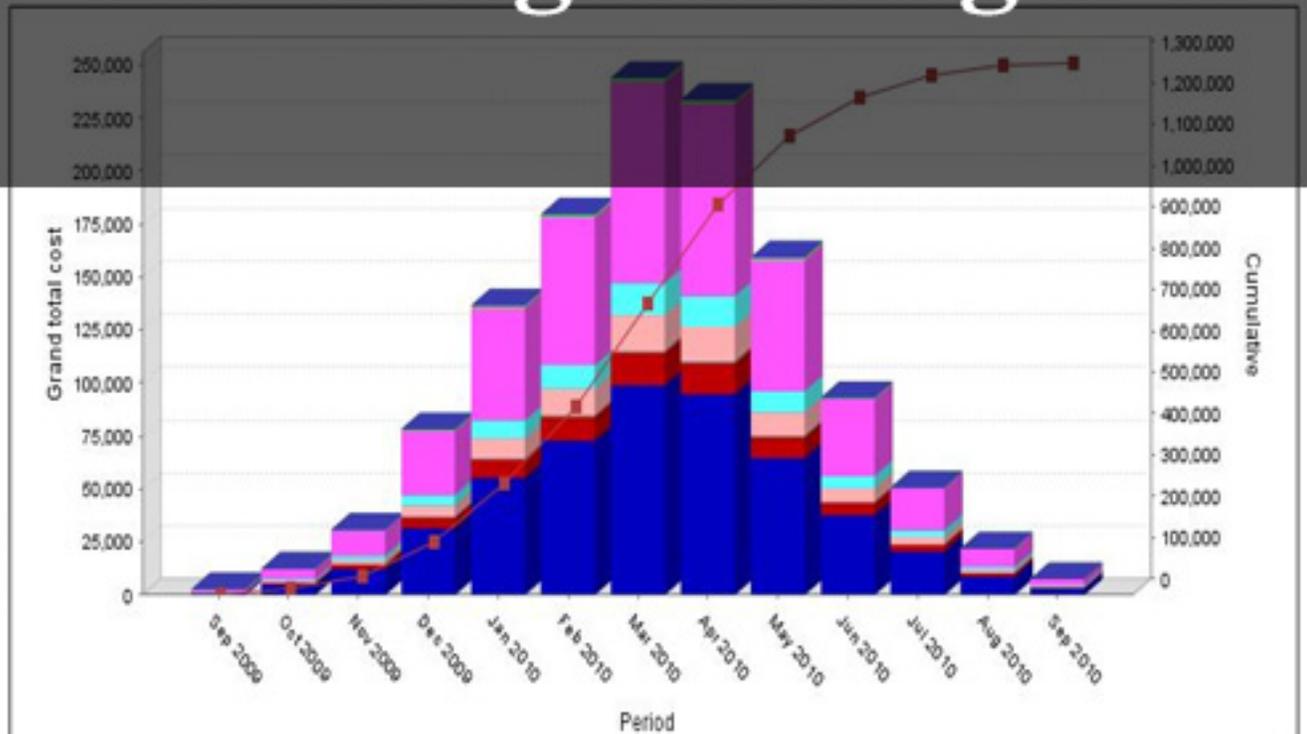


Cost Engineering

- View totals by
- Cost category
 - Discipline
- Filter by
- Cost category
 - Construction tools
 - Consumables
 - Hourly rates
 - Install
 - Material
 - Prefabrication
 - Remove
 - Subcontracting
 - Testing



	Sep 2009	%	Oct 2009	%	Nov 2009	%	Dec 2009	%	Jan 2010	%	Feb 2010	%	Mar 2010
Equipment	1.112,80	0,09%	5.024,98	0,40%	12.518,99	1,00%	31.743,75	2,54%	55.471,87	4,43%	72.934,71	5,83%	99.044,91
Civil	169,54	0,01%	765,58	0,06%	1.907,33	0,15%	4.836,33	0,39%	8.451,43	0,68%	11.111,98	0,89%	15.090,01
Electrical	12,50	0,00%	56,43	0,00%	140,59	0,01%	356,48	0,03%	622,95	0,05%	819,06	0,07%	1.112,27
Instrumentation	190,68	0,02%	861,02	0,07%	2.145,11	0,17%	5.439,24	0,43%	9.505,02	0,76%	12.497,25	1,00%	16.971,20
Insulation	169,32	0,01%	764,58	0,06%	1.904,84	0,15%	4.830,00	0,39%	8.440,38	0,67%	11.097,46	0,89%	15.070,28
Piping	1.066,50	0,09%	4.815,89	0,38%	11.998,07	0,96%	30.422,88	2,43%	53.163,67	4,25%	69.899,87	5,59%	94.923,62
Painting	3,47	0,00%	15,69	0,00%	39,08	0,00%	99,09	0,01%	173,15	0,01%	227,66	0,02%	309,16
Steel structures	14,57	0,00%	65,78	0,01%	163,87	0,01%	415,51	0,03%	726,11	0,06%	954,69	0,08%	1.296,46
General E & I	2,47	0,00%	11,16	0,00%	27,80	0,00%	70,50	0,01%	123,20	0,01%	161,98	0,01%	219,97
Grand total	2.741,85	0,22%	12.381,11	0,99%	30.845,68	2,47%	78.213,78	6,25%	136.677,78	10,93%	179.704,66	14,37%	244.037,88

First Edition, 2012

ISBN 978-81-323-3037-0

© All rights reserved.

Published by:

Research World

4735/22 Prakashdeep Bldg,

Ansari Road, Darya Ganj,

Delhi - 110002

Email: info@wtbooks.com

Table of Contents

Introduction

Chapter 1 - Cost

Chapter 2 - Value Engineering

Chapter 3 - Cost Contingency & Cost Escalation

Chapter 4 - Quantity Surveyor

Chapter 5 - Cost-Benefit Analysis

Chapter 6 - Cost-Effectiveness Analysis & Cost Overrun

Chapter 7 - Opportunity Cost & Cost of Capital

Chapter 8 - Average Cost

Chapter 9 - Cost Accounting

Chapter 10 - Cost Curve

Chapter 11 - Marginal Cost

Chapter 12 - Historical cost

Chapter 13 - Social Cost & Sunk Costs

Chapter 14 - Transaction Cost

Chapter 15 - Cost-Volume-Profit Analysis

Chapter 16 - Isocost

Chapter 17 - Consumer Choice

Introduction

Cost engineering is an area of engineering practice concerned with the "application of scientific principles and techniques to problems of cost estimating, cost control, business planning and management science, profitability analysis, project management, and planning and scheduling."

Overview

According to the American Association of Cost Engineering, cost engineering is defined as that area of engineering practice where engineering judgment and experience are utilized in the application of scientific principles and techniques to the problem of cost estimation, cost control and profitability. Key objectives of cost engineering are to arrive at accurate cost estimates and to avoid cost overruns. The broad array of cost engineering topics represent the intersection of the fields of project management, business management, and engineering. Most people have a limited view of what engineering encompasses. The most obvious perception is that engineering addresses technical issues such as the physical design of a structure or system. However, beyond the physical manifestation of a design of a structure or system (for example, a building), there are other dimensions to consider such as the money, time, and other resources that were invested in the creation of the building. Cost engineers refer to these investments collectively as "costs".

Cost engineering then can be considered an adjunct of traditional engineering. It recognizes and focuses on the relationships between the physical and cost dimensions of whatever is being "engineered". Cost engineering is most often taught at universities as part of construction engineering, engineering management, civil engineering, and related curricula because it is most often practiced on engineering and construction capital projects. Engineering economics is a core skill and knowledge area of cost engineering.

History

Cost engineering is a field of engineering practice that began in the 1950s (AAACE International was founded in 1956). The skills and knowledge areas of Cost Engineers are similar to those of Quantity Surveyors. AAACE International is one of many international engineering organizations representing practitioners in these fields.

In 2006, AAACE published the Total Cost Management (TCM) Framework which outlines an integrated process for applying the skills and knowledge of cost engineering. This has also been called the world's first process for portfolio, program and project management.

Chapter 1

Cost

In business, retail, and accounting, a **cost** is the value of money that has been used up to produce something, and hence is not available for use anymore. In economics, a cost is an alternative that is given up as a result of a decision. In business, the cost may be one of acquisition, in which case the amount of money expended to acquire it is counted as cost. In this case, money is the input that is gone in order to acquire the thing. This acquisition cost may be the sum of the cost of production as incurred by the original producer, and further costs of transaction as incurred by the acquirer over and above the price paid to the producer. Usually, the price also includes a mark-up for profit over the cost of production.

Costs are often further described based on their timing or their applicability.

Accounting vs opportunity costs

In accounting, costs are the monetary value of expenditures for supplies, services, labour, products, equipment and other items purchased for use by a business or other accounting entity. It is the amount denoted on invoices as the price and recorded in bookkeeping records as an expense or asset cost basis.

Opportunity cost, also referred to as *economic cost* is the value of the best alternative that was not chosen in order to pursue the current endeavour—i.e., what could have been accomplished with the resources expended in the undertaking. It represents opportunities forgone.

In theoretical economics, **cost** used without qualification often means opportunity cost.

Comparing private, external, social, and psychic costs

When a transaction takes place, it typically involves both private costs and external costs.

Private costs are the costs that the buyer of a good or service pays the seller. This can also be described as the costs internal to the firm's production function.

External costs (also called externalities), in contrast, are the costs that people other than the buyer are forced to pay as a result of the transaction. The bearers of such costs can be either particular individuals or society at large. Note that external costs are often both non-monetary and problematic to quantify for comparison with monetary values. They include things like pollution, things that society will likely have to pay for in some way or at some time in the future, but that are not included in transaction prices.

Social costs are the sum of private costs and external costs.

For example, the manufacturing cost of a car (i.e., the costs of buying inputs, land tax rates for the car plant, overhead costs of running the plant and labour costs) reflects the *private cost* for the manufacturer (in some ways, normal profit can also be seen as a cost of production; see, e.g., Ison and Wall, 2007, p. 181). The polluted waters or polluted air also created as part of the process of producing the car is an *external cost* borne by those who are affected by the pollution or who value unpolluted air or water. Because the manufacturer does not pay for this external cost (the cost of emitting undesirable waste into the commons), and does not include this cost in the price of the car (a Kaldor-Hicks compensation), they are said to be external to the market pricing mechanism. The air pollution from driving the car is also an externality produced by the car user in the process of using his good. The driver does not compensate for the environmental damage caused by using the car.

A psychic cost is a subset of social costs that specifically represent the costs of added stress or losses to quality of life.

Cost estimates and cost overrun

When developing a business plan for a new or existing company, product, or project, planners typically make cost estimates in order to assess whether revenues/benefits will cover costs. This is done in both business and government. Costs are often underestimated resulting in cost overrun during implementation. Main causes of cost underestimation and overrun are optimism bias and strategic misrepresentation (Flyvbjerg et al. 2002). Reference class forecasting was developed to curb optimism bias and strategic misrepresentation and arrive at more accurate cost estimates.

Cost Plus, is where the Price = Cost plus or minus X%, where x is the percentage of built in overhead or profit margin.

Manufacturing Costs VS. Non-manufacturing costs

Manufacturing Costs are those costs that are directly involved in manufacturing of products. Examples of manufacturing costs include raw materials costs and charges related workers. Manufacturing cost is divided into three broad categories:

1. Direct materials cost.
2. Direct labor cost.
3. Manufacturing overhead cost.

Non-manufacturing Costs are those costs that are not directly incurred to manufacture a product. Examples of such costs are salary of sales personnel and advertising expenses. Generally non-manufacturing costs are further classified into two categories:

1. Marketing and Selling Costs.
2. Administrative Costs.

Biological cost

In biology, the *biological cost* or *metabolic price* is a measure of the increased energy metabolism that is required to achieve a function. Drug resistance in microbiology, for instance, has a very high metabolic price, especially for antibiotic resistance

Chapter 2

Value Engineering

Value engineering (VE) is a systematic method to improve the "value" of goods or products and services by using an examination of function. Value, as defined, is the ratio of function to cost. Value can therefore be increased by either improving the function or reducing the cost. It is a primary tenet of value engineering that basic functions be preserved and not be reduced as a consequence of pursuing value improvements.

In the United States, value engineering is specifically spelled out in Public Law 104-106, which states "Each executive agency shall establish and maintain cost-effective value engineering procedures and processes."

Value engineering is sometimes taught within the project management or industrial engineering body of knowledge as a technique in which the value of a system's outputs is optimized by crafting a mix of performance (function) and costs. In most cases this practice identifies and removes unnecessary expenditures, thereby increasing the value for the manufacturer and/or their customers.

VE follows a structured thought process that is based exclusively on "function", i.e. what something "does" not what it is. For example a screw driver that is being used to stir a can of paint has a "function" of mixing the contents of a paint can and not the original connotation of securing a screw into a screw-hole. In value engineering "functions" are always described in a two word abridgment consisting of an active verb and measurable noun (what is being done - the verb - and what it is being done to - the noun) and to do so in the most non-prescriptive way possible. In the screw driver and can of paint example, the most basic function would be "blend liquid" which is less prescriptive than "stir paint" which can be seen to limit the action (by stirring) and to limit the application (only considers paint.) This is the basis of what value engineering refers to as "function analysis".

Value engineering uses rational logic (a unique "how" - "why" questioning technique) and the analysis of function to identify relationships that increase value. It is considered a

quantitative method similar to the scientific method, which focuses on hypothesis-conclusion approaches to test relationships, and operations research, which uses model building to identify predictive relationships.

Value engineering is also referred to as "value management" or "value methodology" (VM), and "value analysis" (VA). VE is above all a structured problem solving process based on function analysis—understanding something with such clarity that it can be described in two words, the active verb and measurable noun abridgement. For example, the function of a pencil is to "make marks". This then facilitates considering what else can make marks. From a spray can, lipstick, a diamond on glass to a stick in the sand, one can then clearly decide upon which alternative solution is most appropriate.

Origins

Value engineering began at General Electric Co. during World War II. Because of the war, there were shortages of skilled labour, raw materials, and component parts. Lawrence Miles, Jerry Leftow, and Harry Erlicher at G.E. looked for acceptable substitutes. They noticed that these substitutions often reduced costs, improved the product, or both. What started out as an accident of necessity was turned into a systematic process. They called their technique “value analysis”.

The Job Plan

Value engineering is often done by systematically following a multi-stage job plan. Larry Miles' original system was a six-step procedure which he called the "value analysis job plan." Others have varied the job plan to fit their constraints. Depending on the application, there may be four, five, six, or more stages. One modern version has the following eight steps:

1. Preparation
2. Information
3. Analysis
4. Creation
5. Evaluation
6. Development
7. Presentation
8. Follow-up

Four basic steps in the job plan are:

- Information gathering - This asks what the requirements are for the object. Function analysis, an important technique in value engineering, is usually done in this initial stage. It tries to determine what functions or performance characteristics are important. It asks questions like; What does the object do? What must it do? What should it do? What could it do? What must it not do?

- Alternative generation (creation) - In this stage value engineers ask; What are the various alternative ways of meeting requirements? What else will perform the desired function?
- Evaluation - In this stage all the alternatives are assessed by evaluating how well they meet the required functions and how great will the cost savings be.
- Presentation - In the final stage, the best alternative will be chosen and presented to the client for final decision.

How it works

VE follows a structured thought process to evaluate options as follows.

Gather information

1. What is being done now?

Who is doing it?
 What could it do?
 What must it not do?

Measure

2. How will the alternatives be measured?

What are the alternate ways of meeting requirements?
 What else can perform the desired function?

Analyze

3. What must be done?

What does it cost?

Generate

4. What else will do the job?

Evaluate

5. Which Ideas are the best?

6. Develop and expand ideas

What are the impacts?
 What is the cost?
 What is the performance?

7.Present ideas

Sell alternatives

Chapter 3

Cost Contingency & Cost Escalation

Cost Contingency

When estimating the cost for a project, product or other item or investment, there is always uncertainty as to precise content of all items in the estimate, how work will be performed, what work conditions will be like when the project is executed and so on. These uncertainties are risks to the project. Some refer to these risks as "known-unknowns" because the estimator is aware of them, and based on past experience, can even estimate their probable costs. The estimated costs of the known-unknowns is referred to by cost estimators as cost contingency.

AACE International, the Association for the Advancement of Cost engineering, has defined contingency as "An amount added to an estimate to allow for items, conditions, or events for which the state, occurrence, or effect is uncertain and that experience shows will likely result, in aggregate, in additional costs. Typically estimated using statistical analysis or judgment based on past asset or project experience. Contingency usually excludes:

1. Major scope changes such as changes in end product specification, capacities, building sizes, and location of the asset or project;
2. Extraordinary events such as major strikes and natural disasters;
3. Management reserves; and
4. Escalation and currency effects.

Some of the items, conditions, or events for which the state, occurrence, and/or effect is uncertain include, but are not limited to, planning and estimating errors and omissions, minor price fluctuations other than general escalation), design developments and changes within the scope, and variations in market and environmental conditions. Contingency is generally included in most estimates, and **is expected to be expended**".

A key phrase above is that it is "*expected to be expended*". In other words, it is an item in an estimate like any other, and should be estimated and included in every estimate and every budget. Because management often thinks contingency money is "fat" that is not needed if a project team does its job well, it is a controversial topic.

In general, there are four classes of methods used to estimate contingency. . " These include the following:

1. Expert judgment
2. Predetermined guidelines (with varying degrees of judgment and empiricism used)
3. Simulation analysis (primarily risk analysis judgment incorporated in a simulation such as Monte-Carlo)
4. Parametric Modeling (empirically-based algorithm, usually derived through regression analysis, with varying degrees of judgment used).

A fifth method, called Reference Class Forecasting, was recently developed, based on prospect theory and theories of the planning fallacy, that won the Nobel Prize in economics 2002.

While all are valid methods, the method chosen should be consistent with the first principles of risk management in that the method must start with risk identification, and only then are the probable cost of those risks quantified. In best practice, the quantification will be probabilistic in nature (Monte-Carlo is a common method used for quantification).

Typically, the method results in a distribution of possible cost outcomes for the project, product, or other investment. From this distribution, a cost value can be selected that has the desired probability of having a cost underrun or cost overrun. Usually a value is selected with equal chance of over or underrunning. The difference between the cost estimate without contingency, and the selected cost from the distribution is contingency. For more information, AACE International has catalogued many professional papers on this complex topic .

Contingency is included in budgets as a control account. As risks occur on a project, and money is needed to pay for them, the contingency can be transferred to the appropriate accounts that need it. The transfer and its reason is recorded. In risk management, risks are continually reassessed during the course of a project, as are the needs for cost contingency.

Cost Escalation

Cost escalation is defined as changes in the cost or price of specific goods or services in a given economy over a period of time. This is similar to the concepts of inflation and deflation except that escalation is specific to an item or class of items (not as general in nature), it is often not primarily driven by changes in the money supply, and it tends to be

less sustained. While escalation includes general inflation related to the money supply, it is also driven by changes in technology, practices, and particularly supply-demand imbalances that are specific to a good or service in a given economy. For example, while general inflation (e.g., consumer price index) in the US was less than 5% in the 2003-2007 time period, steel prices increased (escalated) by over 50% because of supply-demand imbalance. Cost escalation may contribute to a project cost overrun but it is not synonymous with it.

Over long periods of time, as market supply and demand imbalances are corrected, escalation will tend to more-or-less equal inflation unless there are sustained technology or efficiency changes in a market.

Escalation is usually calculated by examining the changes in price index measures for a good or service. Future escalation can be forecast using econometrics. Unfortunately, because escalation (unlike inflation) may occur in a micro-market, and it may be hard to measure with surveys, indices can be difficult to find. For example, the Bureau of Labor Statistics has a price index for construction wages and compensation (what the construction contractor's labor cost), but has none for the prices that owners must pay the construction contractor for their services.

In cost engineering and project management usage, escalation and cost contingency are both considered risk funds, that should be included in project estimates and budgets. When escalation is minimal, it is sometimes estimated together with contingency. However, this is not a best practice, particularly when escalation is significant.

Chapter 4

Quantity Surveyor

A **quantity surveyor (QS)** is a professional working within the construction industry concerned with building costs.

The profession is one that provides a qualification gained following formal education, specific training and experience that provides a general set of skills that are then applied to a diverse variety of problems. Predominantly these relate to costs and contracts on construction projects. Other areas in which QS find employment include property surveys for hidden defects on behalf of potential purchasers, running estates, valuing the mineral deposits for mining companies, selling property and even Leasehold Reform Act work.

There are around 75,000 professional QSs working in the UK.

History

The profession developed during the 19th century from the earlier "measurer", a specialist tradesman (often a guild member), who prepared standardised schedules for a building project in which all of the construction materials, labour activities and the like were quantified, and against which competing builders could submit priced tenders. Because all tenders were based on the same schedule of information, they could be easily compared so as to identify the best one. As a profession quantity surveying emerged around the 1820s with one of the earliest QSs being Sir Henry Arthur Hunt who was involved in work on the Houses of Parliament. After the fire in 1834 that destroyed the old Palace of Westminster Henry Hunt came up with an estimate cost of £724,984 (changes by Parliament put it up to £1.5m).

The professional institution with which most English-speaking quantity surveyors are affiliated is the UK-based Royal Institution of Chartered Surveyors (RICS). In Australia, the Australian Institute of Quantity Surveyors (AIQS) has over 4300 members, both domestically and overseas. Others are the Chartered Institute of Building (CIOB) and

Institution of Civil Engineering Surveyors (ICES). Those who are qualified members of the RICS are entitled to use the term "Chartered Quantity Surveyor" or simply "Chartered Surveyor".

