COST ALLOCATION:
Methods, Principles, Applications

edited by

H. Peyton YOUNG

School of Public Affairs
University of Maryland
College Park, Maryland 20742
U.S.A.

formerly at the
International Institute for
Applied Systems Analysis

1985
NORTH-HOLLAND
AMSTERDAM - NEW YORK - OXFORD
CONTENTS

Preface vii
Acknowledgments xiii

PART I. THEORY 1
1 Methods and Principles of Cost Allocation 3
   H. Peyton Young
2 Common Cost Allocation in the Firm 31
   Gary C. Diddle and Richard Steinberg
3 On the Use of Game-Theoretic Concepts in Cost Accounting 55
   Leonard J. Mirman, Yair Tauman, and Israel Zang
4 The Cooperative Form, the Value, and the Allocation of Joint 79
   Costs and Benefits
   Martin Shubik
5 The Impossibility of Incentive-Compatible and Efficient Full 95
   Cost Allocation Schemes
   Theodore Groves
6 Bargaining and Fair Allocation 101
   Terje Ivensberg

PART II. APPLICATIONS 117
7 Perceived Economic Justice: The Example of Public Utility 119
   Regulation
   Edward E. Zajac
8 Economic and Game-Theoretic Issues Associated with Cost 155
   Allocation in a Telecommunications Network
   William W. Sharkey
9 Cost Allocation and Pricing Policy: The Case of French 167
   Telecommunications
   Nicolas Curien
10 Auctioning Landing Rights at Congested Airports 179
    Michel I. Ballinski and Francis M. Sand
11 Cost Allocation in Multipurpose Reservoir Development: 193
    The Japanese Experience
    Norio Okada
PREFACE

How should the common costs of an enterprise be shared "fairly" among its beneficiaries? This problem is widespread, both in public enterprises and also within private firms. It arises in the pricing policies of public utilities providing telephone services, electricity, water, and transport. It occurs in the cost-benefit analyses of public works projects designed to serve different constituencies, such as a multipurpose reservoir. It is implicit in the determination of access fees or user charges for common facilities such as airports or waterways. In private corporations it occurs in the form of internal accounting schemes to allocate common and overhead costs among different divisions of the firm. It even crops up in such routine matters as how to allocate travel expenses among the sponsors of a business trip, or how to share the expenses of running a club among its members.

Such cost allocation problems typically exhibit two features:

(i) costs must be allocated exactly, with no profit or deficit;
(ii) there is no objective basis at hand for attributing costs directly to specific products or services.

The goal of analysis is to devise criteria and methods for solving these problems in a just, equitable, fair, and reasonable manner. Cost allocation is thus ultimately concerned with fairness. The methods and principles of cost allocation that are likely to find acceptance must somehow be grounded in primitive, common-sense ideas of fairness and equity.

But precisely what is meant by the word fair? According to Webster (1981), it stems from fagar, an Old High German term meaning "beautiful". Fair means, firstly, "attractive in appearance: pleasant to view". Significantly, a secondary meaning is "pleasing to hear: inspiring hope or confidence often deceptively ... specious". It is closely connected to such ideas as just, equitable, impartial, unbiased, objective. "Fair ... implies a disposition in a person or group to achieve a fitting and right balance of claims or considerations that is free from undue favoritism even to oneself ... Just stresses, more than fair, a disposition to conform with, or conformity with the standard of what is right, true, or lawful, despite strong, especially personal, influences tending to subvert that conformity" (Id).
Fairness involves a balancing of demands, equal treatment, a concern with legitimacy, lack of coercion. In effect, it provides a basis for voluntary consent. Indeed, one could propose the following empirical test of a "fairness principle": is it sufficiently compelling to cause parties with diverse interests to voluntarily agree to its application? Whether the reader finds the cost allocation principles proposed in subsequent chapters compelling will of course depend to some degree on taste, for fairness is partly in the eye of the beholder.

Two approaches to achieving fair allocations may be distinguished. One is strictly normative: all of the objective data are at hand and the problem is to devise an appropriate formula for making an allocation based on these data. Such techniques are typically encountered in cost–benefit analysis or internal accounting schemes in corporations. The second approach is to design a procedure – e.g., a court trial, an arbitration rule, an auction, or a competitive market – that seems fair and impartial a priori, and by its functioning produces an allocation (which might or might not seem fair a posteriori). Most of the chapters in this volume deal with normative criteria and solutions, but some are concerned with procedural allocation mechanisms. In both cases game theory plays an important role: cooperative game theory in the design of normative formulas, and noncooperative game theory in the design of procedures.

Despite its numerous practical applications, there has until recently been relatively little theory, and even less empirical observation, about how cost allocations are, or ought to be, accomplished. Traditionally the subject has been regarded by economists and accountants as at best a necessary evil. In the economics literature one solution is prominent – "Ramsey pricing" (Ramsey, 1927; Baumol and Bradford, 1970) – but, despite over 50 years of writing on the subject by economists, Ramsey pricing has not found general acceptance by rate-makers and cost–benefit analysts. Similarly, the accounting literature has, until recently, given the problem short shrift. It is even claimed by some that cost allocation is an irrational activity that should be indulged in rarely if at all (Thomas, 1974). (For a more balanced range of views within the accounting profession, see Moriarty, 1981, and also Chapter 2 below.)

The aim of this volume is to explore three issues from diverse points of view. First, what methods are in current use by firms, public utilities, government agencies, accountants, and managers of public facilities? Second, are there general principles of allocative fairness that are supported by observation, common sense, and logic? Third, can theory characterize different methods in terms of their fairness properties and identify new ones?

The volume is divided into two parts. Part I deals with the mathematical formulation of cost allocation problems, the definition of solution concepts, and their characterization axiomatically by general principles of equity. Part II deals with applications to airports, reservoirs, telephone networks, and public rate-making in general. The first part is fairly technical mathematically, while the second is more accessible to the general reader.

Chapter 1 focuses on game-theoretic methods of cost allocation. Four cases are used to illustrate how cost allocation problems can be modeled as cooperative games: sharing municipal water costs, imputing costs to purposes in multipurpose
reservoirs, setting airport landing fees, and allocating business trip expenses among sponsors. The most important allocation methods are defined, including the Shapley value, the core, the nucleolus, the separable costs remaining benefits method, Aumann–Shapley pricing, and Ramsey pricing. The pros and cons of these methods are weighed from a normative standpoint, and then certain principles of allocative fairness distilled that axiomatically characterize particular methods. The argument suggests that the Shapley value (or Aumann–Shapley pricing) and the nucleolus are the most robust and satisfactory methods. Which is most appropriate depends on the structure of the problem at hand.

Chapter 2, by Gary Biddle and Richard Steinberg, gives a scholarly survey and critique of the literature on cost allocation in the firm. They observe that objectivity and verifiability of the data on which allocations are based is of paramount importance to accountants, who can be held legally liable for the accuracy of financial statements. However, firms also employ internal allocation schemes for purposes of accounting and control that are subject to less stringent accounting standards. Biddle and Steinberg cite a 1981 study which found that 80 percent of the firms surveyed allocated some portion of the costs associated with corporate management and service departments among the firm’s constituent divisions. They describe several methods that figure prominently in the accounting literature, including those of Morarit, Louderback, Balachandran–Ramarshnan, and Gangolly. The core and the Shapley value are also analyzed from an accounting perspective. They conclude that, while cost allocation forms an integral part of managerial incentive structures and financial accounting systems in firms, the methods advanced in the theoretical literature have not had much practical impact as yet on accounting practices.

In Chapter 3, Mirman, Tauman, and Zang treat the problem of allocating the joint costs of production among the outputs of a firm. They demonstrate why the Aumann–Shapley pricing mechanism is particularly desirable from a normative standpoint. They explain the mechanics of Aumann–Shapley pricing using examples, and show the difficulties inherent in other approaches, such as modified forms of marginal cost pricing. They also point out that Aumann–Shapley prices can be computed relatively easily when the firm’s cost function is estimated by linear programming techniques. Finally, they extend their axiomatic framework to handle the problem of allocating fixed costs of production among different product lines.

Martin Shubik, who in the early 1960s originated the idea of applying game-theoretic ideas to cost accounting, raises in Chapter 4 a number of important questions about the modeling of cost allocation problems. He points out both the advantages and the pitfalls of the characteristic function as a modeling device: while economical, it ignores crucial distinctions in the information available to the parties, the specifics of coalition formation, the sequencing of moves, and the possibility of threats. In some cases it may be more appropriate to model the problem by a strategic or extensive form game, although there are often difficulties in specifying such models. Shubik suggests a novel approach to bounding the set of reasonable allocations in a cooperative game by defining the notion of an upper and a lower characteristic function, based on the partition function form. The central point, however, is that cost allocation is an exercise in modeling; thus the
full diversity of interested parties, their purposes, and the goals of allocation must be recognized before meaningful solutions can be obtained. He predicts that accounting for the combinatorics of joint costs and revenues may prove to be the next major breakthrough in accounting – analogous to the development of input–output and national income accounts – and, as in those cases, the potential usefulness of the theory will stimulate the collection of necessary data to carry out the calculations.

In Chapter 5 Ted Groves addresses the issue of how cost allocation in the firm may affect the reliability of information reported to central management by decentralized divisions. Groves treats the case of a firm providing some common good or service (e.g., computing facilities) for the nonexclusive use of its autonomous divisions, the full cost of which is allocated by a prescribed method among the various divisions. It is assumed management does not know the true demands of the divisions for the service. He shows the impossibility of concocting a method that allocates costs exactly and implements the optimal level of service by inducing the divisions to report their true demands. In other words, no full cost allocation scheme is both efficient and incentive-compatible.

This result stands in contrast to a somewhat different situation portrayed in Chapter 1 (Section 5). Suppose that management can designate specific amounts of the service for the exclusive use of each division. In this case a mechanism can be designed in which divisions bid the amounts they are willing to pay for the designated levels of service. It can be shown that a noncooperative equilibrium exists that results in an efficient level of service and covers all costs. Moreover, if the cost game has a nonempty core, then full costs can be allocated exactly.

