-and-effect relationships. It enables scale and substitution effect experiments to be identified separately without succumbing to the proverbial question, "Which came first, the chicken or the egg?". In particular, the chicken-and-egg production function is blind to whoever owns the chickens or the eggs,

As the line from the origin intersects the line graph of equation (9.49) at a point whose coordinates lie closer to the Δ_C/Δ_E axis, it takes an increasing proportion of chickens to offset the loss of egg output caused by a decreased proportion of days. This result exemplifies a qualitative principle in physical chemistry known as the theorem of Le Chatelier. According to this theorem, if an external force is applied to a system in equilibrium, the system will shift in a direction to minimize the effect of the applied force. The same qualitative principle is known in economics as the law of variable proportions or diminishing returns. Thus, as chickens increase in intensity relative to days, a larger increase in the proportion of chickens are needed to substitute for the loss of egg output caused by a decrease in the proportion of days.

Appendix 9C - Parameter Estimates of the Cobb-Douglas Production Function

The parameters of the Cobb-Douglas production function was originally fitted by log-linear regression analysis to annual indices of industrial output P, labor input L and capital input C of United States from 1899 to 1922. The indices of P, L, and C were all unitized at the base year of 1899. Constant returns to scale, $\alpha+\beta = 1$, was assumed during the 24-year study period making the unknown parameters A, α and $1-\alpha$. The results were as follows:

$$P = 1.01 * L^{0.75} * C^{0.25} \qquad \text{....(9.5.2)}$$

Constant returns to scale was assumed because variations of P, L and C indices are closely correlated which made estimates of parameters A, α and β by log-linear regression analysis unreliable. Therefore, in order to improve the reliability of the estimates, unknown parameter β was replaced by $1-\alpha$ by assuming $\alpha+\beta = 1$. The parameter A was then estimated as a least-squares residual representing technological change.

An alternative method of estimating parameters A, α and β of the Cobb-Douglas production function assumes the variables P, L and C are each exponential functions of time whose straight-line slopes m_P, m_L and m_C on semi-logarithmic graph paper are independent of their given base-period quantities P_0, L_0 and C_0 at the zero point of time. Consequently, the parameters m_P, m_L and m_C can be determined independently by log-linear regression analysis from the following three equations.

$$P = P_0 e^{m_P T}; \qquad \ln(P/P_0) = m_P T \qquad \text{....(9C.1)}$$

$$L = L_0 e^{m_L T}; \qquad \ln(L/L_0) = m_L T \qquad \text{....(9C.2)}$$

$$C = C_0 e^{m_C T}; \qquad \ln(C/C_0) = m_C T \qquad \text{....(9C.3)}$$

These three equations can be combined into one production function by eliminating the time parameter T as follows. Let x and y denote undetermined multipliers of equations (9C.2) and (9C.3), and let $-(xm_L+ym_C)/m_P$ denote the undetermined multiplier of equation (9C.1).

$$-[(xm_L+ym_C)/m_P]\ln(P/P_0) = -(xm_L+ym_C)T \qquad \text{....(9C.4)}$$

$$x\ln(L/L_0) = xm_L T \qquad \text{....(9C.5)}$$

$$y\ln(C/C_0) = ym_C T \qquad \text{....(9C.6)}$$

Upon adding equations (9C.4), (9C.5) and (9C.6) together, the right hand sides of the equations cancel resulting in equation (9C.7) which is exponentiated to give equation (9C.8).

$$-[(xm_L+ym_C)/m_P]\ln(P/P_0) + x\ln(L/L_0) + y\ln(C/C_0) = 0 \qquad \text{....(9C.7)}$$

$$P = P_0(L/L_0)^{xm_P/(xm_L+ym_C)}(C/C_0)^{ym_P/(xm_L+ym_C)} \qquad \text{....(9C.8)}$$

The ratio of the L and C exponents in equation (9C.8) is x/y which represents the marginal substitution of capital for labor. Equation (9C.8) has two undetermined parameters x and y which are determinable by log-linear regression analysis.

Appendix 9D - Euler's Theorem on Homogeneous Functions

When describing increasing, decreasing, or constant returns to scale, all input factors are permitted to change in the same proportion, and the corresponding change in output is observed. In this regard, the definition of homogeneous production functions is most useful. If P = f(L,C) and $\lambda^n P = f(\lambda L,\lambda C)$, where $\lambda>0$ is any proportional increase ($\lambda>1$) or decrease ($0>\lambda<1$) of all input factors, then P = f(L,C) is defined as a homogeneous production function of degree n. A fundamental property of homogeneous functions is given in Euler's theorem.

<u>Theorem</u>: If $P = f(L,C) = \lambda^{-n}f(\lambda L,\lambda C)$ for all $\lambda>0$, then $nP = L(\partial P/\partial L)_C + C(\partial P/\partial C)_L$ and conversely.

<u>Proof of sufficiency</u>: We need to prove if $P = f(L,C) = \lambda^{-n}f(\lambda L,\lambda C)$ is a homogeneous function of degree n, then partial differential equation $nP = L(\partial P/\partial L)_C + C(\partial P/\partial C)_L$ is always satisfied. Let us define $F(L,C,\lambda) \equiv \lambda^{-n}f(\lambda L,\lambda C) = f(L,C)$ and differentiate F with respect to λ.

$$\frac{\partial F}{\partial \lambda} = -n\lambda^{-n-1}f(\lambda L,\lambda C) + \lambda^{-n}\left[\frac{\partial f(\lambda L,\lambda C)}{\partial \lambda L}.\frac{\partial \lambda L}{\partial \lambda} + \frac{\partial f(\lambda L,\lambda C)}{\partial \lambda C}.\frac{\partial \lambda C}{\partial \lambda}\right] \text{for all } \lambda > 0$$

$$\frac{\partial F}{\partial \lambda} = -n\lambda^{-n-1}F(L,C,\lambda) + \lambda^{-n}\left[\frac{\partial f(\lambda L,\lambda C)}{\partial \lambda L}L + \frac{\partial f(\lambda L,\lambda C)}{\partial \lambda C}C\right] \text{for all } \lambda > 0$$

Since $f(L,C) = F(L,C,\lambda)$, it follows that $\frac{\partial F}{\partial \lambda} = 0$. Also, if $\lambda = 1$, then

$$\frac{\partial F}{\partial \lambda} = 0 = -nf(L,C) + \frac{\partial f(L,C)}{\partial L}L + \frac{\partial f(L,C)}{\partial C}C$$

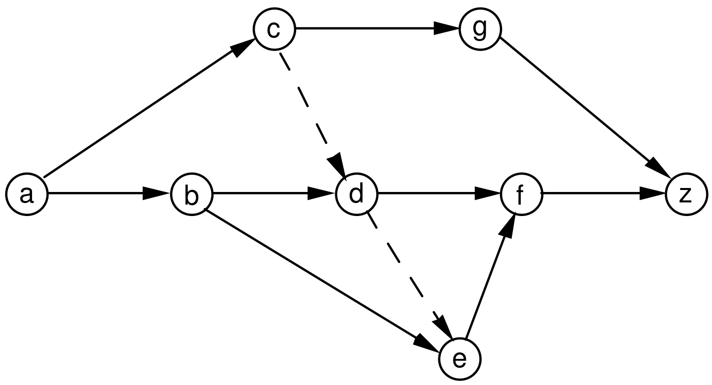
$$nf(L,C) = nP = \frac{\partial f(L,C)}{\partial L}L + \frac{\partial f(L,C)}{\partial C}C$$

<u>Proof of necessity</u>: We need to prove that partial differential equation $nP = L(\partial P/\partial L)_C + C(\partial P/\partial C)_L$ implies production function P = f(L,C) is homogeneous of degree n (i.e., $P = f(L,C) = \lambda^{-n}f(\lambda L,\lambda C) = F(L,C,\lambda)$). The equation $nP = L(\partial P/\partial L)_C + C(\partial P/\partial C)_L$ implies that $\partial F/\partial \lambda = 0$ and that $F(L,C,\lambda)$ is only a function of L and C. This function can be found by putting $\lambda = 1$ in the definition of $F(L,C,\lambda) \equiv \lambda^{-n}f(\lambda L,\lambda C)$ which yields $F(L,C,1) \equiv f(L,C) = P$.

Euler's theorem on homogeneous functions of degree n has many important economic implications because empirical production functions often have homogeneous formulations. Dimensional analysis indicates the production function should be a product of the powers of the input quantities such as the Cobb-Douglas production function. If P = f(L,C) is homogeneous of degree n and V is the price per unit of P, then nPV is the total output value distributed to L and C per unit of time. If W is the wage rate per unit of L and I is the interest rate per unit of C, then $nPV = LW + CI = LV(\partial P/\partial L)_C + CV(\partial P/\partial C)_L$. Consequently, when $W = V(\partial P/\partial L)_C$, then $I = V(\partial P/\partial C)_L$. However, when $W < V(\partial P/\partial L)_C$, then $I > V(\partial P/\partial C)_L$, and vice versa.

Chapter Nine - Exercises

9-1a Complete the boundary time table below for the following network. Identify the critical path(s) by shading in the appropriate arrows of the diagram.



Boundary Time Table: Forward-Float Calculations Backward-Float Calculations

Activity (i,j)	Duration T(i,j)	Earliest Starting ES(i)	Early Finish EF(i,j)	Forward Float FF(i,j)	Latest Finishing LF(j)	Late Start LS(i,j)	Backward Float BF(i,j)	Total Float TF(i,j)
a,b	5							
a,c	3							
b,d	3							
b,e	3							
c,d	0							
c,g	6							
d,e	0							
d,f	5							
e,f	9							
f,z	3							
g,z	10							

9-1b Complete the following real-time forward-float diagram (see Table 9.12).
[Real-Time Forward-Float FF(i,j) … +]: ES(i) + T(i,j) ‘ EF(i,j) + FF(i,j) ‘ ES(j)

```
Time Scale:  0 . . . . 5 . . . .10 . . . .15 . . . .20 . . . .25
(acgz)       . . . . . . . . . . . . . . . . . . . . . . . . . .
(acdfz)      . . . . . . . . . . . . . . . . . . . . . . . . . .
(abdefz)     . . . . . . . . . . . . . . . . . . . . . . . . . .
(abefz)      . . . . . . . . . . . . . . . . . . . . . . . . . .
```

9-1c Complete the following real-time backward-float diagram (see Table 9.13).
[Real-Time Backward-Float BF(i,j) … -]: LF(i) ’ -BF(i,j) + LS(i,j) ’ -T(i,j) + LF(j)

```
Time Scale:  0 . . . . 5 . . . .10 . . . .15 . . . .20 . . . .25
(acgz)       . . . . . . . . . . . . . . . . . . . . . . . . . .
(acdfz)      . . . . . . . . . . . . . . . . . . . . . . . . . .
(abdefz)     . . . . . . . . . . . . . . . . . . . . . . . . . .
(abefz)      . . . . . . . . . . . . . . . . . . . . . . . . . .
```

9-2a The output of a brass mill is limited by the operating capacities of its extrusion press (275 hours per month) and annealing furnace (325 hours per month). In the short run, the problem is to manufacture a mix of products A, B, C and D which utilize existing facilities in the most profitable manner. The mill is small enough so that its output will hardly have any effect on the prices or the size of the market for those products. The mill has two types of cost, namely, incremental or marginal costs which are proportional to the volume of output, and fixed overhead costs which are independent of the volume of output. In order to earn an incremental profit contribution of $1,000 towards the payment for fixed overhead expenses from the production of A, B, C and D, the input requirements of extrusion press and annealing furnace times are given in the table below. Determine which products should be produced and the monthly contribution of each of these products to the maximum incremental profit per month of the brass mill from the operation of its extrusion press and annealing furnace.

9-2b Determine the shadow prices or implicit costs per hour of extrusion press and annealing furnace time.

9-2c Determine the implicit costs of producing $1,000 incremental profit contributions from products A, B, C and D in terms of their extrusion press and annealing furnace time requirements.