The QS usually reports to Project Manager or Project Director and provides advice in the decision-making process throughout the management of a project from initial inception to final completion. The QS handles estimating and cost control, the tendering process and, after contract award, the commercial interface. Qs should be able to carry out estimating and measurement of construction works prior to tender, producing the bill of quantities; produce tender documentation and manage the tender process; clarify and evaluate tenders; and manage the resultant contract through monthly valuations, variations control, contract administration and assessment of claims.

Some Qs are trained in techniques of cost control. Those Qs who emphasise the cost discipline often use the term "Construction Cost Consultant". They ensure that projects are designed and constructed in such a manner as to secure value for money, cost certainty and programme dates.

Others emphasise contracts management. Trained to draft, interpret and administer complex contracts, those Qs who operate in the broader field of project management often adopt other titles such as "Contracts manager" or "Construction surveyor". A number of Qs work in procurement in the oil & gas industry, process and power industries, and civil engineering. Their preferred title, in countries where the QS profession is less known, is "Contracts engineer".

Some Qs specialise in project management and running multi-disciplinary projects, the QS background being a good foundation for understanding the complexities of modern large-scale projects.

As well as in professional quantity surveying practices, the QS finds employment in all parts of industry and government including primary and secondary industry, national and local government bodies and agencies, contractors and subcontractors, developers, and financial and legal companies.

Although all Qs will have followed a similar course of education and training (for those entering the profession today, this is usually to degree level), there are many areas of specialisation in which a QS may concentrate. The main distinction amongst Qs is between those who carry out work on behalf of a client organisation, often known as a "professional quantity surveyor", and those who work for construction companies, often known as a "main contractor's quantity surveyor".

The functions of a consultant quantity surveyor

Traditionally referred to as a Professional Quantity Surveyor or Private Practice Quantity Surveyor they are broadly concerned with contracts and costs on construction projects. The methods employed, however, cover a range of activities which may include cost

planning, value management, feasibility study|feasibility studies, cost benefit analysis, lifecycle costing, tendering, valuation, change control, dispute resolution, claims management and cost estimation.

The QS's traditional independent role on the team comprising client, architect, engineer, QS and contractor has given him a reputation and appreciation for fairness. This, combined with his expertise in drafting and interpretation of contract documents, enables him to settle issues, avoid disputes and ensure the effective progress of a project.

Quantity surveyors control construction costs by accurate measurement of the work required, the application of expert knowledge of costs and prices of work, labour, materials and plant required, an understanding of the implications of design decisions at an early stage to ensure that good value is obtained for the money to be expended.

The technique of measuring quantities from drawings, sketches and specifications prepared by designers, principally architects and engineers, in order to prepare tender/contract documents, is known in the industry as taking off. The quantities of work taken off typically are used to prepare bills of quantities, which usually are prepared in accordance with a published standard method of measurement (SMM) as agreed to by the QS profession and representatives of the construction industry.

In Australia, the Australian Institute of Quantity Surveyors (AIQS) is the peak body for the Quantity Surveying profession. All Members of the AIQS are assessed for membership against strict criteria, and must adhere to a Code of Ethics and Code of Professional Conduct to ensure the highest standards of professional excellence.

The benchmark for quality for a Private Practice Surveyor is the RICS's Chartered Membership MRICS & FRICS. The RICS also has the entry level non Chartered Membership AssocRICS. AssocRICS acts as a qualification in its own right however also offers a progressive route to Chartered RICS membership for able and willing candidates.

Contractor's quantity surveyor

A contractor's QS is responsible for the performance of operations that mirror those of the owner's QS; i.e., the measurement and pricing of construction work, but specifically that actually performed by the contractor (and the contractor's subcontractors) as opposed to the construction work described and measured in the construction contract between the owner and the contractor. Such a difference in quantity of work may arise from changes required by an owner, or by an architect or engineer on an owner's behalf. Typically, the settlement of a change (often referred to in a contract as a 'variation').

The role of a contractor's QS will extend further than the day to day running of building projects and will cover such other areas as sub-contract formation, forecasting of costs and values of the project, cash flow forecasts and the collation of the operation and maintenance manuals of the project (O&M manuals). This increase in the capacity of the

surveying profession has led to an increased demand for qualified personnel and goes some way to explaining the popularity of related degrees at university.

Some contractors and others may attempt to rely on a general accountant to deal with construction costs, but usually this is not effective, primarily because an accountant does not have the technical knowledge to accurately allocate costs to specific items of work performed, especially at times prior to the particular work's completion as required to make accurate assessment of the amounts to be paid to the contractor during the course of the work

Chapter 5

Cost-Benefit Analysis

Cost benefit analysis is a term that refers both to:

- helping to appraise, or assess, the case for a project, programme or policy proposal;
- an approach to making economic decisions of any kind.

Under both definitions the process involves, whether explicitly or implicitly, weighing the total expected costs against the total expected benefits of one or more actions in order to choose the best or most profitable option. The formal process is often referred to as either CBA (Cost-Benefit Analysis) or BCA (Benefit-Cost Analysis).

Benefits and costs are often expressed in money terms, and are adjusted for the time value of money, so that all flows of benefits and flows of project costs over time (which tend to occur at different points in time) are expressed on a common basis in terms of their “present value.” Closely related, but slightly different, formal techniques include cost-effectiveness analysis, economic impact analysis, fiscal impact analysis and Social Return on Investment (SROI) analysis. The latter builds upon the logic of cost-benefit analysis, but differs in that it is explicitly designed to inform the practical decision-making of enterprise managers and investors focused on optimizing their social and environmental impacts.

Theory

Cost–benefit analysis is often used by governments to evaluate the desirability of a given intervention. It is heavily used in today's government. It is an analysis of the cost effectiveness of different alternatives in order to see whether the benefits outweigh the costs. The aim is to gauge the efficiency of the intervention relative to the status quo. The costs and benefits of the impacts of an intervention are evaluated in terms of the public's willingness to pay for them (benefits) or willingness to pay to avoid them (costs). Inputs are typically measured in terms of opportunity costs - the value in their best alternative

use. The guiding principle is to list all parties affected by an intervention and place a monetary value of the effect it has on their welfare as it would be valued by them.

The process involves monetary value of initial and ongoing expenses vs. expected return. Constructing plausible measures of the costs and benefits of specific actions is often very difficult. In practice, analysts try to estimate costs and benefits either by using survey methods or by drawing inferences from market behavior. For example, a product manager may compare manufacturing and marketing expenses with projected sales for a proposed product and decide to produce it only if he expects the revenues to eventually recoup the costs. Cost–benefit analysis attempts to put all relevant costs and benefits on a common temporal footing. A discount rate is chosen, which is then used to compute all relevant future costs and benefits in present-value terms. Most commonly, the discount rate used for present-value calculations is an interest rate taken from financial markets. This can be very controversial; for example, a high discount rate implies a very low value on the welfare of future generations, which may have a huge impact on the desirability of interventions to help the environment. Empirical studies suggest that in reality, people's discount rates *do* decline over time. Because cost–benefit analysis aims to measure the public's true willingness to pay, this feature is typically built into studies.

During cost–benefit analysis, monetary values may also be assigned to less tangible effects such as the various risks that could contribute to partial or total project failure, such as loss of reputation, market penetration, or long-term enterprise strategy alignments. This is especially true when governments use the technique, for instance to decide whether to introduce business regulation, build a new road, or offer a new drug through the state healthcare system. In this case, a value must be put on human life or the environment, often causing great controversy. For example, the cost–benefit principle says that we should install a guardrail on a dangerous stretch of mountain road if the dollar cost of doing so is less than the implicit dollar value of the injuries, deaths, and property damage thus prevented.

Cost–benefit calculations typically involve using time value of money formulas. This is usually done by converting the future expected streams of costs and benefits into a present value amount.

Application and history

The practice of cost–benefit analysis differs between countries and between sectors (e.g., transport, health) within countries. Some of the main differences include the types of impacts that are included as costs and benefits within appraisals, the extent to which impacts are expressed in monetary terms, and differences in the discount rate between countries. Agencies across the world rely on a basic set of key cost–benefit indicators, including the following:

- NPV (net present value)
- PVB (present value of benefits)
- PVC (present value of costs)

- BCR (benefit cost ratio = PVB / PVC)
- Net benefit (= $PVB - PVC$)
- NPV/k (where k is the level of funds available)

The concept of CBA dates back to an 1848 article by Dupuit and was formalized in subsequent works by Alfred Marshall. The practical application of CBA was initiated in the US by the Corps of Engineers, after the Federal Navigation Act of 1936 effectively required cost–benefit analysis for proposed federal waterway infrastructure. The Flood Control Act of 1939 was instrumental in establishing CBA as federal policy. It specified the standard that "the benefits to whomever they accrue [be] in excess of the estimated costs.

Subsequently, cost–benefit techniques were applied to the development of highway and motorway investments in the US and UK in the 1950s and 1960s. An early and often-quoted, more developed application of the technique was made to London Underground's Victoria Line. Over the last 40 years, cost–benefit techniques have gradually developed to the extent that substantial guidance now exists on how transport projects should be appraised in many countries around the world.

In the UK, the New Approach to Appraisal (NATA) was introduced by the then Department for Transport, Environment and the Regions. This brought together cost–benefit results with those from detailed environmental impact assessments and presented them in a balanced way. NATA was first applied to national road schemes in the 1998 Roads Review but subsequently rolled out to all modes of transport. It is now a cornerstone of transport appraisal in the UK and is maintained and developed by the Department for Transport.

The EU's 'Developing Harmonised European Approaches for Transport Costing and Project Assessment' (HEATCO) project, part of its Sixth Framework Programme, has reviewed transport appraisal guidance across EU member states and found that significant differences exist between countries. HEATCO's aim is to develop guidelines to harmonise transport appraisal practice across the EU.

Transport Canada has also promoted the use of CBA for major transport investments since the issuance of its Guidebook in 1994.

More recent guidance has been provided by the United States Department of Transportation and several state transportation departments, with discussion of available software tools for application of CBA in transportation, including HERS, BCA.Net, StatBenCost, CalBC, and TREDIS. Available guides are provided by the Federal Highway Administration, Federal Aviation Administration, Minnesota Department of Transportation, California Department of Transportation (Caltrans), and the Transportation Research Board Transportation Economics Committee .

In the early 1960s, CBA was also extended to assessment of the relative benefits and costs of healthcare and education in works by Burton Weisbrod. Later, the United States Department of Health and Human Services issued its CBA Guidebook. .

Accuracy problems

The accuracy of the outcome of a cost–benefit analysis depends on how accurately costs and benefits have been estimated.

A peer-reviewed study of the accuracy of cost estimates in transportation infrastructure planning found that for rail projects actual costs turned out to be on average 44.7 percent higher than estimated costs, and for roads 20.4 percent higher (Flyvbjerg, Holm, and Buhl, 2002). For benefits, another peer-reviewed study found that actual rail ridership was on average 51.4 percent lower than estimated ridership; for roads it was found that for half of all projects estimated traffic was wrong by more than 20 percent (Flyvbjerg, Holm, and Buhl, 2005). Comparative studies indicate that similar inaccuracies apply to fields other than transportation. These studies indicate that the outcomes of cost–benefit analyses should be treated with caution because they may be highly inaccurate. Inaccurate cost–benefit analyses likely to lead to inefficient decisions, as defined by Pareto and Kaldor-Hicks efficiency. These outcomes (almost always tending to underestimation *unless significant new approaches are overlooked*) are to be expected because such estimates:

1. Rely heavily on past like projects (often differing markedly in function or size and certainly in the skill levels of the team members)
2. Rely heavily on the project's members to identify (*remember* from their collective past experiences) the significant cost drivers
3. Rely on very crude heuristics to estimate the money cost of the intangible elements
4. Are unable to completely dispel the usually unconscious biases of the team members (who often have a vested interest in a decision to go ahead) and the natural psychological tendency to "*think positive*" (whatever that involves)

Reference class forecasting was developed to increase accuracy in estimates of costs and benefits.

Another challenge to cost–benefit analysis comes from determining which costs should be included in an analysis (the significant cost drivers). This is often controversial because organizations or interest groups may think that some costs should be included or excluded from a study.

In the case of the Ford Pinto (where, because of design flaws, the Pinto was liable to burst into flames in a rear-impact collision), the Ford company's decision was not to issue a recall. Ford's cost–benefit analysis had estimated that based on the number of cars in use and the probable accident rate, deaths due to the design flaw would run about \$49.5 million (the amount Ford would pay out of court to settle wrongful death lawsuits). This

was estimated to be less than the cost of issuing a recall (\$137.5 million) [17]. In the event, Ford overlooked (or considered insignificant) the costs of the negative publicity so engendered, which turned out to be quite significant (because it led to the recall anyway *and* to measurable losses in sales).

In the field of health economics, some analysts think cost–benefit analysis can be an inadequate measure because willingness-to-pay methods of determining the value of human life can be subject to bias according to income inequity. They support use of variants such as cost-utility analysis and quality-adjusted life year to analyze the effects of health policies.

Use in regulation

Cost-benefit analysis was widely in the United States under the Bush administration to prevent regulatory initiatives, and there is some debate about whether it is neutral to regulatory initiatives or whether it anti-regulatory and undervalues human life, health, and the environment. In the case of environmental and occupational health regulation, it has been argued that if modern cost-benefit analyses had been applied prospectively to proposed regulations such as removing lead from gasoline, not turning the Grand Canyon into a hydroelectric dam, and regulating workers' exposure to vinyl chloride, these regulations would not have been implemented even though they are considered to be highly successful in retrospect. The Clean Air Act has been cited in retrospective studies as a case where benefits exceeded costs, but the knowledge of the benefits (attributable largely to the benefits of reducing particulate pollution) was not available until many years later.

Chapter 6

Cost-Effectiveness Analysis & Cost Overrun

Cost-Effectiveness Analysis

Cost-effectiveness analysis (CEA) is a form of economic analysis that compares the relative costs and outcomes (effects) of two or more courses of action. Cost-effectiveness analysis is distinct from cost-benefit analysis, which assigns a monetary value to the measure of effect. Cost-effectiveness analysis is often used in the field of health services, where it may be inappropriate to monetize health effect. Typically the CEA is expressed in terms of a ratio where the denominator is a gain in health from a measure (years of life, premature births averted, sight-years gained) and the numerator is the cost associated with the health gain. The most commonly used outcome measure is quality-adjusted life years (QALY). Cost-utility analysis is similar to cost-effectiveness analysis.

General application

The concept of cost effectiveness is applied to the planning and management of many types of organized activity. In the acquisition of military tanks, for example, competing designs are compared not only for purchase price, but also for such factors as their operating radius, top speed, rate of fire, armor protection, and caliber and armor penetration of their guns. If a tank's performance in these areas is equal or even slightly inferior to its competitor, but substantially less expensive and easier to produce, military planners may select it as more cost effective than the competitor. Conversely, if the difference in price is near zero, but the more costly competitor would convey an enormous battlefield advantage through special ammunition, radar fire control and laser range finding, enabling it to destroy enemy tanks accurately at extreme ranges, military planners may choose it instead—based on the same cost effectiveness principle.

Cost effectiveness analysis is also applied to many other areas of human activity, including the economics of automobile usage.

CEA in pharmacoeconomics

In the context of pharmacoeconomics, the cost-effectiveness of a therapeutic or preventive intervention is the ratio of the cost of the intervention to a relevant measure of its effect. Cost refers to the resource expended for the intervention, usually measured in monetary terms such as dollars or pounds. The measure of effects depends on the intervention being considered. Examples include the number of people cured of a disease, the mm Hg reduction in diastolic blood pressure and the number of symptom-free days experienced by a patient. The selection of the appropriate effect measure should be based on clinical judgement in the context of the intervention being considered.

A special case of CEA is cost-utility analysis, where the effects are measured in terms of years of full health lived, using a measure such as quality-adjusted life years or disability-adjusted life years.

Cost-effectiveness is typically expressed as an incremental cost-effectiveness ratio (ICER), the ratio of change in costs to the change in effects.

A complete compilation of cost-utility analyses in the peer reviewed medical literature is available from the Cost-Effectiveness Analysis Registry website.

A 1995 study of the cost-effectiveness of over 500 life-saving medical interventions found that the median cost per intervention was \$42,000 per life-year saved. A 2006 systematic review found that industry-funded studies often concluded with cost effective ratios below \$20,000 per QALY and low quality studies and those conducted outside the US and EU were less likely to be below this threshold. While the two conclusions of this may indicate that industry-funded ICER measures are lower methodological quality than those published by non-industry sources, there is also a possibility that, due to the nature of retrospective or other non-public work, publication bias may exist rather than methodology biases. There may be incentive for an organization not to develop or publish an analysis that does not demonstrate the value of their product. Additionally, peer reviewed journal articles should have a strong and defensible methodology, as that is the expectation of the peer-review process.

Cost Overrun

A **cost overrun**, also known as a **cost increase** or **budget overrun**, is an unexpected cost incurred in excess of a budgeted amount due to an under-estimation of the actual cost during budgeting. Cost overrun should be distinguished from cost escalation, which is used to express an *anticipated* growth in a budgeted cost due to factors such as inflation.

Cost overrun is common in infrastructure, building, and technology projects. A comprehensive study of cost overrun published in the Journal of the American Planning Association in 2002 found that 9 out of ten construction projects had underestimated costs. Overruns of 50 to one hundred percent were common. Cost underestimation was found in each of 20 nations and five continents covered by the study, and cost

underestimation had not decreased in the 70 years for which data were available. For IT projects, an industry study by the Standish Group found that the average cost overrun was 43 percent; 71 percent of projects were over budget, exceeded time estimates, and had estimated too narrow a scope; and total waste was estimated at \$55 billion per year in the US alone.

Many major construction projects have incurred cost overruns. The Suez Canal cost 20 times as much as the earliest estimates; even the cost estimate produced the year before construction began underestimated the project's actual costs by a factor of three. The Sydney Opera House cost 15 times more than was originally projected, and the Concorde supersonic aeroplane cost 12 times more than predicted. When Boston's "Big Dig" tunnel construction project was completed, the project was 275 percent (\$11 billion) over budget. The Channel Tunnel between the UK and France had a construction cost overrun of 80 percent, and a 140-percent financing cost overrun.

Causes

Three types of explanation for cost overrun exist: technical, psychological, and political-economic. Technical explanations account for cost overrun in terms of imperfect forecasting techniques, inadequate data, etc. Psychological explanations account for overrun in terms of optimism bias with forecasters. Finally, political-economic explanations see overrun as the result of strategic misrepresentation of scope or budgets.

All three explanations can be considered forms of risk. A project's budgeted costs should always include cost contingency funds to cover risks (other than scope changes imposed on the project). As has been shown in cost engineering research, poor risk analysis and contingency estimating practices account for many project cost overruns. Numerous studies have found that the greatest cause of cost growth was poorly-defined scope at the time that the budget was established. The *cost growth*, or overrun of the budget before cost contingency is added, can be predicted by rating the extent of scope definition, even on complex projects with new technology.

Professor Bent Flyvbjerg of Oxford University believes that big public-works projects almost always have cost overruns due to strategic misrepresentation—"that is, lying", as he defines the term.

Cost overrun is typically calculated in one of two ways: either as a percentage, namely actual cost minus budgeted cost, in percent of budgeted cost; or as a ratio of actual cost divided by budgeted cost. For example, if the budget for building a new bridge was \$100 million, and the actual cost was \$150 million, then the cost overrun may be expressed by the ratio 1.5, or as 50 percent.

Reference class forecasting was developed to eliminate or reduce cost overrun.

Chapter 7

Opportunity Cost & Cost of Capital

Opportunity Cost

Opportunity cost is the cost related to the second best choice available to someone who has picked among several mutually exclusive choices. It is a key concept in economics. It has been described as expressing "the basic relationship between scarcity and choice." The notion of opportunity cost plays a crucial part in ensuring that scarce resources are used efficiently. Thus, opportunity costs are not restricted to monetary or financial costs: the real cost of output forgone, lost time, pleasure or any other benefit that provides utility should also be considered opportunity costs.

The concept of an opportunity cost was first developed by John Stuart Mill.

Non-monetary applications

Opportunity cost is assessed in not only monetary or material terms, but also in terms of anything which is of value. For example, a person who desires to watch each of two television programs being broadcast simultaneously, and does not have the means to make a recording of one, can watch only one of the desired programs. Therefore, the opportunity cost of watching *Dallas* could be not enjoying the other program i.e. *Dynasty*. Of course, if an individual records one program while watching the other, the opportunity cost will be the time that the individual spends watching one program versus the other. In a restaurant situation, the opportunity cost of eating steak could be trying the salmon. For the diner, the opportunity cost of ordering both meals could be twofold - the extra \$20 to buy the second meal, and his reputation with his peers, as he may be thought gluttonous or extravagant for ordering two meals. A family might decide to use a short period of vacation time to visit Disneyland rather than doing household improvements. The opportunity cost of having happier children could therefore be a remodeled bathroom.

In environmental protection, the opportunity cost is also applicable. This has been demonstrated in the legislation that required the carcinogenic aromatics (mainly reformate) to be largely eliminated from gasoline. Unfortunately, this required refineries to install equipment at a cost of hundreds of millions of dollars - and pass the cost to the consumer. The absolute number of cancer cases following from exposure to gasoline, however, is low, estimated a few cases per year in the U.S. Thus, the decision to legally require less aromatics has been criticized on the grounds of opportunity cost: the hundreds of millions spent on process redesign could have been spent on other, more fruitful ways of reducing deaths caused by cancer or automobiles. These actions (or strictly, the best one of them) are the opportunity cost of reduction of aromatics in gasoline.

