DOCUMENT RESUME

ED 121 332 IR 003 330

AUTHOR Danilov, V. J.; And Others
TITLE Report of the Panel on Economics of the Science Information Council.
INSTITUTION National Science Foundation, Washington, D.C. Science Information Council.
PUB DATE Jan 73
NOTE 119p.

EDRS PRICE MF-$0.83 HC-$6.01 Plus Postage
DESCRIPTORS Abstracting; *Cost Effectiveness; *Economic Factors; Federal Aid; Federal Programs; Indexing; *Information Services; Libraries; Mathematical Models; Publishing Industry; Scholarly Journals; *Sciences; Special Libraries; Technology; *Use Studies

ABSTRACT

Scientific and technical information services cost the federal government close to a billion dollars a year and cost further large sums to other sectors of the economy. Unless the relationship of benefit to cost for these services is quantitatively understood, there are dangers of enormous waste in some areas or of the passing up of enormous net benefits in others. Very little has been done toward the assignment of a monetary value to these benefits. It should be possible to develop rough but useful quantitative measures for the benefits delivered by many kinds of information services by combining theoretical analysis with suitably gathered empirical data on the response of the user community to these or related services. The two most useful sources of information on quantified judgments of value by users are market response--what users, as individuals or through their organizations, are willing to pay for an information product or service--and investment of time by individual users. To obtain these value measures, extensive data on the markets for information services and on the habits of their users will be needed, as well as sophisticated theoretical analysis of markets and of the interaction of different modes of information transfer. (Author/PF)

REPORT OF THE PANEL ON ECONOMICS OF THE
SCIENCE INFORMATION COUNCIL

Members of the Panel:

V. J. Danilov
C. Herring (Chairman)
D. J. Hillman

January, 1973

CONTENTS

REPORT OF THE PANEL ON ECONOMICS TO
THE SCIENCE INFORMATION COUNCIL

DIGEST

What is at stake. Scientific and technical information services cost the Federal government close to a billion dollars a year, and cost further large sums to non-Federal sectors of the economy. They undoubtedly deliver benefits of the same order as their cost, but unless the relation of benefit to cost is quantitatively understood, there are dangers of enormous waste in some areas, or of the passing up of enormous net benefits in others. (Section I)

Issues needing economic judgments. The information services for which sizable expenditures are made are very diverse (Section I.3 Fig. 1, Appendix A). They cover: written, oral, and photographic media; research results, secondary information, evaluated information, and engineering data. They may be provided at the producer or the user end of the chain. There is need for techniques of economic analysis applicable to as many of these services as possible.

Factors in economic decisions: gaps in our knowledge. For intelligent decision-making, a broad spectrum of factors need to be borne in mind, extending from costs at the one end to quantified benefits at the other end, and including such intermediate considerations as stability, motivation, etc., and also measures of performance in non-monetary terms. (Section II.1, Fig. 3.) There is a sizable literature on the costs of information services, and on the evaluation of their performance in various kinds of nondollar terms. (Sections II.2, 3.) Very little has so far been done toward the assignment of a monetary value to the benefits they deliver (Sections II.2, 4, 5, 6). While some useful types of economic analysis can be made without such assignment (Section II.3), most economic decisions require a weighing of benefits against costs, and so require that benefits be quantified. Thus it is this component of our understanding of economics that we feel is most in need of attention now.

Techniques for quantifying benefits. We feel that it should be possible to develop rough but useful quantitative measures for the benefits delivered by many kinds of information services, by combining theoretical analysis with suitably gathered empirical data on the response of the user community to these or related services. Though it may in rare instances be possible to get objective measures of benefits, or lower bounds to them (Section II.4), in most cases the ultimate judgment of value will have to come from the multitudes of users or purchasers of the services: these will usually have more expert and first-hand knowledge of the benefits delivered than anyone else, and reliance on the collective judgments of many people is a further safeguard. (See Section II.5). However, these small-scale judgments need to be corrected in several ways if they are to provide a measure of benefit to the whole society, or even to a large organization: they are self-interest judgments, and do not usually take account of externalities, such as the benefit received by one person or a group through the use of an information service by another. To make such corrections it is necessary to construct reasonably adequate economic models and models of the way in which different channels for the communication of information interact. (Sections II.7, 9, III.1, 2.)