Input requirements (hours/$1,000 profit increment)	Product A	B	C	D	Input capacity (hours/month)
Extrusion press time	4.0	2.0	1.5	3.0	275
Annealing furnace time	0.0	1.5	3.0	1.0	325

Chapter Ten - Statistical Expectations

Section 10.1 - Conditions of Economic Risk and Uncertainty

Risk generally means the chance of loss. However, the chance of loss in economic ventures is invariably coupled to some chance of profit or gain. Therefore, economic risk is defined as ventures in which a given input could result in more than one distinctly outcomes involving either loss or gain. If the probabilities or relative frequencies of the outcomes resulting from a given input are either known or estimable with some measure of statistical reliability, then the problems are called decision-making under *conditions of economic risk*. But if the probabilities or relative frequencies of the outcomes resulting from a given input are unknown due to a lack of empirical observations or an applicable probability model, then the problems are called decision-making under *conditions of economic uncertainty*.

Structured methods of handling problems of decision-making under conditions of economic risk began to make real progress about 300 years ago. Foundations of empirical methods to uncover statistical regularities in large bodies of data were developed in England by John Gaunt (1620-1674) and Edmund Halley (1656-1742) in connection with life insurance annuities. In a different approach designed to compute likely outcomes of games of chance, Pierre de Fermat (1601-1665) and Blaise Pascal (1623-1662) devised a system of probability calculus in France. The practice of mathematical statistics at present combines empirical approaches of Graunt and Halley and theoretical methods of Fermat and Pascal.

Known probability distributions of possible outcomes are often reduced to single measures of central tendency such as their arithmetic averages or other statistically expected values. This method of analysis is generally referred to as the expectation principle. When probability distributions are summarized in the form of statistically expected values, the problems of decision-making under conditions of economic risk become the same as the deterministic problems previously handled under conditions of economic certainty. Section 10.2 outlines methods of statistical averaging that are commonly used in economic decision-making.

In order to organize empirical observations into probability distributions, it is necessary for the outcomes from a given input to occur with statistical regularity so that their relative frequencies can be estimated in an objectively verifiable manner. For this purpose, it is useful to derive empirical and theoretical frequency distributions to serve as norms of comparison. Section 10.3 reviews rules of adding probabilities of single events in order to summarize laws of chance into empirical probability distributions that are useful for comparing expected values of perfect and imperfect information.

Rules of multiplying probabilities are outlined in Section 10.4 in order to define conditional probabilities and statistical independence. Problems of diagnosing the economic risks of accepting or rejecting statistical inferences of Types I and II successes and failures are then presented in a number of practical cases. Life insurance annuities are discussed in Section 10.5 relative to the Commissioners 1980 Standard Ordinary Mortality Table.

Problems of decision-making under conditions of economic uncertainty are largely outside the scope of this book. Such problems pose a very difficult and important area of economic decision-making, and they are not noted for standardized practical applications. Many problems are handled by means of *subjective probabilities* based on the judgement and experience of knowledgeable persons. Subjective probabilities are then treated in the same manner as objective probabilities in decision-making under conditions of economic risk.

Section 10.2 - Statistical Averaging

The problem of condensing a large array of numbers into a single figure which is representative of the entire array is usually solved by taking the average or *arithmetic mean* of the array. To calculate the arithmetic mean, we add all the figures in the array and divide the total by the number of figures in the array. However, the arithmetic mean is only one of several different ways of replacing an array of numbers by a single representative figure. Although the arithmetic mean is the most frequently used method of averaging, it is by no means the most representative and may even be misleading in some circumstances.

For example, consider a hamburger manufacturer who pays equal amounts of money to purchase two grades of beef, one of which costs p_1 = $1 per pound and the other costs p_2 = $2 per pound. What is the average price per pound? This question cannot be answered until the average price per pound is defined properly. To say the average price is $(p_1+p_2)/2$ = $1.50 per pound is not necessarily incorrect, but it is an average which may not explain the total cost of the hamburgers correctly. For example, suppose $100 is spent on each grade of beef. When 150 total pounds of beef is multiplied by the arithmetic mean price of $1.50 per pound, we obtain a total cost of $225 instead of the actual cost of $200.

Let us now obtain a proper solution to the hamburger problem. Let A represent the amount of money spent on each grade of beef. Then the number of pounds of beef that are bought is $A/p_1 + A/p_2$ for a total cost of 2A. The average price paid per pound of beef is

$$\text{Average price/lb} = \frac{2A}{Ap_1^{-1} + Ap_2^{-1}} = \frac{2}{p_1^{-1} + p_2^{-1}} = \left[\frac{p_1^{-1} + p_2^{-1}}{2}\right]^{-1} \quad \text{....(10.2.1)}$$

The above average is known as the harmonic mean and it represents the reciprocal of the average reciprocal of the prices. Substituting p_1 = $1 and p_2 = $2 in equation (10.2.1), we get $1.33 for the average price per pound and verify 150 * $1.33 = $200 for the total cost. It should be noted equation (10.2.1) is independent of amount A spent on each grade of beef. More generally, if $n = n_1 + n_2 + ... + n_m$ dollars were spent on m grades of beef, we get

$$\text{Average price/lb} = \{[n_1(p_1)^{-1} + n_2(p_2)^{-1} + ... + n_m(p_m)^{-1}]/n\}^{-1} \quad \text{...(10.2.2)}$$

Equation (10.2.2) is particular example of a general expression for the exponential mean c_K of order K shown in equation (10.2.3) for averaging m prices per unit on which $n = n_1 + n_2 + ... + n_m$ dollars were spent. If K = 1, c_1 is the *arithmetic mean*. If K = -1, c_{-1} is the *harmonic mean*. If K = 2, c_2 is the *root-mean-square*. In the limit as $K \to 0$, c_0 is the *geometric mean* as shown in equation (10.2.4). It can be shown that $c_K \geq c_J$ for all K > J, and the equality sign of $c_K \geq c_J$ holds if, and only if, $(p_1) = (p_2) = ... = (p_m)$.

$$\text{Exponential mean } c_K = \{[n_1(p_1)^K + n_2(p_2)^K + ... + n_m(p_m)^K]/n\}^{1/K} \quad \text{...(10.2.3)}$$

$c_K = \{[n_1 e^{K \ln p_1} + n_2 e^{K \ln p_2} + \ldots + n_m e^{K \ln p_m}]/n\}^{1/K}$. Expand exponentials for $K \approx 0$.

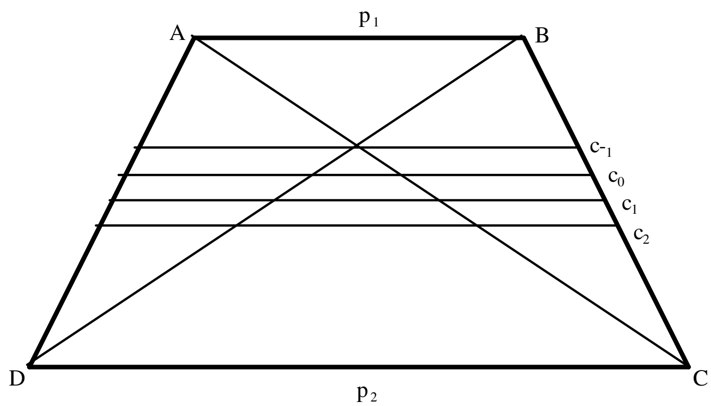
$$c_K = \{[n_1(1 + K \ln p_1) + n_2(1 + K \ln p_2) + \ldots + n_m(1 + K \ln p_m)]/n\}^{1/K}$$

$c_K = \{1 + K(\ln p_1^{n_1/n} + \ln p_2^{n_2/n} + \ldots + \ln p_m^{n_m/n})\}^{1/K}$. In the limit as $K \to 0$, we get

$$\text{Geometric mean } c_0 = \left[(p_1)^{n_1} \times (p_2)^{n_2} \times \ldots \times (p_m)^{n_m}\right]^{1/n} \quad \text{...(10.2.4)}$$

In the special case where m = 2 and $n_1 = n_2 = n/2$, the c_1c_{-1} product of arithmetic and harmonic means equals the $(c_0)^2$ square of the geometric mean. The relationship between c_2, c_1, c_0 and c_{-1} for m = 2 and $n_1 = n_2 = n/2$ is shown geometrically in Figure 10.2.1 by the line segments which are parallel to and part way between the bases of trapezoid ABCD. The root-mean square c_2 of the bases p_1 and p_2 of ABCD is represented by the line segment which divides ABCD into two trapezoids of equal area. The arithmetic mean c_1 of the bases is represented by the line segment halfway between the bases. The geometric mean c_0 of the bases is represented by the line segment which divides ABCD into two similar trapezoids. The harmonic mean c_{-1} of the bases is represented by the line segment passing through the intersection of the diagonals AC and BD. (proofs are given in "An Introduction to Inequalities" by Edwin Beckenbach and Richard Bellman, Random House, 1961)

Figure 10.2.1 - Geometric illustration of means c_2, c_1, c_0 and c_{-1} for m = 2 and $n_1 = n_2 = n/2$.



The tendency to use the arithmetic mean in all cases of averaging stems from a desire to be unbiased and treat all data equally. However, in many cases, every number in the data should not be treated on a par with every other number. In the case of the hamburger problem, there were twice as many pounds of the lower grade of beef as there were of the better grade. Consequently, the price p_1 of the lower grade of beef should influence the average price more than the price p_2 of the better grade of beef. This is taken into account by the harmonic mean.

It is necessary to see how data are generated before using an average. For example, a city with stable birth and death rates and no migration grew from 250,000 in 1940 to 490,000 persons in 1950. Let us determine the annual growth rate and population in 1945 (see *Averages and Scatter* by M. J. Moroney, Volume 3, World of Mathematics, J. R. Newman editor, Simon & Schuster, 1956). The arithmetic mean implies the annual growth rate is 24,000 persons per year and the 1945 population is 370,000. But the 24,000 annual growth rate is 9.6%/year of the 1940 population and 4.9%/year of the 1950 population.

However, the geometric mean implies the 1945 population is $[(250{,}000)\cdot(490{,}000)]^{1/2} = 500\cdot 700 = 350{,}000$ instead of 370,000 in 1945, and the average growth rate is

$(490{,}000/250{,}000)^{1/10}$=1.0696 or 6.96% per year of the population from 1940 to 1950

The 6.96% annual growth rate was 0.0696•250,000 = 17,400 persons in 1940 and 0.0696•490,000 = 34,104 persons in 1950. From a strict mathematical viewpoint, the arithmetic mean is no better or worse than the geometric mean. However, from the statement of the problem that the city has stable birth and death rates with no migration, the population will increase at a rate proportional to the current number of people in the city. Therefore, the arithmetic mean solves a different question than the one which should be answered by the geometric mean.

In accounting for inflation as discussed in Appendix 7C, the weighted-average design of the consumer's price index (CPI) is to represent the change in annual earning power which an average family needs for maintaining a constant annual purchasing power. For this purpose, it is necessary to invent an "average family" and a "standard basket of goods" which it purchases. Because the makeup of average families and their purchases change in the course of time, the averages underlying the CPI must be specially weighted to reflect such changes. Problems of weighting averages arise in all composite price indices such as those used in the formulation of the Cobb-Douglas production function in Section 9.5.

Problems of decision-making under conditions of economic risk are commonly handled by reducing probability distributions to measures of central tendency such as their averages or expected values. This method of analysis is generally referred to as the expectation principle. When probability distributions of outputs or inputs are reduced to expected values, the problems of decision-making under conditions of economic risk become the same as the deterministic problems previously handled under conditions of economic certainty. Hence, decision-making under conditions of economic risk does not need any new economic principles, but it does require the possible outcomes to occur with statistical regularity so that their probabilities can be estimated in an objectively verifiable manner.

The common practice of discounting future input costs and output revenues with high constant interest rates for the purpose of reducing economic risk and uncertainty is a misuse of the methods of statistical averaging. The probability of completing a project alternative with a given set of input costs differs greatly from the probability of realizing sales revenues from the output. The probability that market rates of interest will either rise or fall is much different than the probability that interest rates will be constant. Economic risks connected with input costs, output revenues and market rates of interest are essentially different, and they cannot be taken into account with high constant discount rates by statistical averaging.

Although high discount rates favor alternatives with short-term profits, they also limit the spectrum of available alternatives and increase risk by rejecting alternatives which provide long-term productivity gains. There is also room for questioning the use of high discount rates to account for uncertainty. Uncertainty stems primarily from a lack of understanding which can be resolved only by more research and development.