Evaluation

The consideration of opportunity costs is one of the key differences between the concepts of economic cost and accounting cost. Assessing opportunity costs is fundamental to assessing the true cost of any course of action. In the case where there is no explicit accounting or monetary cost (price) attached to a course of action, or the explicit accounting or monetary cost is low, then, ignoring opportunity costs may produce the illusion that its benefits cost nothing at all. The unseen opportunity costs then become the implicit hidden costs of that course of action.

Note that opportunity cost is not the sum of the available alternatives when those alternatives are, in turn, mutually exclusive to each other. The opportunity cost of the city's decision to build the hospital on its vacant land is the loss of the land for a sporting center, or the inability to use the land for a parking lot, or the money which could have been made from selling the land, as use for any one of those purposes would preclude the possibility to implement any of the other. It is also the cost of the forgone products after making a choice.

However, most opportunities are difficult to compare. Opportunity cost has been seen as the foundation of the marginal theory of value as well as the theory of time and money. In some cases it may be possible to have more of everything by making different choices; for instance, when an economy is within its production possibility frontier. In microeconomic models this is unusual, because individuals are assumed to maximise utility, but it is a feature of Keynesian macroeconomics. In these circumstances opportunity cost is a less useful concept.

In a 2005 survey at the annual meeting of American Economic Association, 21.6% of professional economists surveyed chose the correct answer to a question on opportunity cost. The researchers later asked a similar but differently phrased question, to which a majority of the economists surveyed gave an incorrect answer. When the researchers posed the original question to a larger group of college students, 7.4% of those who had taken a course in economics answered correctly, compared to 17.2% of those who had never taken one. The researchers, Paul J. Ferraro and Laura O. Taylor of Georgia State University, labeled the results "a dismal performance from the dismal science."

Cost of Capital

The **cost of capital** is a term used in the field of financial investment to refer to the cost of a company's funds (both debt and equity), or, from an investor's point of view "the shareholder's required return on a portfolio of all the company's existing securities". It is used to evaluate new projects of a company as it is the minimum return that investors expect for providing capital to the company, thus setting a benchmark that a new project has to meet.

Summary

For an investment to be worthwhile, the expected return on capital must be greater than the cost of capital. The cost of capital is the rate of return that capital could be expected to earn in an alternative investment of equivalent risk. If a project is of similar risk to a company's average business activities it is reasonable to use the company's average cost of capital as a basis for the evaluation. A company's securities typically include both debt and equity, one must therefore calculate both the cost of debt and the cost of equity to determine a company's cost of capital. However, a rate of return larger than the cost of capital is usually required.

The **cost of debt** is relatively simple to calculate, as it is composed of the rate of interest paid. In practice, the interest-rate paid by the company can be modelled as the risk-free rate plus a risk component (risk premium), which itself incorporates a probable rate of default (and amount of recovery given default). For companies with similar risk or credit ratings, the interest rate is largely exogenous (not linked to the company's activities).

The **cost of equity** is more challenging to calculate as equity does not pay a set return to its investors. Similar to the cost of debt, the cost of equity is broadly defined as the risk-weighted projected return required by investors, where the return is largely unknown. The cost of equity is therefore *inferred* by comparing the investment to other investments (comparable) with similar risk profiles to determine the "market" cost of equity. It is commonly equated using the CAPM formula (below), although articles such as Stulz 1995 question the validity of using a local CAPM versus an international CAPM- also considering whether markets are fully integrated or segmented (if fully integrated, there would be no need for a local CAPM).

Once cost of debt and cost of equity have been determined, their blend, the weighted-average cost of capital (WACC), can be calculated. This WACC can then be used as a discount rate for a project's projected cash flows.

Cost of debt

The cost of debt is computed by taking the rate on a risk free bond whose duration matches the term structure of the corporate debt, then adding a default premium. This default premium will rise as the amount of debt increases (since, *ceteris paribus*, "all other

things being equal", the risk rises as the amount of debt rises). Since in most cases debt expense is a deductible expense, the cost of debt is computed as an after tax cost to make it comparable with the cost of equity (earnings are after-tax as well). Thus, for profitable firms, debt is discounted by the tax rate. The formula can be written as **(Rf + credit risk rate)(1-T)**, where T is the corporate tax rate and Rf is the risk free rate.

The yield to maturity can be used as cost of capital.

Cost of equity

Cost of equity = Risk free rate of return + Premium expected for risk

Cost of equity = Risk free rate of return + Beta x (market rate of return- risk free rate of return) Where Beta= sensitivity to movements in the relevant market:

$$E_s = R_f + \beta_s(R_m - R_f).$$

Where:

E_s

The expected return for a security

R_f

The expected risk-free return in that market (government bond yield)

β_s

The sensitivity to market risk for the security

R_M

The historical return of the stock market/ equity market

$(R_M - R_f)$

The risk premium of market assets over risk free assets.

The risk free rate is taken from the lowest yielding bonds in the particular market, such as government bonds.

Expected return

The expected return (or required rate of return for investors) can be calculated with the "dividend capitalization model", which is

$$K_{cs} = \frac{\text{Dividend}_{\text{Payment/Share}}}{\text{Price}_{\text{Market}}} + \text{Growth}_{\text{rate}}.$$

Comments

The models state that investors will expect a return that is the risk-free return plus the security's sensitivity to market risk times the market risk premium.

The risk premium varies over time and place, but in some developed countries during the

twentieth century it has averaged around 5%. The equity market real capital gain return has been about the same as annual real GDP growth. The capital gains on the Dow Jones Industrial Average have been 1.6% per year over the period 1910-2005. The dividends have increased the total "real" return on average equity to the double, about 3.2%.

The sensitivity to market risk (β) is unique for each firm and depends on everything from management to its business and capital structure. This value cannot be known "ex ante" (beforehand), but can be estimated from *ex post* (past) returns and past experience with similar firms.

Cost of retained earnings/cost of internal equity

Note that retained earnings are a component of equity, and therefore the cost of retained earnings (internal equity) is equal to the cost of equity as explained above. Dividends (earnings that are paid to investors and not retained) are a component of the return on capital to equity holders, and influence the cost of capital through that mechanism.

Weighted average cost of capital

The Weighted Average Cost of Capital (WACC) is used in finance to measure a firm's cost of capital.

The total capital for a firm is the value of its equity (for a firm without outstanding warrants and options, this is the same as the company's market capitalization) plus the cost of its debt (the cost of debt should be continually updated as the cost of debt changes as a result of interest rate changes). Notice that the "equity" in the debt to equity ratio is the market value of all equity, not the shareholders' equity on the balance sheet. To calculate the firm's weighted cost of capital, we must first calculate the costs of the individual financing sources: Cost of Debt Cost of Preference Capital Cost of Equity Capital

Calculation of WACC is an iterative procedure which requires estimation of the fair market value of equity capital.

Capital structure

Because of tax advantages on debt issuance, it will be cheaper to issue debt rather than new equity (this is only true for profitable firms, tax breaks are available only to profitable firms). At some point, however, the cost of issuing new debt will be greater than the cost of issuing new equity. This is because adding debt increases the default risk - and thus the interest rate that the company must pay in order to borrow money. By utilizing too much debt in its capital structure, this increased default risk can also drive up the costs for other sources (such as retained earnings and preferred stock) as well. Management must identify the "optimal mix" of financing - the capital structure where the cost of capital is minimized so that the firm's value can be maximized.

The Thomson Financial league tables show that global debt issuance exceeds equity issuance with a 90 to 10 margin. weighted average cost of capital

Modigliani-Miller theorem

If there were no tax advantages for issuing debt, and equity could be freely issued, Miller and Modigliani showed that, under certain assumptions, the value of a leveraged firm and the value of an unleveraged firm should be the same.

Chapter 8

Average Cost

In economics, **average cost** or **unit cost** is equal to total cost divided by the number of goods produced (the output quantity, Q). It is also equal to the sum of average variable costs (total variable costs divided by Q) plus average fixed costs (total fixed costs divided by Q). Average costs may be dependent on the time period considered (increasing production may be expensive or impossible in the short term, for example). Average costs affect the supply curve and are a fundamental component of supply and demand.

$$AC = \frac{TC}{Q}$$

Short-run average cost

Average cost is distinct from the price, and depends on the interaction with demand through elasticity of demand and elasticity of supply. In cases of perfect competition, price may be lower than average cost due to marginal cost pricing.

Short-run average cost will vary in relation to the quantity produced unless fixed costs are zero and variable costs constant. A cost curve can be plotted, with cost on the y-axis and quantity on the x-axis. Marginal costs are often shown on these graphs, with marginal cost representing the cost of the last unit produced at each point; marginal costs are the first derivative of total or variable costs.

A typical average cost curve will have a U-shape, because fixed costs are all incurred before any production takes place and marginal costs are typically increasing, because of diminishing marginal productivity. In this "typical" case, for low levels of production marginal costs are below average costs, so average costs are decreasing as quantity increases. An increasing marginal cost curve will intersect a U-shaped average cost curve at its minimum, after which point the average cost curve begins to slope upward. For further increases in production beyond this minimum, marginal cost is above average

costs, so average costs are increasing as quantity increases. An example of this typical case would be a factory designed to produce a specific quantity of widgets per period: below a certain production level, average cost is higher due to under-utilised equipment, while above that level, production bottlenecks increase the average cost.

Long-run average cost

The long run is a time frame in which the firm can vary the quantities used of all inputs, even physical capital. A long-run average cost curve can be upward sloping, downward sloping, or downward sloping at relatively low levels of output and upward sloping at relatively high levels of output, with an in-between level of output at which the slope of long-run average cost is zero. The typical long-run average cost curve is U-shaped, by definition reflecting increasing returns to scale where negatively-sloped and decreasing returns to scale where positively sloped.

If the firm is a perfect competitor in all input markets, and thus the per-unit prices of all its inputs are unaffected by how much of the inputs the firm purchases, then it can be shown that at a particular level of output, the firm has economies of scale (i.e., is operating in a downward sloping region of the long-run average cost curve) if and only if it has increasing returns to scale. Likewise, it has diseconomies of scale (is operating in an upward sloping region of the long-run average cost curve) if and only if it has decreasing returns to scale, and has neither economies nor diseconomies of scale if it has constant returns to scale. In this case, with perfect competition in the output market the long-run market equilibrium will involve all firms operating at the minimum point of their long-run average cost curves (i.e., at the borderline between economies and diseconomies of scale).

If, however, the firm is not a perfect competitor in the input markets, then the above conclusions are modified. For example, if there are increasing returns to scale in some range of output levels, but the firm is so big in one or more input markets that increasing its purchases of an input drives up the input's per-unit cost, then the firm could have diseconomies of scale in that range of output levels. Conversely, if the firm is able to get bulk discounts of an input, then it could have economies of scale in some range of output levels even if it has decreasing returns in production in that output range.

In some industries, the LRAC is always declining (economies of scale exist indefinitely). This means that the largest firm tends to have a cost advantage, and the industry tends naturally to become a monopoly, and hence is called a natural monopoly. Natural monopolies tend to exist in industries with high capital costs in relation to variable costs, such as water supply and electricity supply.

Long run average cost is the unit cost of producing a certain output when all inputs are variable. The behavioral assumption is that the firm will choose that combination of inputs that will produce the desired quantity at the lowest possible cost.

Relationship to marginal cost

When average cost is declining as output increases, marginal cost is less than average cost. When average cost is rising, marginal cost is greater than average cost. When average cost is neither rising nor falling (at a minimum or maximum), marginal cost equals average cost.

Other special cases for average cost and marginal cost appear frequently:

- Constant marginal cost/high fixed costs: each additional unit of production is produced at constant additional expense per unit. The average cost curve slopes down continuously, approaching marginal cost. An example may be hydroelectric generation, which has no fuel expense, limited maintenance expenses and a high up-front fixed cost (ignoring irregular maintenance costs or useful lifespan). Industries where fixed marginal costs obtain, such as electrical transmission networks, may meet the conditions for a natural monopoly, because once capacity is built, the marginal cost to the incumbent of serving an additional customer is always lower than the average cost for a potential competitor. The high fixed capital costs are a barrier to entry.
- Minimum efficient scale / maximum efficient scale: marginal or average costs may be non-linear, or have discontinuities. Average cost curves may therefore only be shown over a limited scale of production for a given technology. For example, a nuclear plant would be extremely inefficient (very high average cost) for production in small quantities; similarly, its maximum output for any given time period may essentially be fixed, and production above that level may be technically impossible, dangerous or extremely costly. The long run elasticity of supply will be higher, as new plants could be built and brought on-line.
- Zero fixed costs (long-run analysis) / constant marginal cost: since there are no economies of scale, average cost will be equal to the constant marginal cost.

Relationship between AC, AFC, AVC and MC

1. The Average Fixed Cost curve starts from a height and goes on declining continuously as production increases.

2. The Average Variable Cost curve, Average Cost curve and the Marginal Cost curve start from a height, reach the minimum points, then rise sharply and continuously.

3. The Average Fixed Cost curve approaches zero asymptotically. The Average Variable Cost curve is never parallel to or as high as the Average Cost curve due to the existence of positive Average Fixed Costs at all levels of production; but the Average Variable Cost curve asymptotically approaches the Average Cost curve from below.

4. The Marginal Cost curve always passes through the minimum points of the Average Variable Cost and Average Cost curves, though the Average Variable Cost curve attains the minimum point prior to that of the Average Cost curve.

Chapter 9

Cost Accounting

In management accounting, **cost accounting** establishes budget and actual cost of operations, processes, departments or product and the analysis of variances, profitability or social use of funds. Managers use cost accounting to support decision-making to cut a company's costs and improve profitability. As a form of management accounting, cost accounting need not follow standards such as GAAP, because its primary use is for internal managers, rather than outside users, and what to compute is instead decided pragmatically.

Costs are measured in units of nominal currency by convention. Cost accounting can be viewed as translating the supply chain (the series of events in the production process that, in concert, result in a product) into financial values.

There are various managerial accounting approaches:

- standardized or standard cost accounting
- lean accounting
- activity-based costing
- resource consumption accounting
- throughput accounting
- marginal costing/cost-volume-profit analysis

Classical cost elements are:

1. raw materials
2. labor
3. indirect expenses/overhead

Origins

Cost accounting has long been used to help managers understand the costs of running a business. Modern cost accounting originated during the industrial revolution, when the complexities of running a large scale business led to the development of systems for recording and tracking costs to help business owners and managers make decisions.

In the early industrial age, most of the costs incurred by a business were what modern accountants call "variable costs" because they varied directly with the amount of production. Money was spent on labor, raw materials, power to run a factory, etc. in direct proportion to production. Managers could simply total the variable costs for a product and use this as a rough guide for decision-making processes.

Some costs tend to remain the same even during busy periods, unlike variable costs, which rise and fall with volume of work. Over time, the importance of these "fixed costs" has become more important to managers. Examples of fixed costs include the depreciation of plant and equipment, and the cost of departments such as maintenance, tooling, production control, purchasing, quality control, storage and handling, plant supervision and engineering. In the early twentieth century, these costs were of little importance to most businesses. However, in the twenty-first century, these costs are often more important than the variable cost of a product, and allocating them to a broad range of products can lead to bad decision making. Managers must understand fixed costs in order to make decisions about products and pricing.

For example: A company produced railway coaches and had only one product. To make each coach, the company needed to purchase \$60 of raw materials and components, and pay 6 laborers \$40 each. Therefore, total variable cost for each coach was \$300. Knowing that making a coach required spending \$300, managers knew they couldn't sell below that price without losing money on each coach. Any price above \$300 became a contribution to the fixed costs of the company. If the fixed costs were, say, \$1000 per month for rent, insurance and owner's salary, the company could therefore sell 5 coaches per month for a total of \$3000 (priced at \$600 each), or 10 coaches for a total of \$4500 (priced at \$450 each), and make a profit of \$500 in both cases.

Elements of cost

- 1. Material(Material is a very important part of business)
 - A. Direct material
- 2. Labor
 - A. Direct labor
- 3. Overhead
 - A. Indirect material
 - B. Indirect labor

(In some companies, machine cost is segregated from overhead and reported as a separate element)

They are grouped further based on their functions as,

- 1. Production or works overheads
- 2. Administration overheads
- 3. Selling overheads
- 4. Distribution overheads

Classification of costs

Classification of cost means, the grouping of costs according to their common characteristics. The important ways of classification of costs are:

- By nature or element: materials, labor, expenses
- By functions: production, selling, distribution, administration, R&D, development,
- By traceability: direct and indirect
- By variability: fixed, variable, semi-variable
- By controllability: controllable, uncontrollable
- By normality: normal, abnormal

Standard cost accounting

In modern cost accounting, the concept of recording historical costs was taken further, by allocating the company's fixed costs over a given period of time to the items produced during that period, and recording the result as the total cost of production. This allowed the *full cost* of products that were not sold in the period they were produced to be recorded in inventory using a variety of complex accounting methods, which was consistent with the principles of GAAP (Generally Accepted Accounting Principles). It also essentially enabled managers to ignore the fixed costs, and look at the results of each period in relation to the "standard cost" for any given product.

For example: if the railway coach company normally produced 40 coaches per month, and the fixed costs were still \$1000/month, then each coach could be said to incur an overhead of \$25 ($\$1000 / 40$). Adding this to the variable costs of \$300 per coach produced a full cost of \$325 per coach.

This method tended to slightly distort the resulting unit cost, but in mass-production industries that made one product line, and where the fixed costs were relatively low, the distortion was very minor.

For example: if the railway coach company made 100 coaches one month, then the unit cost would become \$310 per coach ($\$300 + (\$1000 / 100)$). If the next

month the company made 50 coaches, then the unit cost = \$320 per coach (\$300 + (\$1000 / 50)), a relatively minor difference.

An important part of standard cost accounting is a variance analysis, which breaks down the variation between actual cost and standard costs into various components (volume variation, material cost variation, labor cost variation, etc.) so managers can understand *why costs were different from what was planned* and take appropriate action to correct the situation.

The development of throughput accounting

As business became more complex and began producing a greater variety of products, the use of cost accounting to make decisions to maximize profitability came under question. Management circles became increasingly aware of the Theory of Constraints in the 1980s, and began to understand that "every production process has a limiting factor" somewhere in the chain of production. As business management learned to identify the constraints, they increasingly adopted throughput accounting to manage them and "maximize the *throughput dollars*" (or other currency) from each unit of constrained resource.

For example: The railway coach company was offered a contract to make 15 open-topped streetcars each month, using a design that included ornate brass foundry work, but very little of the metalwork needed to produce a covered rail coach. The buyer offered to pay \$280 per streetcar. The company had a firm order for 40 rail coaches each month for \$350 per unit.

The company accountant determined that the cost of operating the foundry vs. the metalwork shop each month was as follows:

Overhead Cost by Department	Total Cost	Hours Available per month	Cost per hour
Foundry	\$ 7,300.00	160	\$45.63
Metal shop	\$ 3,300.00	160	\$20.63
Total	\$10,600.00	320	\$33.13

The company was at full capacity making 40 rail coaches each month. And since the foundry was expensive to operate, and purchasing brass as a raw material for the streetcars was expensive, the accountant determined that the company would lose money on any streetcars it built. He showed an analysis of the estimated product costs based on **standard cost accounting** and recommended that the company decline to build any streetcars.

Standard Cost Accounting Analysis	Streetcars	Rail coach
Monthly Demand	15	40
Price	\$280	\$350
Foundry Time (hrs)	3.0	2.0
Metalwork Time (hrs)	1.5	4.0

Total Time	4.5	6.0
Foundry Cost	\$136.88	\$ 91.25
Metalwork Cost	\$ 30.94	\$ 82.50
Raw Material Cost	\$120.00	\$ 60.00
Total Cost	\$287.81	\$233.75
Profit per Unit	\$ (7.81)	\$116.25

However, the company's operations manager knew that recent investment in automated foundry equipment had created idle time for workers in that department. The constraint on production of the railcoaches was the metalwork shop. She made an analysis of profit and loss if the company took the contract using throughput accounting to determine the profitability of products by calculating "throughput" (revenue less variable cost) in the metal shop.

Throughput Cost Accounting Analysis Decline Contract Take Contract

Coaches Produced	40	34
Streetcars Produced	0	15
Foundry Hours	80	113
Metal shop Hours	160	159
Coach Revenue	\$14,000	\$11,900
Streetcar Revenue	\$ 0	\$ 4,200
Coach Raw Material Cost	\$(2,400)	\$(2,040)
Streetcar Raw Material Cost	\$ 0	\$(1,800)
Throughput Value	\$11,600	\$12,260
Overhead Expense	\$(10,600)	\$(10,600)
Profit	\$1,000	\$1,660

After the presentations from the company accountant and the operations manager, the president understood that the metal shop capacity was limiting the company's profitability. The company could make only 40 rail coaches per month. But by taking the contract for the streetcars, the company could make nearly all the railway coaches ordered, and also meet all the demand for streetcars. The result would increase **throughput** in the metal shop from \$6.25 to \$10.38 per hour of available time, and increase profitability by 66 percent.

Activity-based costing

Activity-based costing (ABC) is a system for assigning costs to products based on the activities they require. In this case, activities are those regular actions performed inside a company. "Talking with customer regarding invoice questions" is an example of an activity inside most companies.

Accountants assign 100% of each employee's time to the different activities performed inside a company (many will use surveys to have the workers themselves assign their

time to the different activities). The accountant then can determine the total cost spent on each activity by summing up the percentage of each worker's salary spent on that activity.