The two most useful sources of information on quantified judgments of value by users are:

(a) Market response, i.e., what users, as individuals or through their organizations, are willing to pay for an information product or service. Existing market data can often provide an adequate measure of value for a hypothetical change that will alter the price at which a given information service is marketed, without altering its characteristics. However, market data usually tell only the number buyers who assess the value of the product at more than the offering price, but do not give good information regarding how much more the value to certain buyers may be. Therefore such data are rather unsatisfactory for assigning monetary value to changes that will alter the nature of a product or service. (Sections II.5, III.3, Appendix B.)

(b) Investment of time by individual users. Each user of an information service derives benefits from it which increase more and more slowly with increase in the amount of time he devotes to use of it. In practice he will try to devote only as much time to such use as will yield him more benefit than the same amount of time devoted to his other activities. Thus a knowledge of the investment of time by users provides a measure of value; it must of course be calibrated from a knowledge of the dollar value of their time and corrected for externalities having to do with the availability of alternate channels for information. (Section II.6, III.4, Appendix D.) Sometimes opportunities can be found for checking the validity of the hypothesis that users allocate their time according to intelligent self-interest judgments; when the assumption fails, such failure can point the way to much-needed user education. Even in such cases, a lower bound on the value of a service, or a change in one, may be obtainable.

A number of examples are discussed (Section IV, Appendix E) showing how data of either the market-response or the user-time kind might be applied to assign a monetary benefit to each of a number of typical decisions regarding an innovation in an information service. These examples, though very sketchily discussed, make more explicit the kinds of empirical information about markets, users, and the functioning of information channels, which would need to be gathered.

<u>Recommendations</u>. (Section V.) To obtain quantitative value measures by the techniques just sketched, extensive data on the markets for information services and on the habits of their users will be needed, as well as sophisticated theoretical analysis of markets and of the interaction of different modes of information transfer. Studies of all these types can appropriately be sponsored by OSIS, and have indeed been so conducted in the past; however, what is needed now is a closer integration of theoretical and empirical studies, so that theoretical models are constructed with due account of real-life circumstances, and so that "user studies" do indeed collect the kind of data needed for quantification of value, and in particular give adequate attention to the interplay of different modes of information transfer. In the course of such studies, any opportunities for evaluation of factors affecting the stability, continuity, and feasibility, of information services should be identified and exploited. Acting through COSATI, OSIS should keep other agencies of the Federal government informed of the state of development of techniques for quantification of benefits, and should assist them in making economic decisions on their information programs. To fulfill this role adequately, it will need to develop substantial expertise on information economics within its own staff.

I. INTRODUCTION

1. Background

The Federal government spends, directly or through its contractors, something like a billion dollars a year for the dissemination of scientificand technical information, according to recent studies.[1,2] Nongovernmental organizations spend further large sums for the same purpose. This impressive total is compounded of many diverse activities, each of which is presumably undertaken in the expectation that our society, or at least some component of it, will receive benefits greater than the expenditure involved. To justify such expectations one needs an at least roughly quantitative measure of benefits. Here is the difficulty. It is much harder to express in dollar terms the benefits provided by information services than those resulting from most other activities. Yet without such quantification one runs the risk of spending hundreds of millions of dollars wastefully or of commiting the equally grievous sin of holding back on expenditures that could produce net benefits of this order or greater.

Recognizing this need, and taking note of recent innovative applications of operational reserach and economic theory to information services, the Science Information Council at its February 1971 meeting decided to set up a Panel on Economics from its membership, with a charge to consider the economic aspects of the role of government, in both its operational and its supportive functions, in the development and operation of information systems. The Panel, whose findings are described in the present report, has chosen to interpret this charge rather narrowly, in that it has devoted most of its attention to the question: Could governmental or other agencies, that operate or directly or indirectly support information activities, get more value for their money with the help of economic research in certain as yet unexploited areas? The conclusion will be that this is indeed the case, and that the principal such areas meriting further study are those having to do with the assignment of a dollar value to the benefits derived from information activities. In the following paragraph, and in more detail in Section II below, we shall sketch how the quantification of value fits into the overall picture of information economics, and shall indicate why we feel that our concentration on this aspect is in accord with the motive of the Science Information Council in setting up the Panel.