Problems of decision-making under conditions of economic uncertainty is a most difficult and important area of economic decision making which is not noted for standardized practical applications. Many problems are handled by means of subjective probabilities based on the judgement and experience of knowledgeable persons. The subjective probabilities are then treated as objective probabilities in economic decision-making under conditions of economic risk.

Section 10.3 - Statistical Summarization and Specification

The problem of statistical averaging concerns an attempt to boil down a number of possible outcomes into a single figure which represents the totality of possible outcomes. However, under ordinary conditions of economic risk, it is essential to distinguish between risks of profit and loss as two distinct possibilities. In dealing with economic risk, there is a basic tradeoff between limiting losses (downside risk) and obtaining profits (upside risk). In the final analysis, if one does not risk any loss, there is little chance of making any profit. This does not mean that upside and downside risk are always equal or correlated. To the contrary, not only is it important to distinguish between upside and downside risk, but it is also important to identify and measure their magnitudes independently.

The totality of possible outcomes is summarized either by empirical or theoretical probability distributions. In this connection, it is important to understand what the word 'probability' means in different contexts, and how probabilities of single events are added and multiplied to form different probability distributions. The probability of a particular event A is defined as the fraction of the total number of relevant events which correspond to event A. Thus, letting P(A) denote the probability of event A, we have

$$P(A) = \frac{\text{number of ways A occurs}}{\text{total number of relevant outcomes}} \qquad \text{....(10.3.1)}$$

If a company sells 50 items, 2 of which are defective, then the probability of getting a defective item is 2/50 = 0.04 or 4%, assuming all 50 items are sold to the customer.

The definition above is the relative-frequency meaning of probability. Probability can also be defined as a degree of belief of important human observers. Such probabilities are called *subjective* or *personalistic* probabilities. In essence, subjective probabilities differ from the relative frequency meaning of probability in that it relaxes the need for measurement or counting. In order to estimate subjective probabilities, one only needs to pose the question to a keen observer of facts, or bracket in the opinions of top management. The ambiguity of subjective probabilities can be eliminated by asking one person only once for the probability estimate. The difficulty with subjective probabilities is that even if one stumbles on a true probability estimate, its truth remains uncertain without obtaining other lines of evidence to substantiate the result.

In the relative frequency definition of probability, the ways that event A occurs must be compared to the total number of relevant outcomes. It is important to specify which outcomes are relevant to the occurrence of A in order to make its probability meaningful. If we sample only from a small region of possible outcomes, the probabilities of particular events can be highly misleading. In effect, all probabilities are conditional probabilities which spell out the outcomes that are relevant to the favorable occurrence of each event.

The total number of relevant outcomes may be described as a sample space S of an experiment. Each outcome of the experiment corresponds to exactly one sample point, and all points of the sample space are equally likely. The sample points may be classified into *events* or *subsets* A, B, C, ... of sample space S. Each sample point in S may belong to one or more of these subsets. Probability P(A) is defined as the fraction of sample points in S that belong to A. The basic probability axioms are $P(A) \geq 0$ and $P(S) = 1$.

Addition Theorem: $P(A+B) = P(A) + P(B) - P(A \cdot B)$ ...(10.3.2)

Probability P(A+B) is the fraction of sample points in S belonging to the (A+B) *union* or *outline* of sample points A or B or both. Probability P(A•B) is the fraction of sample points in S belonging to the (A•B) *intersection* of sample points A and B in Figure 10.3.1. Sample points ~(A+B) in S are outside the (A+B) union. DeMorgan's law equates sample points ~(A+B) to the (~A•~B) intersection of sample points in S belonging to neither A nor B.

Figure 10.3.1 - Venn diagram of the union (A+B) and intersection (A•B) of subsets A and B.

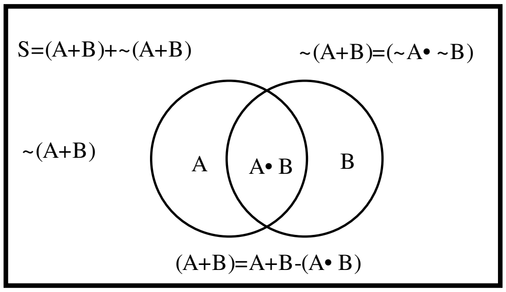


Figure 10.3.1 shows the sample points in A cannot be simply added to those in B to get the number of sample points in the union (A+B) because sample points of A and B in the intersection (A•B) would be double-counted. Therefore, we need to subtract the sample points in (A•B) from the sum of sample points in (A) and (B) to obtain the number of sample points in the union (A+B). When the sample points in the sets (A+B) = (A) + (B) - (A•B) are divided by the number of sample points in S, we obtain the addition theorem of probabilities in equation (10.3.1). If the intersection (A•B) is empty, then A and B are said to be *mutually exclusive*. If A is contained in B, then A is said to *imply* B.

The inclusion-exclusion principle underlies the addition theorem. Thus,

$$P(A+B+C) = P(A) + P(B) + P(C) - P(A \bullet B) - P(B \bullet C) - P(A \bullet C) + P(A \bullet B \bullet C) \quad ...(10.3.3)$$

The first three terms P(A) + P(B) + P(C) include all sample points of A, B and C regardless of double-counting sample points belonging to (A•B), (B•C) and (A•C), and triple-counting sample points belonging to (A•B•C). Then we exclude the double-counting of sample points by subtracting P(A•B), P(B•C) and P(A•C) from P(A) + P(B) + P(C). But in the proces of excluding double-counted sample points, we have also subtracted off triply-counted sample points belonging to (A•B•C). Therefore, we must add back P(A•B•C) in order to get P(A+B+C).

The inclusion-exclusion principle provides a system for recording observations which exhausts all logical possibilities. For example, suppose a survey was made of personal drinking preferences for bourbon (B), rum (R), and vodka (V). It is possible that some persons would like all three, some persons might like none, and some will like rum and vodka but not bourbon, etc. The results of the survey are as follows:

(B) = 762 like bourbon	(B•R) = 560 like bourbon and rum
(R) = 720 like rum	(R•V) = 369 like rum and vodka
(V) = 615 like vodka	(B•V) = 465 like bourbon and vodka
(B•R•V) = 297 like all three	–(B+R+V) = (–B•–R•–V) = 200 don't like any

The problem is to determine how many persons were surveyed (S), and how many persons who like vodka do not like bourbon and rum (V•–B•–R). The solutions below show that 1,200 persons were surveyed and 78 persons who like vodka do not like bourbon and rum.

$$(S) = (B+R+V) + –(B+R+V) = (B)+(R)+(V) - (B•R) - (R•V) - (B•V)+ (B•R•V) + –(B+R+V)$$
$$(S) = 762 + 720 + 615 - 560 - 369 - 465 + 297 + 200 = 1,200$$

$$(V•–B•–R) = (V) - (R•V) - (B•V) + (B•R•V) = 615 - 369 - 465 + 297 = 78$$

Let us now focus on sample spaces of *mutually exclusive* and *collectively exhaustive* subsets (i.e., non-intersecting subsets whose union exhausts all sample points). Such experiments enable us to define *random variables* and *probability functions*. A random variable is a discrete or continuous variable whose value is determined by the outcome of an experiment. A probability function of a random variable is a function whose value equals the probability a random variable will have one of the possible values in its range.

Let random variable X have discrete values $x_1 < x_2 < ... < x_m$. Suppose a sample space of n sample points has n_1 values of x_1, n_2 values of x_2, and n_m values of x_m. The probability function $f\{X=x_1\} = n_1/n$, $f\{X=x_2\} = n_2/n$, and $f\{X=x_m\} = n_m/n$. Since random variable X is always implied in a probability function, the notation $f\{x_1\} = n_1/n$, $f\{x_2\} = n_2/n$, and $f\{x_m\} = n_m/n$ is preferred. In all such experiments, $f\{x_K\} \geq 0$ for all K and $\sum f(x_K) = 1$ summed over all K.

The expectation of X, denoted by E(X) or μ, is defined in equation (10.3.4). E(X) or μ is also called the mean value or arithmetic mean of random variable X. The expectation of $(X-\mu)^2$, denoted by $E(X-\mu)^2$ or σ^2, is defined in equation (10.3.5). $E(X-\mu)^2$ or σ^2 is also called the *variance* of X, and its square root σ is called the *standard deviation* of X.

$$E(X) \text{ or } \mu \; ... \; \sum x_K f(x_K) \quad \text{summed over all K} \qquad ....(10.3.4)$$

$$E(X-\mu)^2 \text{ or } \sigma^2 \equiv \sum (x_K-\mu)^2 f(x_K) = \sum x_K^2 f(x_K) - \mu^2 = E(X^2) - \mu^2 \qquad ....(10.3.5)$$

If X is a continuous random variable from $x = -\infty$ to $x = \infty$, then g(x)dx is the probability that X will lie between x and x+dx. Probability density function g(x) also requires $g(x) \geq 0$ for all x, and $\int g(x)dx = 1$ integrated over all x. Similarly, the expectations of X and $(X-\mu)^2$ are given in equations (10.3.4’) and (10.3.5’). In mechanical analogs, μ is the center of gravity of the probability distribution, and σ corresponds to the radius of gyration of the probability distribution about its center of gravity as the axis.

$$E[X] \text{ or } \mu \; ... \; \int xg(x)dx \quad \text{for } -\infty < x < \infty \qquad ....(10.3.4')$$

$$E(X-\mu)^2 \text{ or } \sigma^2 \; ... \; \int (x-\mu)^2 g(x)dx = E(X^2)-\mu^2 \quad \text{for } -\infty < x < \infty \qquad ....(10.3.5')$$

The demand for a product may be a random variable which is known in the form of an empirical probability distribution. For example, a vendor sells programs at football games for $0.50 which were bought before the game for $0.15 per copy. Unsold programs after the game are worth $0.02 per copy. If the vendor buys too few programs before the game, all programs may be sold but potential customers would be turned away. If the vendor buys too many programs, all customers may be sold a program but unsold programs could result in substantial losses. The problem is to determine how many programs should be bought before the game in order to maximize the expected value of the vendor's profit.

Fermat's method of solving this problem calculates the expected value of profit for specific amounts of programs that the vendor could buy before the game. Pascal's method uses incremental changes in the profit expectation to obtain an answer. Both Fermat's and Pascal's methods of solution require a known probability distribution of program sales at a game which can be derived from records of 40 previous football games:

S = number of programs sold at a game.
s = S/500 = program lots of 500 sold per game.
f(s) = probability of s sales per game, f(s) = 0 for $0 \le s < 5$ and $9 < s < \infty$.
F(s) = cumulative probability of s sales or less per game (i.e., f(0) + f(1) + ... +f(s)).
F(s) = 0 for $0 \le s < 5$, F(s) = 1 for $9 \le s < \infty$.

S sales per game	s sales per game	No. of games	f(s)	F(s)
2500	5	4	0.10	0.10
3000	6	6	0.15	0.25
3500	7	10	0.25	0.50
4000	8	16	0.40	0.90
4500	9	4	0.10	1.00
		Total=40	Total=1.00	

In general, let m = market price of a program lot sold at a game, c = cost price of a program lot before a game, and r = return price of a program lot left unsold after a game. For the problem on hand, m = 500*0.50 = \$250, c = 500*0.15 = \$75, and r = 500*0.02 = \$10. Also, (m-c) = \$175 is the profit on each program lot sold at a game, and (r-c) = -\$65 is the loss on each program lot left unsold after a game.