The concluding chapter of Part I, by Terje Lensberg, views cost allocation within the broader framework of economic welfare. He poses the following question: under what circumstances is the decentralized allocation of costs and benefits within public enterprises and firms "consistent" with an allocation that is fair for society as a whole? In other words, what characterizes societal allocations that have the property that they seem fair when viewed by any subgroup of society (assuming the others' allocations are fixed)? This "consistency" or "stability" principle (together with several regularity properties) implies that the allocation must maximize some additively separable social welfare function on the space of feasible alternatives. In other words, in order for an allocation to be locally stable, it must meet some global optimization criterion. Commonly advocated social welfare functions of this type include classical utilitarianism, the Nash social welfare function, and a refinement (due to Sen) of Rawls's maximin criterion. Within this framework, Lensberg provides new axiomatizations of particular allocation rules, such as those of Nash and Rawls, and a normative framework for treating decentralized allocation problems in relation to global concepts of social welfare.

The first chapter of Part II represents a scholarly, fascinating exploration by Edward Zajac into the public's perception of what constitutes justice and fairness in economic allocation. Drawing on his experience with the outcomes of public utility rate cases in the United States, combined with examples culled from economic history, politics, and the law, Zajac formulates six Propositions of perceived economic injustice – on the theory that examples of injustice are easier to pinpoint then justice per se. He finds that these Propositions not only conflict
in some cases with economists' notions of economic efficiency, but to some extent with each other. Zajac's point is that justice (or the lack of justice) means balancing conflicting principles, the quest for economic efficiency being only one ingredient in achieving that balance.

Chapters 8 and 9 focus on the specific issue of allocating costs and setting prices in the telecommunications industry. In Chapter 8, William Sharkey summarizes the economic, technological, and regulatory aspects of telecommunications in the United States. He then shows how concepts from cooperative game theory can be used to clarify several important issues, including cross-subsidization between markets, and inefficiencies that result from the fragmentation of markets by the competitive entry of other firms. He also discusses the special structure of cost allocation problems on fixed networks.

Chapter 9, by Nicolas Curien, summarizes recent French telecommunications pricing policy. Using estimates of marginal costs as a reference point, he describes a methodology for computing the amount of cross-subsidization between different services such as long-distance traffic, local traffic, and network access, and between different classes of users such as businesses, households, and coin telephones. He concludes, based on recent French data, that local traffic tends to subsidize long-distance traffic, that general traffic tends to subsidize network access, and that businesses subsidize household and coin telephone use. He also draws attention to cross-subsidizations between peak and off-peak users, between urban and rural customers, and between different uses of the same network such as voice and data transmission. He concludes by pointing out that the goals of economic efficiency and subsidy-free pricing may be outweighed by other political, economic, and equity considerations in setting French pricing policy.

In Chapter 10, Michel Balinski and Francis Sand describe a procedure for allocating landing rights at congested airports. The number of scheduled landings or take-offs in a particular hourly period (called "slots") are limited by the capacity of the airport. In the United States, the Federal Aviation Administration (FAA) establishes quotas on the number of these slots at certain airports, and parcels them out among the different airlines by "scheduling committees" composed of industry representatives. Balinski and Sand investigate an auctioning procedure for achieving a more economically efficient allocation of these slots. In this method, repeated bidding for slots occurs simultaneously across different markets, which allows airlines to take into account the complex interdependence among the slots that they require for scheduling flights. It thus represents an approach to allocating public rights of access by competitive bidding.

In the final chapter, Norio Okada describes in detail how the "separable costs remaining benefits method" is employed in Japan to attribute costs to the different purposes served by multipurpose reservoirs: flood control, irrigation, power, and industrial and municipal supplies. Using Japanese engineering data from the Sameura Dam project, which was built between 1963 and 1970, he gives a detailed account of the cost allocation process, including all of the codicils and ad hoc assumptions needed to make the analysis complete. It thus forms a sobering antidote to the earlier theory, which often takes the problem of estimating the relevant data for granted.

This diverse compendium suggests that theory has moved rapidly in recent years, but is still far in advance of actual practice. Indeed the challenge to theory is to become both simple and robust enough to be practical. For this
Preface

effort to be successful more careful experimentaion and observation in conjunc-
tion with theorizing seems to be called for. If this volume serves no other pur-
pose, it is hoped that it may stimulate further research of this type.

H. Peyton Young
Washington, D.C.

References

Press).
47–61.
No. 9, National Accounting Association.
Webster's Third New International Dictionary of the English Language Unabridged
CHAPTER 1

METHODS AND PRINCIPLES OF COST ALLOCATION

H. Peyton Young

1 Introduction

A central problem in planning the provision of goods or services by a public enterprise is determining a "fair" or "just" allocation of the common costs of production. Examples include setting fees for the use of a common facility such as an airport, a transit system, a communications network, a canal, or a reservoir. Such enterprises, whether strictly public (like a reservoir), or publicly regulated (like telephones) are generally required to "stand on their own bottoms": prices must be set to exactly cover costs, possibly including a mark-up to cover the costs of capital. The economic ideal — marginal cost pricing — does not have this property except in very special cases.

In practice, agencies and firms use simple costing formulas and criteria to solve such problems. A case in point is the "separable costs remaining benefits" method employed by water resource planning agencies to cost out multipurpose reservoir projects. This method is defined in Section 3. Other less sophisticated devices include allocating costs in proportion to some criterion such as number of customers, revenues, profits, or usage rates if the latter can be unambiguously defined. The defect with these methods is that they almost completely ignore the problem of motivation: why, for example, should agents accept an allocation that exceeds their opportunity costs or willingness to pay?

The aim of this chapter is to survey cost allocation methods from an axiomatic perspective. What are the essential principles and properties that characterize different methods? In answering this query we must bear in mind that principles which seem compelling in one context may not be so in another. There is no "method for all seasons". Nevertheless, the axiomatic method does narrow down the plethora of possibilities to a handful of reasonable choices: the Shapley value, the core and certain particular core solutions like the nucleolus,
Ramsey prices when demands are known, and outcomes of demand revelation mechanisms when demands are not known. General context, the level of information, and — not least — precedent, all play a role in determining which principles and methods seem most apt for a particular problem.

2 Problem Formulation and Examples

Many of the salient features of cost allocation can be captured in the following simple format. Let \( N = \{1, 2, \ldots, n\} \) represent a set of potential customers of a public service or public facility. Each customer will either be served at some targeted level or not served at all. In other words, a customer \( i \in N \) will either get a telephone or not, take a train ride or not, hook up to the local water supply or not. The problem is to determine how much to charge for the service, based on the costs of providing it.

The cost data are summarized by a joint cost function \( c(S) \), which is defined for all subsets \( S \subseteq N \) of potential customers. \( c(S) \) represents the least cost of serving the customers in \( S \) by the most efficient means. The cost of serving no one is assumed to be zero: \( c(\emptyset) = 0 \). \( c \) is called the characteristic function of a cost game.

A cost allocation method is a function \( \varphi \) defined for all \( N \) and all joint cost functions \( c \) on \( N \) such that

\[
\varphi(c) = (x_1, \ldots, x_n) \in \mathbb{R}^N \quad \text{and} \quad \sum_{i=1}^{n} x_i = c(N)
\]

where \( x_i \) is the charge assessed customer \( i \).

Example 1: Multipurpose reservoirs. One of the most perplexing examples of joint cost allocation is the multipurpose reservoir. Suppose that a dam on a river is planned to serve several different regional interests, such as flood control, hydro-electric power, navigation, irrigation, and municipal supply. The dam can be built to different heights, depending on which purposes are to be included. The cost function associated with such a problem typically exhibits decreasing marginal costs per acre-foot of water impounded up to some critical height of the dam, after which increasing marginal costs set in due to technological limitations. The water resource planning problem is how to apportion the costs among the different purposes.

This problem has a rich history dating back to the creation of the Tennessee Valley Authority (TVA) in the 1930s (see the historical accounts by Ransmeier, 1942 and Parker, 1943). Certain cost allocation formulas suggested for the TVA system are still in use today (in modified form) by water resource agencies, including the Bureau of Reclamation in the United States Department of the Interior.

Table 1.1 shows a joint cost function, based on actual TVA data, for three "purposes": navigation (\( n \)), flood control (\( f \)), and power (\( p \)). (Ransmeier, 1942, p 329).
Methods and Principles of Cost Allocation

Table 1.1 TVA cost data for navigation (\(n\)), flood control (\(f\)), and power (\(p\)) (thousands of dollars).

<table>
<thead>
<tr>
<th>Coalition</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\emptyset)</td>
<td>0</td>
</tr>
<tr>
<td>({n})</td>
<td>163,520</td>
</tr>
<tr>
<td>({f})</td>
<td>140,926</td>
</tr>
<tr>
<td>({p})</td>
<td>250,086</td>
</tr>
<tr>
<td>({n,f})</td>
<td>301,607</td>
</tr>
<tr>
<td>({n,p})</td>
<td>378,621</td>
</tr>
<tr>
<td>({f,p})</td>
<td>367,370</td>
</tr>
<tr>
<td>({n,f,p})</td>
<td>412,054</td>
</tr>
</tbody>
</table>

**Example 2: Municipal cost sharing.** The second example is one of investment planning, and is based on actual cost data derived from an engineering study (Young et al., 1982). The Skåne region of southern Sweden consists of 19 municipalities, including the city of Malmö (see Figure 1.1). Each municipality requires a certain minimum supply of water that it can either buy from outside sources, pump from its own wells, or obtain by cooperating with some or all of the other municipalities in a regional system. Each municipality \(i\) has a minimum alternative cost or opportunity cost of supply \(c(i)\), assuming that no agreement with the others is reached. Similarly, each potential coalition \(S\) of municipalities has a minimum alternative cost \(c(S)\) of supplying just the members of \(S\) by the most efficient means available, independently of the others.

![Figure 1.1 Skåne, Sweden, and its division into groups of municipalities.](image)

The value \(c(S)\) is determined by engineering considerations, i.e., by estimating the least-cost routing of pipes and pumping stations. \(c(S)\) is defined so that it includes the possibility that some or all members of \(S\) develop independent on-site
sources if this is the least-cost alternative of supplying $S$. Under these circumstances $c$ will be **subadditive**:

$$c(S) + c(T) \geq c(S \cup T)$$

for all disjoint $S, T$.

since the cost of serving two disjoint groups includes the possibility of serving them separately. This amounts to a reasonable convention in defining the cost function, but we shall not assume it in the sequel without special notice.