2. The competition between costs and benefits

While the problem of economic management is basically one of subtracting costs from benefits to get net benefits, it has many more ramification than are superficially apparent in these simple terms. Some arbitrary decisions must be made: Costs to whom? Benefits to whom? What things does one include as costs, rather than as detractors from benefits? Even for a well-defined operation, cost accounting is a rather subtle science. Moreover, in many cases one ought to take account of hidden costs to elements of society outside the organization performing the service. Sound economics in management involves many further considerations: such things as stability in the face of fluctuations, motivations for good performance, political feasibility, etc. In the area of results, the easiest data to come by are usually performance measures that are only presumed to be related to benefits -- such things as volume and speed of publication (for primary literature), completeness and depth of indexing (for secondary services), attendance (for meetings), etc. While one can sometimes make decisions that are obviously good in that they reduce costs without changing performance or improve performance without changing costs, most decisions turn out to affect both costs and performance, and to make them wisely, one must make some sort of guess at the dollar value, to society as a whole or at least to some enterprise within it, of the benefit that the users of an information service receive from it.

While there is an extensive literature on costs, and also on measures of performance for information services, the literature on the quantification of benefits actually delivered is extremely meager. There is a gaping hole in our knowledge at this crucial point. It is for this reason that we have given most of our attention to this topic; we shall argue below that there is hope for real progress on it. Besides this central concern, we shall have a few words to say on some hitherto overly-neglected aspects affecting the practicality of information services; costs and performance, however, we shall not touch in any detail.

3. Range of activities to which this Report is addressed

All these remarks have referred to information serivces in very general terms. Most of the conceptual framework we have sketched, and many of the ideas to be developed in the body of this Report, could be applied to information activities as diverse as schools, product advertising, public libraries, etc. Though such applications might be very important, we have thought only in terms of scientific and technical information activities, and in fact not quite all of these. But even so, the activities to which this Report is addressed may be extremely diverse, and it will be illuminating to examine this diversity and list of few typical examples of how decisions about information services may depend upon the quantification of value in specific situations.

Figure 1 essays to classify information services according to the type of communication and the identity of the performer; some eleven areas, covering the most important types of information activities, are identified on the grid and will be used in our discussion and listing of examples.