Let E{Z(b)} be the expected-value profit of buying b program lots before the game. Suppose b = 7. Since f(s) = 0 for $0 \le s < 5$ and $9 < s < \infty$, the only possibilities for f(s) > 0 are {s = 5 sold, b-s = 2 unsold}, {s = 6 sold, b-s = 1 unsold}, and {s = 7 sold, b-s = 0 unsold}. According to Fermat's method, the expectation E{Z(7)} is

$$E\{Z(7)\} = \{5*175-2*65\}f(5) + \{6*175-1*65\}f(6) + 7*175[f(7)+f(8)+f(9)] = \$1,141.00$$

In general, $$E\{Z(b)\} = \sum_{S=0}^{b-1} [\overset{\text{sold}}{s(m-c)} + \overset{\text{unsold}}{(b-s)(r-c)}]f(s) + \overset{\text{sold}}{b(m-c)} \sum_{S=b}^{\infty} f(s) \qquad ...(10.3.6)$$

Evaluating E{Z(b)} for b = 1, ... , 9 shows E{Z(8)}=\$1,187.50 is the maximum expected profit.
E{Z(1)}= 1*175=\$175; E{Z(2)}=2*175=\$350; E{Z(3)}=3*175=\$525; E{Z(4)}=4*175=\$700;
E{Z(5)}=5*175=\$875; E{Z(6)} = (5*175-1*65)f(5) + 6*175(1-0.10) = \$1,026
E{Z(7)} = (5*175-2*65)f(5) + (6*175-1*65)f(6) + 7*175(1-0.25) = \$1,141
E{Z(8)} = (5*175-3*65)f(5) + (6*175-2*65)f(6) + (7*175-1*65)f(7) + 8*175(1-0.50) = \$1,196
E{Z(9)} = (5*175-4*65)f(5) + (6*175-3*65)f(6) + (7•175-2*65)f(7) + (8*175-1*65)f(8)
+ 9*175(1-0.90) = \$1,155

If the vendor could predict from "Perfect Information" how many program lots would be sold at each game, then b = s at every game. The expected-value profits of such "Perfect Information" is called EVPI (i.e., Expected Value of Perfect Information) which is given by

$$EVPI = \sum_{S=0}^{\infty} s(m-c)f(s) = (m-c)\sum_{S=0}^{\infty} sf(s) = (m-c) \bullet E(s) \qquad ...(10.3.7)$$

Thus, EVPI = 175*[5*f(5) + 6*f(6) + 7*f(7) + 8*f(8) + 9*f(9)] = 175•[7.25] = \$1,268.75 . Therefore, the vendor should not spend more than \$1,268.75 - \$1,196 = \$72.75 for perfect information to forecast accurately the number of program sales before every football game.

Pascal's method of solving this problem calculates the change of expected-value profit $\Delta E\{Z(b)\}$ when buying the bth lot before the game. The chance of selling the bth lot or more at the game is f(b) + f(b+1) + ... = 1- F(b-1). The chance of not selling the bth lot is the probability F(b-1) of b-1 sales or less at the game. Thus, the change of expected value of profit which results from buying the bth lot is given in equation (10.3.8) which is simplified in equation (10.3.9). The sum of $\Delta E\{Z(b)\}$ to obtain E{Z(b)} is given in equation (10.3.10). We may now evaluate $\Delta E\{Z(b)\}$ for $1 \le b \le 9$ with equation (10.3.9) on the left of Table 10.3.1 below, and evaluate E{Z(b)} for $1 \le b \le 9$ with equation (10.3.10) on the right of Table 10.3.1 to show the equivalence of the methods of Fermat and Pascal.

$$\Delta E\{Z(b)\} = \underset{\Delta(\text{expected profit})}{(m-c)\{1- F(b-1)\}} + \underset{\Delta(\text{expected loss})}{(r-c)F(b-1)} \quad \text{for } b \ge 1 \qquad ....(10.3.8)$$

$$\Delta E\{Z(b)\} = (m-c) - (m-r)F(b-1) \quad \text{for } b \ge 1 \qquad ....(10.3.9)$$

$$E\{Z(b)\} = b(m-c) - (m-r)\sum_{s=1}^{s=b} F(s-1) \quad \text{for } b \ge 1 \qquad ....(10.3.10)$$

Table 10.3.1 - $\Delta E\{Z(b)\} = (m-c) - (m-r)F(b-1)$ $\quad E\{Z(b)\} = b(m-c) - (m-r)\sum_{s=1}^{s=b} F(s-1)$

b=1:	ΔE{Z(1)} = 175 - 240•0.00 = $175;	E{Z(1)} = 175 - 240•0.00 = $175
b=2:	ΔE{Z(2)} = 175 - 240•0.00 = $175;	E{Z(2)} = 350 - 240•0.00 = $350
b=3:	ΔE{Z(3)} = 175 - 240•0.00 = $175;	E{Z(3)} = 525 - 240•0.00 = $525
b=4:	ΔE{Z(4)} = 175 - 240•0.00 = $175;	E{Z(4)} = 700 - 240•0.00 = $700
b=5:	ΔE{Z(5)} = 175 - 240•0.00 = $175;	E{Z(5)} = 875 - 240•0.00 = $875
b=6:	ΔE{Z(6)} = 175 - 240•0.10 = $151;	E{Z(6)} = 1050 - 240•0.10 = $1,026
b=7:	ΔE{Z(7)} = 175 - 240•0.25 = $115;	E{Z(7)} = 1225 - 240•0.35 = $1,141
b=8:	ΔE{Z(8)} = 175 - 240•0.50 = $55;	E{Z(8)} = 1400 - 240•0.85 = $1,196
b=9:	ΔE{Z(9)} = 175 - 240•0.90 = -$41;	E{Z(9)} = 1575 - 240•1.75 = $1,155

The maximum E{Z(b=8)} = $1196 occurs at the last value b=8 for which $\Delta E\{Z(b)\}$ is positive. This solution can also be derived from (10.3.9) by setting $\Delta E\{Z(b)\} = 0$ and solving for F(b-1).

$$\Delta E\{Z(b)\} = (m-c) - (m-r)\,F(b-1) = 0 \;\text{ or }\; F(b-1) = (m-c)/(m-r) \qquad ....(10.3.11)$$

Upon substituting m-c = 175 and m-r = 240 in (10.3.11), we find F(b-1) = 175/240 = 0.729 . The table of cumulative probabilities F(s) indicates $7 < b-1 < 8$, or that $8 < b < 9$. From this information, we again conclude the optimal buying policy is b = 8. Pascal's method for calculating EVPI is based on summing Δ(expected profit) in equation (10.3.8) over the entire range of the random variable as shown numerically below. Equation (10.3.12) then shows Pascal's method is equivalent to equation (10.3.7) previously derived by Fermat's method.

EVPI = $175º[5*0.10 + 6*0.15 + 7*0.25 + 8*0.40 + 9*0.10] = $1268.75 as before.

$$EVPI = \sum_{S=1}^{\infty}(m-c)\{1- F(s-1)\} = (m-c)\sum_{S=1}^{\infty}\sum_{k=S}^{\infty}f(k) = (m-c)\sum_{S=0}^{\infty}sf(s) = (m-c)\bullet E(s) \qquad ...(10.3.12)$$

The methods of Fermat and Pascal are readily extended from discrete to continuous random variables. In this regard, let probability density and cumulative distribution functions g(s) and G(s) replace the roles of f(s) and F(s) respectively. Fermat's method of determining E{Z(b)} and EVPI in equations (10.3.13) and (10.3.14) for continuous random variables are analogous to equations (10.3.6) and (10.3.7) for discrete random variables.

$$E\{Z(b)\} = \int_0^b \{s(m - c) + (b - s)(r - c)\}g(s)ds + b(m - c)\int_b^\infty g(s)ds \quad ....(10.3.13)$$

$$EVPI = \int_0^\infty s(m - c)g(s)ds = (m - c)\int_0^\infty sg(s)ds = (m\text{-}c)\cdot E(s) \quad ....(10.3.14)$$

For continuous random variables, Pascal's method of determining the change of expected profits with respect to an infinitesmal increase in the buying policy b is shown in equation (10.3.15) which is analogous to differentiating E{Z(b)} in (10.3.8) partially with respect to b. Equation (10.3.15) is simplified in equation (10.3.16) in a similar manner as was done in equation (10.3.9). Upon setting $(\partial/\partial b)E\{Z(b)\}$ equal to zero, we may solve for G(b) as shown in equation (10.3.16) which is analogous to the solution for F(b-1) in equation (10.3.11).

$$\frac{\partial}{\partial b} E\{Z(b)\} = (m\text{-}c)\{1\text{-} G(b)\} + (r\text{-}c)G(b) \quad ....(10.3.15)$$

$$\frac{\partial}{\partial b} E\{Z(b)\} = (m\text{-}c) - (m\text{-}r)G(b) = 0, \text{ or } G(b) = (m\text{-}c)/(m\text{-}r) \quad ....(10.3.16)$$

Pascal's method of determining E{Z(b)} and EVPI for continuous probability distributions can be carried out with formulas (10.3.17) and (10.3.18) below. Equations (10.3.13) and (10.3.14) are equivalent to equations (10.3.17) and (10.3.18) respectively, and it is only a matter of convenience in the presentation of the relative frequency data and their evaluation as to which set of equations should be used in calculating the results.

$$E\{Z(b)\} = \int_0^b [(m - c)\{1 - G(s)\} + (r - c)G(s)]ds = b(m - c) - (m - r)\int_0^b G(s)ds \quad ....(10.3.17)$$

$$EVPI = \int_0^\infty (m\text{-}c)\{1\text{-} G(s)\}ds = (m\text{-}c)\int_0^\infty \int_r^\infty g(r)drds = (m\text{-}c)\int_0^\infty sg(s)ds = (m\text{-}c)E(s) \quad ....(10.3.18)$$

It does not appear necessary to introduce new economic principles when solving problems under conditions of risk and uncertainty. The only requirement seems to be the estimation of probability distributions which could convert the problem into one which could be solved as any other problem under conditions of economic certainty. The example of the vendor who sells programs at football games is analogous to a more general economic problem of deciding how much of a perishable item should be stocked, when reordering during the season is impractical, and items left unsold at the end of the season suffer an appreciable loss in value. Because empirical probability distributions are often lacking in such situations, several exercises at the end of Chapter Ten introduce theoretical probability distributions which may be used to simulate special circumstances.

Section 10.4 - Statistical Inference and Diagnosis

In many situations, sample points of relevant outcomes can only be partially classified into subsets with identifiable characteristics. In order to take greater advantage of available data, the sample space can be reduced by means of the multiplication theorem:

Multiplication Theorem: $P(A{\bullet}B) = P(A){\bullet}P(B/A) = P(B){\bullet}P(A/B)$ ....(10.4.1)

Joint probability P(A•B) is the probability of an outcome from an experiment to be a sample point belonging to both A and B. Thus, P(A•B) is the fraction of sample space S in the *intersection* (A•B) of the two circles in Figure 10.3.1. *Conditional probability* P(B/A) is the probability of an outcome from an experiment to be a sample point of B when we already know the outcome is a sample point of A. Hence, P(B/A) is the fraction of sample points in subset A belonging to *intersection* (A•B). Sample space S is reduced to sample space A when determining P(B/A) to account for already knowing the outcome is from subset A.

Conditional probability P(A/B) is the probability of an outcome from an experiment to be a sample point of A when we already know the outcome is a sample point of B. Hence, P(A/B) is the fraction of sample points in subset B belonging to *intersection* (A•B). Sample space S is reduced to sample space B when determining P(A/B) to account for already knowing the outcome is from subset B. Thus, P(B/A) and P(A/B) are ratios of sample points in (A•B) to those in reduced sample spaces A and B as shown in equation (10.4.2).

$$P(B/A) \equiv \frac{P(A\cdot B)}{P(A)} \quad \text{and} \quad P(A/B) \equiv \frac{P(A\cdot B)}{P(B)} \qquad ....(10.4.2)$$

The multiplication theorem may be extended by the chain rule as follows:

$$P(A{\bullet}B{\bullet}C) = P(A{\bullet}B)\cdot P(C/A{\bullet}B) = P(B{\bullet}C)\cdot P(A/B{\bullet}C) = P(A{\bullet}C)\cdot P(B/A{\bullet}C) \qquad ....(10.4.3)$$

This extension of the multiplication theorem serves to define conditional probabilities P(C/A•B), P(A/B•C) and P(B/A•C) as shown in equation (10.4.4).

$$P(C/A{\bullet}B) \equiv \frac{P(A\cdot B\cdot C)}{P(A\cdot B)}\ ;\ P(A/B{\bullet}C) \equiv \frac{P(A\cdot B\cdot C)}{P(B\cdot C)}\ ;\ P(B/A{\bullet}C) \equiv \frac{P(A\cdot B\cdot C)}{P(A\cdot C)} \qquad ...(10.4.4)$$

In some experiments, the conditional probability P(B/A) equals the probability P(B). For example, P(R/V) = P(R) and P(V/R) = P(V) in the 1200 person survey of drinking preferences for bourbon (B), rum (R) and vodka (V) in Section 10.3 as shown below.

$$P(R/V)\ \ldots\ \frac{P(R\cdot V)}{P(V)} = \frac{369}{615} = 0.6000;\ \text{and}\ \ P(R)\ \ldots\ \frac{P(R)}{P(S)} = \frac{720}{1200} = 0.6000$$

$$P(V/R)\ \ldots\ \frac{P(V\cdot R)}{P(R)} = \frac{369}{720} = 0.5125;\ \text{and}\ \ P(V)\ \ldots\ \frac{P(V)}{P(S)} = \frac{615}{1200} = 0.5125$$

Since P(R/V)=P(R) and P(V/R)=P(V), the multiplication theorem becomes P(R•V) = P(R)•P(V) = P(V)•P(R). If measurements of P(R/V) and P(V/R) are unchanged when sample space S of all relevant events is reduced to sample spaces R or V, then R and V are defined to be *statistically independent*.