In practice estimating the costs of $2^{16} - 1 = 256,143$ coalitions is impossible. However, the municipalities fall into natural groupings based on geographical location, existing water transmission systems, and hydrological features that determine the best routes for transmission networks. These conditions can be used to aggregate the 18 municipalities into six units, denoted by A, H, K, L, M, T, as shown in Figure 1.1. The cost function is given in Table 1.2.

**Table 1.2** Costs of alternative supply systems (millions of Swedish crowns). Coalitions are separated by commas if there are no economies of scale from their combination.

<table>
<thead>
<tr>
<th>Group</th>
<th>Total Cost</th>
<th>Group</th>
<th>Total Cost</th>
<th>Group</th>
<th>Total Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>21.95</td>
<td>AHK</td>
<td>40.74</td>
<td>AHKL</td>
<td>46.95</td>
</tr>
<tr>
<td>H</td>
<td>17.08</td>
<td>AHL</td>
<td>43.22</td>
<td>AHKLM</td>
<td>60.25</td>
</tr>
<tr>
<td>K</td>
<td>10.91</td>
<td>AHM</td>
<td>55.50</td>
<td>AHKT</td>
<td>62.72</td>
</tr>
<tr>
<td>L</td>
<td>15.88</td>
<td>AHT</td>
<td>56.67</td>
<td>AHLM</td>
<td>64.03</td>
</tr>
<tr>
<td>M</td>
<td>20.82</td>
<td>AKL</td>
<td>48.74</td>
<td>AHLT</td>
<td>65.20</td>
</tr>
<tr>
<td>T</td>
<td>21.98</td>
<td>AKM</td>
<td>53.40</td>
<td>AHMT</td>
<td>74.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AKT</td>
<td>54.84</td>
<td>AKLMT</td>
<td>63.96</td>
</tr>
<tr>
<td>AH</td>
<td>34.69</td>
<td>ALM</td>
<td>53.05</td>
<td>AKLMT</td>
<td>70.72</td>
</tr>
<tr>
<td>A.K</td>
<td>32.86</td>
<td>A.LT</td>
<td>59.81</td>
<td>A.KMT</td>
<td>72.27</td>
</tr>
<tr>
<td>A.L</td>
<td>37.83</td>
<td>A.MT</td>
<td>61.36</td>
<td>A.LMT</td>
<td>73.41</td>
</tr>
<tr>
<td>A.M</td>
<td>42.76</td>
<td>HML</td>
<td>27.26</td>
<td>HKLM</td>
<td>48.07</td>
</tr>
<tr>
<td>A.T</td>
<td>43.93</td>
<td>HKL</td>
<td>42.55</td>
<td>HKLT</td>
<td>49.24</td>
</tr>
<tr>
<td>HK</td>
<td>22.98</td>
<td>HKT</td>
<td>44.94</td>
<td>HKMT</td>
<td>60.95</td>
</tr>
<tr>
<td>HL</td>
<td>25.00</td>
<td>HLM</td>
<td>45.81</td>
<td>HLMT</td>
<td>64.41</td>
</tr>
<tr>
<td>H,M</td>
<td>37.89</td>
<td>HLT</td>
<td>48.98</td>
<td>KLM</td>
<td>56.61</td>
</tr>
<tr>
<td>H,T</td>
<td>39.05</td>
<td>HMT</td>
<td>56.49</td>
<td>AKLMT</td>
<td>70.93</td>
</tr>
<tr>
<td>K,L</td>
<td>28.79</td>
<td>KLM</td>
<td>42.91</td>
<td>AHKLMT</td>
<td>69.76</td>
</tr>
<tr>
<td>KM</td>
<td>31.45</td>
<td>KLMT</td>
<td>48.77</td>
<td>AKHM</td>
<td>77.42</td>
</tr>
<tr>
<td>K,T</td>
<td>32.89</td>
<td>KMT</td>
<td>50.32</td>
<td>AML</td>
<td>83.00</td>
</tr>
<tr>
<td>LM</td>
<td>31.10</td>
<td>LMT</td>
<td>51.46</td>
<td>AKLM</td>
<td>73.97</td>
</tr>
<tr>
<td>L,T</td>
<td>37.86</td>
<td>KLMT</td>
<td></td>
<td>HKLM</td>
<td>66.46</td>
</tr>
<tr>
<td>MT</td>
<td>39.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AHKLMT</td>
<td>83.92</td>
<td></td>
</tr>
</tbody>
</table>

**Example 3: Airport landing fees.** Landing fee schedules at airports are often established to cover the costs of building and maintaining the runways. The capital cost of a runway is essentially determined by the size of the largest aircraft using it. Suppose that there are $m$ different types of aircraft using an airport and that $c_k$ ($1 \leq k \leq m$) is the cost of building a runway to accommodate an aircraft of type $k$. Index the types so that $0 < c_1 < \cdots < c_m$, and let $c_0 = 0$. Let $N_k$ be the set of all aircraft landings of type $k$ in a given year (say $n_k$ in number) and let
Methods and Principles of Cost Allocation

\[ N = \sum_{k=1}^{m} N_k, \quad n_k = \sum_{k=1}^{m} n_k. \]

Each "player" \( i \) in \( N \) represents an aircraft using the airport exactly once. The cost game is defined as follows:

\[ c(S) = \max \{ c_k : S \cap N_k \neq \emptyset \}. \]

Table 1.3 gives cost and landing data for Birmingham airport in 1969–69, as reported by Littlechild and Thompson (1977).

**Table 1.3** Aircraft landings, runway costs, and charges at Birmingham airport, 1969–69.

<table>
<thead>
<tr>
<th>Aircraft Type</th>
<th>Number of Aircraft Landings</th>
<th>Annual Capital Cost</th>
<th>Charges (Shapley Value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fokker Friendship 27</td>
<td>42</td>
<td>65,899</td>
<td>4.86</td>
</tr>
<tr>
<td>Viscount 800</td>
<td>9,555</td>
<td>76,725</td>
<td>5.66</td>
</tr>
<tr>
<td>Hawker Siddeley Trident</td>
<td>298</td>
<td>95,200</td>
<td>10.30</td>
</tr>
<tr>
<td>Britannia</td>
<td>303</td>
<td>97,200</td>
<td>10.85</td>
</tr>
<tr>
<td>Caravelle VI R</td>
<td>151</td>
<td>97,436</td>
<td>10.92</td>
</tr>
<tr>
<td>BAC 111 (500)</td>
<td>1,315</td>
<td>98,142</td>
<td>11.13</td>
</tr>
<tr>
<td>Vanguard 953</td>
<td>505</td>
<td>102,496</td>
<td>13.40</td>
</tr>
<tr>
<td>Comet 4B</td>
<td>1,128</td>
<td>104,849</td>
<td>15.07</td>
</tr>
<tr>
<td>Britannia 300</td>
<td>151</td>
<td>113,322</td>
<td>44.80</td>
</tr>
<tr>
<td>Corvair Corronado</td>
<td>112</td>
<td>115,440</td>
<td>60.61</td>
</tr>
<tr>
<td>Boeing 707</td>
<td>22</td>
<td>117,676</td>
<td>162.24</td>
</tr>
</tbody>
</table>

**Example 4: Travel expenses.** An itinerant mathematician is invited on a lecture tour originating in Washington, with stops in New York, Boston, and Chicago. The total fare is $540. The problem is how to allocate the fare "fairly" among the trip's three sponsors. The data of the problem shown below are the alternative costs that would be incurred if the traveler had to make either three separate trips, or one trip involving two of the sponsors and a separate trip for the third, etc.

\[
\begin{align*}
\{NY\} + \{BOS\} + \{CHI\} & = 120 + 180 + 420 = 720 \\
\{NY, BOS\} + CHI & = 180 + 420 = 600 \\
\{NY\} + \{BOS, CHI\} & = 120 + 510 = 630 \\
BOS + \{NY, CHI\} & = 180 + 480 = 660 \\
\{NY, BOS, CHI\} & = 540
\end{align*}
\]

Suppose instead that the traveler gives three lectures in Boston, two in New York, and one in Chicago, the cost data remaining as before. A problem in modeling arises: Are the different lectures the primary entities to which costs are assigned? Or is it the sponsors? (Or is it the minutes of lecture time?) The answer depends on the context. If, for example, the lectures are deemed to be the primary objects of allocation, then the cost function would be defined on the set \( N \) of all six lectures as follows. \( c(i) \) would be $180 if \( i \) is a lecture in Boston,
$120 if i is a lecture in New York and $420 if it is the lecture in Chicago. \( c(i, j) \) would be $460 if i is in New York and j is in Chicago, and $120 if both i and j are in New York, and so forth. In short, \( c(S) \) is the least cost of supplying each set S of lectures, and the allocation of costs is made (by whatever method) on a per-lecture basis.

3 Cost Allocation and Cooperative Game Theory

3.1 The Core

The foundations of cooperative game theory were laid down in the treatise of Von Neumann and Morgenstern, *The Theory of Games and Economic Behavior* (1944). The idea of the "core" of a game, which received only passing mention from these two founding fathers, was later developed by Gillies (1953) and Shapley. Interestingly, the core was foreshadowed in the early literature on cost–benefit analysis. A case in point is the Tennessee Valley Authority Act of 1933 as analyzed by Ransmeier (1942) (see also the excellent review by Straffin and Heaney, 1991). The Act stipulated that the costs of TVA projects be specifically allocated among the purposes involved, the principal ones being navigation, flood control, and power. Ransmeier suggested several criteria for judging cost allocation methods:

The method should have a reasonable logical basis ... It should not result in charging any objective with a greater investment than would suffice for its development at an alternate single purpose site. Finally, it should not charge any two or more objectives with a greater investment than would suffice for alternate dual or multiple purpose development. (p 229)

In terms of the joint cost function \( c(S) \) these requirements state that, if \( x_i \) is the charge to purpose i, then in addition to the break-even requirement \( \sum x_i = c(N) \) the following inequality should hold for every subset \( S \) of purposes \( N \) (including singlons).

\[
\sum_{S} x_i \leq c(S) 
\]

(2)

This condition is known as the stand-alone cost test. Its rationale is evident: if cooperation among the parties is to be voluntary, then the calculus of self-interest dictates that no participant – or group of participants – be charged more than their "stand-alone" (opportunity) costs. Otherwise they would have no incentive to agree to the proposed allocation.