A company can use the resulting activity cost data to determine where to focus their operational improvements. For example, a job-based manufacturer may find that a high percentage of its workers are spending their time trying to figure out a hastily written customer order. Via ABC, the accountants now have a currency amount pegged to the activity of "Researching Customer Work Order Specifications". Senior management can now decide how much focus or money to budget for resolving this process deficiency. Activity-based management includes (but is not restricted to) the use of activity-based costing to manage a business.

While ABC may be able to pinpoint the cost of each activity and resources into the ultimate product, the process could be tedious, costly and subject to errors.

As it is a tool for a more accurate way of allocating fixed costs into product, these fixed costs do not vary according to each month's production volume. For example, an elimination of one product would not eliminate the overhead or even direct labor cost assigned to it. ABC better identifies product costing in the long run, but may not be too helpful in day-to-day decision-making.

Lean accounting

Lean accounting has developed in recent years to provide the accounting, control, and measurement methods supporting lean manufacturing and other applications of lean thinking such as healthcare, construction, insurance, banking, education, government, and other industries.

There are two main thrusts for Lean Accounting. The first is the application of lean methods to the company's accounting, control, and measurement processes. This is not different from applying lean methods to any other processes. The objective is to eliminate waste, free up capacity, speed up the process, eliminate errors & defects, and make the process clear and understandable. The second (and more important) thrust of Lean Accounting is to fundamentally change the accounting, control, and measurement processes so they motivate lean change & improvement, provide information that is suitable for control and decision-making, provide an understanding of customer value, correctly assess the financial impact of lean improvement, and are themselves simple, visual, and low-waste. Lean Accounting does not require the traditional management accounting methods like standard costing, activity-based costing, variance reporting, cost-plus pricing, complex transactional control systems, and untimely & confusing financial reports. These are replaced by:

- lean-focused performance measurements
- simple summary direct costing of the value streams
- decision-making and reporting using a *box score*

- financial reports that are timely and presented in "plain English" that everyone can understand
- radical simplification and elimination of transactional control systems by eliminating the need for them
- driving lean changes from a deep understanding of the value created for the customers
- eliminating traditional budgeting through monthly sales, operations, and financial planning processes (SOFP)
- value-based pricing
- correct understanding of the financial impact of lean change

As an organization becomes more mature with lean thinking and methods, they recognize that the combined methods of lean accounting in fact creates a lean management system (LMS) designed to provide the planning, the operational and financial reporting, and the motivation for change required to prosper the company's on-going lean transformation.

Marginal costing

This method is used particularly for short-term decision-making. Its principal tenets are:

- Revenue (per product) – variable costs (per product) = contribution (per product)
- Total contribution – total fixed costs = (total profit or total loss)

Thus, it does not attempt to allocate fixed costs in an arbitrary manner to different products. The short-term objective is to maximize contribution per unit. If constraints exist on resources, then Managerial Accounting dictates that marginal cost analysis be employed to maximize contribution per unit of the constrained resource.

Chapter 10

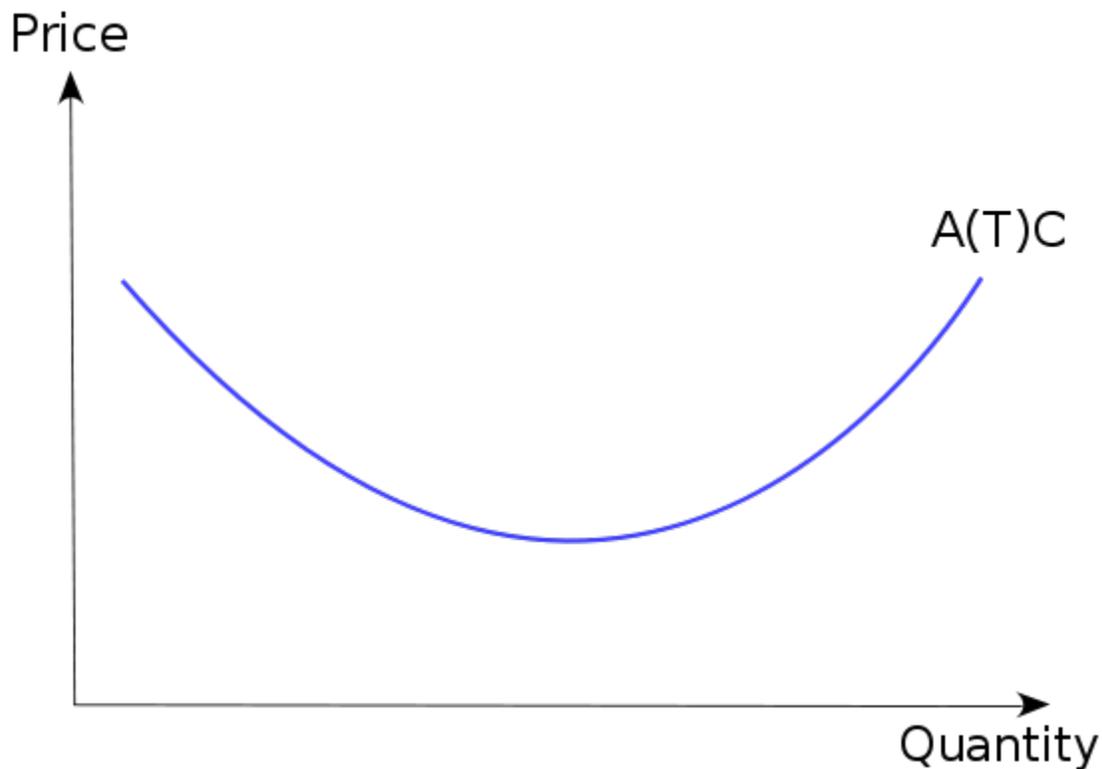
Cost Curve

In economics, a **cost curve** is a graph of the costs of production as a function of total quantity produced. In a free market economy, productively efficient firms use these curves to find the optimal point of production (minimising cost), and profit maximizing firms can use them to decide output quantities to achieve those aims. There are various types of cost curves, all related to each other, including total and average cost curves, and marginal ("for each additional unit") cost curves, which are the equal to the differential of the total cost curves. Some are applicable to the short run, others to the long run.

Short-run average variable cost curve (SRAVC)

Average variable cost (which is a short-run concept) is the variable cost (typically labor cost) per unit of output: $SRAVC = wL / Q$ where w is the wage rate, L is the quantity of labor used, and Q is the quantity of output produced. The SRAVC curve plots the short-run average variable cost against the level of output, and is typically U-shaped.

Short-run average total cost curve (SRATC or SRAC)



Typical short run average cost curve

The average total cost curve is constructed to capture the relation between cost per unit of output and the level of output, *ceteris paribus*. A perfectly competitive and productively efficient firm organizes its factors of production in such a way that the average cost of production is at the lowest point. In the short run, when at least one factor of production is fixed, this occurs at the output level where it has enjoyed all possible average cost gains from increasing production. This is at the minimum point in the diagram on the right.

Short-run total cost is given by

$$STC = P_K K + P_L L,$$

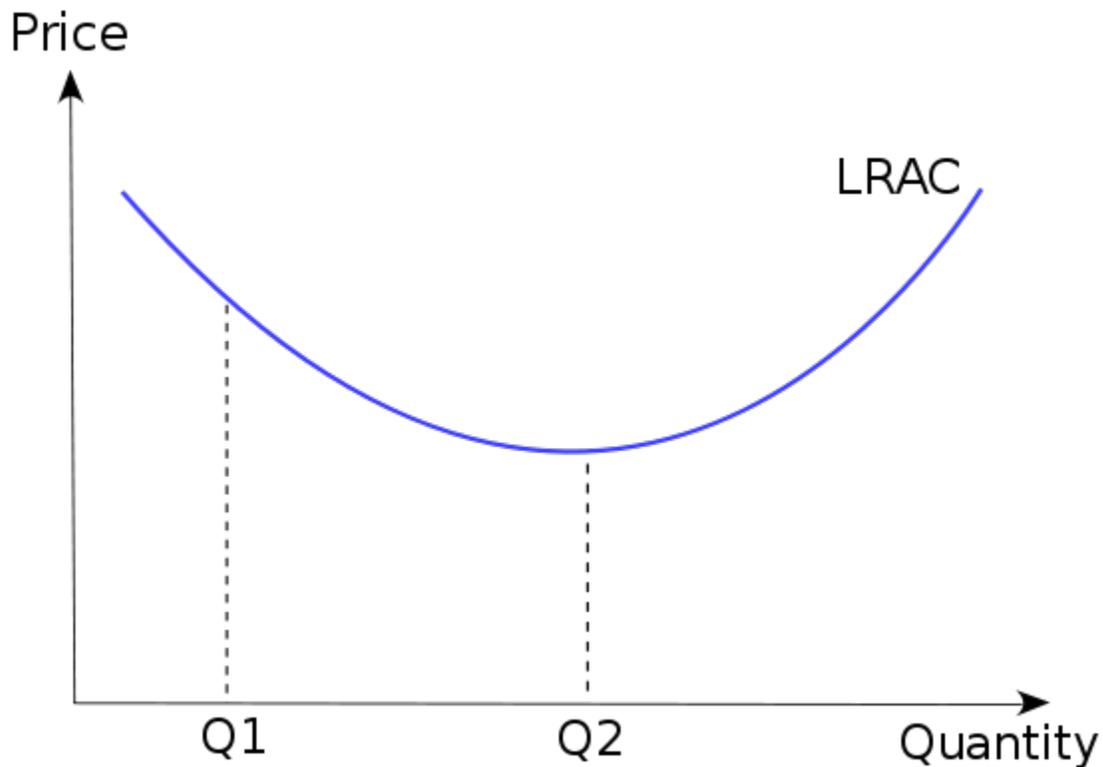
where P_K is the unit price of using physical capital per unit time, P_L is the unit price of labor per unit time (the wage rate), K is the quantity of physical capital used, and L is the quantity of labor used. From this we obtain short-run average cost, denoted either SATC or SAC, as STC / Q :

$$SRATC \text{ or } SRAC = P_K K / Q + P_L L / Q = P_K / AP_K + P_L / AP_L,$$

where $AP_K = Q/K$ is the average product of capital and $AP_L = Q/L$ is the average product of labor.^{:191}

Short run average cost equals average fixed costs plus average variable costs. Average fixed cost continuously falls as production increases in the short run, because K is fixed in the short run. The shape of the average variable cost curve is directly determined by increasing and then diminishing marginal returns to the variable input (conventionally labor).^{:210}

Long-run average cost curve (LRAC)



Typical long run average cost curve

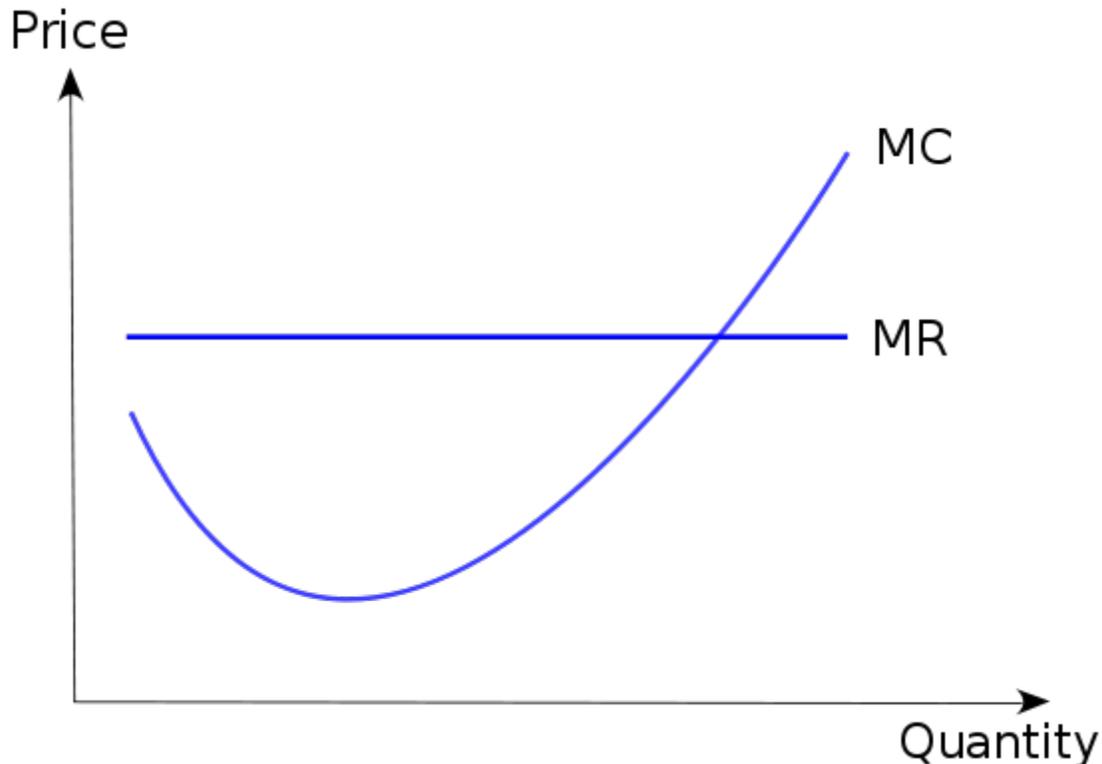
The long-run average cost curve depicts the cost per unit of output in the long run—that is, when all productive inputs' usage levels can be varied. All points on the line represent least-cost factor combinations; points above the line are attainable but unwise, while points below are unattainable given present factors of production. The behavioral assumption underlying the curve is that the producer will select the combination of inputs that will produce a given output at the lowest possible cost. Given that LRAC is an *average* quantity, one must not confuse it with the long-run *marginal* cost curve, which is the cost of one more unit.^{:232} The LRAC curve is created as an envelope of an infinite number of short-run average total cost curves, each based on a particular fixed level of capital usage.^{:235} The typical LRAC curve is U-shaped, reflecting increasing returns of scale where negatively-sloped, constant returns to scale where horizontal and decreasing

returns (due to increases in factor prices) where positively sloped.^{:234} Contrary to Viner, the envelope is not created by the minimum point of each short-run average cost curve.^{:235} This mistake is recognized as Viner's Error.

In a long-run perfectly competitive environment, the equilibrium level of output corresponds to the minimum efficient scale, marked as Q2 in the diagram. This is due to the zero-profit requirement of a perfectly competitive equilibrium. This result, which implies production is at a level corresponding to the lowest possible average cost,^{:259} does not imply that production levels other than that at the minimum point are not efficient. All points along the LRAC are productively efficient, by definition, but not all are equilibrium points in a long-run perfectly competitive environment.

In some industries, the bottom of the LRAC curve is large in comparison to market size (that is to say, for all intents and purposes, it is always declining and economies of scale exist indefinitely). This means that the largest firm tends to have a cost advantage, and the industry tends naturally to become a monopoly, and hence is called a natural monopoly. Natural monopolies tend to exist in industries with high capital costs in relation to variable costs, such as water supply and electricity supply.^{:312}

Short-run marginal cost curve (SRMC)



Typical marginal cost curve

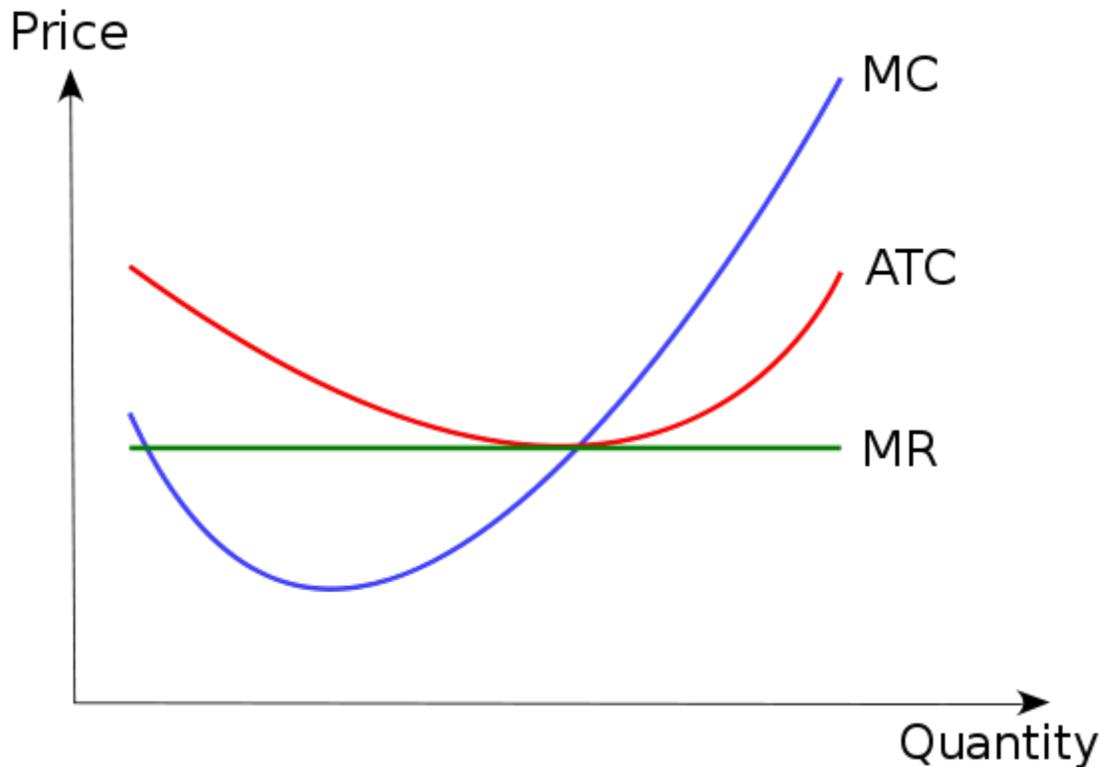
A short-run marginal cost curve graphically represents the relation between marginal (i.e., incremental) cost incurred by a firm in the short-run production of a good or service and the quantity of output produced. This curve is constructed to capture the relation between marginal cost and the level of output, holding other variables, like technology and resource prices, constant. The marginal cost curve is U-shaped. Marginal cost is relatively high at small quantities of output; then as production increases, marginal cost declines, reaches a minimum value, then rises. The marginal cost is shown in relation to marginal revenue, the incremental amount of sales revenue that an additional unit of the product or service will bring to the firm. This shape of the marginal cost curve is directly attributable to increasing, then decreasing marginal returns (and the law of diminishing marginal returns). Marginal cost equals w/MP_L .¹⁹¹ For most production processes the marginal product of labor initially rises, reaches a maximum value and then continuously falls as production increases. Thus marginal cost initially falls, reaches a minimum value and then increases.²⁰⁹ The marginal cost curve intersects both the average variable cost curve and (short-run) average total cost curve at their minimum points. When the marginal cost curve is above an average cost curve the average curve is rising. When the marginal cost curve is below an average curve the average curve is falling. This relation holds regardless of whether the marginal curve is rising or falling.²²⁶

Long-run marginal cost curve (LRMC)

The long-run marginal cost curve shows for each unit of output the added total cost incurred in the long run, that is, the conceptual period when all factors of production are variable so as to minimize long-run average total cost. Stated otherwise, LRMC is the minimum increase in total cost associated with an increase of one unit of output when all inputs are variable.

The long-run marginal cost curve is shaped by economies and diseconomies of scale, a long-run concept, rather than the law of diminishing marginal returns, which is a short-run concept. The long-run marginal cost curve tends to be flatter than its short-run counterpart due to increased input flexibility as to cost minimization. The long-run marginal cost curve intersects the long-run average cost curve at the minimum point of the latter.²⁰⁸ When long-run marginal costs are below long-run average costs, long-run average costs are falling (as to additional units of output).²⁰⁷ When long-run marginal costs are above long-run average costs, average costs are rising. Long-run marginal cost equals short-run marginal cost at the least-long-run-average-cost level of production. LRMC is the slope of the LR total-cost function.

Graphing cost curves together



Cost curves in perfect competition compared to marginal revenue

Cost curves can be combined to provide information about firms. In this diagram for example, firms are assumed to be in a perfectly competitive market. In a perfectly competitive market the price that firms are faced with would be the price at which the marginal cost curve cuts the average cost curve.

Cost curves and production functions

Assuming that factor prices are constant, the production function determines all cost functions. The variable cost curve is the inverted short-run production function or total product curve and its behavior and properties are determined by the production function.²⁰⁹ Because the production function determines the variable cost function it necessarily determines the shape and properties of marginal cost curve and the average cost curves.

If the firm is a perfect competitor in all input markets, and thus the per-unit prices of all its inputs are unaffected by how much of the inputs the firm purchases, then it can be shown that at a particular level of output, the firm has economies of scale (i.e., is operating in a downward sloping region of the long-run average cost curve) if and only if it has increasing returns to scale. Likewise, it has diseconomies of scale (is operating in an upward sloping region of the long-run average cost curve) if and only if it has

decreasing returns to scale, and has neither economies nor diseconomies of scale if it has constant returns to scale. In this case, with perfect competition in the output market the long-run market equilibrium will involve all firms operating at the minimum point of their long-run average cost curves (i.e., at the borderline between economies and diseconomies of scale).

If, however, the firm is not a perfect competitor in the input markets, then the above conclusions are modified. For example, if there are increasing returns to scale in some range of output levels, but the firm is so big in one or more input markets that increasing its purchases of an input drives up the input's per-unit cost, then the firm could have diseconomies of scale in that range of output levels. Conversely, if the firm is able to get bulk discounts of an input, then it could have economies of scale in some range of output levels even if it has decreasing returns in production in that output range.