Statistically independent events always occur in pairs. However, it does not follow that other combinations of events in sample space S will also be statistically independent as shown by the following examples:

$$P(B/R) \ldots \frac{P(B \cdot R)}{P(R)} = \frac{560}{720} = 0.778 \text{ and } P(B) \ldots \frac{P(B)}{P(S)} = \frac{762}{1200} = 0.635$$

$$P(B/V) \ldots \frac{P(B \cdot V)}{P(V)} = \frac{465}{615} = 0.756 \text{ and } P(B) \ldots \frac{P(B)}{P(S)} = \frac{762}{1200} = 0.635$$

It is also wrong to interpret statistical independence as evidence of events not having anything to do with each other or of being mutually exclusive. Neither of these two interpretations is correct. Statistically independent events must contain at least one sample point in their intersection as compared to mutually exclusive events which, by definition, cannot have any sample points in their intersections. Moreover, statistically independent events must have something to do with each other because they have one or more sample points in common. If two events are statistically independent, it simply means that experimental evidence on the relative frequency of occurrence of one event would not change the estimate of the relative frequency of occurrence of the other event.

Event ~A (not A) contains all points of the sample space not in A and is called the *complement* of event A. Since event A and its complement ~A are mutually exclusive, their union is the whole sample space. As a consequence, P(~A) = 1 - P(A) and P(A) = 1 - P(~A). If A and B are statistically independent (i.e., P(A•B) = P(A)•P(B)), then ~A and B, A and ~B, and ~A and ~B are also statistically independent. Proof:

Figure 10.4.1 - Two-way Contingency Table of Statistically Independent Events

	B	~B	Row Sums
A	P(A)• P(B)	P(A• ~B)	P(A)
~A	P(~A• B)	P(~A• ~B)	P(~A)
Column Sums	P(B)	P(~B)	1

In Figure 10.4.1, we are given joint probabilibity P(A•B) = P(A)•P(B) and we need to prove joint probabilities P(~A•B), P(~A•~B) and P(A•~B) are also statistically independent. For this purpose, let us sum the first column, second row and second column as follows:

P(B) = P(~A•B)+P(A)•P(B); P(~A•B) = P(B) - P(A)•P(B) = [1-P(A)]•P(B) = P(~A)•P(B)
P(~A) = P(~A•~B)+P(A•~B); P(~A•~B) = P(~A)-P(~A)•P(B) = P(~A)[1-P(B)]• = P(~A)•P(~B)
P(~B) = P(~A•~B)+P(A•~B); P(A•~B) = P(~B)-P(~A)•P(~B) = [1-P(~A)]•P(~B) = P(A)•P(~B)

Similarly, it can be shown if events A and B are statistically dependent (i.e., P(A•B) ≠ P(A)•P(B)), then joint probabilities P(–A•B), P(A•–B) and P(–A•–B) are also statistically dependent (i.e., P(–A•B)≠P(–A)•P(B), P(A•–B)≠P(A)•P(–B) and P(–A•–B)≠P(–A)•P(–B)).

Statistical dependence occurs with most, if not all, diagnostic tools. A perfect diagnostic tool has only true positive and negative responses. If what you are looking for is present, a perfect diagnostic tool has only positive responses, and if the thing you are looking for isn't present, a perfect diagnostic tool has only negative responses. Due to design limitations, diagnostic tools in use and those that are being developed have to be calibrated and tested for accuracy. Virtually all practical diagnostic tools give some false positive and false negative responses.

In choosing a tool, true positive readings (i.e., Type I successes) must be weighed against true negative readings (i.e., Type II successes) in order to measure how accurately the tool discriminates between the presence and absence of the thing you seek. But it is often even more important to weigh false negative responses of a tool (i.e., Type I errors) against false positive responses (i.e., Type II errors) because of the practical consequences resulting from wrong diagnoses. The statistical and economic evaluation of successes and errors of diagnostic tools affects virtually every activity in science and industry. The reliability of diagnostic tools in either finding or avoiding what you are looking for can have far-reaching economic and practical consequences, even so far as matters of life and death. This presentation outlines the practical significance of statistical and economic decision theory in measuring the reliability of diagnostic tools.

First let us examine more closely the meaning of Type I and Type II successes and errors. Every diagnostic tool is designed to find the presence of an event, regardless of whether you want or don't want that event to be there. It is impossible to design a tool to look for the absence of an event without first defining the event you want to be absent. The statistical evaluation of a diagnostic tool deals with its physical reliability in detecting whether an event is or is not present (i.e., signal detection), regardless of the economic and practical consequences when the event does or does not occur. In order to bring out the distinction between the statistical and economic evaluation of diagnostic tools, let us use an example of looking for the presence of an event you don't want (Cancer), and compare it to looking for the presence of an event you do want (Oil and Gas Deposits).

Figure 10.4.2 - Two-way Contingency Table of Tool Response and Presence of an Event

	Presence of an Event: Positive	Presence of an Event: Negative	Row Sums
Tool Response: Positive	a True Positive Type I Success	b False Positive Type II Error	a + b
Tool Response: Negative	c False Negative Type I Error	d True Negative Type II Success	c + d
Column Sums	a + c	b + d	a+b+c+d = N

Primary Diagonal: True Positive and Negative, Type I and Type II Success
Secondary Diagonal: False Positive and Negative, Type II and Type I Error

CANCER - As 1st Preference, you want a tool to give TRUE NEGATIVE responses (Type II Successes) which assure you of not having cancer. If the response is not a true negative, then as 2nd Preference, you want the tool to give a FALSE POSITIVE response (a Type II Error) because you really don't want to have cancer. But a false positive tool response could result in costly, disabling and potentially dangerous cancer treatments even though you don't have the disease. As a 3rd Preference, you want the tool to give a TRUE POSITIVE response (a Type I Success) so that cancer treatments can be started before it is too late. As 4th Preference, the tool gives a FALSE NEGATIVE response (a Type I Error). Cancer is present, but by the time it becomes obvious, it is incurable. These preferences of tool responses in looking for cancer trace a counter clockwise pattern on the two-way contingency table, starting with a true negative response (Type II Success) and ending with a false negative response (Type I Error).

OIL & GAS DEPOSITS - As 1st Preference, you want the tool to give a TRUE POSITIVE response (a Type I Success) which provides you with a discovery of an oil or gas field. If the response is not a true positive, then as 2nd Preference, you want the tool to give a TRUE NEGATIVE response (a Type II Success) so that you won't drill a dry hole. As a 3rd Preference, you want the tool to give a FALSE POSITIVE response (a Type II Error) because the risk is limited to the cost of drilling a dry hole from which important geological information may be obtained. As 4th Preference, the tool gives a FALSE NEGATIVE response (a Type I Error). As a consequence, you won't drill, but the risk of not finding a major oil field could be a serious loss of opportunity. These preferences of tool responses in looking for oil trace the diagonals of the two-way contingency table, starting with a true positive response (Type I Success) of the primary diagonal and ending again with a false negative response (Type I Error) of the secondary diagonal.

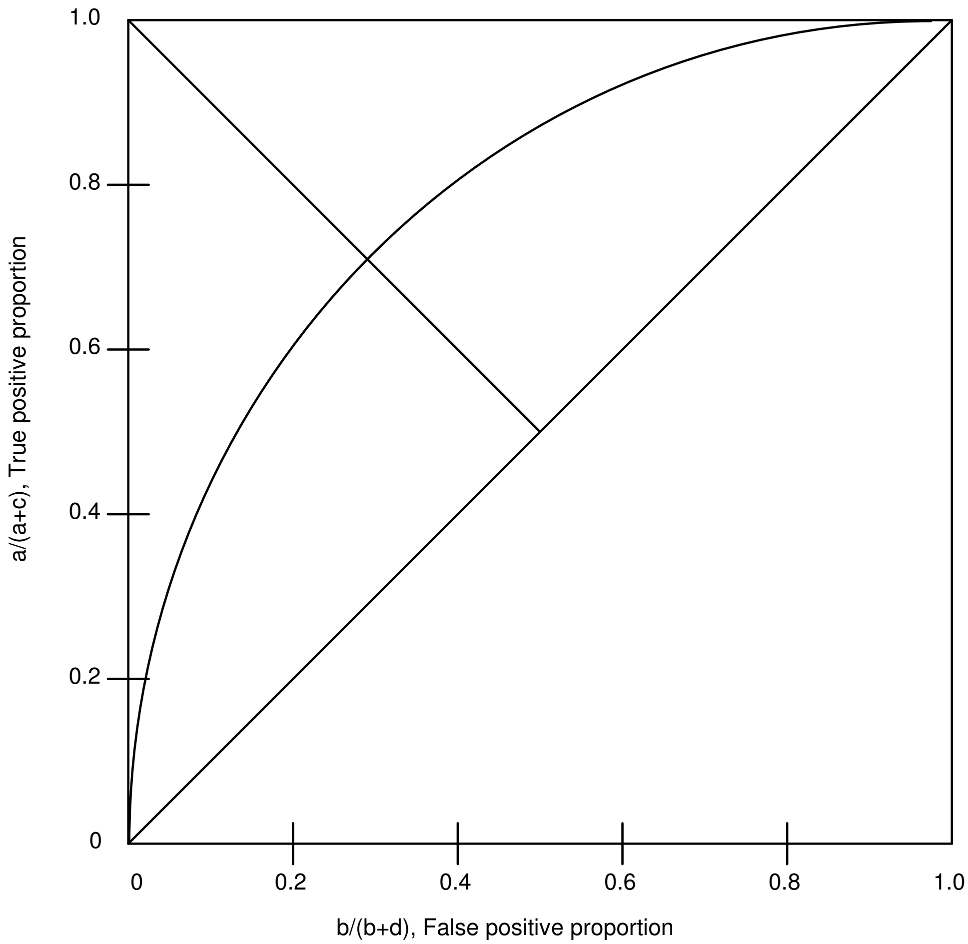
The comparison between looking for cancer and looking for oil and gas deposits shows that the preferences of outcomes strongly depend upon the particular event under consideration. For this reason, each tool used has to be statistically evaluated according to its physical characteristics before any economic or other practical considerations are taken into account.

The sample size for the presence of the event is a+c, and the sample size for the absence of the event is b+d. If we consider proportions rather than raw frequencies of the four outcomes, then the results in each cell of the two-way table will be independent of the absolute difference in the magnitudes of a+c and b+d. Whenever the known presence of an event is positive, the tool diagnosis can either be positive or negative. Hence, the proportion of true- positive diagnoses a/(a+c) will always complement the proportion of false-negative diagnoses c/(a+c). Similarly, when the known presence of an event is negative, the tool diagnosis can either be positive or negative. Hence, the proportion of false-positive diagnoses b/(b+d) will always complement the proportion of true-negative diagnoses d/(b+d). Therefore, all of the relevant information with regard to tool accuracy can be obtained from one proportion taken from each of the two columns of the contingency table. The usual choices are those of the top row, namely, the true-positive and false-positive proportions a/(a+c) and b/(b+d), commonly called "hits" and "false alarms" respectively.