A related principle of cost allocation is known as the incremental cost test. It states that no participant should be charged less than the marginal cost of including him. For example, in Table 1.1, the cost of including \( n \) at the margin is

\[
c(n, f, p) - c(f, p) = 45,214
\]

In general, the incremental or marginal cost of any set \( S \) is defined to be \( c(N) - c(N-\bar{S}) \), and the incremental cost test requires that the allocation \( z \subset R^N \) satisfy
\[ \sum_{S} x_{I} \geq c(N) - c(N-S) \quad \text{for all } S \subseteq N. \] (3)

Whereas (2) provides incentives for voluntary cooperation, (3) arises from considerations of equity. For, if (3) were violated for some \( S \), then it could be said that the coalition \( N-S \) is subsidizing \( S \). In fact, (2) and (3) are equivalent given the assumption of full cost allocation \( \sum_{N} x_{I} = c(N) \).

The core of \( c \), written Core (c), is the set of all allocations \( x \in R^{N} \) such that (2) [equivalently (3)] holds for all \( S \subseteq N \). The core is a closed, compact, convex subset of \( R^{N} \). Unfortunately, it may be empty, even if \( c \) is subadditive.

The core of the TVA cost game is illustrated in Figure 1.2. In this figure, the top vertex \( x_{n} \) represents the situation where all costs are allocated to \( n \); the right-hand vertex the case where all costs are allocated to \( p \), etc. The core is fairly large, reflecting the rapidly decreasing marginal costs of building higher dams. To illustrate other possibilities, let us modify the TVA cost data in one respect: imagine that total costs \( c(n,f,p) \) increase to 515,000 due to a cost overrun (the other costs remaining as before). The core of this modified TVA cost game is shown in enlarged form in Figure 1.3.

![Figure 1.2 The core of the TVA cost game.](image)
3.2 Nonemptiness of the core

When is the core nonempty? It is not enough that \( c \) be subadditive. For example, the following 3-person cost game is subadditive but has an empty core.

\[
\begin{align*}
c(1) &= c(2) = c(3) = 6 \\
c(1,2) &= c(1,3) = c(2,3) = 7 \\
c(1,2,3) &= 11.
\end{align*}
\]

The reason is that the inequalities on the charges

\[
x_1 + x_2 \leq 7, \quad x_1 + x_3 \leq 7, \quad x_2 + x_3 \leq 7,
\]

when summed, imply that \( 2(x_1 + x_2 + x_3) \leq 21 \), which contradicts the break-even requirement \( x_1 + x_2 + x_3 = 11 \).

A natural but quite strong condition guaranteeing the nonemptiness of the core is that \( c \) be concave (or submodular)

\[
c(S \cup T) + c(S \cap T) \leq c(S) + c(T) \quad \text{for all } S, T \subseteq N.
\]

For each \( i \in N \) and \( S \subseteq N \) define \( i \)'s marginal cost contribution relative to \( S \) by

\[
c^i(S) = \begin{cases} 
c(S) - c(S - i) & \text{if } i \in S \\ c(S + i) - c(S) & \text{if } i \not\in S\end{cases}.
\]
The function $c^i(S)$ is the derivative of $c$ with respect to $i$. It may be verified that $c$ is concave if and only if for all $i$ the derivative of $c$ with respect to $i$ is a monotonically decreasing function of $S$.

$$S \subseteq S' \implies c^i(S) \geq c^i(S')$$  \hspace{1cm} (7)

**Theorem 1** (Shapley, 1971): If $c$ is concave then the core of $c$ is nonempty.

### 3.4 Costs versus benefits

In cooperative game theory it is customary to focus on the gains each coalition of players can realize rather than on the costs directly. Given a cost function $c$, one way of defining the potential gain of a coalition $S$ is the savings its members can achieve by cooperating instead of going alone:

$$v(S) = \sum_{S} c(\{i\}) - c(S) \text{ for all } S \subseteq N$$ \hspace{1cm} (8)

$v(S)$ is called the **value** of $S$, and $v$ is the characteristic function of the cost-savings game. Note that $v(\emptyset) = 0$.

In some situations net profits or benefits, rather than costs, are the primary objects of allocation. For example, if $c(S)$ is the cost of a firm producing a subset of outputs $S \subseteq N$, and $r_i$ is the revenue from product $i \in N$, then the net profits from $S$ are given by the characteristic function

$$v(S) = \sum_{S} r_i - c(S)$$ \hspace{1cm} (9)

The net profitability of different combinations of divisions is a natural concern for corporate management interested in acquiring or disposing of certain divisions.

Whether costs or benefits are the primary focus of attention depends on the context. An allocation method $\varphi$ may be applied to any characteristic function ($c$ or $v$) whether it represents costs or benefits. In principle, of course, it would be desirable if the two approaches give equivalent results. In general, say that two characteristic functions $u$ and $v$ are **strategically equivalent** if for some scalar $\alpha \neq 0$ and vector $b \in R^N$, $u = \alpha v + b$; that is

$$u(S) = \alpha v(S) + \sum_{S} b_i \text{ for all } S \subseteq N$$  \hspace{1cm} (10)

An allocation method $\varphi$ is **covariant** if

$$\varphi(\alpha u + b) = \alpha \varphi(u) + b$$  \hspace{1cm} (11)

Most (though not all) of the methods we will discuss are covariant. As a practical matter, in most applications it is the costs that are known. Benefits are often conjectural and subject to manipulation or distortion. Hence the main focus will be on cost functions. Methods for determining demand are treated in Section 6.
4 Methods

4.1 The separable costs remaining benefits method

The TVA asserted that its allocation of joint costs was not based on any one mathematical formula, but on judgment (TVA, 1938). As Ransmeier (1942) wryly observes, "there is little to recommend the pure judgment method for allocation. In many regards it resembles what Professor Lewis has called the 'trance method' of utility valuation" (p. 386). Nevertheless, according to Ransmeier, the TVA did in fact use a method and merely "rounded off" the resulting allocations in the light of judgment. This method, called the "alternative justifiable expenditure method", is a variant of an earlier proposal called the "alternate cost avoided method", due to Martin Glaeser, professor of economics at the University of Wisconsin (see Ransmeier, 1942, pp. 270-5). It has become, after further refinements, the principal textbook method used by civil engineers to allocate the costs of multipurpose reservoirs, and is known as the "separable costs remaining benefits method" (SCRB); see James and Lee (1971).

The separable cost of a purpose \( i \in N \) is its marginal cost \( s_i = c(N)-c(N-i) \). The alternate cost for \( i \) is \( c(i) \), and the remaining benefit \( \tau_i \) to \( i \) (after deducting its "separable cost") is

\[
\tau_i = c(i) - s_i.
\]

The SCRB method assigns costs according to the formula

\[
x_i = s_i + \frac{\tau_i}{\sum_{j} \tau_j} \left[ c(N) - \sum_{j} s_j \right].
\]

(12)

In other words, each purpose pays its separable cost and the "nonseparable costs" \( c(N) - \sum_{j} s_j \) are then allocated in proportion to the remaining benefits. The implicit assumption is that all \( \tau_i \geq 0 \), which is the case if \( c \) is subadditive.

For the TVA modified cost data of Figure 1.3, the separable costs, remaining benefits, and corresponding allocation of total costs are shown in Table 1.4.

<p>| Table 1.4 The SCRB method applied to the modified TVA data of Figure 1.3. |
|-----------------|-----------------|-----------------|------------------|</p>
<table>
<thead>
<tr>
<th></th>
<th>( n )</th>
<th>( f )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternate cost</td>
<td>163,520</td>
<td>140,826</td>
<td>250,096</td>
</tr>
<tr>
<td>Separable cost</td>
<td>147,630</td>
<td>136,179</td>
<td>213,393</td>
</tr>
<tr>
<td>Remaining benefit, ( \tau_i = c(i) - s_i )</td>
<td>15,890</td>
<td>4,647</td>
<td>36,703</td>
</tr>
<tr>
<td>Allocation of nonsep. cost</td>
<td>4,941</td>
<td>1,445</td>
<td>11,412</td>
</tr>
<tr>
<td>Allocation</td>
<td>152,571</td>
<td>137,624</td>
<td>224,805</td>
</tr>
</tbody>
</table>

A simple manipulation of (12) reveals that the SCRB method can be given a succinct formulation in terms of the cost-savings game \( \nu \). For each agent \( i \in N \) let \( \nu^i(N) = \nu(N) - \nu(N-i) \) represent the marginal cost savings attributable to \( i \). Given the charges \( x_1, \ldots, x_n \), let \( y_i = c(i) - x_i \) represent the savings imputed to \( i \). The SCRB imputes savings according to the formula

\[
y_i = \frac{\nu^i(N)}{\sum_{j} \nu^j(N)} \nu(N) \quad \text{for all } i.
\]

(13)
In other words, the SCRBN allocates cost savings in proportion to i’s marginal contribution to cost savings (see Straffin and Heaney, 1981). This solution has also been proposed for general games v as a means of minimizing players’ “propensity to disrupt” the solution (Gately, 1974).

4.2 The Shapley value

Imagine that the participants in a cost allocation problem are rational agents who view the outcome as being subject to uncertainty. They might reason about their prospects as follows. Everyone is thought of as “signing up”, or committing themselves, in some random order. At each stage of the sign-up the allocation rule is myopic: each player must pay the incremental cost of being included at the moment of signing. The assessments will therefore depend on the particular order in which the players join.

Instead of actually proceeding in this way, rational agents might simply evaluate their prospects from the comfort of their armchairs by calculating their expected payoffs from such a scheme. Assume that all orderings are a priori equally likely. The “expected” cost assessment for i is then

\[ x_i = \sum_{S \subseteq N, i \in S} \frac{|S - i| |N - S|}{|N|!} c^i(S) , \]

where \( c^i(S) \) is the marginal cost of i relative to S, and the sum is over all subsets S containing i.