Relationship between different curves

- Total Cost = Fixed Costs (FC) + Variable Costs (VC)
- Marginal Cost (MC) = dC/dQ ; MC equals the slope of the total cost function and of the variable cost function
- Average Total Cost (ATC) = Total Cost/Q
- Average Fixed Cost (AFC) = FC/Q
- Average Variable Cost = VC/Q.
- $ATC = AFC + AVC$
- The MC curve is related to the shape of the ATC and AVC curves:^{:212}
 - At a level of Q at which the MC curve is above the average total cost or average variable cost curve, the latter curve is rising.^{:212}
 - If MC is below average total cost or average variable cost, then the latter curve is falling.
 - If MC equals average total cost, then average total cost is at its minimum value.
 - If MC equals average variable cost, then average variable cost is at its minimum value.

Relationship between short run and long run cost curves

Basic: For each quantity of output there is one cost minimizing level of capital and a unique short run average cost curve associated with producing the given quantity.

- Each STC curve can be tangent to the LRTC curve at only one point. The STC curve cannot cross (intersect) the LRTC curve.^{:230:228-229} The STC curve can lie wholly “above” the LRTC curve with no tangency point.^{:256}
- One STC curve is tangent to LRTC at the long-run cost minimizing level of production. At the point of tangency $LRTC = STC$. At all other levels of production STC will exceed LRTC.^{:292-299}
- Average cost functions are the total cost function divided by the level of output. Therefore the SATC curve is also tangent to the LRATC curve at the cost-

- minimizing level of output. At the point of tangency $LRATC = SATC$. At all other levels of production $SATC > LRATC$.²⁹²⁻²⁹⁹ To the left of the point of tangency the firm is using too much capital and fixed costs are too high. To the right of the point of tangency the firm is using too little capital and diminishing returns to labor are causing costs to increase.
- The slope of the total cost curves equals marginal cost. Therefore when STC is tangent to LTC , $SMC = LRMC$.
 - At the long run cost minimizing level of output $LRTC = STC$; $LRATC = SATC$ and $LRMC = SMC$.²⁹²⁻²⁹⁹
 - The long run cost minimizing level of output may be different from minimum $SATC$.^{229 :186}
 - With fixed unit costs of inputs, if the production function has constant returns to scale, then at the minimal level of the $SATC$ curve we have $SATC = LRATC = SMC = LRMC$.²⁹²⁻²⁹⁹
 - With fixed unit costs of inputs, if the production function has increasing returns to scale, the minimum of the $SATC$ curve is to the right of the point of tangency between the $LRAC$ and the $SATC$ curves.²⁹²⁻²⁹⁹ Where $LRTC = STC$, $LRATC = SATC$ and $LRMC = SMC$.
 - With fixed unit costs of inputs and decreasing returns the minimum of the $SATC$ curve is to the left of the point of tangency between $LRAC$ and $SATC$.²⁹²⁻²⁹⁹ Where $LRTC = STC$, $LRATC = SATC$ and $LRMC = SMC$.
 - With fixed unit input costs, a firm that is experiencing increasing (decreasing) returns to scale and is producing at its minimum SAC can always reduce average cost in the long run by expanding (reducing) the use of the fixed input.^{292-99 :186}
 - $LRATC$ will always equal to or be less than $SATC$.²¹¹
 - If production process is exhibiting constant returns to scale then minimum $SRAC$ equals minimum long run average cost. The $LRAC$ and $SRAC$ intersect at their common minimum values. Thus under constant returns to scale $SRMC = LRMC = LRAC = SRAC$.
 - If the production process is experiencing decreasing or increasing, minimum short run average cost does not equal minimum long run average cost. If increasing returns to scale exist long run minimum will occur at a lower level of output than $SRAC$. This is because there are economies of scale that have not been exploited so in the long run a firm could always produce a quantity at a price lower than minimum short run average cost simply by using a larger plant.
 - With decreasing returns, minimum $SRAC$ occurs at a lower production level than minimum $LRAC$ because a firm could reduce average costs by simply decreasing the size or its operations.
 - The minimum of a $SRAC$ occurs when the slope is zero. Thus the points of tangency between the U-shaped $LRAC$ curve and the minimum of the $SRAC$ curve would coincide only with that portion of the $LRAC$ curve exhibiting constant economies of scale. For increasing returns to scale the point of tangency between the $LRAC$ and the $SRAC$ would have to occur at a level of output below level associated with the minimum of the $SRAC$ curve.

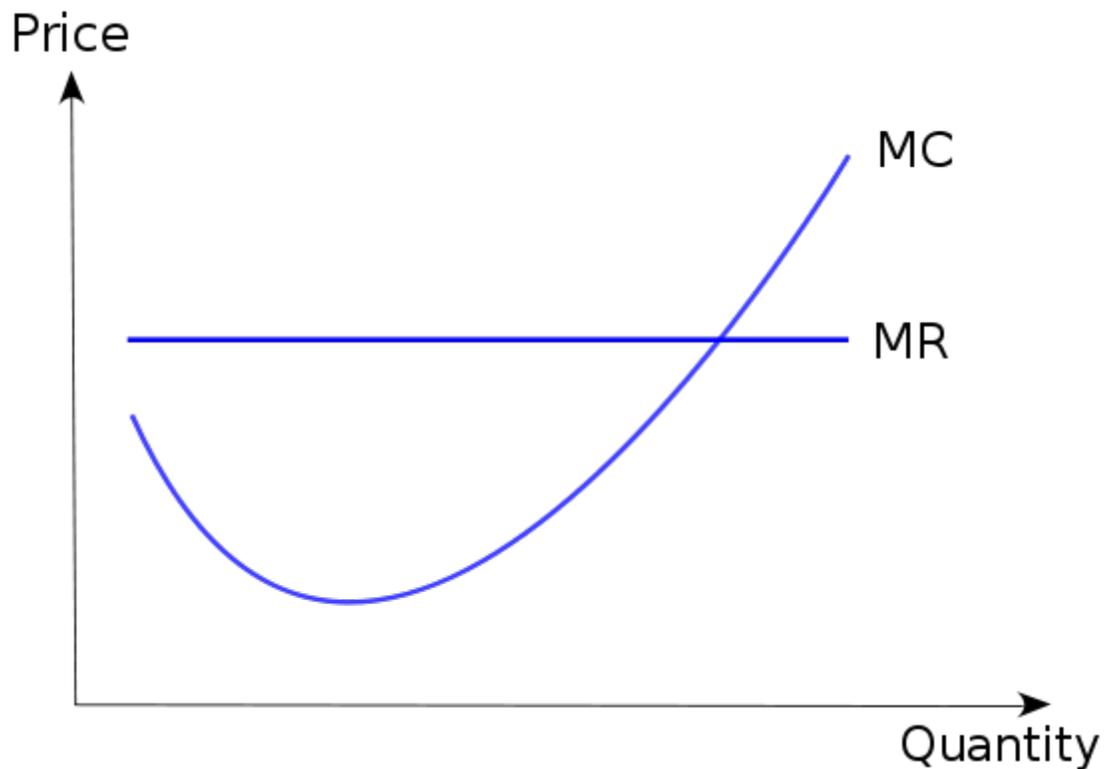
These statements assume that the firm is using the optimal level of capital for the quantity produced. If not, then the SRAC curve would lie "wholly above" the LRAC and would not be tangent at any point.

U-shaped curves

Both the SRAC and LRAC curves are typically expressed as U-shaped.^{:211; 226 :182;187-188} However, the shapes of the curves are not due to the same factors. For the short run curve the initial downward slope is largely due to declining average fixed costs.^{:227} Increasing returns to the variable input at low levels of production also play a role, while the upward slope is due to diminishing marginal returns to the variable input.^{:227} With the long run curve the shape by definition reflects economies and diseconomies of scale.^{:186} At low levels of production long run production functions generally exhibit increasing returns to scale, which, for firms that are perfect competitors in input markets, means that the long run average cost is falling;^{:227} the upward slope of the long run average cost function at higher levels of output is due to decreasing returns to scale at those output levels.^{:227}

Chapter 11

Marginal Cost



A typical marginal cost curve with marginal revenue overlaid

In economics and finance, **marginal cost** is the change in total cost that arises when the quantity produced changes by one unit. That is, it is the cost of producing one more unit of a good. If the good being produced is infinitely divisible, so the size of a "unit" is infinitesimal, then assuming the cost function is differentiable the marginal cost (MC) function is the first derivative of the total cost (TC) function with respect to quantity (Q).

Note that the marginal cost will change with volume, as a non-linear and non-proportional cost function includes

- variable terms dependent to volume,
- constant terms independent to volume and occurring with the respective lot size,
- jump fix cost increase or decrease dependent to steps of volume increase.

In general terms, marginal cost at each level of production includes any additional costs required to produce the next unit. If producing additional vehicles requires, for example, building a new factory, the marginal cost of those *extra* vehicles includes the cost of the new factory. In practice, the analysis is segregated into short and long-run cases, and over the longest run, all costs are marginal. At each level of production and time period being considered, marginal costs include all costs which vary with the level of production, and other costs are considered fixed costs.

If the cost function is differentiable, the marginal cost is the cost of the next unit produced referring to the basic volume.

$$MC = \frac{dTC}{dQ}$$

If the cost function is not differentiable, the marginal cost can be expressed as follows.

$$MC = \frac{\Delta TC}{\Delta Q}$$

A number of other factors can affect marginal cost and its applicability to real world problems. Some of these may be considered market failures. These may include information asymmetries, the presence of negative or positive externalities, transaction costs, price discrimination and others.

Cost functions and relationship to average cost

In the simplest case, the total cost function and its derivative are expressed as follows, where Q represents the production quantity, VC represents variable costs, FC represents fixed costs and TC represents total costs.

$$MC = \frac{dTC}{dQ} = \frac{d(FC + VC)}{dQ} = \frac{dVC}{dQ}$$

Since (by definition) fixed costs do not vary with production quantity, it drops out of the equation when it is differentiated. The important conclusion is that marginal cost *is not related to* fixed costs. This can be compared with average total cost or ATC, which is the total cost divided by the number of units produced and *does* include fixed costs.

$$ATC = \frac{FC + VC}{Q}$$

For discrete calculation without calculus, marginal cost equals the change in total (or variable) cost that comes with each additional unit produced. In contrast, incremental cost is the composition of total cost from the surrogate of contributions, where any increment is determined by the contribution of the cost factors, not necessarily by single units.

For instance, suppose the total cost of making 1 shoe is \$30 and the total cost of making 2 shoes is \$40. The marginal cost of producing the second shoe is \$40 - \$30 = \$10.

Marginal cost is not the cost of producing the "next" or "last" unit. As Silberberg and Suen note the cost of the last unit is the same as the cost of the first unit and every other unit. In the short run increasing production requires using more of the variable input - conventionally assumed to be labor. Adding more labor to a fixed capital stock reduces the marginal product of labor because of the diminishing marginal returns. This reduction in productivity is not limited to the additional labor needed to produce the marginal unit - the productivity of every unit of labor is reduced. Thus the costs of producing the marginal unit of output has two components: the cost associated with producing the marginal unit and the increase in average costs for all units produced due to the "damage" to the entire productive process ($\partial AC/\partial q$)q. The first component is the per unit or average cost. The second unit is the small increase in costs due to the law of diminishing marginal returns which increases the costs of all units of sold. **Therefore, the precise formula is:**

$$MC = AC + (\partial AC/\partial q)q.$$

Marginal costs can also be expressed as the cost per unit of labor divided by the marginal product of labor.

$$MC = \Delta VC/\Delta q;$$

$$\Delta VC = w\Delta L;$$

$$MC = w\Delta L/\Delta q;$$

$\Delta L/\Delta q$ the change in quantity of labor to affect a one unit change in output = $1/MPL$.

Therefore $MC = w/MPL$ Since the wage rate is assumed constant marginal cost and marginal product of labor have an inverse relationship - if marginal cost is increasing (decreasing) the marginal product of labor is decreasing (increasing).

Economies of scale

Economies of scale is a concept that applies to the long run, a span of time in which all inputs can be varied by the firm so that there are no fixed inputs or fixed costs.

Production may be subject to economies of scale (or diseconomies of scale). Economies of scale are said to exist if an additional unit of output can be produced for less than the average of all previous units— that is, if long-run marginal cost is below long-run average cost, so the latter is falling. Conversely, there may be levels of production where

marginal cost is higher than average cost, and average cost is an increasing function of output. For this generic case, minimum average cost occurs at the point where average cost and marginal cost are equal (when plotted, the marginal cost curve intersects the average cost curve from below); this point will *not* be at the minimum for marginal cost if fixed costs are greater than zero.

Perfectly competitive supply curve

The portion of the marginal cost curve above its intersection with the average variable cost curve is the supply curve for a firm operating in a perfectly competitive market. (the portion of the MC curve below its intersection with the AVC curve is not part of the supply curve because a firm would not operate at price below the shut down point) This is not true for firms operating in other market structures. For example, while a monopoly "has" an MC curve it does not have a supply curve. In a perfectly competitive market, a supply curve shows the quantity a seller's willing and able to supply at each price - for each price there is a unique quantity that would be supplied. The one-to-one relationship simply is absent in the case of a monopoly. With a monopoly there could be an infinite number of prices associated with a given quantity. It all depends on the shape and position of the demand curve and its accompanying marginal revenue curve.

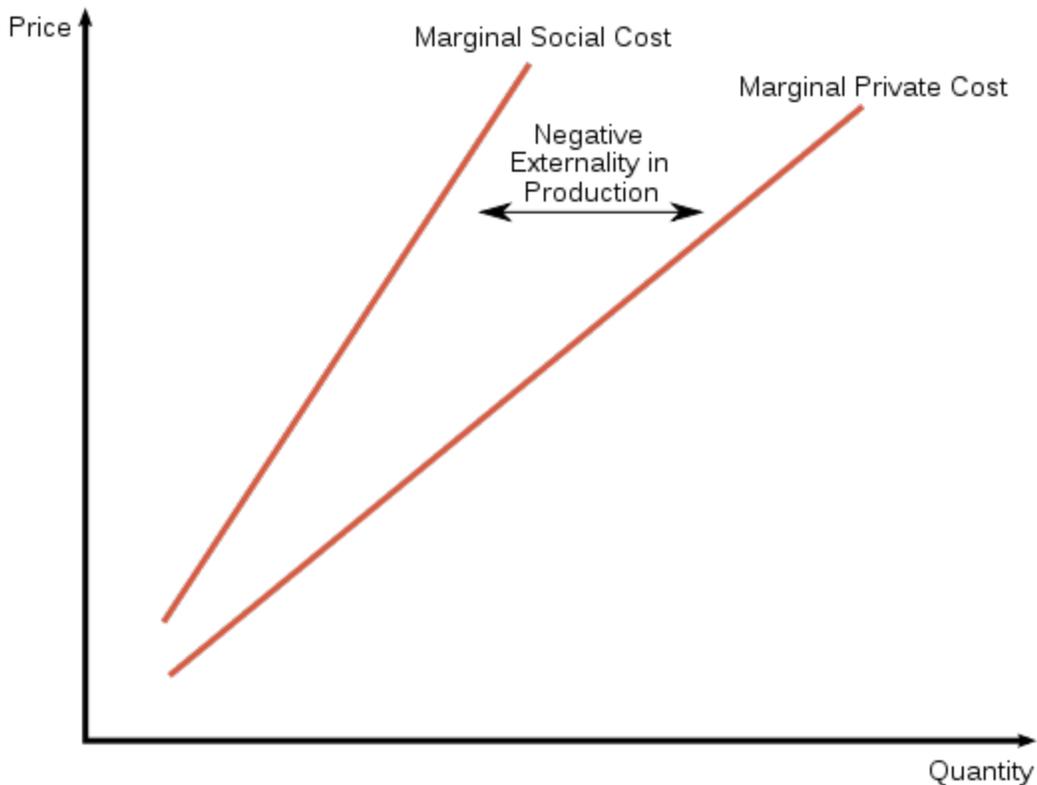
Relationship to fixed costs

Marginal Costs are not affected by changes in fixed cost. Marginal costs can be expressed as $\Delta C(q)/\Delta Q$. Since fixed costs do not vary with (depend on) changes in quantity, MC is $\Delta VC/\Delta Q$. Thus if fixed cost were to double MC would not be affected and consequently the profit maximizing quantity and price would not change. this can be illustrated by graphing the short run total cost curve and the short run variable cost curve. The shape of the curves are identical. Each curve initially increases at a decreasing rate reaches and inflection point then increases at a decreasing rate. the only difference between the curves is that the SRVC curve begins from the origin while the SRTC curve originates on the y-axis. The distance of the origin of the SRTC above the origin represents the fixed cost - the vertical distance between the curves. This distant remains constant as the quantity produced Q increases. MC is the slope of the SRVC curve. A change in fixed cost would be reflected by a change in the vertical distance between the SRTC and SRVC curve. Any such change would have no effect on the shape of the SRVC curve and therefore its slope at any point - MC.

Externalities

Externalities are costs (or benefits) that are not borne by the parties to the economic transaction. A producer may, for example, pollute the environment, and others may bear those costs. A consumer may consume a good which produces benefits for society, such as education; because the individual does not receive all of the benefits, he may consume less than efficiency would suggest. Alternatively, an individual may be a smoker or alcoholic and impose costs on others. In these cases, production or consumption of the good in question may differ from the optimum level.

Negative externalities of production



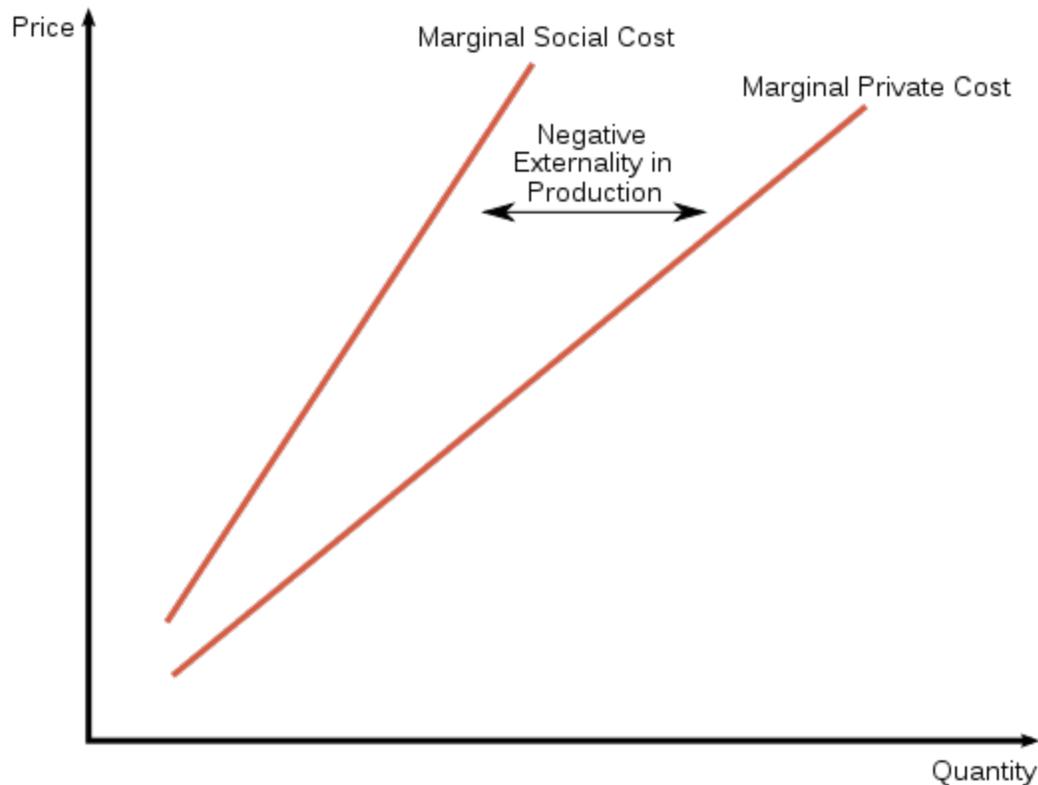
Negative Externalities of Production

Much of the time, private and social costs do not diverge from one another, but at times social costs may be either greater or less than private costs. When marginal social costs of production are greater than that of the private cost function, we see the occurrence of a negative externality of production. Productive processes that result in pollution are a textbook example of production that creates negative externalities.

Such externalities are a result of firms externalising their costs onto a third party in order to reduce their own total cost. As a result of externalising such costs we see that members of society will be negatively affected by such behavior of the firm. In this case, we see that an increased cost of production on society creates a social cost curve that depicts a greater cost than the private cost curve.

In an equilibrium state we see that markets creating negative externalities of production will overproduce that good. As a result, the socially optimal production level would be lower than that observed.

Positive externalities of production



Positive Externalities of Production

When marginal social costs of production are less than that of the private cost function, we see the occurrence of a positive externality of production. Production of public goods are a textbook example of production that create positive externalities. An example of such a public good, which creates a divergence in social and private costs, includes the production of education. It is often seen that education is a positive for any whole society, as well as a positive for those directly involved in the market.

Examining the relevant diagram we see that such production creates a social cost curve that is less than that of the private curve. In an equilibrium state we see that markets creating positive externalities of production will under produce that good. As a result, the socially optimal production level would be greater than that observed.

Social costs

Of great importance in the theory of marginal cost is the distinction between the marginal *private* and *social* costs. The marginal private cost shows the cost associated to the firm in question. It is the marginal private cost that is used by business decision makers in their profit maximization goals, and by individuals in their purchasing and consumption choices. Marginal social cost is similar to private cost in that it includes the cost functions of private enterprise but *also* that of society as a whole, including parties that have no

direct association with the private costs of production. It incorporates all negative and positive externalities, of both production and consumption.