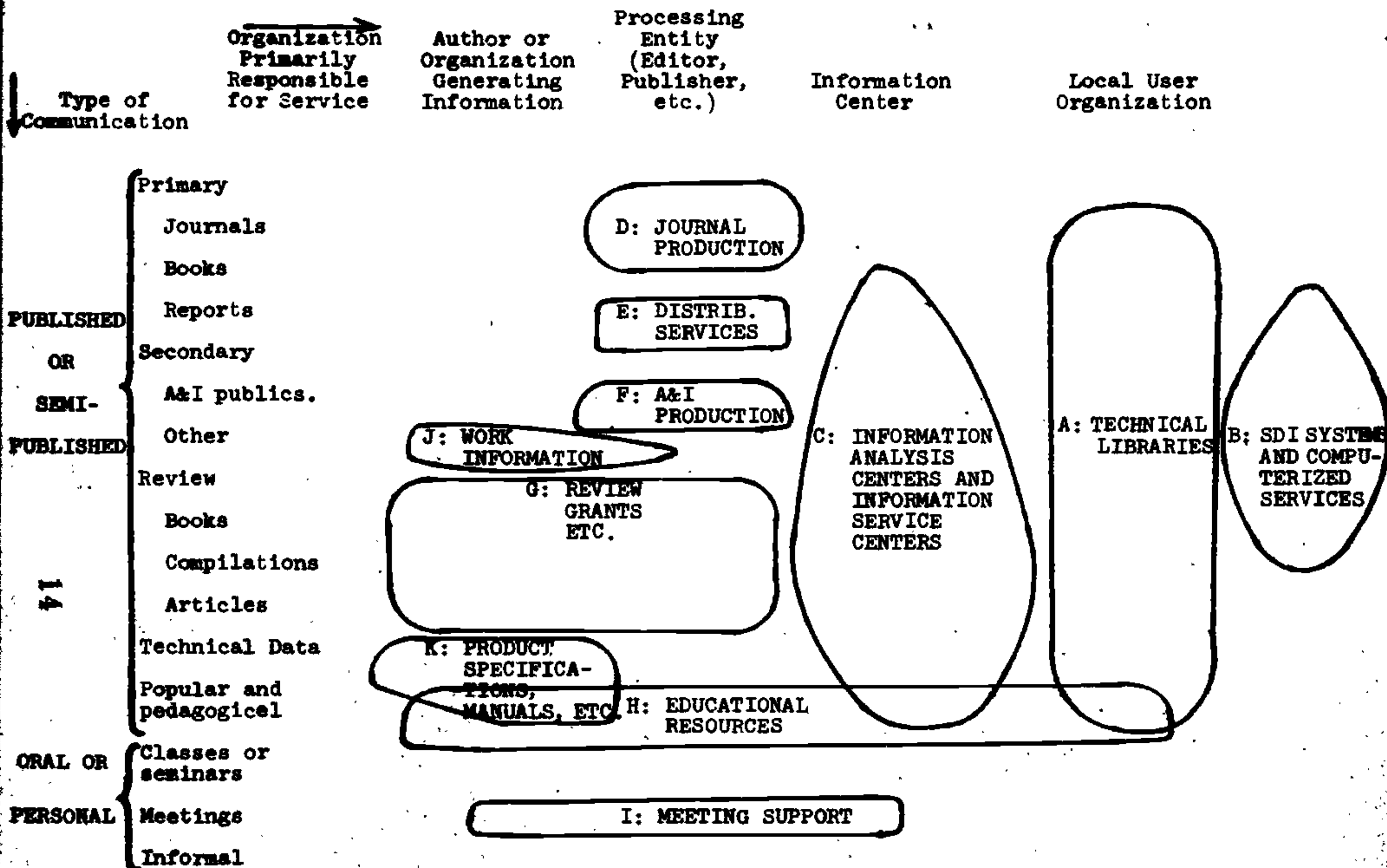

Fig. 1. Major categories of scientific and technical information services for which economic decisions must be made. See Appendix A for further discussion.

Appendix A, which the reader is urged to consult at this point, contains a few words about the scale and economic importance of each of the eleven areas, followed by one or more examples of decisions a governmental agency, a scientific society, or a research organization might have to make regarding the operation of some service of this type. For brevity we quote here only one example for each major category in Fig. 1, omitting categories H, J, and K: the list in Appendix A includes a few further examples. Our numbering here is the same as that used there, hence the missing numbers:

A. Technical libraries.

(1) When should a university or other research organization split off local departmental libraries from a centralized library?

B. Systems for selective dissemination of information and computerized services.

(4) When should a research organization put into operation a selective-dissemination system for its staff?

C. Information-analysis centers.

(6) When is it worthwhile to start an information-analysis center in a new field, or to significantly expand an existing center?

D. Journal production.

(8) When should a scientific society adopt a larger part format for its journals?

E. Schemes for distribution for reports and preprints.

(9) When, and on what scale, should preprint exchanges be subsidized?

F. Abstracting and indexing publications.

(13) Should an abstract journal computerize its annual-index production to enable the index to be produced more promptly?

G. Review grants, etc.

(14) Should an agency supporting research divert some of its research funds to a program of grants to authors for the preparation of reviews and compilations?

I. Meeting support.

(15) How much money should a research organization set aside for travel of its staff members to meetings?

4. Organization of the Report

We shall return to a discussion of some of these in Section IV, where we shall suggest ways of quantifying benefits for particular sample situations. We shall commence our report, however, in Section II, with a broader look at the components entering into decision-making regarding the economics of information services, and a brief indication of how much is now known about them. Section III will elaborate on the prospects for improving our knowledge at the points where it is now most critically lacking. Section V, finally, will offer some recommendations for steps OSIS might take to facilitate more cost-effective decisions on information services.

II. COMPONENTS OF THE PROBLEM, AND SOURCES OF KNOWLEDGE ABOUT THEM

1. Economic decisions: quantifiable factors, arbitrary definitions, and imponderables

It will be well to commence our discussion with a brief elaboration of the remarks made in Section I.2 above regarding the subtleties of the net-benefit assessment. As indicated there, one must first ask: net benefit to whom? A corporation may be interested mainly in its own profits, a scientific society mainly in its membership, or perhaps the enterprises in which its members work, etc. For at least some of the agencies of the Federal government, the concern may be for the welfare of the whole United States. We shall give especial attention to this case, while recognizing that it is by no means the only one of importance. Right now what we want to stress is merely that anyone making economic decisions must clearly identify at the outset for what entity -- we shall call it a "universe of concern" -- he wants to know net benefit.