Formula (14) is known as the Shapley value of the characteristic function c (Shapley, 1953). It may be interpreted as the average marginal contribution each player would make to the grand coalition if it were to form one player at a time.

The Shapley value for the revised TVA cost game is:

\[ x_n = 151,967 - 2/3, \quad x_f = 134,895 - 1/6, \quad x_p = 228,137 - 1/6 \]

This allocation is not in the core (see Figure 1.4) since

\[ x_n + x_p = 380,104 - 5/6 > 378,821 = c(n,p) \]

It can be shown (Shapley, 1971) that if c is concave, then the Shapley value is in Core (c).

4.3 The nucleolus and its relatives

If the core conditions are considered of primary importance (as in public utility pricing, for example), the Shapley value may not do. An allocation method \( \varphi \) is a core allocation method if \( \varphi(c) \in \text{Core}(c) \) whenever Core \( (c) \neq \emptyset \). What constitutes a reasonable and consistent way of selecting a unique point from the (nonempty) core of a game?

A standard answer to this question is to select an allocation that makes the least-well-off coalition as well-off as possible. The problem is to agree on a meaning of “well-off”. One tack is to say that coalition \( S \) is better off than \( T \), relative to an allocation \( x \), if

\[ c(S) - \sum_{i \in S} x_i > c(T) - \sum_{i \in T} x_i \]

(15)
The quantity \( e(x, S) = c(S) - \sum_{S \in \mathcal{S}} x_i \) is called the **excess of** \( S \) **relative to** \( x \).

To find an allocation \( x \) that minimizes the maximum excess \( e(x, S) \) over all proper subsets \( \emptyset \neq S \subseteq N \) is a problem in linear programming:

\[
\begin{align*}
\text{max} & \quad e \\
\text{subject to} & \quad e(x, S) \geq e \quad \text{for all} \quad S \neq \emptyset, N \\
\text{and} & \quad \sum_{N} x_i = c(N)
\end{align*}
\]  
(16)

If there is a unique optimal solution \( x^* \) to (16), this is the nucleolus of \( c \). If not, use the following tie-breaking rule. Order the excesses \( e(x, S), \emptyset \neq S \subseteq N \), from lowest to highest, and denote this \( (2^n - 2) \) vector by \( e(x) \).

The **nucleolus** \( x^* \) (Schmeidler, 1969) is the vector \( x \) that maximizes \( e(x) \) lexicographically, i.e., for which the value of the smallest excess is as large as possible and is attained on as few sets as possible, the next smallest excess is as large as possible, and is attained on as few sets as possible, etc. It can be demonstrated (Schmeidler, 1969) that there is a unique \( x \) that minimizes \( e(x) \). The proof is based on the observation that if \( x' \) and \( x'' \) maximize \( e(x) \), then \( e((x' + x'') / 2) \) is strictly larger lexicographically than either \( e(x') \) or \( e(x'') \) unless \( x' = x'' \).

The idea of the nucleolus is to find a solution in the core that is "central" in the sense of being as far away from the boundaries as possible (see Figure 1.4). There is some arbitrariness in the definition of the metric, however. A reasonable variant of the nucleolus is to define the excess of a coalition on a per capita basis: \( e(x, S) = [c(S) - \sum_{S} x_i] / |S| \). The above construction then defines the **per capita nucleolus**; (or "normalized" nucleolus; Grotte, 1970). As will be seen in Section 5 below, the nucleolus is probably superior to the per capita nucleolus from an axiomatic perspective.

Solutions for the Swedish municipal cost-sharing game by various methods are compared in Table 1.5.

**Table 1.5** Cost allocations for the Swedish municipal cost-sharing game by the SCRIB, Shapley value, nucleolus, and per capita nucleolus.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCRIB</td>
<td>19.54</td>
<td>13.29</td>
<td>5.62</td>
<td>10.90</td>
<td>16.66</td>
<td>17.92</td>
</tr>
<tr>
<td>Shapley</td>
<td>20.01</td>
<td>10.71</td>
<td>6.61</td>
<td>10.37</td>
<td>16.94</td>
<td>19.18</td>
</tr>
<tr>
<td>Nuc.</td>
<td>20.35</td>
<td>12.06</td>
<td>5.00</td>
<td>8.61</td>
<td>18.32</td>
<td>19.49</td>
</tr>
<tr>
<td>PC. Nuc.</td>
<td>20.03</td>
<td>12.52</td>
<td>3.94</td>
<td>9.07</td>
<td>18.54</td>
<td>19.71</td>
</tr>
</tbody>
</table>

### 5 Principles of Cost Allocation

Faced with a host of competing methods, what basis is there for choosing among them? Each is computationally seductive; each has a certain mathematical charm. The question remains: what are the fundamental properties that an allocation method should enjoy? We focus on three general types of allocative principles that apply not only to cost sharing but to other allocation problems as well. These are: (1) **additivity**: if an allocation problem decomposes naturally into
subproblems, can their solutions be added?; (2) monotonicity: as the data of the problem change, do solutions change in parallel fashion?; (3) consistency: are solutions invariant when restricted to subgroups of agents? Not surprisingly, these properties cannot all be satisfied simultaneously. Nevertheless, taken in various combinations and strengths, they provide a framework for determining which methods are most appropriate to the situation at hand.

Throughout, several basic properties of an allocation method $\phi$ will be taken for granted. These are the break-even (or "efficiency") constraint: $\sum \varphi_i(c) = c(N)$ for all $c$ and $N$. A second and very significant assumption is that all of the data relevant to the allocation problem are contained in the cost function $c$. This leads to the natural requirement that, if $c$ is symmetric with respect to some two players $i$ and $j$ (i.e., if interchanging $i$ and $j$ leaves $c$ invariant), then $\varphi_i(c) = \varphi_j(c)$. This symmetry assumption rules out "biased" methods that allocate everything to player number 1, for example. It also rules out methods that allocate on the basis of information contained in $c$ and some other criterion (such as size, usage rate, etc.). For a discussion of asymmetric solutions, see Shapley (1981).

5.1 Additivity

For accounting purposes it is often convenient to assign costs to different "cost categories" such as operations, plant maintenance, interest expense, and marketing. In theory, each such category (denoted $k = 1, \ldots, m$) gives rise to a different joint cost function $c_k(S)$ on the given set $N$ of activities or products. The sum of these cost functions

$$\sum_{k=1}^m c_k(S) = c(S)$$

represents the total joint cost function. From an accounting standpoint it would be desirable if the allocation process could be carried out separately for each of the cost categories. The total allocation of costs would be the sum of the cost assignments in each category, and we would like this total allocation to be independent of the particular way the costs are categorized.

A cost allocation method $\phi$ is additive if for any joint cost functions $c$ and $c'$ on $N$, $\varphi(c + c') = \varphi(c) + \varphi(c')$, where $c + c'$ is defined by $(c + c')(S) = c(S) + c'(S)$ for all $S \subseteq N$.

In any cost game $c$ on $N$, a player $i$ is a dummy if $i$ contributes nothing to any coalition; i.e., if $c^i(S) = 0$ for all $S \subseteq N$. The dummy axiom states that if $i$ is a dummy in $c$, then $\varphi_i(c) = 0$.

Theorem 2 (Shapley, 1953): There is a unique allocation method that satisfies the dummy and additivity axioms, namely the Shapley value.

The proof of this result relies on the following lemma, which shows that any cost function can be decomposed into a linear combination of "primitive" ones. For each nonempty subset $R \subseteq N$ define the primitive cost function $c_R$ as follows:

$$c_R(S) = \begin{cases} 1 & \text{if } S \supseteq R \\ 0 & \text{if } S \nsubseteq R \end{cases}$$
Lemma (Shapley, 1953): Any cost function \( c \) can be expressed as a weighted combination of primitive ones:

\[
c = \sum_{\phi \in R \subseteq N} w_R c_R.
\]  

(17)

The proof is obtained by letting

\[
w_R = \sum_{T \subseteq R} (-1)^{|R|-|T|} c(T)
\]

and verifying that (17) is an identity.

To prove Theorem 2, let \( \varphi \) be a method that satisfies dummy and additivity. For any nonempty \( R \subseteq N \), the dummy axiom implies that \( \varphi_i(c_R) = 0 \) for all \( i \not\in R \). Since \( c_R \) is symmetric with respect to all \( i \in R \) and since \( \sum_{N} \varphi_i(c_R) = 1 \), it follows that

\[
\varphi_i(c_R) = \begin{cases} 
1/|R| & \text{if } i \in R \\
0 & \text{if } i \not\in R
\end{cases}
\]  

(18)

In other words, \( \varphi \) is uniquely determined on all primitive cost functions. By (17) and additivity, \( \varphi \) is uniquely determined on all games. Since the Shapley value obviously has the claimed properties, it follows that \( \varphi \) is the Shapley value.

For the airport game (Section 2, Example 3), the Shapley value has a particularly simple interpretation and is easy to compute (Littlechild and Owen, 1973). All aircraft of type \( k \) are charged the same amount \( x_k \) where

\[
x_k = \sum_{i=1}^{k} \left( \frac{c_i - c_{i-1}}{\sum_{j=1}^{m} n_j} \right).
\]

The cost of serving the smallest types of aircraft is divided equally among the aircraft of all types, then the incremental cost of serving the second smallest type is divided equally among all aircraft except those of the smallest type, and so forth. The proof follows by observing that for all \( S \),

\[
c(S) = \sum_{k=1}^{m} c_k(S),
\]

where each cost function \( c_k(S) \) has the form \( c_k(S) = c_k - c_{k-1} \) if \( S \) contains a player of type \( j \geq k \), and \( c_k(S) = 0 \) otherwise. The Shapley solution for Birmingham airport is shown in Table 1.3.

5.2 Monotonicity

Stated broadly, monotonicity means that if some player’s contribution to all coalitions to which he belongs increases or stays fixed, then that player’s allocation should not decrease. This concept is relevant to situations in which an allocation is not a one-shot affair, but is reassessed periodically as new information emerges. For example, operations and maintenance costs will typically be allocated annually or even quarterly; investment costs may have to be reassessed ex post if there is a cost overrun.