Hence, when deciding whether or how much to buy, buyers take account of the cost to society of their actions *if* private and social marginal cost coincide. The equality of price with social marginal cost, by aligning the interest of the buyer with the interest of the community as a whole is a necessary condition for economically efficient resource allocation.

Chapter 12

Historical cost

In accounting, **historical cost** is the original monetary value of an economic item. Historical cost is based on the stable measuring unit assumption. In some circumstances, assets and liabilities may be shown at their historical cost, as if there had been no change in value since the date of acquisition. The balance sheet value of the item may therefore differ from the "true" value.

While historical cost is criticised for its inaccuracy (deviation from "true" value), it remains in use in most accounting systems. Various corrections to historical cost are used, many of which require the use of management judgment and may be difficult to implement or verify. The trend in most accounting standards is a move to more accurate reflection of the fair or market value, although the historical cost principle remains in use, particularly for assets of little importance.

Depreciation affects the carrying value of an asset on the balance sheet. The historical cost will equal the carrying value if there has been no change recorded in the value of the asset since acquisition. Improvements may be added to the cost basis of an asset.

Historical cost does not generally reflect current market valuation. Alternative measurement bases to the historical cost measurement basis, which may be applied for some types of assets for which market values are readily available, require that the carrying value of an asset (or liability) be updated to the market price (mark-to-market valuation) or some other estimate of value that better approximates the real value. Accounting standards may also have different methods required or allowed (even for different types of balance sheet variable real value non-monetary assets or liabilities) as to how the resultant change in value of an asset or liability is recorded, as a part of income or as a direct change to shareholders' equity.

The Constant Item Purchasing Power Accounting model is an International Accounting Standards Board approved alternative basic accounting model to the traditional Historical Cost Accounting model.

Historical cost basis (original cost)

Under the historical cost basis of accounting, assets and liabilities are recorded at their values when first acquired. They are not then generally restated for changes in values.

Costs recorded in the Income Statement are based on the historical cost of items sold or used, rather than their replacement costs.

For example –

- a company acquires an asset in year 1 for \$100;
- the asset is still held at the end of year 1, when its market value is \$120;
- the company sells the asset in year 2 for \$115.

At the end year 1 the asset is recorded in the balance sheet at cost of \$100. No account is taken of the increase in value from \$100 to \$120 in year 1. In year 2 the company records a sale of \$115. The cost of sales is \$100, being the historical cost of the asset. This gives rise to a profit of \$15 which is wholly recognised in year 2.

Measurement under the historical cost basis

Inventory

It is standard under the historical cost basis to write down the value of inventory (stock) to a lower cost and net realisable value. As a result:-

- A downward movement in the realisable value of inventory below cost is recognised immediately
- An upward movement in the realisable value of inventory is not recognised until the inventory is sold

Property, plant and equipment

Property, plant and equipment is recorded at cost under the historical cost basis. Cost includes:-

- Purchase price, including import duties and non-refundable purchase taxes, after deducting trade discounts and rebates;
- Any costs directly attributable to bringing the asset to the location and condition necessary for it to be capable of operating. These can include site preparation, delivery and handling costs, installation, assembly, testing, professional fees and the costs of employees directly involved in these activities.

In IFRS, cost also includes the initial estimate of the costs of dismantling and removing the item and restoring it. Cost may include the cost of borrowing to finance construction if this policy is consistently adopted. Cost is then subject to **depreciation** with to write off the cost of the asset over its estimated useful life down to the recoverable amount. In most cases the method is "straight line", with the same depreciation charge from the date when an asset is brought into use until it is expected to be sold or no further economic benefits obtained from it, but other patterns of depreciation are used if assets are used proportionately more in some periods than others.

Assets and liabilities denominated in foreign currency

Monetary items such as cash balances, receivables and payables which are denominated in foreign currency are reported using the closing exchange rate under IFRS.

Exceptions to the historical cost basis of accounting

Revaluation of property, plant and equipment

Under IFRS it is acceptable, but not required, to restate the values of property, plant and equipment to fair value. 'Fair value' is the amount for which an asset could be exchanged, or a liability settled, between knowledgeable, willing parties in an arm's length transaction. Such a policy must be applied to all assets of a particular class. It would therefore be acceptable for an entity to revalue freehold properties every three years. The revaluations must be made with sufficient regularity to ensure that the carrying value does not differ materially from market value in subsequent years. A surplus on revaluation would be recorded as a reserve movement, not as income.

Derivative financial instruments

Under IFRS and US GAAP derivative financial instruments are stated at fair value ("mark to market") with movements recorded in the income statement.

Financial reporting in hyperinflationary economies

IFRS requires a separate method of accounting in currencies deemed to be hyperinflationary. The characteristics of a hyperinflation include the population keeping its wealth in non-monetary assets or relatively stable foreign currencies, prices quoted in foreign currencies or widespread indexation of prices. This might arise if cumulative inflation reaches or exceeds 100% over three years. An entity operating in a hyperinflationary economy:-

- Records a gain or loss on its 'net monetary position' in its income statement.
- Records non-monetary items (for example, property, plant & equipment) in the balance sheet by applying indexation to their historical cost.

Management accounting techniques

In management accounting there are a number of techniques used as alternatives to historical cost accounting including:-

- measuring profit on sale of inventory by reference to its replacement cost. If inventory with a historical cost of \$100 is sold for \$115 when it costs \$110 to replace it, the profit recorded would be \$5 only based on replacement cost, not \$15;
- charging economic rent for assets, particularly property. If a business uses a 20-year old property which it owns, depreciation on a historical cost basis might be insignificant. However, the management accounts could show a notional rent payable, being perhaps opportunity cost - the amount the business could receive if it let the property to a third party.

IASB approved alternative to Historical Cost Accounting

The IASB's Framework introduced Constant Item Purchasing Power Accounting as an alternative to Historical Cost Accounting in 1989 in Par. 104 (a) where it states that financial capital maintenance - not variable real value non-monetary items, e.g. property, plant, equipment, inventory, intangible assets, etc., - can be measured in either nominal monetary units - the traditional HCA model - or in units of constant purchasing power: the CIPPA model.

The specific choice of measuring financial capital maintenance in units of constant purchasing power (the CIPPA model) at all levels of inflation and deflation as contained in the Framework for the Preparation and Presentation of Financial Statements, was approved by the International Accounting Standards Board's predecessor body, the International Accounting Standards Committee Board, in April 1989 for publication in July 1989 and adopted by the IASB in April 2001.

“In the absence of a Standard or an Interpretation that specifically applies to a transaction, management must use its judgement in developing and applying an accounting policy that results in information that is relevant and reliable. In making that judgement, IAS 8.11 requires management to consider the definitions, recognition criteria, and measurement concepts for assets, liabilities, income, and expenses in the Framework. This elevation of the importance of the Framework was added in the 2003 revisions to IAS 8.”

IAS8, 11:

“In making the judgement, management shall refer to, and consider the applicability of, the following sources in descending order: (a) the requirements and guidance in Standards and Interpretations dealing with similar and related issues; and (b) the definitions, recognition criteria and measurement concepts for assets, liabilities, income and expenses in the Framework.”

There is no applicable International Financial Reporting Standard or Interpretation regarding the valuation of constant real value non-monetary items, e.g. issued share capital, retained earnings, capital reserves, all other items in Shareholders Equity, trade debtors, trade creditors, deferred tax assets and liabilities, taxes payable and receivable, all other non-monetary receivables and payables, Profit and Loss account items such as salaries, wages, rents, etc. The Framework is thus applicable.

The CIPPA model is chosen by hardly any accountant in non-hyperinflationary economies even though it would maintain the real values of constant real value non-monetary items - e.g. issued share capital, retained income, other shareholder equity items, trade debtors, trade creditors, etc. for an unlimited period of time. This is because the CIPPA model is generally viewed by accountants as a 1970's failed inflation accounting model that requires all non-monetary items - variable real value non-monetary items and constant real value non-monetary items - to be inflation-adjusted by means of the Consumer Price Index.

The IASB did not approve CIPPA in 1989 as an inflation accounting model. CIPPA by measuring financial capital maintenance in units of constant purchasing power incorporates an alternative capital concept, financial capital maintenance concept and profit determination concept to the Historical Cost capital concept, financial capital maintenance concept and profit determination concept. CIPPA only requires all constant real value non-monetary items, e.g. issued share capital, retained income, all other items in Shareholders Equity, trade debtors, trade creditors, deferred tax assets and liabilities, taxes payable and receivable, all items in the profit and loss account, etc. to be valued in units of constant purchasing power. Variable real value non-monetary items, e.g. property, plant, equipment, listed and unlisted shares, inventory, etc. are valued in terms of IFRS and are not required in terms of the Framework, Par. 104 (a) to be valued in units of constant purchasing power.

The IASB requires entities to implement IAS 29 which is a Constant Purchasing Power Accounting model during hyperinflation.

Advantages and disadvantages of historical cost accounting

Advantages

- Historical cost accounts are straightforward to produce
- Historical cost accounts do not record gains until they are realized
- Historical cost accounts are still used in most accounting systems

Disadvantages

- Historical cost accounts give no indication of current values of the assets of a business

- Historical cost accounts do not record the opportunity costs of the use of older assets, particularly property which may be recorded at a value based on costs incurred many years ago
- Historical cost accounts do not measure the loss of value of monetary assets as a result of inflation.

Chapter 13

Social Cost & Sunk Costs

Social Cost

Social cost, in economics, is generally defined in opposition to "private cost". In economics, theorists model individual decision-making as measurement of costs and benefits. Rational choice theory often assumes that individuals consider only the costs they themselves bear when making decisions, not the costs that may be borne by others.

With pure private goods, the costs carried by the individuals involved are the only economically meaningful costs. The choice to purchase a glass of lemonade at a lemonade stand has little consequence for anyone other than the seller or the buyer. The costs involved in this economic activity are the costs of the lemons and the sugar and the water that are ingredients to the lemonade, the opportunity cost of the labour to combine them into lemonade, as well as any transaction costs, such as walking to the stand.

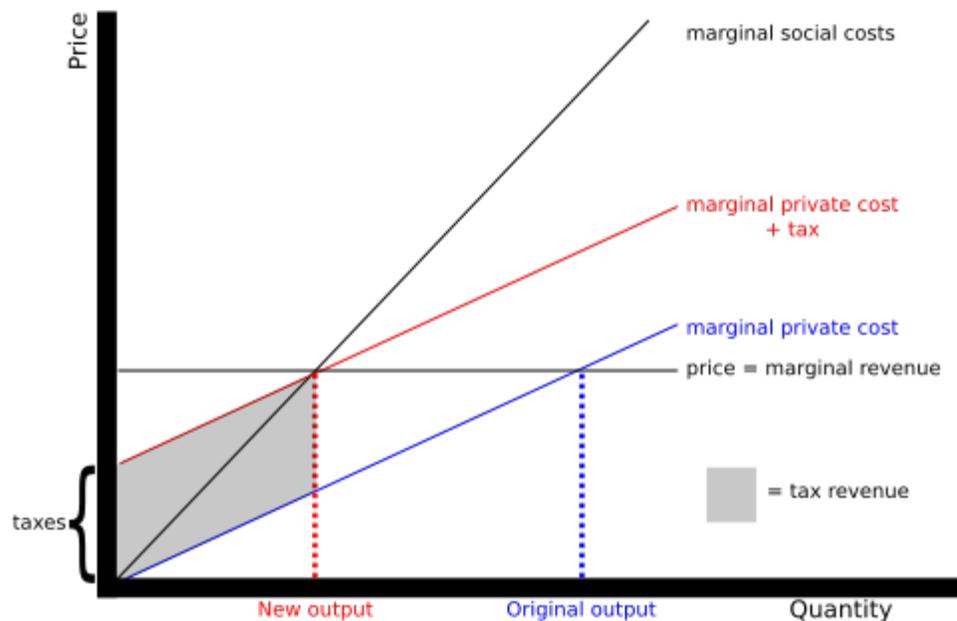
Implications

If there is a negative externality, then social costs will be greater than private costs. Environmental pollution is an example of a social cost that is seldom borne completely by the polluter, thereby creating a negative externality. If there is a positive externality, then one will have higher social *benefits* than private *benefits*. For example, when a supplier of educational services indirectly benefits society as a whole but only receives payment for the direct benefit received by the recipient of the education: the benefit to society of an educated populace is a positive externality. In either case, economists refer to this as market failure because resources will be allocated inefficiently. In the case of negative externalities, private agents will engage in too much of the activity; in the case of positive externalities, they will engage in too little. (The marginal rate of transformation in production will not be equal to the marginal rate of substitution in consumption due to the effect of the externality and as a result Pareto optimality will not occur.)

Theory

The ideas of social cost, externalities, and market failure are often used as an argument for government intervention in the form of regulations. Libertarians who believe in a free market respond that the existence of market failure should not lead to government intervention. They prefer to rely on tradition, community pressure, and dollar voting.

Negative externalities (external costs) lead to an over-production of those goods that have a high social cost. For example, the logging of trees for timber may result in society losing a recreation area, shade, beauty, good quality soil to grow crops on, and air quality but this loss is usually not quantified and included in the price of the timber that is made from the trees. As a result, individual entities in the marketplace have no incentive to factor in these externalities. More of this activity is performed than would be if its cost had a true accounting.



This can be illustrated with a diagram. Profit-maximizing organizations will set output at Q_p where marginal private costs (MPC) is equal to marginal revenue (MR). (This diagram assumes perfect competition, under which price (P) equals MR.) This will yield a profit shown by the triangular area $0,C,F$.

But if externalities are present, the attainment of social optimality requires that the full social costs must be considered. The socially optimum level of output is Q_s where marginal social costs (MSC) is equal to marginal revenue (MR). The amount of output, Q_p minus Q_s , indicates the excess output due to the externality. Profits will decrease also, from $0,C,F$ to $0,A,F$. It is clearly profitable for the firm to pollute, since "internalizing the externality" hurts profits. The amount of the externality will decrease from C,D to B,A .

Because the marginal social cost curve (**MSC**) is *above* the marginal private cost curve (**MPC**), this diagram illustrates the case of a negative externality. If the marginal social cost curve was below the marginal private cost curve, it would be a positive externality and social optimality would require a greater output than Q_p rather than a reduction of output.

Institutional ecological economists in the tradition of Karl William Kapp provide a different definition of social costs, i.e. that share of the total costs of production that is not born by producers but is shifted to 3rd parties, future generations or society at large. Kapp, hence, rejected Pigou's confusing terminology of externalities and provides several hundred pages of empirical data to support his argument that social costs are systemic, i.e. rooted in profit maximizing behavior of businesses, and an enormous problem of modern civilization. In the real world, they are usually not or cannot be internalized and must not be considered as accidental minor aberration from the "optimal norm" that can be fixed with ad hoc measures.

Sunk Costs

In economics and business decision-making, **sunk costs** are retrospective (past) costs that have already been incurred and cannot be recovered. Sunk costs are sometimes contrasted with *prospective costs*, which are future costs that may be incurred or changed if an action is taken. Both retrospective and prospective costs may be either fixed (that is, they are not dependent on the volume of economic activity, however measured) or variable (dependent on volume).

In traditional microeconomic theory, only prospective (future) costs are relevant to an investment decision. Traditional economics proposes that an economic actor not let sunk costs influence one's decisions, because doing so would not be rationally assessing a decision exclusively on its own merits. The decision-maker may make rational decisions according to their own incentives; these incentives may dictate different decisions than would be dictated by efficiency or profitability, and this is considered an *incentive problem* and distinct from a sunk cost problem.

Evidence from behavioral economics suggests this theory fails to predict real-world behavior. Sunk costs greatly affect actors' decisions, because many humans are loss-averse and thus normally act irrationally when making economic decisions.

Sunk costs should not affect the rational decision-maker's best choice. However, until a decision-maker irreversibly commits resources, the prospective cost is an avoidable future cost and is properly included in any decision-making processes. For example, if one is considering preordering movie tickets, but has not actually purchased them yet, the cost remains avoidable. If the price of the tickets rises to an amount that requires him to pay more than the value he places on them, he should figure the change in prospective cost into the decision-making and re-evaluate his decision.

Description

The sunk cost is distinct from economic loss. For example, when a car is purchased, it can subsequently be resold; however, it will probably not be resold for the original purchase price. The economic loss is the difference (including transaction costs). The sum originally paid should not affect any rational future decision-making about the car, regardless of the resale value: if the owner can derive more value from selling the car than not selling it, it should be sold, regardless of the price paid. In this sense, the sunk cost is not a precise quantity, but an economic term for a sum paid, in the past, which is no longer relevant to decisions about the future; it may be used inconsistently in quantitative terms as the original cost or the expected economic loss. It may also be used as shorthand for an error in analysis due to the sunk cost fallacy, irrational decision-making or, most simply, as irrelevant data.

Features characterizing the sunk cost heuristic

Two specific features characterizing the sunk cost heuristic worth mentioning are:

1. An overly optimistic probability bias, whereby after an investment the evaluation of one's investment-reaping dividends is increased.
2. The requisite of personal responsibility. Sunk cost appears to operate chiefly in those who feel personal responsibility for the investments that are to be viewed as sunk.

Overly optimistic probability bias

In 1968 Knox and Inkster, in what is perhaps the classic sunk cost experiment, approached 141 horse bettors: 72 of the people had just finished placing a \$2.00 bet within the past 30 seconds, and 69 people were about to place a \$2.00 bet in the next 30 seconds. Their hypothesis was that people who had just committed themselves to a course of action (betting \$2.00) would reduce post-decision dissonance by believing more strongly than ever that they had picked a winner. Knox and Inkster asked the bettors to rate their horse's chances of winning on a 7-point scale. What they found was that people who were about to place a bet rated the chance that their horse would win at an average of 3.48 which corresponded to a "fair chance of winning" whereas people who had just finished betting gave an average rating of 4.81 which corresponded to a "good chance of winning". Their hypothesis was confirmed: after making a \$2.00 commitment, people became more confident their bet would pay off. Knox and Inkster performed an ancillary test on the patrons of the horses themselves and managed (after normalization) to repeat their finding almost identically.

Additional evidence of inflated probability estimations can be found in Arkes and Blumer (1985) and Arkes & Hutzel (2000).

Requisite of personal responsibility

In a study of 96 business students in 1976 Staw and Fox gave the subjects a choice between making an R&D investment in either an underperforming company department, or in other sections of the hypothetical company. Staw and Fox divided the participants into two groups; a low responsibility condition and a high responsibility condition. In the high responsibility condition the participants were told that they as manager had made an earlier, disappointing R&D investment. In the low responsibility condition, subjects were told that a former manager had made a previous R&D investment in the underperforming division and were given the same profit data as the other group. In both cases subjects were then asked to make a new \$20 million investment. There was a significant interaction between assumed responsibility and average investment, with the high responsibility condition averaging \$12.97 million and the low condition averaging \$9.43 million.

Similar results have been obtained in earlier studies by Staw (1974, 1976) and by Arkes and Blumer (1985) and Whyte (1986).

Loss aversion and the sunk cost fallacy

Many people have strong misgivings about "wasting" resources (loss aversion). In the above example involving a non-refundable movie ticket, many people, for example, would feel obliged to go to the movie despite not really wanting to, because doing otherwise would be wasting the ticket price; they feel they've passed the point of no return. This is sometimes referred to as the *sunk cost fallacy*. Economists would label this behavior "irrational": it is inefficient because it misallocates resources by depending on information that is irrelevant to the decision being made. Colloquially, this is known as "throwing good money after bad".

This line of thinking, in turn, may reflect a non-standard measure of utility, which is ultimately subjective and unique to the consumer. A ticket-buyer who purchases a ticket to a bad movie in advance makes a semi-public commitment to watching it. To leave early is to make this lapse of judgment manifest to strangers, an appearance he might otherwise choose to avoid. Alternatively, he may take pride in having recognized the opportunity cost of the alternative use of time.

The idea of sunk costs is often employed when analyzing business decisions. A common example of a sunk cost for a business is the promotion of a brand name. This type of marketing incurs costs that cannot normally be recovered. It is not typically possible to later "demote" one's brand names in exchange for cash. A second example is R&D costs. Once spent, such costs are sunk and should have no effect on future pricing decisions. So a pharmaceutical company's attempt to justify high prices because of the need to recoup R&D expenses is fallacious. The company will charge the same price whether R&D had cost one dollar or one million dollars. R&D costs do count when deciding whether to spend the money on R&D but once spent they have no effect on firm decisions.

The sunk cost fallacy is in game theory sometimes known as the "Concorde Fallacy", referring to the fact that the British and French governments continued to fund the joint development of Concorde even after it became apparent that there was no longer an economic case for the aircraft. The project was regarded privately by the British government as a "commercial disaster" which should never have been started, and was almost cancelled, but political and legal issues had ultimately made it impossible for either government to pull out.

Sunk cost dilemma

The economic approach that sunk costs should not be considered when decisions are being made can lead to a situation where the sum of a number of good decisions can lead to one big disaster. This dilemma situation can be described using a game theory approach for 1-player games.

The sunk cost dilemma with its sequence of good decisions should not be confused with the sunk cost fallacy, where a misconception of sunk costs can lead to bad decisions.

Bygones principle

The *bygones principle* is an economic theory used in business. Economists stress the "extra" or "marginal" costs and benefits of every decision. The theory emphasizes the importance of ignoring past costs and only taking into account the future costs and benefits when making decisions. It states that when making a decision, one should make a hard-headed calculation of the extra costs one will incur and weigh these against its extra advantages.

Example

An important example of this is related to nuclear power. In the late 1980s, about two dozen partially complete nuclear power plants dotted the USA's landscape. Some had already absorbed billions of dollars of investment but were not yet ready to operate.