The information content of two-way contingency tables are frequently converted into graphs that are referred to as Receiver Operating Characteristics or ROC curves. To convert the two-way contingency table into an ROC curve, the decision threshold, x_t , of the diagnostic tool response is varied from one end of the scale where all diagnoses are either true-positive or false-positive to the other end of the scale where all diagnoses are either false-negative or true-negative.

As the decision threshold is varied over its entire range, the ROC curve is traced out as shown in the diagram. If the tool does not discriminate between the presence and absence of an event, then the ROC curve will trace out the 45 degree line, and the area under the ROC curve will be one-half. If the tool discriminates perfectly between the presence and absence of an event, then the ROC curve will trace out the upper left-hand corner of the diagram, and the area under the ROC curve will then be unity. The areas under the ROC curves of diagnostic tools used in many diverse fields have been reported extensively in the literature. In particular, an article in the June 3rd, 1988 issue of Science magazine entitled "Measuring the Accuracy of Diagnostic Systems" by J. A. Swets is highly recommended.

Figure 10.4.3 - Receiver Operating Characteristics (ROC) of Diagnostic Tools.

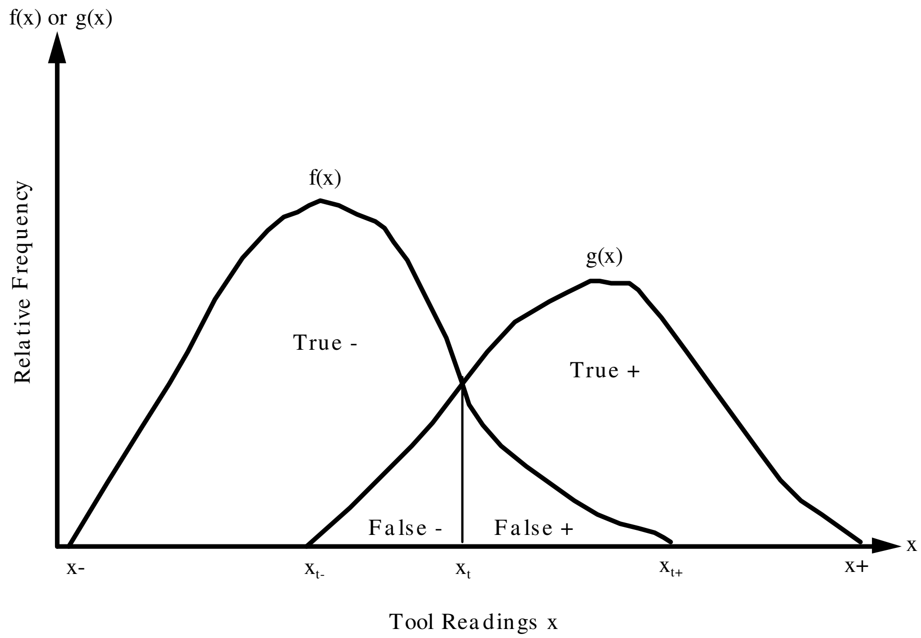


Let us now consider diagnostic tools for use in oil and gas exploration. Before using a tool in an unexplored region, it must be calibrated over areas H+ where subsurface hydrocarbon accumulations are known to be present as well as over neighboring background areas H^- where subsurface hydrocarbon accumulations are known to be absent. Let us consider the readings, x, from a tool of positive polarity. Such a tool will generally give higher readings over areas of H+ than over areas of H^-.

Let f(x) denote the relative frequency of tool readings, x, over areas of H-, and let g(x) denote the relative frequency of tool readings, x, over areas of H+. For any position of the decision threshold, x_t, the probability of true-positive Type I Successes is given by the area under the g(x) curve to the right of x_t, the probability of false-negative Type II Errors is given by the area under the f(x) curve to the right of x_t, the probability of true-negative Type II Successes is given by the area under the f(x) curve to the left of x_t, and the probability of false-negative Type I Errors is given by the area under the g(x) curve to the left of x_t.

The minimum sum of Type I and Type II Error probabilities occurs when x_t lies at the intersection of the f(x) and g(x) curves. However, owing to the greater economic importance of Type I Errors, the economic tradeoff favors using a smaller value of x_t which decreases Type I errors and increases Type II errors.

Figure 10.4.4 - Probability Density Functions of tool readings

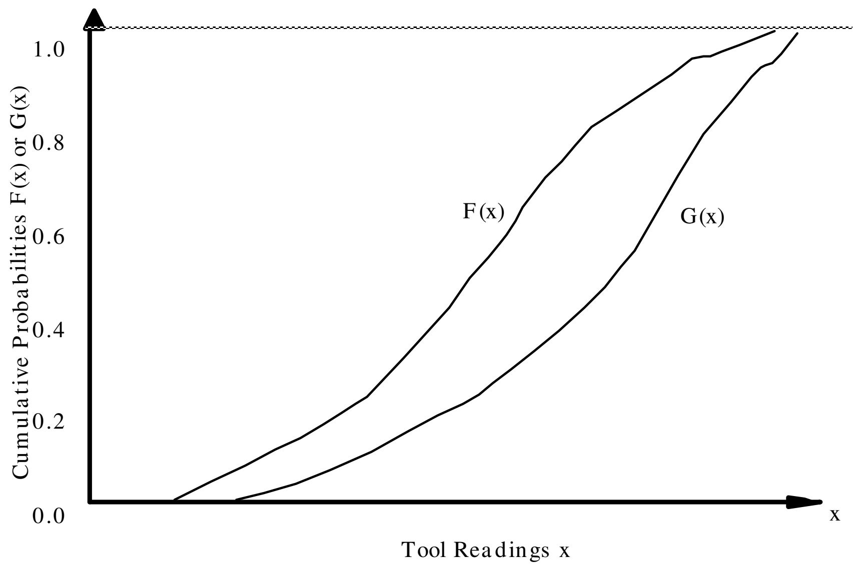


f(x) = Probability densities of tool readings from H- areas.
g(x) = Probability densities of tool readings from H+ areas.

False - = Type I Error   True + = Type I Success
False + = Type II Error   True - = Type II Success

In the process of calibrating an exploration tool, we are confronted with the ability of H- areas to mimic the behavior of H+ areas. Consequently, we have to test whether the differences between the two sample distributions over H+ and H- areas are statistically significant. For this purpose, the Kolmogorov-Smirnov Two-sample Test can be used because it is a nonparametric test which does not assume that the two distribution functions are known, only that they are continuous population distributions. The Kolmogorov-Smirnov is carried out as a function of the maximum probability difference between the two cumulative distribution functions F(x) and G(x) shown below.

Figure 10.4.5 - Cumulative Distribution Functions of Tool Readings



F(x) = Cumulative probabilities of readings x or less from H- areas.
G(x) = Cumulative probabilities of readings x or less from H+ areas.

Section 10.5 - Life Expectancy

The theory of life insurance is based on the probability that a certain individual will survive a certain amount of time. The measurement of this probability is based on the assumption that the individual is classified in a category which has been under observation for a long time. The probability the individual will survive a certain amount of time is defined as the fraction of the individuals in the category that can be expected to survive the amount of time in question. For the purposes of life insurance, healthy individuals are usually classified in categories of their ages.

The fundamental problem of life insurance is to determine the probability of a healthy individual of age x to survive one or more years. The probability of a person surviving two years is the product of the probability of surviving one year multiplied by the probability of a person one year older will also survive one year, and so on for a number of years. These probabilities are derived from a mortality table which is a record of the observed number of deaths in a large population. The mortality table that we will use is the 1980 Commissioners Standard Ordinary Mortality Table based on a population of North American life insurance policyholders. Separate tables are given for females and males in Appendices II-1 to II-4.

Column 1 has an individual's age denoted by x which ranges from zero to 99 years. This assumes that no individual will attain 100 years of age because the percentage of individuals who attain or live beyond the age of 100 is too small to affect insurance rates. Each individual is given a fictitious birthday, the first of July nearest to the date of actual birth. This assumes that those who announce a certain age on the first of January have their birthdays scattered evenly over a twelvemonth period (see An Introduction to Mathematical Probability by J. L. Coolidge, Dover Publications, Inc., New York 1962).

Column 2 is the l_x-column which gives the number of individuals who attain age x out of the normalized original l_0 = 10,000,000 individuals aged zero. The d_x-Column 3 gives the number of deaths of individuals aged x who will die before age $x+1$ as shown in equation (10.5.1). The probability that an individual aged x will not live a full year is defined as q_x in equation (10.5.2). The $1000q_x$-column (i.e., Column 4) multiplies q_x by 1,000. Thus, for a female aged 30, we have $1000q_{30}$ = 1.35 which means life insurance data indicate that 1.35 out of 1,000 females aged 30 are expected to die before age 31.

$$d_x = l_x - l_{x+1} \qquad \text{....(10.5.1)}$$

$$q_x = d_x/l_x \quad \text{and} \quad 1000q_x = 1000d_x/l_x \qquad \text{....(10.5.2)}$$

The p_x-probability that an individual aged x will survive for at least one year as defined in equation (10.5.3) is the complement of the q_x-probability of dying between ages x and $x+1$. Consequently, the probability ${}_np_x$ that an individual aged x will survive for at least n years is given in equation (10.5.4). It follows the probability ${}_nq_x$ that an individual aged x will die before age $x+n$ is given by equation (10.5.5).

$$p_x = l_{x+1}/l_x \quad \text{and} \quad p_x = 1 - q_x \qquad \text{....(10.5.3)}$$

$${}_np_x = l_{x+n}/l_x \quad \text{and} \quad {}_np_x = 1 - {}_nq_x \qquad \text{....(10.5.4)}$$

$${}_nq_x = (l_x - l_{x+n})/l_x \quad \text{and} \quad {}_nq_x = 1 - {}_np_x \qquad \text{....(10.5.5)}$$

Longevity L_x is defined as the expected lifespan of an individual aged x which is listed in Column 8 of Appendices II-1 to II-4. By definition, L_{99} has a lifespan of 99 years. The longevity of females aged 98 is $L_{98} = 98 + {}_1p_{98} = 98 + (1 - {}_1q_{98}) = 98 + (1\text{-}0.65585) = 98.3$ years. The longevity of females aged 97 is $L_{97} = 97 + {}_1p_{97} + {}_2p_{97} = 97.7$ years, and so on. In general, the longevity of an individual aged x as derived from the mortality tables in Appendices II-1 to II-4 is given by equation (10.5.6).

$$L_x = x + \sum_{k=1}^{k=99-x} {}_k p_x \quad \text{for} \quad 0 \leq x < 99 \qquad \text{....(10.5.6)}$$

Columns 5, 6 and 7 of the mortality tables with headings D_x, N_x and M_x respectively pertain to benefits offered by life insurance companies in consideration of premiums paid by policyholders. There are two kinds of benefits which life insurance companies pay, namely, annuities if people survive, and insurance benefits if they die. The present value of the amount to be paid to an individual at some specified future date if that individual is still alive, is called an endowment. The D_x, N_x and M_x series in Appendices II-1 to II-4, called "Commutation Columns at 4%", facilitate present-value calculations of annuities and endowments discounted at 4% compound interest per year.

The *pure endowment factor*, ${}_nE_x$, denotes the present value of \$1 to be paid to an individual now aged x if and when that individual attains age $x+n$. The probability an individual now aged x will survive to age $x+n$ is ${}_np_x$, so that

$${}_nE_x = (1+i)^{-n} \cdot {}_np_x = v^n \cdot l_{x+n}/l_x \qquad \text{....(10.5.7)}$$

In column 5 of Appendices II-1 to II-4 we tabulated commutation symbol D_x, defined by equation (10.5.8). In terms of the D_x symbol, equation (10.5.8) can be simplified to equation (10.5.9) which can utilize the tabulated values of D_x if i = 4% compound interest per year.

$$D_x = (1+i)^{-x} \cdot l_x = v^x l_x \quad \text{where } v \equiv (1+i)^{-1} \qquad \text{....(10.5.8)}$$

$${}_nE_x = (1+i)^{-n} \cdot l_{x+n}/l_x = v^{x+n} l_{x+n}/v^x l_x = D_{x+n}/D_x \qquad \text{....(10.5.9)}$$

In applying equation (10.5.9), suppose l_x individuals all aged x agree to contribute equally to a fund which after n years will be sufficient to pay a \$1 endowment to each of the group who attains age $x+n$ at i% compound interest per year. Since l_{x+n} of those individuals will survive n years from now, the fund will need l_{x+n} dollars, the present value of which is $(1+i)^{-n}\, l_{x+n}$. Hence, each one of the l_x individuals must pay a *net premium* of ${}_nE_x$ = $(1+i)^{-n}\, l_{x+n}/l_x = D_{x+n}/D_x$ for the l_{x+n} survivors to receive a \$1 endowment n years later. The net premium does not include ordinary expenses to take care of insurance company profit, agent's commissions and other contingencies. Because *loading factors* added to net premiums vary from company to company, only net premiums are discussed here.