In its most elementary form, monotonicity says that an increase in \( c(N) \) alone does not cause any player’s allocation to decrease. We say that \( \varphi \) is monotonic in the aggregate if for all \( c, \widehat{c} \) and \( N \)
Methods and Principles of Cost Allocation

\[ c(N) \geq \bar{c}(N) \text{ and } c(S) = \bar{c}(S) \text{ for all } S \subsetneq N \]
\[ \text{implies } \varphi_i(c) \geq \varphi_i(\bar{c}) \text{ for all } i \in N . \]

This version of monotonicity was first stated for cooperative games by Megiddo (1974). It is analogous to "house monotonicity" in apportionment (Balinski and Young, 1962) and to monotonicity in bargaining problems (Kalai, 1977).

It is a simple exercise to show that both the Shapley value and the per capita nucleolus are monotonic in the aggregate and that they allocate any change in \( c(N) \) equally among the players.

Both the SCRB and the nucleolus are egregiously non-monotonic. To illustrate, consider the SCRB allocation of 515,000 in total costs given in Table 1.4. Suppose that total costs increase by 3,000 due to a cost overrun. Then the separable cost of each player increases by 3,000, and the total nonseparable costs decrease by 6,000. The new allocation is \( x_m = 153,781, x_f = 139,581, x_p = 224,636 \). Thus the overrun is allocated as \( \Delta x_m = 1,210, \Delta x_f = 1,957, \text{ and } \Delta x_p = -169 \). In other words, \( p \) enjoys a rebate from a cost overrun.

The nucleolus suffers from the same defect. In the Swedish municipal cost-sharing game, if total costs increase from 83.82 to 87.82, the nucleolus allocates the cost overrun in the amounts:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>H</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.41</td>
<td>1.19</td>
<td>-0.49</td>
<td>1.19</td>
<td>0.64</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Hence \( K \) reaps a gain out of the group's loss.

Monotonicity in the aggregate can be generalized to changes in the value of any coalition. A method \( \varphi \) is monotonic if an increase in the cost of a particular coalition implies, \textit{ceteris paribus}, no decrease in the allocation to any member of that coalition. That is, for all \( c, \bar{c} \) and \( T \subset N \),

\[ c(T) \geq \bar{c}(T) \text{ and } c(S) = \bar{c}(S) \text{ for all } S \neq T \]
\[ \text{implies } \varphi_i(c) \geq \varphi_i(\bar{c}) \text{ for all } i \in T . \]

It is readily verified that (20) is equivalent to the following definition: \( \varphi \) is monotonic if for all \( c, \bar{c} \), and every fixed \( i \in N \),

\[ c(S) \geq \bar{c}(S) \text{ whenever } i \in S, \text{ and } c(S) = \bar{c}(S) \text{ whenever } i \notin S \]
\[ \text{implies } \varphi_i(c) \geq \varphi_i(\bar{c}) . \]

The following "impossibility" theorem shows that monotonicity is fundamentally incompatible with staying in the core.

\textbf{Theorem 3} (Young, 1985a): For \( |N| \geq 5 \) there exists no monotonic core allocation method.

The proof is by example. Consider the cost function \( c \) defined on \( N = \{1,2,3,4,5\} \) as follows:
$c(S_1) = c(3.5) = 3$
$c(S_2) = c(1.2, 3) = 3$
$c(S_3) = c(1.3, 4) = 9$
$c(S_4) = c(2.4, 5) = 9$
$c(S_5) = c(1.2, 3, 4, 5) = 11$

For $S \neq S_1, \ldots, S_5, S_6, \emptyset$, define $c(S) = \min_{S_k \supset S} c(S_k)$, and let $c(\emptyset) = 0$.

If $x$ is in the core of $c$, then

$\sum_{S_k \supset S} x_i \leq c(S_k)$ for $1 \leq k \leq 5$.

Adding these five relations we find that $3 \sum N x_i \leq 33$ whence $\sum N x_i \leq 11$. But $\sum N x_i = 11$ by definition, so all inequalities $\sum S_k x_i \leq c(S_k)$ must be equalities. These have a unique solution, $x = (0, 1, 2, 7, 1)$, which constitutes the core of $c$.

Compare the game $\tilde{c}$, which is identical to $c$ except that $\tilde{c}(S_5) = \tilde{c}(S_6) = 12$. A similar argument shows that the unique core element is now $\tilde{x} = (3, 0, 0, 6, 3)$. Thus the allocation to both 2 and 4 decreases when the value of some of the sets containing them monotonically increases. This shows that no core allocation procedure is monotonic for $|N| = 5$, and by extension for $|N| \geq 5$.

Monotonicity refers to monotonic changes in the value of a single coalition, or in the value of coalitions containing a single player. It is more likely in practice that, over a period of time, some coalitions will increase in value and others will decrease. These changes may occur by chance, or they may be under the control of the players themselves.

A classical instance of a cost allocation problem in which the players can manipulate the value of the coalitions is the apportionment of overhead costs among the divisions of a firm (Shubik, 1962). Corporate management may employ an internal cost accounting scheme to provide incentives for divisions to perform more efficiently. Yet non-monotonic allocation methods have just the opposite effect: by providing perverse incentives they reward inefficiency and sloth.

To illustrate, let $N$ now denote a set of divisions (or product lines) in a firm, each represented by a manager. For each $S \subseteq N$, $c(S)$ represents the costs of those divisions doing business without benefit of the others. The numbers under the (partial) control of a division manager $i$ are the costs of the coalitions $i$ belongs to. Suppose that manager $i$ improves efficiency by lowering the costs of all sets $S$ of which he is a member. Simultaneously, other product managers may be engaged in decreasing certain of the numbers $c(T)$, or through mismanagement increasing them. This adjustment process yields a new cost function $c$ on $N$.

We can unequivocally say that $i$'s contribution to costs is lower in $\tilde{c}$ than in $c$ if

$$c^i(S) \geq \tilde{c}^i(S) \text{ for all } S \subseteq N.$$  \hspace{1cm} (21)

A method $\phi$ is strongly monotonic (does not create perverse incentives) if whenever (21) holds for $c, \tilde{c},$ and $i$, then $\phi_i(c) \geq \phi_i(\tilde{c})$. Strong monotonicity obviously implies monotonicity.

**Theorem 4** (Young, 1986a): The Shapley value is the unique allocation method that is strongly monotonic.
Note that the Shapley value is not the only monotonic allocation method, since equal division \( x_i = c(N) / |N| \) for all \( i \) has this property and is additive too.

5.3 Consistency

If attention is restricted to concave cost functions, then the Shapley value is in the core, is additive, and strongly monotonic. In short, it is a uniquely desirable method for such problems. If we consider cost functions that are not concave and the allocation problem is one-shot, it may be more desirable to choose a method that guarantees a solution in the core whenever the core is nonempty.

The core has important applications to the pricing policies of regulated monopolies (Faulhaber, 1975; Zajac, 1976; Sharkey, 1982). Let \( (r_1, r_2, \ldots, r_n) \) denote the revenues from products \( \{1, 2, \ldots, n\} = N \) of a public utility. If these violate the core conditions for some subset \( S \) of products, \( \sum_S r_i > c(S) \), then a competitor might be able to enter the market, undercut the prices of the regulated firm and steal a portion of his business (namely, the products in \( S \)) while making a positive profit. Moreover, if \( c \) is subadditive, the regulated firm would be left in the unenviable position of producing the set \( N-S \) at a loss, since \( c(N) - c(S) \leq c(N-S) \) and \( \sum_S r_i > c(S) \) imply \( \sum_{N-S} r_i < c(N-S) \).

For such problems the nucleolus has a strong claim to being the preferred method. The argument for the nucleolus is based on the following general allocative principle. A globally valid, acceptable, or "fair" allocation should also be seen as valid, acceptable, or "fair" when viewed by any subgroup of the claimants. In effect, no subgroup should want to "re-contract". This concept is known variously in the literature as consistency or stability. It has been applied to a wide variety of allocation problems, including the apportionment of representation (under the name "uniformity", Balinski and Young, 1982); taxation rules, Young, 1995a), bankruptcy allocation methods (Aumann and Maschler, 1985), cooperative games (Sobolev, 1979), surplus sharing (Moulin, 1985), and bargaining problems (Harsanyi, 1959; Lensberg, 1982, 1983).

To illustrate the consistency concept for cost games, consider the nucleolus \( \bar{z} \) of the revised TVA cost game as pictured in Figure 1.4. Players \( n \) and \( p \) might reason about their situation as follows. If \( f \)'s allocation is granted to be \( \bar{x}_f = 138,502.5 \), this leaves a total of 376,407.5 to be divided between \( n \) and \( p \). The range of possible divisions (holding \( \bar{x}_f \) fixed) is represented by the dotted line segment through \( \bar{z} \) labeled \( L \). In effect, \( L \) is the core of a smaller (or "reduced") game on the two-player set \( \{n, p\} \). The question is: does \( \bar{x}_n, \bar{x}_p \) represent a fair division relative to this reduced game? The answer is yes if we accept the nucleolus as a fair division concept. The reason is that the nucleolus of any two-person game with a nonempty core (which is simply a line segment) is the midpoint of the core. As shown in Figure 1.4, the nucleolus bisects each line segment through it in which the coordinate of one player is held fixed. In other words, the nucleolus is still the nucleolus when viewed by any pair of players. The same conclusion holds for every subgroup of players.

This leads to the following general definition. Let \( c \) be any cost game on a set \( N \) (with or without a core) and let \( z \) be an allocation of \( c \), \( \frac{1}{N} \sum x_i = c(N) \). For
any proper subset $T \subset N (T \neq \emptyset)$, the reduced cost game for $T$ relative to $z$ is the
game $c_{T,z}$ defined on all subsets $S$ of $T$ as follows:

$$c_{T,z}(S) = \min_{S' \subset N-T} \{c(S \cup S') - \sum_{S'} z_i\}, \text{ for } \phi \subset S \subset T,$$

$$c_{T,z}(T) = \sum_{T} z_i$$

$$c_{T,z}(\emptyset) = 0.$$

An allocation method $\varphi$ is consistent if for every set of players $N$, every
cost game $c$ on $N$, and every proper subset $T$ of $N$, $\varphi(c) = \bar{x}$ implies $\varphi(c_{T,\bar{x}}) = \bar{x}_T$.