One particularly difficult case was the Shoreham plant on Long Island Sound, New York. By 1987 the owner had spent \$5.5 billion on bricks, mortar, fuel rods, and interest, but the operating license had not been granted. From an economic point of view, the \$5.5 billion of past investments should not be weighed in decision-making processes.

The bygones principle would state that the \$5.5 billion of past cost is irrelevant. From an economic point of view, the only relevant issue concerns future costs and benefits – that is, the economic benefits of the electricity that Shoreham would produce.

The key to observe in making this calculation is that the sunk cost of \$5.5 billion is irrelevant to future costs and benefits. Studies indicated that, if the \$5.5 billion were ignored, the future costs of the nuclear power plant would be slightly less than the next-best alternative, even though the total cost was far higher than the alternative. A purely

economic analysis would conclude that the most efficient outcome would be to finish the construction and open the Shoreham plant. However, citing many reasons, including the sunk costs, the plant was closed by protests in 1989 without generating any commercial electrical power.

Chapter 14

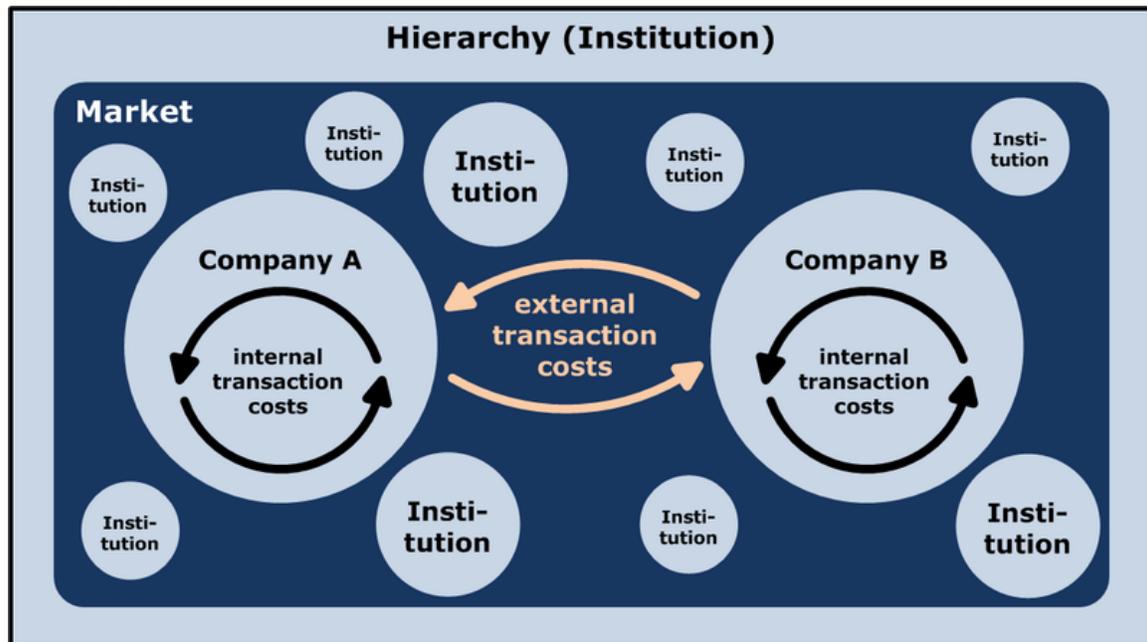
Transaction Cost

In economics and related disciplines, a **transaction cost** is a cost incurred in making an economic exchange (restated: the cost of participating in a market). For example, most people, when buying or selling a stock, must pay a commission to their broker; that commission is a transaction cost of doing the stock deal. Or consider buying a banana from a store; to purchase the banana, your costs will be not only the price of the banana itself, but also the energy and effort it requires to find out which of the various banana products you prefer, where to get them and at what price, the cost of traveling from your house to the store and back, the time waiting in line, and the effort of the paying itself; the costs above and beyond the cost of the banana are the transaction costs. When rationally evaluating a potential transaction, it is important to consider transaction costs that might prove significant.

A number of kinds of transaction cost have come to be known by particular names:

- *Search and information costs* are costs such as those incurred in determining that the required good is available on the market, which has the lowest price, etc.
- *Bargaining costs* are the costs required to come to an acceptable agreement with the other party to the transaction, drawing up an appropriate contract and so on. In game theory this is analyzed for instance in the game of chicken. On asset markets and in market microstructure, the transaction cost is some function of the distance between the bid and ask.
- *Policing and enforcement costs* are the costs of making sure the other party sticks to the terms of the contract, and taking appropriate action (often through the legal system) if this turns out not to be the case.

History of development



The model shows institutions and market as a possible form of organization to coordinate economic transactions. When the external transaction costs are higher than the internal transaction costs, the company will grow. If the internal transaction costs are higher than the external transaction costs the company will be downsized by outsourcing, for example.

The idea that transactions form the basis of an economic thinking was introduced by the institutional economist John R. Commons (1931). He says that, "These individual actions are really trans-actions instead of either individual behavior or the "exchange" of commodities. It is this shift from commodities and individuals to transactions and working rules of collective action that marks the transition from the classical and hedonic schools to the institutional schools of economic thinking. The shift is a change in the ultimate unit of economic investigation. The classic and hedonic economists, with their communistic and anarchistic offshoots, founded their theories on the relation of man to nature, but institutionalism is a relation of man to man. The smallest unit of the classic economists was a commodity produced by labor. The smallest unit of the hedonic economists was the same or similar commodity enjoyed by ultimate consumers. One was the objective side, the other the subjective side, of the same relation between the individual and the forces of nature. The outcome, in either case, was the materialistic metaphor of an automatic equilibrium, analogous to the waves of the ocean, but personified as "seeking their level." But the smallest unit of the institutional economists is a unit of activity -- a transaction, with its participants. Transactions intervene between the labor of the classic economists and the pleasures of the hedonic economists, simply because it is society that controls access to the forces of nature, and transactions are, not the "exchange of commodities," but the alienation and acquisition, between individuals, of the rights of property and liberty created by society, which must therefore be

negotiated between the parties concerned before labor can produce, or consumers can consume, or commodities be physically exchanged".

The term "transaction cost" is frequently thought to have been coined by Ronald Coase, who used it to develop a theoretical framework for predicting when certain economic tasks would be performed by firms, and when they would be performed on the market. However, the term is actually absent from his early work up to the 1970s. While he did not coin the specific term, Coase indeed discussed "costs of using the price mechanism" in his 1937 paper *The Nature of the Firm*, where he first discusses the concept of transaction costs, and refers to the "Costs of Market Transactions" in his seminal work, *The Problem of Social Cost* (1960). The term "Transaction Costs" itself can instead be traced back to the monetary economics literature of the 1950s, and does not appear to have been consciously 'coined' by any particular individual.

Arguably, transaction cost reasoning became most widely known through Oliver E. Williamson's *Transaction Cost Economics*. Today, transaction cost economics is used to explain a number of different behaviours. Often this involves considering as "transactions" not only the obvious cases of buying and selling, but also day-to-day emotional interactions, informal gift exchanges, etc. Oliver E. Williamson was awarded the 2009 Nobel Prize in Economics

According to Williamson, the determinants of transaction costs are frequency, specificity, uncertainty, limited rationality, and opportunistic behavior.

At least two definitions of the phrase "transaction cost" are commonly used in literature. Transaction costs have been broadly defined by Steven N. S. Cheung as any costs that are not conceivable in a "Robinson Crusoe economy"—in other words, any costs that arise due to the existence of institutions. To Cheung, "transaction costs", if the term is not so popular in economics literatures, should be called "institutional costs". But many economists seem to restrict the definition to exclude costs internal to an organization. The latter definition parallels Coase's early analysis of "costs of the price mechanism" and the origins of the term as a market trading fee.

Starting with the broad definition, many economists then ask what kind of institutions (firms, markets, franchises, etc.) minimize the transaction costs of producing and distributing a particular good or service. Often these relationships are categorized by the kind of contract involved. This approach sometimes goes under the rubric of New Institutional Economics.

A simple example

A supplier may bid in a competitive environment with a customer to build a widget. However, to make the widget, the supplier will be required to build specialized machinery which cannot be easily redeployed to make other products. Once the contract is awarded to the supplier, the relationship between customer and supplier changes from a competitive environment to a monopoly/monopsony relationship, known as a bilateral

monopoly. This means that the customer has greater leverage over the supplier such as when price cuts occur. To avoid these potential costs, "hostages" may be swapped to avoid this event. These hostages could include partial ownership in the widget factory; revenue sharing might be another way.

Car companies and their suppliers often fit into this category, with the car companies forcing price cuts on their suppliers. Defense suppliers and the military appear to have the opposite problem, with cost overruns occurring quite often.

Technologies like enterprise resource planning (ERP) can provide technical support for these strategies.

Chapter 15

Cost-Volume-Profit Analysis

Cost-Volume-profit (CVP), in managerial economics is a form of cost accounting. It is a simplified model, useful for elementary instruction and for short-run decisions.

Cost-volume-profit (CVP) analysis expands the use of information provided by breakeven analysis. A critical part of CVP analysis is the point where total revenues equal total costs (both fixed and variable costs). At this breakeven point (BEP), a company will experience no income or loss. This BEP can be an initial examination that precedes more detailed CVP analysis.

Cost-volume-profit analysis employs the same basic assumptions as in breakeven analysis. The assumptions underlying CVP analysis are:

The behavior of both costs and revenues is linear throughout the relevant range of activity. (This assumption precludes the concept of volume discounts on either purchased materials or sales.) Costs can be classified accurately as either fixed or variable. Changes in activity are the only factors that affect costs. All units produced are sold (there is no ending finished goods inventory). When a company sells more than one type of product, the sales mix (the ratio of each product to total sales) will remain constant.

The components of Cost-Volume-Profit Analysis are:

- Level or volume of activity
- Unit Selling Prices
- Variable cost per unit
- Total fixed costs
- Sales mix

Assumptions

CVP assumes the following:

- Constant sales price;
- Constant variable cost per unit;
- Constant total fixed cost;
- Constant sales mix;
- Units sold equal units produced.

These are simplifying, largely linearizing assumptions, which are often implicitly assumed in elementary discussions of costs and profits. In more advanced treatments and practice, costs and revenue are nonlinear and the analysis is more complicated, but the intuition afforded by linear CVP remains basic and useful.

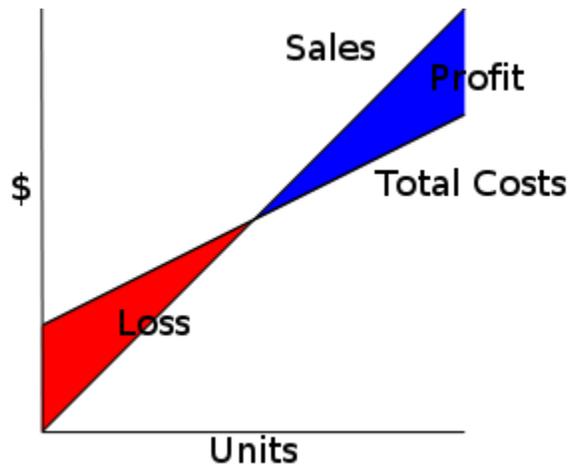
One of the main Methods of calculating CVP is Profit volume ratio: which is $(\text{contribution} / \text{sales}) * 100 =$ this gives us profit volume ratio.

- contribution stands for Sales minus variable costs.

Therefore it gives us the profit added per unit of variable costs.

Model

Basic graph



Basic graph of CVP, demonstrating relation of Total Costs, Sales, and Profit and Loss.

The assumptions of the CVP model yield the following linear equations for total costs and total revenue (sales):

$$\text{Total Costs} = \text{Fixed Costs} + (\text{Unit Variable Cost} \times \text{Number of Units})$$

$$\text{Total Revenue} = \text{Sales Price} \times \text{Number of Units}$$

These are linear because of the assumptions of constant costs and prices, and there is no distinction between Units Produced and Units Sold, as these are assumed to be equal. Note that when such a chart is drawn, the linear CVP model is assumed, often implicitly.

In symbols:

$$\begin{aligned} TC &= TFC + V \times X \\ TR &= P \times X \end{aligned}$$

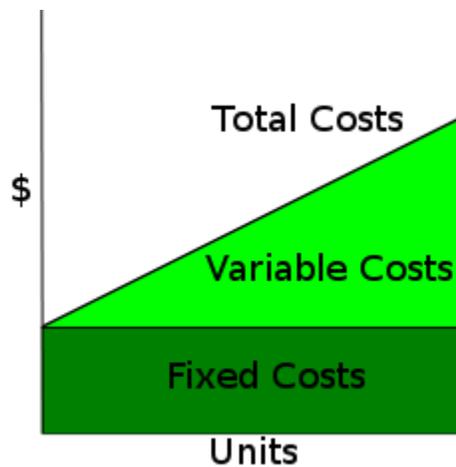
where

- **TC = Total Costs**
- **TFC = Total Fixed Costs**
- **V = Unit Variable Cost (Variable Cost per Unit)**
- **X = Number of Units**
- **TR = S = Total Revenue = Sales**
- **P = (Unit) Sales Price**

Profit is computed as $TR - TC$; it is a profit if positive, a loss if negative.

Break down

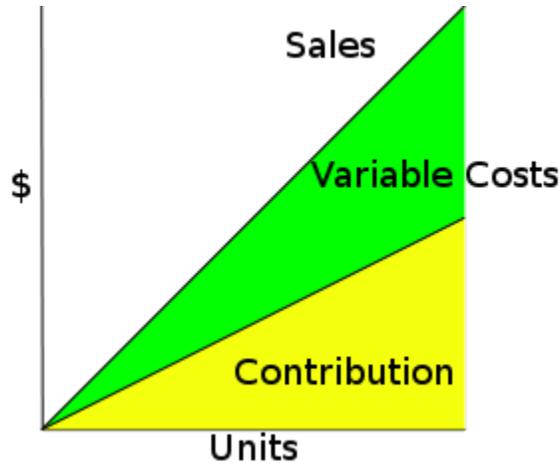
Costs and Sales can be broken down, which provide further insight into operations.



Decomposing Total Costs as Fixed Costs plus Variable Costs.

One can decompose Total Costs as Fixed Costs plus Variable Costs:

$$TC = TFC + V \times X$$



Decomposing Sales as Contribution plus Variable Costs.

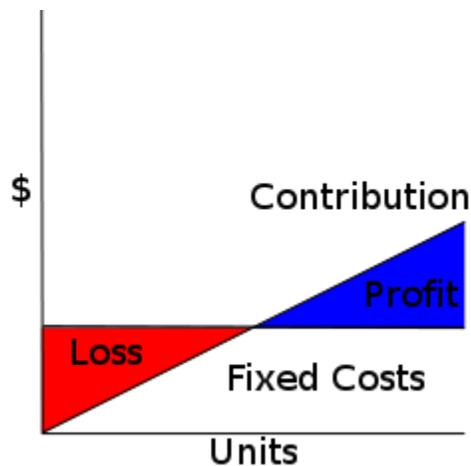
Following a matching principle of matching a portion of sales against variable costs, one can decompose Sales as Contribution plus Variable Costs, where **contribution** is "what's left after deducting variable costs". One can think of contribution as "the marginal contribution of a unit to the profit", or "contribution towards offsetting fixed costs".

In symbols:

$$\begin{aligned}
 TR &= P \times X \\
 &= ((P - V) + V) \times X \\
 &= (C + V) \times X \\
 &= C \times X + V \times X
 \end{aligned}$$

where

- **C = Unit Contribution (Margin)**



Profit and Loss as Contribution minus Fixed Costs.

Subtracting Variable Costs from both Costs and Sales yields the simplified diagram and equation for Profit and Loss.

In symbols:

$$\begin{aligned}
 PL &= TR - TC \\
 &= (C + V) \times X - (TFC + V \times X) \\
 &= C \times X - TFC
 \end{aligned}$$

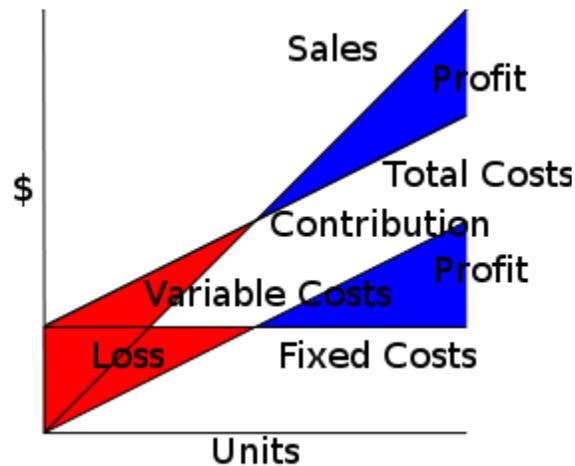


Diagram relating all quantities in CVP.

These diagrams can be related by a rather busy diagram, which demonstrates how if one subtracts Variable Costs, the Sales and Total Costs lines shift down to become the Contribution and Fixed Costs lines. Note that the Profit and Loss for any given number of unit sales is the same, and in particular the break-even point is the same, whether one computes by Sales = Total Costs or as Contribution = Fixed Costs. Mathematically, the contribution graph is obtained from the sales graph by a shear, to be precise $\begin{pmatrix} 1 & 0 \\ -V & 1 \end{pmatrix}$, where V are Unit Variable Costs.

Applications

CVP simplifies the computation of breakeven in break even analysis, and more generally allows simple computation of Target Income Sales. It simplifies analysis of short run trade-offs in operational decisions.

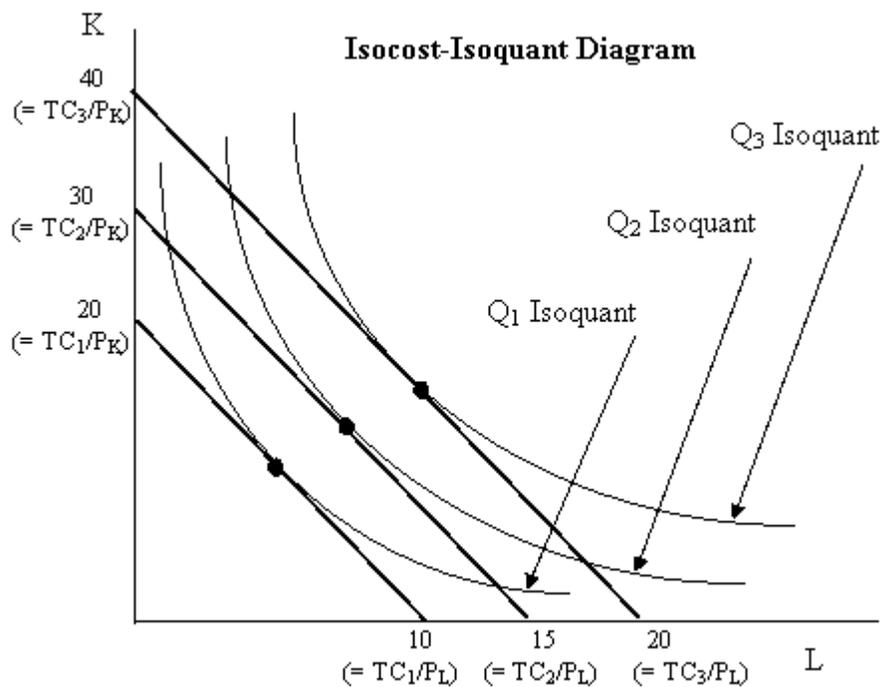
Limitations

CVP is a **short run, marginal** analysis: it assumes that unit variable costs and unit revenues are constant, which is appropriate for small deviations from current production

and sales, and assumes a neat division between fixed costs and variable costs, though in the long run all costs are variable. For longer-term analysis that considers the entire life-cycle of a product, one therefore often prefers activity-based costing or throughput accounting.

Chapter 16

Isocost



Isocost v. Isoquant Graph

In economics an **isocost** line shows all combinations of inputs which cost the same total amount. Although similar to the budget constraint in consumer theory, the use of the isocost line pertains to cost-minimization in production, as opposed to utility-maximization. For the two production inputs labour and capital, with fixed unit costs of the inputs, the equation of the isocost line is

$$rK + wL = C$$

where w represents the wage rate of labour, r represents the rental rate of capital, K is the amount of capital used, L is the amount of labour used, and C is the total cost of acquiring those quantities of the two inputs.

The absolute value of the slope of the isocost line, with capital plotted vertically and labour plotted horizontally, equals the ratio of unit costs of labour and capital. The slope is:

$$-w/r.$$

The isocost line is combined with the isoquant map to determine the optimal production point at any given level of output. Specifically, the point of tangency between any isoquant and an isocost line gives the lowest-cost combination of inputs that can produce the level of output associated with that isoquant. Equivalently, it gives the maximum level of output that can be produced for a given total cost of inputs.

The cost-minimization problem

The cost-minimization problem of the firm is to choose an input bundle (K,L) feasible for the output level y that costs as little as possible. A cost-minimizing input bundle is a point on the isoquant for the given y that is on the lowest possible isocost line. Put differently, a cost-minimizing input bundle must satisfy two conditions:

1. it is on the y -isoquant
2. no other point on the y -isoquant is on a lower isocost line.

The case of smooth isoquants convex to the origin

If the y -isoquant is smooth and convex to the origin and the cost-minimizing bundle involves a positive amount of each input, then at a cost-minimizing input bundle an isocost line is tangent to the y -isoquant. Now since the absolute value of the slope of the isocost line is the input cost ratio w/r , and the absolute value of the slope of an isoquant is the marginal rate of technical substitution (MRTS), we reach the following conclusion: If the isoquants are smooth and convex to the origin and the cost-minimizing input bundle involves a positive amount of each input, then this bundle satisfies the following two conditions:

- It is on the y -isoquant (i.e. $F(K, L) = y$ where F is the production function), and
- the MRTS at (K, L) equals w/r .