Having identified this universe of concern, one must next decide how to define costs and benefits for it. The separation of costs from benefits is purely a matter of convenience, as a cost can be regarded as a negative benefit, and vice versa. What one usually does is exemplified by Fig. 2: One isolates, among the activities of the universe of concern, a fairly accurately monitorable subset associated with provision of an information service, such that the benefits attributable to these activities

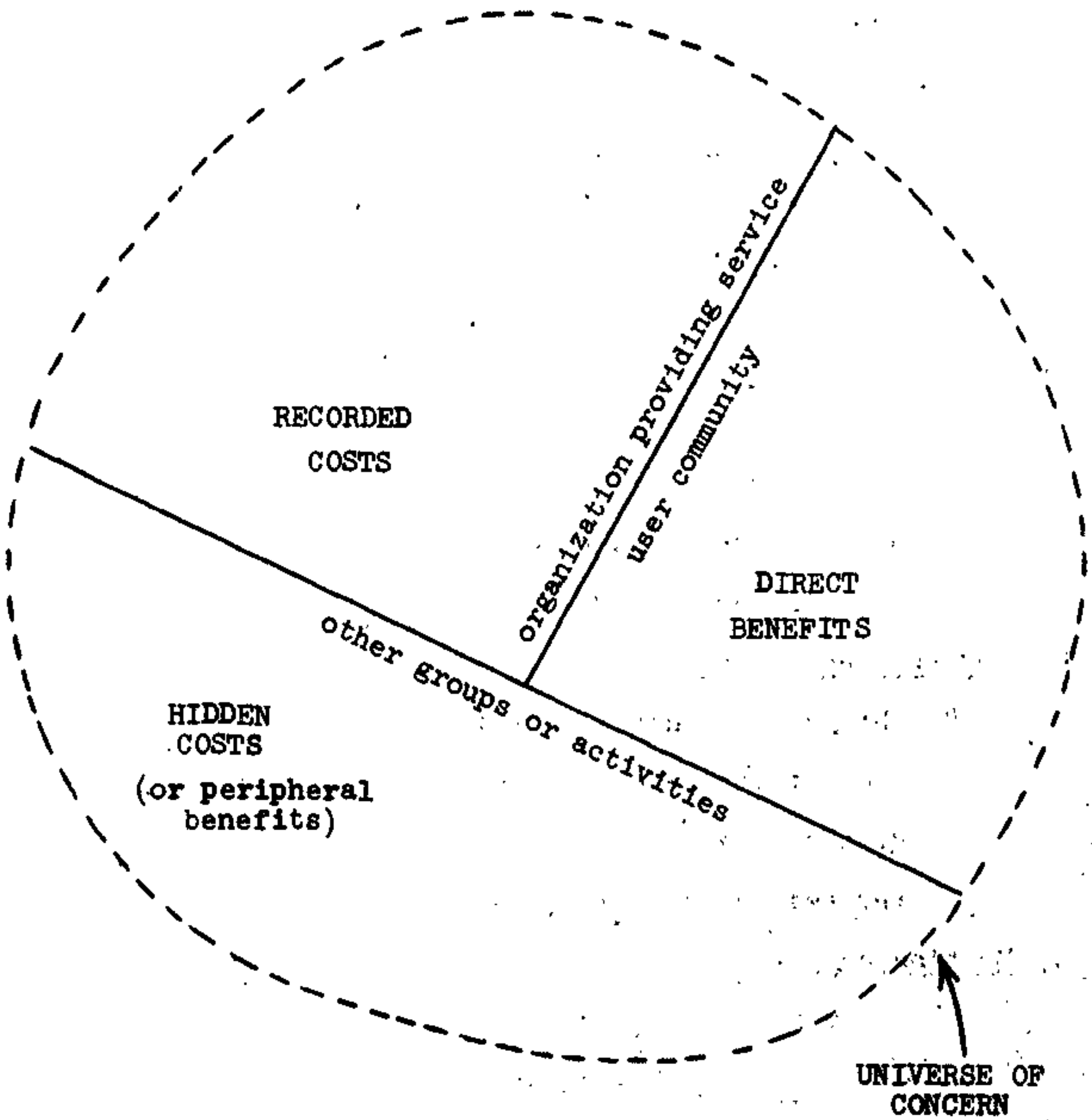

Fig. 2. Separation of benefits from costs. (Costs=negative benefits.) Note that there is some arbitrariness in the locations both of the outer boundary and of the internal dividing lines.