Note that this definition applies to all cost games $c$, whether or not they have a
nonempty core. It is a rather remarkable fact (and straightforward to check) that the nucleolus is consistent.

**Theorem 5** (Sobolev, 1975): The nucleolus is the unique allocation method
that is covariant and consistent.

Recall that a method $\varphi$ is covariant if $\varphi(\alpha c + b) = \alpha \varphi(c) + b$ for every non-zero scalar $\alpha$ and every $n$-vector $b$. In other words, changing the units in which
costs are measured, or translating the baselines from which they are measured, does not materially affect the outcome. The nucleolus is clearly covariant. The
property is an important one from a practical standpoint, because there is often a
question as to which cost components ought to be included in the joint cost function $c(S)$, and which should be attributed directly to particular players. In the
mathematician's lecture tour, for example, one might include in $c$ the costs of local
travel (e.g. taxi fares), or attribute them directly to particular lectures and leave
them out of $c$. In the Swedish cost-sharing problem one could either attribute
local pumping costs to individual municipalities *separately* or else include them in the joint cost function. The covariance property says that, insofar as such costs are additive and separable, the resulting allocations should be equivalent.

6 Economic Efficiency and Demand Revelation

Formulating the cost allocation problem as a cost game sweeps several important issues under the rug. In particular, the cost game says nothing about agents' willingness to pay. If a cost allocation method assigns charges that exceed the benefits received, the agents might refuse to pay. There is a further problem: if benefits are not incorporated into the analysis, how is one to choose the optimal set of customers to serve? In specifying the set $N$, the cost game takes the optimal set for granted. Assuming that benefits are known, how should they be incorporated into the allocation process?

To answer this question, consider a cost game $c(S)$, $S \subset N$. Suppose that each agent $i \in N$ has a known utility of being served, $u_i$. To keep matters simple, think of $u_i$ as $i$'s willingness to pay for receiving a fixed level of service (e.g., for having a telephone installed). The *net benefit* of serving the set $S$ of customers can be defined as

$$\beta(S) = \sum_{S} u_i - c(S)$$

An economically *efficient* set is one that maximizes $\beta(S)$, say $S^*$. The *value* of a coalition $S \subset N$ can be defined as the maximum net benefit obtainable by serving some, or all, of its members:

$$\nu(S) = \max_{R \subseteq S} \beta(R)$$

(23)

$\nu$ is called the *benefit game*. Note that by definition $\nu(S) \geq \nu(T) \geq \nu(\emptyset) = 0$, whenever $S \supset T$; also $\nu(\{i\}) = \max\{u_i - c(i), 0\}$. If benefits are known, $\nu$ can be computed and an allocation $y = (y_1, \ldots, y_n)$ defined, where $\sum_{N} y_i = \nu(N)$. For example, $y$ might be obtained by applying the Shapley value, the SCR method, the nucleolus, or any other method directly to the game $\nu$. For every allocation of benefits $y$, and every choice of efficient set $S^* \subset N$, there is a concomitant allocation of costs, namely

$$x_i = \begin{cases} u_i - y_i & \text{if } i \in S^* \\ 0 & \text{if } i \not\in S^* \end{cases}$$

(24)

Unfortunately, benefits do not have the same aura of reliability and objectivity that costs do. They are too easily subject to manipulation: agents will have an incentive to misreport their true benefits if this strategy results in lower assessed costs.

Suppose, then, that each agent's willingness to pay is known only to himself, but that the cost function is common knowledge. Is there a bidding procedure or incentive system that implements an efficient decision and allocates costs exactly? Essentially the answer is no. For any general allocation mechanism that exactly covers costs there is always some situation in which it pays some individual to distort the result by sending a false signal.
Nevertheless positive results can be obtained by relaxing these requirements and exploiting the special structure of cost allocation problems. Namely, it is possible to design a straightforward demand revelation mechanism that covers costs and selects an economically efficient set \( S^* \). Moreover, for certain important classes of cost functions, one can do this without generating any surplus, i.e., so that assessments exactly equal total costs.

Given a subadditive cost game \( c \), the idea is to let each agent announce his willingness to pay — not necessarily the true one — and to select an "apparently" efficient set on this basis. (If there are ties the auctioneer has discretion in which efficient set to select.) Each member in the selected set pays his bid; the others pay nothing and are not served. This is called the demand revelation game. There always exists a set of bids such that no set of agents can deviate from these bids and each improve his gain. At such an equilibrium (called a strong equilibrium) the set of bids accepted (i.e., the set of agents served) will be efficient relative to the true demands, and costs will be covered (though not necessarily exactly). The payoff to agent \( i \) is \( y_i = u_i - b_i \) if \( i \) is served and \( b_i \) is his bid, and \( y_i = 0 \) if \( i \) is not served.

We can illustrate this with an example. Consider again the TVA cost data of Table 1.1, and suppose that the benefit to each of the three purposes is \( u_n = 125,000, u_f = 150,000, u_p = 275,000 \). The costs \( c(S) \), net benefits \( \beta(S) \), and values \( v(S) \) are shown in Table 1.6.

The "true" bids \( b_n = 125,000, b_f = 150,000, \) and \( b_p = 275,000 \) are not an equilibrium: at these bids the set \{\( n, f, p \)\} uniquely maximizes bids less costs, hence all bids will be accepted; but each of the players can lower his bid (say by 10,000) and thereby raise his payoff by 10,000, because even at these lower bids the set \{\( n, f, p \)\} still uniquely maximizes bids less costs, so all bids will be accepted.

Is there a set of bids, and a selection of bids to accept, from which there is no incentive to move? The answer is yes. For example, suppose that \( n \) lowers his bid from 125,000 to 45,214, the others staying put. Then both of the sets \{\( n, f, p \)\} and \{\( f, p \)\} maximize bids net of costs at 57,630. Next suppose that \( f \) lowers his bid from 150,000 to 117,274, and \( p \) lowers from 275,000 to 250,096. Then total bids equal total costs. Moreover, no agent can lower his bid further without being excluded and receiving a payoff of zero. If any group of agents should lower their bids, at least one of them will receive a payoff of zero. Hence the bids \( b_n = 45,214, b_f = 117,274, b_p = 250,096 \), with all bids accepted, constitutes a strong equilibrium with payoffs.

<table>
<thead>
<tr>
<th>( S )</th>
<th>( c(S) )</th>
<th>( \beta(S) )</th>
<th>( v(S) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \emptyset )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( {n} )</td>
<td>163,520</td>
<td>-38,520</td>
<td>0</td>
</tr>
<tr>
<td>( {f} )</td>
<td>140,826</td>
<td>9,174</td>
<td>9,174</td>
</tr>
<tr>
<td>( {p} )</td>
<td>250,096</td>
<td>24,904</td>
<td>24,904</td>
</tr>
<tr>
<td>( {n, f} )</td>
<td>301,607</td>
<td>-28,607</td>
<td>9,174</td>
</tr>
<tr>
<td>( {n, p} )</td>
<td>378,821</td>
<td>21,179</td>
<td>24,904</td>
</tr>
<tr>
<td>( {f, p} )</td>
<td>367,370</td>
<td>57,630</td>
<td>57,630</td>
</tr>
<tr>
<td>( {n, f, p} )</td>
<td>412,584</td>
<td>137,416</td>
<td>137,416</td>
</tr>
</tbody>
</table>
\[ u_n = 125,000 - 45,214 = 79,786 \]
\[ u_f = 150,000 - 117,274 = 32,726 \]
\[ u_p = 275,000 - 250,096 = 24,904 \]

This allocation \( y \) is in the core of the benefit game (see Table 1.6) in the sense that \( \sum_{s \subseteq S} y_s \geq v(S) \) for all \( S \subseteq N \). This is a special case of the following result.

**Theorem 6** (Young, 1980): Let \( c \) be a subadditive cost game on \( N \), and \( u \in \mathbb{R}^N \) a vector of benefits. The associated demand revelation game has a strong equilibrium, and for any such equilibrium the set selected is efficient and the total of accepted bids covers total costs. Moreover, if the benefit game \( v \) defined as in (23) has a nonempty core, then there exists a strong equilibrium at which the total of accepted bids exactly equals total costs, and the payoff vectors from such equilibria correspond one-to-one with the allocations in Core \( (v) \).

### 7 Aumann–Shapley and Ramsey Pricing

An alternative formulation of the cost allocation problem familiar to economists involves a firm producing \( n \) homogeneous products in quantities \( q_1, ..., q_n \geq 0 \). Costs are given by a **production function** \( F: \mathbb{R}_+^n \rightarrow \mathbb{R} \), where \( F(q) \) is the joint cost of producing the "bundle" \( q = (q_1, ..., q_n) \). Assume that \( F \) has continuous first partial derivatives everywhere on the domain \( \mathbb{R}_+^n \) (one-sided on the boundary), and that \( F(0) = 0 \). Given a target level of production \( q^* > 0 \), the goal is to find **unit prices** \( p = (p_1, ..., p_n) \) for the \( n \) goods such that costs are exactly covered:

\[
\sum p_i q_i^* = F(q^*) \quad .
\]  
(25)

This is the **break-even constraint**. By suitably defining \( F \), one can include in this formulation the cost of capital, or a percentage mark-up over cost.

An \( n \)-vector \( p \) satisfying (25) is a **solution of the problem** \( (F,q^*) \). A **method** is a function \( \psi \) defined on all problems \( (F,q^*) \), and all \( n \geq 1 \), such that \( \psi(F,q^*) = p \) is a solution of \( (F,q^*) \).