The condition that the MRTS be equal to w/r can be given the following intuitive interpretation. We know that the MRTS is equal to the ratio of the marginal products of the two inputs. So the condition that the MRTS be equal to the input cost ratio is equivalent to the condition that the marginal product per dollar is equal for the two inputs. This condition makes sense: at a particular input combination, if an extra dollar

spent on input 1 yields more output than an extra dollar spent on input 2, then more of input 1 should be used and less of input 2, and so that input combination cannot be optimal. Only if a dollar spent on each input is equally productive is the input bundle optimal.

Chapter 17

Consumer Choice

Consumer choice is a theory of microeconomics that relates preferences to consumer demand curves. The link between personal preferences, consumption, and the demand curve is one of the most complex relations in economics. Implicitly, economists assume that anything purchased will be consumed, unless the purchase is for a productive activity.

Preferences are the desires by each individual for the consumption of goods and services, and ultimately translate into employment choices based on abilities and the use of the income from employment for purchases of goods and services to be combined with the consumer's time to define consumption activities.

Consumption is separated from production, logically, because two different consumers are involved. In the first case consumption is by the primary individual; in the second case, a producer might make something that he would not consume himself. Therefore, different motivations and abilities are involved.

The models that make up consumer theory are used to represent prospectively observable demand patterns for an individual buyer on the hypothesis of constrained optimization.

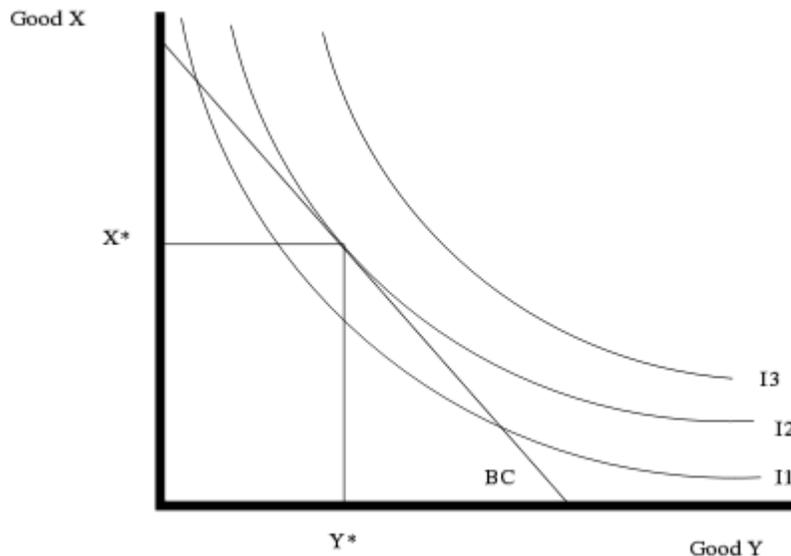
Prominent variables used to explain the rate at which the good is purchased (demanded) are the price per unit of that good, prices of related goods, and wealth of the consumer.

The fundamental theorem of demand states that the rate of consumption falls as the price of the good rises. This is called the *substitution effect*. Clearly if one does not have enough money to pay the price then they cannot buy any of that item. As prices rise, consumers will substitute away from higher priced goods and services, choosing less costly alternatives. Subsequently, as the wealth of the individual rises, demand increases, shifting the demand curve higher at all rates of consumption. This is called the *income*

effect. As wealth rises, consumers will substitute away from less costly inferior goods and services, choosing higher priced alternatives.

Model setup

For an individual, indifference curves and an assumption of constant prices and a fixed income in a two-good world will give the following diagram. The consumer can choose any point on or below the budget constraint line BC. This line is diagonal since it comes from the equation $xp_X + yp_Y \leq \text{income}$. In other words, the amount spent on both goods together is less than or equal to the income of the consumer. The consumer will choose the indifference curve with the highest utility that is within his budget constraint. Every point on I₃ is outside his budget constraint so the best that he can do is the single point on I₂ that is tangent to his budget constraint. He will purchase x^* of good X and y^* of good Y.



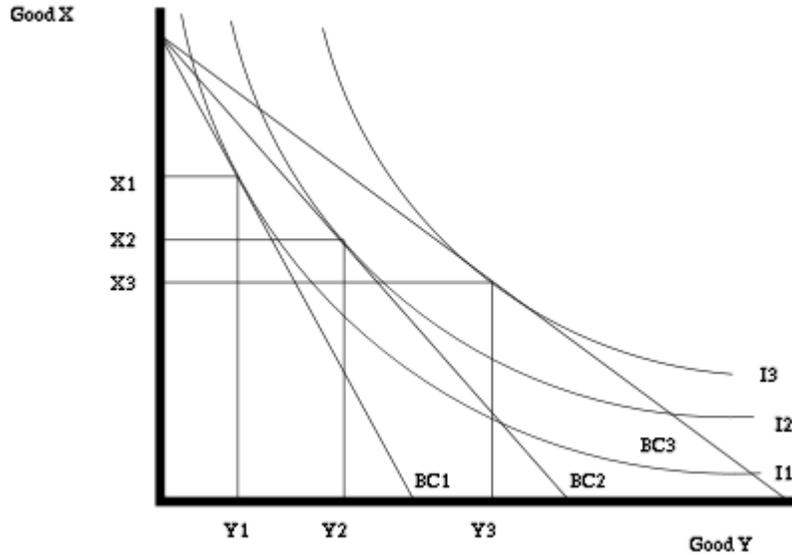
Income effect and price effect deal with how the change in price of a commodity changes the consumption of the good. The theory of consumer choice examines the trade-offs and decisions people make in their role as consumers as prices and their income changes.

Substitution effect

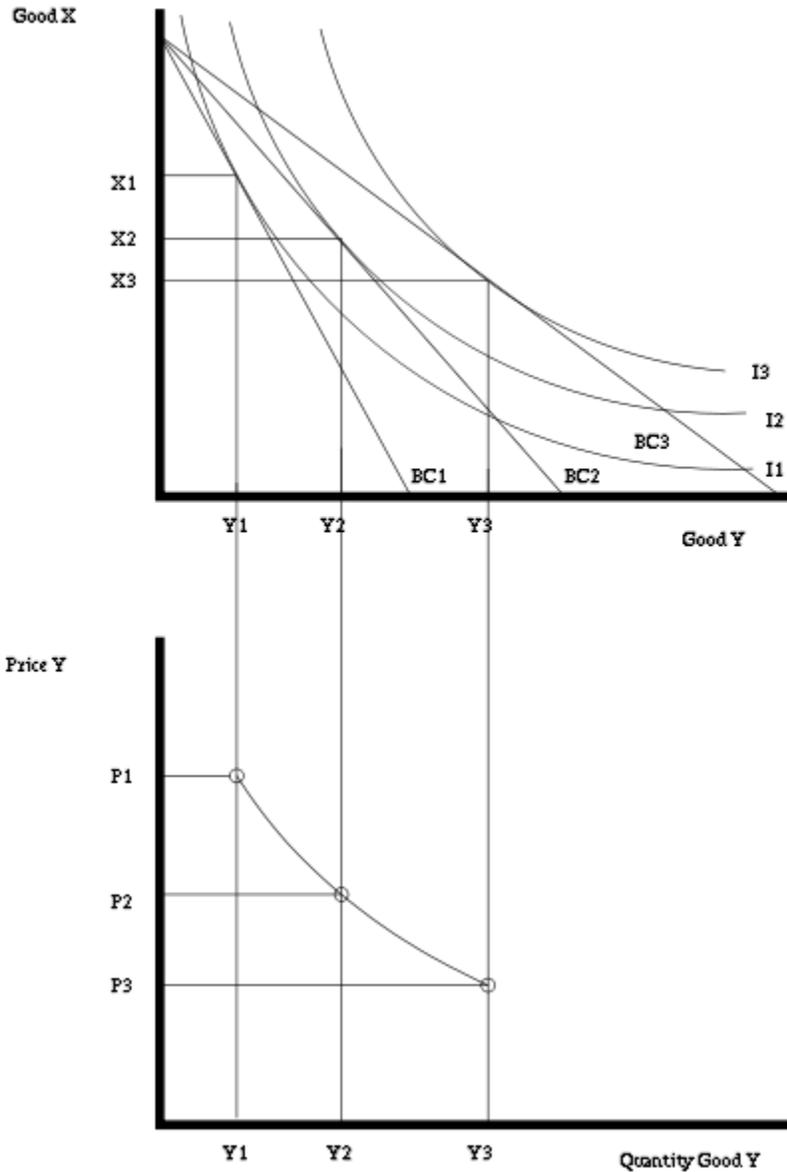
The substitution effect is the effect observed with changes in relative price of goods. This effect basically affects the movement along the curve.

These curves can be used to predict the effect of changes to the budget constraint. The graphic below shows the effect of a price increase for good Y. If the price of Y increases, the budget constraint will pivot from BC₂ to BC₁. Notice that because the price of X does not change, the consumer can still buy the same amount of X if he or she chooses to buy only good X. On the other hand, if the consumer chooses to buy only good Y, he or she will be able to buy less of good Y because its price has increased.

To maximize the utility with the reduced budget constraint, BC1, the consumer will re-allocate consumption to reach the highest available indifference curve which BC1 is tangent to. As shown on the diagram below, that curve is I1, and therefore the amount of good Y bought will shift from Y2 to Y1, and the amount of good X bought to shift from X2 to X1. The opposite effect will occur if the price of Y decreases causing the shift from BC2 to BC3, and I2 to I3.

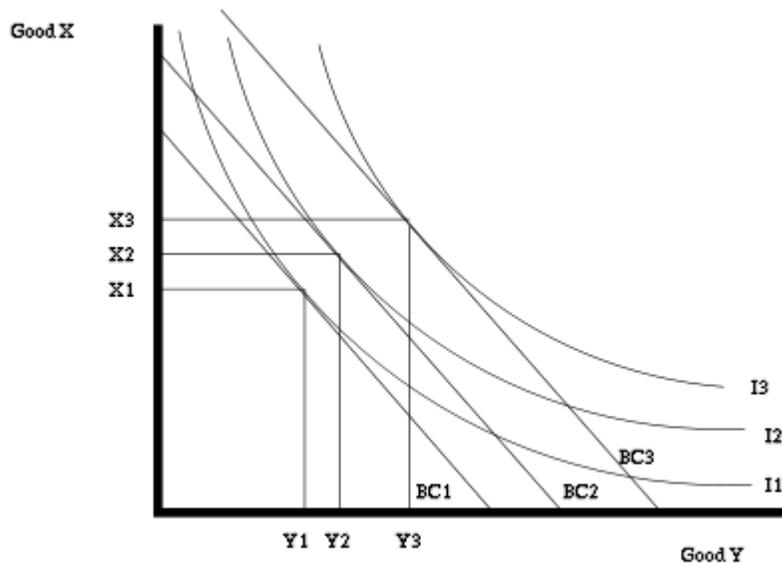


If these curves are plotted for many different prices of good Y, a demand curve for good Y can be constructed. The diagram below shows the demand curve for good Y as its price varies. Alternatively, if the price for good Y is fixed and the price for good X is varied, a demand curve for good X can be constructed.

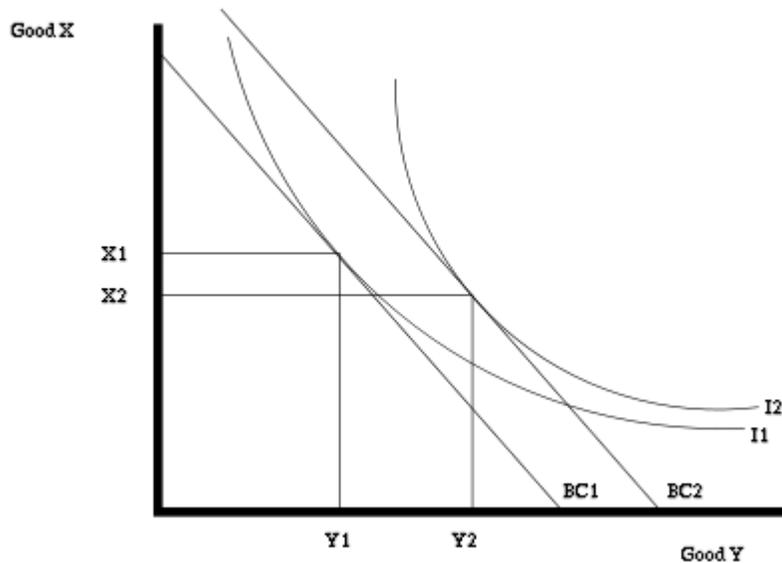


Income effect

Another important item that can change is the money income of the consumer. The income effect is the phenomenon observed through changes in purchasing power. It reveals the change in quantity demanded brought by a change in real income (utility). Graphically, as long as the prices remain constant, changing the income will create a parallel shift of the budget constraint. Increasing the income will shift the budget constraint right since more of both can be bought, and decreasing income will shift it left.



Depending on the indifference curves, as income increases, the amount purchased of a good can either increase, decrease or stay the same. In the diagram below, good Y is a normal good since the amount purchased increased as the budget constraint shifted from BC1 to the higher income BC2. Good X is an inferior good since the amount bought decreased as the income increases.



Δy_1^n is the change in the demand for good 1 when we change income from m' to m , holding the price of good 1 fixed at p_1' :

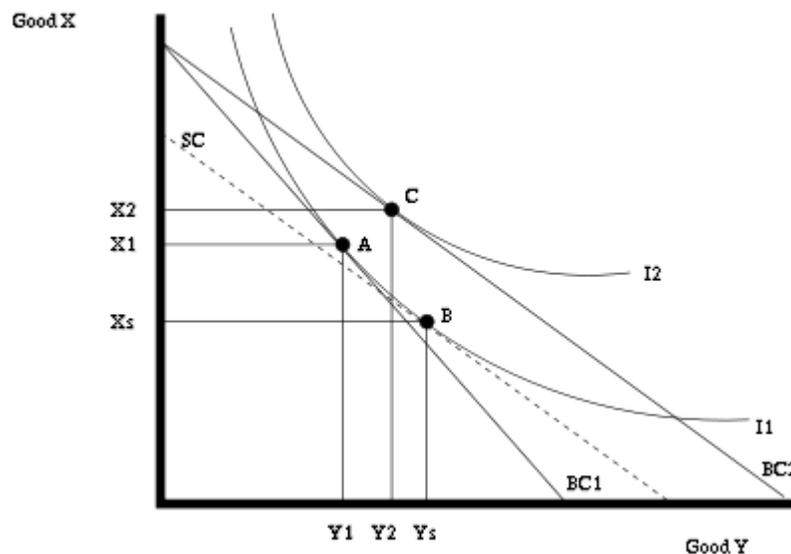
$$\Delta y_1^n = y_1(p_1', m) - y_1(p_1', m').$$

Price effect as sum of substitution and income effects

Every price change can be decomposed into an income effect and a substitution effect; the price effect is the sum of substitution and income effects.

The substitution effect is a price change that alters the slope of the budget constraint but leaves the consumer on the same indifference curve. In other words, it illustrates the consumer's new consumption basket after the price change while being compensated as to allow the consumer to be as happy as he or she was previously. By this effect, the consumer is posited to substitute toward the good that becomes comparatively less expensive. In the illustration below this corresponds to an imaginary budget constraint denoted SC being tangent to the indifference curve I1.

If the good in question is a *normal good*, then the income effect from the rise in purchasing power from a price fall reinforces the substitution effect. If the good is an *inferior good*, then the income effect will offset in some degree the substitution effect. If the income effect for an inferior good is sufficiently strong, the consumer will buy less of the good when it becomes less expensive, a Giffen good (commonly believed to be a rarity).



In the figure, the substitution effect, Δy_1^s , is the change in the amount demanded for y when the price of good y falls from p_1 to p'_1 (increasing purchasing power for y) and, at the same time, the money income falls from m to m' to keep the consumer at the same level of utility on I_1 :

$$\Delta y_1^s = y_1(p'_1, m') - y_1(p_1, m).$$

The substitution effect increases the amount demanded of good y from y_1 to y_s . In the example, the income effect of the price fall in y partly offsets the substitution effect as

the amount demanded of y goes from y_2 to y_3 . Thus, the price effect is the algebraic sum of the substitution effect and the income effect.

Assumptions of consumer preference theory

The behavioral assumption of consumer theory is that all consumers are rational decision makers who seek to maximize utility. More specifically all consumers seek to maximize a utility function subject to a budgetary constraint. This means that consumer theory asserts that all consumers purchase that combination of goods and services that will make them happiest given the amount of income they have to spend. The fundamental problem of consumer theory is to derive a demand function from the behavioral postulates of the theory. Economists' modern solution to this problem is indifference curve analysis. Indifference curve analysis begins with the utility function. The utility function is treated as an index of utility. All that is necessary is that the utility index change as more preferred bundles are consumed. Indifference curves are typically numbered with the number increasing as more preferred bundles are consumed. However, it is not necessary that numbers be used - any indexing system would suffice - colors for example. The advantage of numbers is that their use makes the math simpler. Numbers used to index indifference curves have no cardinal significance. For example if three indifference curves are labeled 1, 4, and 16 respectively that means nothing more than the bundles "on" indifference curve 4 are more preferred than the bundles "on" indifference curve 1. The fact that the index number is a multiple of another is of no significance. For example, the bundles of good on 4 does not mean that they are four times more satisfying than those on 1. As noted they merely mean they are more satisfying.

1. Preferences are complete -

- Assume that there are two consumption bundles A and B each containing two commodities x and y. A consumer can unambiguously make the following comparisons:
 - A is preferred to B
 - B is referred to A
 - A is indifferent to B
 - Note that this axiom precludes the possibility that the consumer cannot decide. The consumer will chose A, B, or simply not care which she receives; "Six of one or half a dozen of another." Indifference is different from indecision, as "I don't care" is different from "I can't make up my mind".
 - Note this assumption means that a consumer is able to make this comparison with respect to every conceivable bundle of goods.
1. Preferences are reflexive
- Means that if A and B are in all respect identical the consumer will recognize this fact and be indifferent in comparing A and B

1. Preference are transitive

- If A is preferred to B and B is preferred to C then A must be preferred to C.
- this also means that if the consumer is indifferent between A and B and is indifferent between B and C she will be indifferent between A and C.
- This is the consistency assumption. A real world example: You enter a restaurant where the available soups are pea soup, tomato soup and mushroom soup. The waiter asks "What will you have?" You say "I don't know; I prefer pea soup over tomato soup and tomato soup over mushroom soup, and I prefer mushroom soup over pea soup. Listen, why don't you chose for me." The waiter won't be able to help you.
- This assumption eliminates the possibility of intersecting indifference curves.

1. Preferences are continuous

- If A is preferred to B and C is infinitesimally close to B then A is preferred to C.
- "continuous" means infinitely divisible - just like there are an infinity of numbers between 1 and 2, all bundles are infinitely divisible. This assumption allows the use of curves and areas all the basic requirements of differential and integral calculus. Sine qua non.

1. Preferences exhibit non-satiation.

- if A and B have identical amounts of x and A has a little more y than B then A is preferred to B,
- this is the more is better assumption or i never saw a donut i didn't like
- Among other things this assumption precludes the embarrassment of circular indifference curve -
- Non-satiation is not a necessary but a convenient assumption. It avoids unnecessary complications in the mathematical models.

1. Indifference Curves exhibit diminishing marginal rates of substitution

- This assumption assures that indifference curves are smooth and convex to the origin.
- This assumption is implicit in the last assumption.
- This assumption also set the stage for using techniques of constrained optimization. Because the shape of the curve assures that the first derivative is negative and the second is positive.

- The MRS tells how much y a person is willing to sacrifice to get one more unit of x .
- This assumption incorporates the theory of diminishing marginal utility.

The primary reason to have these technical preferences is to replicate the properties of the real number system so the math will work.

Note the assumptions do not guarantee that the demand curve will be negatively sloped. A positively sloped curve is not inconsistent with the assumptions. Binger & Hoffman, *Microeconomics with Calculus* 2d ed. (Addison Wesley 1998) at 141-43.

Use Value

In Marx's critique of political economy, any labor-product has a value and a use value, and if it is traded as a commodity in markets, it additionally has an exchange value, most often expressed as a money-price. Marx acknowledges that commodities being traded also have a general utility, implied by the fact that people want them, but he argues that this by itself tells us nothing about the specific character of the economy in which they are produced and sold.

Labor-leisure tradeoff

Consumer theory can also be used to analyze a consumer's choice between leisure and labor. Leisure is considered one good (often put on the horizontal-axis) and consumption is considered the other good. Since a consumer has a finite and scarce amount of time, he must make a choice between leisure (which earns no income for consumption) and labor (which does earn income for consumption).

The previous model of consumer choice theory is applicable with only slight modifications. First, the total amount of time that an individual has to allocate is known as his time endowment, and is often denoted as T . The amount an individual allocates to labor (denoted L) and leisure (l) is constrained by T such that:

$$l + L = T$$

or

$$l + (T - l) = T$$

A person's consumption is the amount of labor they choose multiplied by the amount they are paid per hour of labor (their wage, often denoted w). Thus, the amount that a person consumes is:

$$C = w(T - l)$$

When a consumer chooses no leisure ($l = 0$) then $T - l = T$ and $C = wT$.

From this labor-leisure tradeoff model, the substitution and income effects of various changes in price caused by welfare benefits, labor taxation, or tax credits can be analyzed.