For example, on his 30th birthday, a person uses \$5,000 of his savings to purchase a pure endowment payable if and when he attains age 55. Assuming he survives to age 55, how much will his endowment be worth at 4% compound interest per year. The net premium for a \$1 endowment 25 years later is

$${}_{25}E_{30} = (1.04)^{-25} \cdot l_{55}/l_{30} = D_{55}/D_{30} = 995{,}972.07/2{,}953{,}692.22 = 0.3371956$$

For \$5,000, the endowment is \$5,000/${}_{25}E_{30}$ = \$5,000/0.3371956 = \$14,828.19

If a person invested \$5,000 for 25 years at 4% compound interest per year, the investment would be worth $\$5{,}000 \cdot (1.04)^{25} = \$13{,}329.18$.

A *life annuity* consists of periodic payments which continue for all or for some portion of the life of an individual, called the *annuitant*. A life annuity whose payments continue so long as the annuitant is alive is called a *whole life annuity*. If payments to an individual aged x are made at the end of each year, the 1st payment at age $x+1$, the 2nd at age $x+2$, and so on, the annuity is called an *ordinary whole life annuity* which is a set of pure endowments payable at the ends of 1, 2, 3, ... years until the annuitant dies. If a_x denotes the present-value of the *net single premium* of an *ordinary whole life annuity* of \$1 per year for an individual aged x, then

$$a_x = {}_1E_x + {}_2E_x + {}_3E_x + \ldots + {}_{99-x}E_x \qquad \text{....(10.5.10)}$$

By means of comutation symbols in equation (10.5.9), the expression for a_x becomes

$$a_x = [D_{x+1} + D_{x+2} + D_{x+3} + \ldots + D_{99}]/D_x \qquad \text{....(10.5.11)}$$

We now introduce the symbol $N_x = D_x + D_{x+1} + D_{x+2} + \ldots + D_{99}$ in column six so that

$$a_x = N_{x+1}/D_x \qquad \text{....(10.5.12)}$$

For example, the net single premium for an *ordinary whole life annuity* of \$1,000/year for a male aged 30 at 4% interest/year is $a_{30} = N_{31}/D_{30} = \$1{,}000 \cdot 57{,}777{,}736/2{,}953{,}692 = \$19{,}561$.

If payments to an individual aged x are made at the <u>beginning</u> of each year, the 1st payment at age x, the 2nd at age $x+1$, and so on, the annuity is called a *whole life annuity due*. This type of whole life annuities is more common because payments are usually due at the beginning of each period. Thus, a *whole life annuity due* of \$1 per year consists of a \$1 payment now plus a \$1 ordinary whole life annuity. If $\boldsymbol{a_x}$ denotes the present-value of the *net single premium* of a *whole life annuity due* of \$1 per year for an individual aged x, then

$$\boldsymbol{a_x} = 1 + a_x = 1 + [N_{x+1}/D_x] = N_x/D_x \qquad \text{....(10.5.13)}$$

For example, the net single premium for a *whole life annuity due* of \$1,000/year for a male aged 30 at 4% interest/year is $\boldsymbol{a_{30}} = N_{30}/D_{30} = \$1{,}000 \cdot 60{,}731{,}428/2{,}953{,}692 = \$20{,}561$.

An *ordinary whole life annuity deferred for k years* is a sequence of pure endowments, the 1st payable at the end of $k+1$ years, the 2nd payable at the end of $k+2$ years, and so on, and ending when the annuitant dies. If ${}_ka_x$ is the present-value of the *net single premium* of an *ordinary whole life annuity of \$1 deferred for k years* for an individual aged x, then

$${}_ka_x = N_{x+k+1}/D_x \qquad \text{....(10.5.14)}$$

For example, the net single premium for an *ordinary whole life annuity* of \$1,000/year for a male aged 30 starting at age 55 is ${}_{24}a_{30} = N_{55}/D_{30} = \$1{,}000 \cdot 14{,}036{,}801/2{,}953{,}692 = \$4{,}752$.

If $\boldsymbol{{}_ka_x}$ denotes the present-value of the *net single premium* of a *whole life annuity due of \$1 deferred for k years* for an individual aged x, then equation (10.5.15) is applicable. The previous example is the same as a whole life annuity due deferred for 25 years.

$$\boldsymbol{{}_ka_x} = N_{x+k}/D_x \qquad \text{....(10.5.15)}$$

A temporary life annuity is a whole life annuity that ends after a specified number of payments even though the annuitant is still alive. An ordinary n-year temporary life annuity is equivalent to an ordinary whole life annuity minus an ordinary whole life annuity deferred for n years. Therefore, if $a_{x:n}$ denotes the present-value of the *net single premium* of an *ordinary n-year temporary life annuity* of \$1 for an individual aged x, then

$$a_{x:n} = a_x - {}_na_x = [N_{x+1} - N_{x+n+1}]/D_x \qquad \text{....(10.5.16)}$$

The *net single premium* of a *n-year temporary life annuity due* of \$1 for a person aged x is

$$\mathbf{a}_{x:n} = \mathbf{a}_x - {}_n\mathbf{a}_x = [N_x - N_{x+n}]/D_x \qquad \text{....(10.5.17)}$$

Annuity policies use equal annual premium payments over a given period to create pensions for individuals whose payments begin on specified dates and continue for life. The premium payments form a temporary life annuity due since the first premium is payable when the policy is purchased. The pension payments form a deferred life annuity due.

For example, a person aged 30 plans to retire at age 55 with an annual pension of \$5,000 starting with his 55th birthday. He purchases this annuity by agreeing to make equal annual premium payments starting now and including his 54th birthday. Let us find the annual premium P required for this pension plan. The premium payments form a 25-year temporary life annuity due whose present value at 4% compound interest is $P\mathbf{a}_{30:25}$. The \$5,000 pension payments form a deferred life annuity due \$5,000 $\cdot{}_{25}\mathbf{a}_{30}$. Therefore,

$$P\mathbf{a}_{30:25} = \mathrm{P}[N_{30} - N_{55}]/D_{30} = \$5{,}000 \cdot {}_{25}\mathbf{a}_{30} = \$5{,}000 \cdot N_{55}/D_{30}$$

$$P = \$5{,}000 \cdot N_{55}/[N_{30} - N_{55}] = \$5{,}000 \cdot 14{,}036{,}801/[60{,}731{,}428 - 14{,}036{,}801] = \$1{,}503\ .$$

A life insurance policy is a contract between a life insurance company and a person who agrees to pay premiums for which the company promises to pay a fixed sum to designated beneficiaries when the insured person dies. Major types of life insurance are:

i - <u>whole life insurance</u> - the company promises to pay the face value of the policy to the beneficiary whenever the insured dies.

ii - <u>*n*-year term insurance</u> - the company promises to pay the face value of the policy to the beneficiary only if the insured dies within n years after the policy was issued.

iii - <u>*n*-year endowment insurance</u> - the company promises to pay the face value of the policy to the beneficiary if the insured dies within n years after the policy was issued, and to pay the face value of the policy to the insured if he survives after n years.

To simplify calculations from mortality tables, we assume beneficiaries are paid at the end of the year of death. Also, as with life annuities, only net premiums will be considered.

i - Let A_x be the net single premium for whole life insurance of \$1 for persons aged x. To determine A_x, let us find how much l_x persons, all aged x, must each contribute to enable the insurance company to pay their beneficiaries \$1 at the end of the year in which the insured persons die. The total contributions to the company would be l_xA_x. In the 1st year, d_x of the policy holders will die and d_x must be paid to beneficiaries at the end of the year, the present value of which is $(1 + \mathrm{i})^{-1} \cdot d_x = vd_x$. In the 2nd year, d_{x+1} will die and the present value paid beneficiaries at the end of the 2nd year is v^2d_{x+1}, and so on. Thus,

$$l_xA_x = vd_x + v^2d_{x+1} + v^3d_{x+2} + \ldots + v^{100-x}d_{99} \qquad \text{....(10.5.18)}$$

Multiplying both sides of equation (10.5.18) by v^x gives us

$$v^xl_xA_x = v^{x+1}d_x + v^{x+2}d_{x+1} + v^{x+3}d_{x+2} + \ldots + v^{100}d_{99} \qquad \text{....(10.5.19)}$$

In terms of the commutation symbols $D_x = v^xl_x$, $C_x = v^{x+1}d_x$ and $M_x = C_x + C_{x+1} + C_{x+2} + \ldots + C_{99}$

$$D_xA_x = C_x + C_{x+1} + C_{x+2} + \ldots + C_{99} = M_x \quad \text{and} \quad A_x = M_x/D_x \qquad \text{....(10.5.20)}$$

Policy holders are rarely able to pay up front the A_x net single premium for whole life insurance of $1. Instead, the present value of equivalent annual premiums are paid at the beginning of each year either (a) over the entire life of an *ordinary* life insurance policy or (b) over the first m years of an m-payment life insurance policy.

(a) Let P_x denote the net annual premium payable at the beginning of the year for an ordinary life insurance policy of $1 issued to a person aged x. Since P_x annual premium payments form a whole life annuity due, we have $P_x \cdot \mathbf{a_x} = A_x$ or $P_x = A_x/\mathbf{a_x}$. Upon substitution of equations (10.5.20) and (10.5.13) for $A_x/\mathbf{a_x}$, we get

$$P_x = A_x/\mathbf{a_x} = [M_x/D_x]/[N_x/D_x] = M_x/N_x \qquad \text{....(10.5.21)}$$

Hence, the net annual premium for ordinary life insurance policies of $1,000 for 30-year males is $\$1{,}000 \cdot P_{30} = \$1{,}000 \cdot M_{30}/N_{30} = \$1{,}000 \cdot 617{,}868/60{,}731{,}428 = \$10.17$ when interest is compounded at 4% per year.

(b) Let ${}_mP_x$ denote the net annual premium payable at the beginning of the year for an m-payment life insurance policy of $1 issued to a person aged x. Since ${}_mP_x$ annual premium payments form an m-year temporary life annuity due, we have ${}_mP_x \cdot \mathbf{a_{x:m}} = A_x$ or ${}_mP_x = A_x/\mathbf{a_{x:m}}$. Upon substituting equations (10.5.20) and (10.5.17) for $A_x/\mathbf{a_{x:m}}$, we get

$${}_mP_x = A_x/\mathbf{a_{x:m}} = [M_x/D_x]/[(N_x-N_{x+m})/D_x] = M_x/(N_x-N_{x+m}) \qquad \text{....(10.5.22)}$$

Hence, the net annual premium for 25-year life insurance policies of $1,000 for 30-year males is $\$1{,}000 \cdot {}_{25}P_{30} = \$1{,}000 \cdot M_{30}/(N_{30}-N_{55}) = \$1{,}000 \cdot 617{,}868/(60{,}731{,}428-14{,}036{,}801) = \$13.23$ when interest is compounded at 4% per year.

ii - <u>n-year term insurance</u> - Let $A_{x:n}$ be the net single premium for n-year term insurance of $1 for persons aged x. Just as A_x was determined from equations (10.5.18) to (10.5.20), let us determine $A_{x:n}$ whose last benefit is paid at the end of n years.

$$l_xA_{x:n} = vd_x + v^2d_{x+1} + v^3d_{x+2} + \ldots + v^nd_{x+n-1} \qquad \text{....(10.5.18')}$$

Multiplying both sides of equation (10.5.18') by v^x gives us

$$v^xl_xA_{x:n} = v^{x+1}d_x + v^{x+2}d_{x+1} + v^{x+3}d_{x+2} + \ldots + v^{x+n}d_{x+n-1} \qquad \text{....(10.5.19')}$$

In terms of the commutation symbols $D_x = v^xl_x$, $C_x = v^{x+1}d_x$ and $M_x = C_x + C_{x+1} + C_{x+2} + \ldots + C_{99}$

$$D_xA_{x:n} = [C_x + C_{x+1} + C_{x+2} + \ldots + C_{99}] - [C_{x+n} + \ldots + C_{99}] = M_x-M_{x+n} \quad \text{and}$$

$$A_{x:n} = [M_x-M_{x+n}]/D_x \qquad \text{....(10.5.20')}$$

Hence, the net single premium for a 10-year term insurance policy of $1,000 for a 30-year male is $\$1,000 \cdot A_{30:10} = \$1,000 \cdot [M_{30}-M_{40}]/D_{30} = \$1,000 \cdot [617,868-568,002]/2,953,692 = \$16.88$ when interest is compounded at 4% per year.