Most of the concepts introduced previously have their analogues in this setting. We briefly mention two of them. \( F \) may decompose naturally into a sum of different cost categories

\[
F(q) = F_1(q) + F_2(q) \text{ for all } q \geq 0 \quad .
\]  
(26)

In this case it would be desirable if the cost allocation process could be similarly decomposed: \( \psi \) is **additive** if (26) implies that for all \( q^* > 0 \),

\[
\psi(F,q^*) = \psi(F_1,q^*) + \psi(F_2,q^*)
\]  
(27)

As shown by Billera and Heath (1982) and Mirman and Tauman (1982) (see also Billera, Heath, and Verrecchia, 1981), additivity, together with several additional regularity conditions, uniquely characterizes the following pricing method, which is based on the Aumann–Shapley value for non-atomic games (Aumann and Shapley, 1974):
\[ p_i = \int_0^1 \frac{\partial F(tq^*)}{\partial q_i} \, dt \quad \text{for all } i. \]  

These Aumann–Shapley prices satisfy the break-even constraint, and are additive. Intuitively, \( p_i \) represents the average marginal cost of \( i \) along the ray from 0 to the target \( q^* \).

Aumann–Shapley prices have an important incentive property analogous to strong monotonicity in cost games. Imagine that each product line \( i \) is supervised by a division manager. The manager's concern is that his division looks profitable – indeed his bonus may depend upon it. Whether it is profitable depends on how costs are imputed to his division, and hence on the cost allocation formula. From the point of view of corporate headquarters, the purpose of an accounting scheme is to encourage product managers to innovate and reduce costs. In short, its interest is in a method that rewards efficiency. How can increases in efficiency be measured with respect to individual products? A natural test is to say that \( i \)'s contribution to costs decreases if \( \frac{\partial F(q)}{\partial q_i} \) decreases for all levels of output \( q \).

A method \( \psi \) is strongly monotonic if for all production functions \( F,G \) and for every fixed \( i \),
\[ \frac{\partial F(q)}{\partial q_i} \geq \frac{\partial G(q)}{\partial q_i} \quad \text{for all } q \geq 0 \]  

implies \( \psi_i(F^*,q^*) \geq \psi_i(G^*,q^*) \quad \text{for all } q^* > 0 \) .

Together with the analog of symmetry, strong monotonicity characterizes the Aumann–Shapley method uniquely (Young, 1985b).

Aumann–Shapley prices may be computed from the piecewise linear production function that arises from solving a cost-minimizing linear program.

Another method of cost allocation, which is popular in the economics literature, is "Ramsey pricing". As above, assume that there are \( n \) homogeneous divisible goods produced in the quantities \( q = (q_1, \ldots, q_n) \) and a joint production function \( F(q) \) having continuous first partial derivatives on the domain \( q \geq 0 \). In Ramsey pricing, the demands for the products are factored into the calculation. To simplify matters, assume that demands for the products are independent: \( q_i = u_i(p_i) \) is the amount demanded of product \( i \) when prices are set at \( p_i \). \( u_i \) is assumed to be strictly monotone decreasing in \( p_i \) and continuously differentiable. The target level of production, \( q^* \), is not given \( a \ priori \) as in the Aumann–Shapley setting, but is derived from the demand and cost information. The idea is to set prices and production levels so as to maximize benefits net of costs ("net social product"). One measure of net social product is total consumers' surplus over all products, minus the costs of production:
\[ \max \frac{1}{2} \sum_{i=1}^{n} \int_0^{q_i} u_i^{-1}(t) \, dt \]  

\[ -F(q) \]  

\[ (30) \]

There are two constraints on this maximum problem. The first is implicit in the formulation: namely, the same price \( p_i \) is charged to all consumers of \( i \) (i.e., pricing is nondiscriminatory). The second is the break-even requirement
\[ \sum p_i q_i - F(q) = 0 \]  

\[ (31) \]
Solving (30) subject to (31) is a standard exercise in constrained optimization. Form the Lagrangian

\[ L(q) = S(q) + \lambda \left[ \sum q_i u_i^{-1}(q_i) - F(q) \right] . \]

A necessary condition for the optimal target \( q \) is that for all \( i \)

\[ u_i^{-1}(q_i) - \frac{\partial F}{\partial q_i}(q) + \lambda \left[ u_i^{-1}(q_i) + q_i \frac{\partial u_i^{-1}}{\partial q_i}(q_i) - \frac{\partial F}{\partial q_i}(q) \right] = 0 . \]  
\[
(32)
\]

Let \( p_i = u_i^{-1}(q_i), c_i = \frac{\partial F}{\partial q_i}(q_i) \), which is the marginal cost of producing \( i \), and \( \eta_i = \frac{p_i}{q_i} - \frac{\partial F}{\partial q_i}(q_i) \), which is the elasticity of demand for \( i \), is always negative since \( u_i \) is decreasing. Then (32) has the simpler expression

\[ (c_i - p_i)/p_i = \frac{\lambda}{1 + \lambda} \frac{1}{\eta_i} . \]  
\[
(33)
\]

In other words, a necessary condition for optimality is that, for each \( i \), the percentage by which marginal cost differs from price is inversely proportional to the elasticity of demand for \( i \).

Prices satisfying (33) are known as Ramsey prices, after Frank Ramsey (1927) who applied similar reasoning to determining optimal methods of taxation. They have been discussed extensively in the economics literature (see Manne, 1952; Baumol and Bradford, 1970; Boiteux, 1971). Their essential property, in industries with declining average costs, is that the percentage mark-up over marginal cost is greater the more inelastic the demand for the good is. As W. Arthur Lewis (1949) put it:

The principle is ... that those who cannot escape must make the largest contribution to indivisible cost, and those to whom the commodity does not matter much may escape. The man who has to cross Dupont's bridge to see his dying father is mulcted thoroughly; the man who wishes only to see the scenery on the other side gets off lightly. (p 21)

The inherent unfairness of such a criterion may strike the reader as debatable. For example, should low-income families who need telephones for emergencies subsidize long-distance usage by businesses? The Ramsey principle might well lead to such a result. Certainly there is no guarantee that Ramsey prices are subsidy-free, i.e., are in the core of a suitably defined cost game. From a practical point of view, Ramsey prices may be difficult to estimate, since they rely on demand elasticities that may not be known. For these and other reasons Ramsey pricing, while interesting in theory, has thus far not found much practical acceptance.

8 Conclusion

There is no shortage of plausible methods for cost allocation. The essence of the problem, however, lies not in defining methods, but in formulating principles and standards that should govern allocations, and then determining which methods satisfy them.

Minimal data of the problem are the total costs to be allocated, and the objects to which costs are to be assigned. In specifying the latter one must be
clear about when two cost objects are comparable, i.e., would be assigned equal costs in the absence of other information. The cost objects could be divisions of a firm, dollars of revenue from different products, individual consumers, classes of consumers, stops on a trip, lectures given on a trip ..., etc. Their identity varies greatly from one situation to the next.

The second element of the problem is estimating the cost associated with each subset of cost objects. The specification of these costs for all subsets defines a cost game on the cost objects, which are called "players". Computing a cost for each of \( 2^n \) subsets of \( n \) cost objects is daunting if \( n \) is larger than 4 or 5. In practice the structure of the problem often allows simplifications. For example, airport-type cost games, in which costs depend only on certain critical thresholds, are manageable even for very large \( n \). Other situations may allow the grouping of players and allocating costs hierarchically, first among groups and then within each group. (There is no guarantee, however, that the outcome will be the same as if costs had been allocated in one fell swoop.)

A third element of the problem is the anticipated benefit from the project. Economic efficiency suggests that the optimal set of cost objects is the one that yields the greatest benefits net of costs. In principle, the allocation of benefits can be carried out using the same methods that apply to the allocation of costs. If demands are not known, however, then serious difficulties arise in trying to design a cost allocation scheme that implements an efficient decision. One reasonable approach is to employ a competitive bidding scheme that generates an efficient decision and at least covers all costs.

From a normative standpoint, four general principles stand out as important criteria for judging cost allocation methods. These are: monotonicity, additivity, consistency, and staying in the core. The last two seem most compelling for one-shot investment problems, or in public utility pricing where cross-subsidization is a major issue. In this case the nucleolus seems to be the best choice, although it should be noted that another core solution — the per capita nucleolus — has the advantage of being monotonic in the aggregate, which the nucleolus is not.

An impossibility theorem states that no core solution method is fully monotonic, and only the Shapley value is monotonic in the strongest sense. The Shapley value can also be characterized as the unique method that is additive and allocates no costs to "dummy" players. On the other hand, the Shapley value is not necessarily in the core. The Shapley value seems well suited to situations in which costs are allocated in parts, or are reallocated periodically, but it is not satisfactory when core solutions are required.

In sum, there is no all-embracing solution to the cost allocation problem. Which method suits best depends on the context, the computational resources, and the amount of cost and benefit information available.

Acknowledgments

This work was supported in part by the National Science Foundation, Grant SES-8319530. The author wishes to thank M.L. Balinski, H. Moulin, L. Mirman, A. Roth, and W.W. Sharkey for constructive comments on the manuscript.
Notes

1. If the benefit $u_i$ from purpose $i$ is known, then the remaining benefits are defined by $\tau_i = \min\{u_i, c(t)\} - s_i$.

2. This is sometimes called the "pre-nucleolus", since Schmeidler's definition required in addition that $x_i \leq c(t)$ for all $i$. If $c$ is subadditive these constraints are automatically satisfied.

3. The definition of the reduced game (and hence of consistency) is inconsistent unless $c$ has a nonempty core and $x \in \text{core}(c)$. The reason is that, by analogy with the definition of $C_T, x_1, \ldots, x_N$ on the other coalitions, we should have

$$C_T, x(T) = \min_{S \subseteq N - T} \left\{ c(T \cup S') - \frac{1}{|S'|} \right\}$$

but this is strictly less than $\sum_{i \in T} x_i$ (for some $T$) unless $x \in \text{core}(c)$. Thus Sobolev's result seems most persuasive when applied only to the class of games $c$ having nonempty cores. It seems reasonable to conjecture that the proof of Theorem 5 can be carried out on this restricted class.

4. The auctioneer is constrained to choose a set that maximizes bids net of costs. In case of ties, some choices may yield an equilibrium, others may not. The claim is that some bids exist, and some way of breaking ties exists, that is a strong equilibrium. Thus in a very limited sense the auctioneer could be considered a player too, although his only function is to break ties efficiently. In practice, ties rarely, if ever, occur; instead some stopping rule is imposed that allows bids to converge. Experiments conducted at IIASA with the six-person Swedish municipal cost-sharing game show that this convergence is remarkably rapid: a solution in the core or nearly in the core was usually found within ten bids.

References