Let $P_{x:n}$ denote the net annual premium payable at the beginning of the year for an n-year term insurance policy of $1 issued to a person aged x. Since $P_{x:n}$ annual premium payments form an n-year temporary annuity due, we have $P_{x:n} \cdot \boldsymbol{a_{x:n}} = A_{x:n}$ or $P_{x:n} = A_{x:n}/\boldsymbol{a_{x:n}}$. Upon substituting equations (10.5.20') and (10.5.17) for $A_{x:n}/\boldsymbol{a_{x:n}}$, we get

$$P_{x:n} = A_{x:n}/\boldsymbol{a_{x:n}} = [(M_x-M_{x+n})/D_x]/[(N_x-N_{x+n})/D_x] = (M_x-M_{x+n})/(N_x-N_{x+n}) \qquad ....(10.5.21')$$

For 30-year males, the net annual premium for 10-year term insurance of $1000 is $\$1000 \cdot P_{30:10} = \$1000 \cdot (M_{30}-M_{40})/(N_{30}-N_{40}) = \$1000 \cdot (617,868-568,002)/(60,731,428-36,014,445) = \$2.02$ when interest is compounded at 4% per year.

iii - <u>n-year endowment insurance</u> - An n-year endowment insurance policy combines the benefits of n-year term insurance and a pure endowment at the end of n years. Let $\boldsymbol{A_{x:n}}$ be the net single premium for n-year endowment insurance of $1 for persons aged x. Then

$$\boldsymbol{A_{x:n}} = A_{x:n} + {}_nE_x = \{[M_x-M_{x+n}]/D_x\} + \{D_{x+n}/D_x\} = [M_x-M_{x+n}+D_{x+n}]/D_x \quad ....(10.5.22)$$

Thus, the net single premium for 25-year endowment insurance of $1000 for a male aged 30 is $\$1000 \cdot \boldsymbol{A_{30:25}} = \$1000 \cdot [M_{30}-M_{55}+D_{55}]/D_{30} = \$1000 \cdot [617,868-456,095+995,972]/2,953,692 = \$391.97$ when interest is compounded at 4% per year.

Let $\boldsymbol{P_{x:n}}$ denote the net annual premium for an n-year endowment insurance policy of $1 for persons aged x. Then

$$\boldsymbol{P_{x:n}} = [M_x-M_{x+n}+D_{x+n}]/[N_x-N_{x+n}] \qquad ....(10.5.23)$$

The net annual premium for 25-year endowment insurance of $1000 for a male aged 30 is $\$1000 \cdot \boldsymbol{P_{30:25}} = \$1000 \cdot [617,868-456,095+995,972]/[60,731,428-14,036,801] = \$24.79$.

Let us compare the net annual premium $P_x = M_x/N_x = \$10.17$ for a $1,000 whole life insurance policy for a male aged 30 (see equation (10.5.21)) to net single premiums $A_{x:1} = [M_x-M_{x+1}]/D_x$ for $1,000 1-year term insurance policies that he could have taken at ages 30, 31, and so on (see equation (10.5.20')) as shown in the table below.

Age	Net Annual Premium at Age 30	One-year Term Premium
30	$10.17	$1.66
31	$10.17	$1.71
40	$10.17	$2.90
50	$10.17	$6.45
60	$10.17	$15.46
70	$10.17	$37.99

At age 30, the insured pays $10.17-$1.66 = $8.51 more than the one-year cost of insurance. The excess is placed in a *reserve fund* that earns the same interest used to compute the annual premium. At age 60, $15.46-$10.17 = $5.29 is withdrawn from the reserve fund. The terminal reserve at the end of any year the policy is in force less a nominal charge for expenses is called the *cash surrender value* of the policy which belongs to the insured. Thus, the terminal reserve at age 40 is $\$1000 \cdot A_{40}-\$10.17 \cdot \boldsymbol{a_{40}} = \$290.81-\$187.52 = \$103.29$.

Section 10.6 - Summary of Chapter Ten

The first nine chapters dealt with *deterministic* decision-making in which inputs and outputs of alternatives are known with certainty. This enabled a provably optimal method of economic decision-making to be developed. In most real-world situations, the outcomes of alternatives are not known with certainty. Decision-making under such circumstances is known as *probabilistic* decision-making which needs clear definitions of risk and uncertainty. Chapter Ten indicates the deterministic decision methods previously developed appear to be the best guide for decision-making wherever probabilities can be estimated reliably.

The probability of event A is defined as the fraction of the total number of relevant events corresponding to event A. Letting P(A) denote the probability of event A, we have

P(A) = number of ways A occurs/total number of relevant events.

If a company sells 50 items, 2 of which are defective, the probability of a customer getting a defective item is 2/50 = .04 or 4% if each item is equally likely to be sold to the customer.

The probability of an event occurring is always greater than or equal to zero and less than or equal to one. If the occurrence of a particular event is *impossible* then its probability equals zero, and if the occurrence of an event is *certain* then its probability equals *one*. For example, the probability of rolling a seven in a fair die is zero, and the probability of rolling any number less than seven in a fair die is one. If only one event can occur on any given trial, then the possible events are known as *mutually exclusive* events. If all the possible events are included in a test, then the events are *collectively exhaustive*. The sum of the probabilities of a set of mutually exclusive and collectively exhaustive events must equal one. All possible outcomes of rolling a fair die represent a set of mutually exclusive and collectively exhaustive events.

Probabilities can be described as being either *a priori* or *empirical* in nature. A priori probabilities can be computed prior to conducting any trials. For example, we know a priori that the probability of a six appearing when a fair die is rolled is 1/6. On the other hand, empirical probabilities are estimated in a relative frequency sense on data gathered after conducting repeated trials of an event, or on historical data of a particular population. If all the people in a village whose population is 1,000 are screened for color-blindness and 60 tested positive, then the probability of a randomly picked person from the village being color-blind is 60/1000 or 6%, assuming all persons have an equal chance of being picked.

In the relative frequency definition of probability, the number of ways that event A occurs must be compared to the total number of relevant events. It is important to specify which events are relevant to make the definition of probability meaningful. If we are sampling from a small region of the population in question, the probabilities of particular events can be highly distorted. In effect, all probabilities are conditional probabilities, and the condition is spelled out by specifying what constitutes relevant events.

Probability is often applied in a loose sense to the measurement of risk and uncertainty. It is important to bring out that in dealing with risk, there is a basic tradeoff between limiting losses (downside risk) and realizing profits (upside risk). In the final analysis, if one does not risk any loss, he does not risk making any profit either. This does not mean that upside and downside risk are always equal or correlated. To the contrary, not only is it important to distinguish between upside and downside risk, but it is also important to estimate their magnitudes independently.

Risk generally means the chance of loss which is invariably coupled to some chance of profit or gain. Therefore, economic risk is defined here as ventures in which a given input could result in more than one distinctly different outputs involving either loss or gain. If the probabilities or relative frequencies of various outputs could be estimated, then the problem is referred to as decision-making under conditions of economic risk. But if the probabilities or relative frequencies of various outputs that could result from a given input are unknown, either due to a lack experience or an inability to identify possible outcomes, then the problem is called decision-making under conditions of economic uncertainty.

A common method of handling problems of decision-making under conditions of economic risk is to reduce probability distributions to measures of central tendency such as their expected values. This method of analysis is generally referred to as the expectation principle. When probability distributions of outputs or inputs are reduced to expected values, the problems of decision-making under conditions of economic risk become the same as the deterministic problems previously handled under conditions of economic certainty. Hence, decision-making under conditions of economic risk does not need any new economic principles, but it does require possible outcomes to occur with statistical regularity so that their probabilities can be estimated or measured in an objectively verifiable manner.

Interest rates used in discounting future cash flows of alternatives are often exaggerated by lumping various risks together. Although high discount rates favor alternatives with early profits, they also increase risk by rejecting alternatives which provide greater long-term profits. Thus, high discount rates could decrease short-term risk but increase long-term risk. There is also considerable room for questioning the use of high discount rates to account for risk, inflation and investor expectations because the results are unrelated to costs of borrowing money that would appear in future financial statements of an economic organization.

It is not at all obvious how high discount rates could account for different major risks encountered in most alternatives. *Engineering risks* of carrying out the input requirements of alternatives and getting the job done may be appreciably smaller than *marketing risks* of selling their outputs in sufficient quantities with adequate profit margins. Moreover, the *financial risks* of debt and equity financing in the money markets to satisfy the input and output requirements of alternatives may differ significantly from their engineering and marketing risks. In many cases, risks have been included in cash flow estimates of alternatives so that discounted cash flows evaluated with high discount rates results in double-counting the underlying risks.

The focus of the tenth chapter is on specific identification of risks and bringing out similarities of probabilistic and deterministic decision-making. Probabilities of major risk categories need to be evaluated on their own merits in order to make better progress in probabilistic decision-making. In this regard, Section 10.1 describes the early development of empirical and theoretical methods to uncover statistical regularities in large bodies of data. Section 10.2 explains the problems of averaging large arrays of numbers into single figures. Section 10.3 introduces the addition theorem of probabilities which is applied to sampling surveys and to problems of deciding how many items to stock, when reordering during the season is impractical, and unsold items at the end of the season lose much of their value. Section 10.4 presents the multiplication theorem of probabilities for handling problems of statistical inference. Lastly, Section 10.5 explains life insurance principles derived from mortality tables in connection both with life annuities if people survive and with life insurance if they die.

Chapter Ten - Exercises

10-1 An electronic device is subject to three types of faults labelled A, B and C. An inspector takes a random sample of these devices and finds:

200 have no faults	85 have faults A and B
200 have fault A	130 have faults B and C
359 have fault B	105 have faults C and A
731 have fault C	30 have all three faults

10-1a How many devices were in the inspector's sample? (Ans. 1200)

10-1b How many of the 200 specimens that have fault A do not have either B or C?

10-1c If the inspector obtained one additional sample at random and found that it does not have either faults B or C, would this information change the probability that the sample has fault A? This question is equivalent to asking if the probability of having fault A only (and does not have faults B and C) is equal to the probability of having fault A. For example, the probability of having fault B only is 174/359 = 46.52%. The probability of having fault B is 359/1200 = 29.91%. Hence, knowledge or information that an additional random sample does not have either faults A or C increases the chances of the sample having fault B.

10-2 A dairy product merchant buys from 30 to 40 units of cheese every week at \$3.00 per unit. At a price of \$7.00 per unit, the demand for cheese each week is a random variable D. Cheese that is unsold at the end of the week is disposed of at \$1.00 per unit. Assume that one week is indistinguishable from any other, and that the probability f(D) of the demand D for cheese each week is as given below.

D=	30	31	32	33	34	35	36	37	38	39	40
f(D)=	.01	.03	.06	.10	.20	.25	.15	.10	.05	.04	.01
F(D)=											

10-2a Determine the cumulative probabilities F(D) of demand D or less per week.

10-2b Find the number of cheese units the dairy merchant should buy each week in order to maximize the expected value of his profit.

10-2c If the dairy merchant could predict from "Perfect Information" the demand for cheese units each week, determine his expected-value profit from such perfect information.

Index