alone are negative and appear in the accounts of some organizational entity as "costs"; these activities are represented by the sector at the top of the figure labeled "organization providing service," and the corresponding costs are the "recorded costs" of this organization. Next one identifies, again in the same universe, a "user community," or more properly, a complex of use activities, through which the principal benefits of the service are delivered. These benefits, unlike the costs just mentioned, will usually not be available in the form of raw monetary records, but will have to be evaluated by a special effort. What one must realize is that the dividing line between production (costs) and utilization (benefits) is somewhat arbitrary, and may be chosen differently in different approaches to the economic analysis of a given situation. For example, if one tries to measure the benefits provided by a scientific journal from data on how much its purchasers are willing to pay for it, one will include in the costs only the production cost proper; but if one approaches benefits from the amount of time users spend reading it, one should include on the cost side such things as library storage costs, xeroxing costs, etc. For a given universe of concern the net benefit will not depend on where the dividing line is drawn, but the partitioning into "costs" and "benefits" will.

After all this has been done there may or may not remain, within the universe of concern, other groups or activities for or through which other advantages or disadvantages accrue, less subject to direct observation than the recorded costs and direct benefits just mentioned, but still deserving some consideration in decisions. For example, producing a scientific journal entails a cost to the scientific community which does not appear on the books of the journal, namely, the time invested by referees; similarly, the practice of publishing articles not only provides information to readers, but motivates authors to complete their work. Such hidden costs or peripheral benefits are shown at the lower left of the figure.

We have so far been talking about the information service and its use as if they were static, in a uniform steady state. But real life is full of fluctuations, innovations, and secular trends. Economic decisions, like any others, must allow for foreseeable changes and preserve options for dealing with the unforeseeable. Changing circumstances can sometimes be allowed for in a quite quantitative way, as when one amortizes a developmental cost or a capital investment. But even in such cases an assessment of net benefit must refer to a specific period of years , the duration of which is chosen somewhat arbitrarily by the planner. Slightly harder to reduce to numbers, though basically amendable to orderly planning, is the question of how much extra cost is incurred, or how much utility lost, when basically desirable changes are made too suddenly. Other factors having to do with adaptability to changes are still more difficult to quantify: What is the value of stability with respect to unforeseen fluctuations in user demand, in unit costs, in amount or kind of information to be handled, etc.? Will the motivation for wise or efficient decisions by the operators of an information service be improved or dampened by a given change in its mode of operation or support? Finally there is always the question: can one win the support of the people whose cooperation is needed in order for an enterprise to work?

Figure 3 summarizes the factors we have mentioned as important for economic decisions, and lists various types of decisions to which they must be applied. (It is interesting to note that the categories used in this figure have been used independently by others:[3,4] thus our first, third, and fourth columns are essentially what are called "cost," "effectiveness," and "benefits" in the opening chapter of the book by King and Bryant[3].) Although all the columns of Fig. 3 are important, this Report will give most of its attention to the topics that have been least adequately dealt with in previous economic studies. As we shall see presently, these are the two entries encircled by the full line in the column "quantified benefits"; we shall discuss their nature, and the distinction between them, in item 5 below. Our recommendations, given in Section V, will

FACTORS IN COST-BENEFIT ANALYSIS

← EFFORT | RESULTS →

COSTS	IMPONDERABLES AFFECTING COSTS OR BENEFITS	PERFORMANCE	QUANTIFIED BENEFITS
Recorded costs:	Stability	(Easily measured characteristics presumed positively correlated with benefits)	Dollar calibration
Development			
Start-up	Motivation		Theory of social value
Input			
Output	Continuity		
Marketing			
Hidden costs	Political harmony		Objective methods

DECISIONS DETERMINED BY SUCH ANALYSIS:

Inaugurate?

Close down?

Modify?

Pricing and promotion

Subsidy?

Fig. 3. Components of economic decision-making about information services. The full and dashed contours identify the major and subsidiary concerns, respectively, of the present Report.

deal largely with these topics, but will also touch on the entries "stability" and "motivation," and "objective methods" which are accordingly encircled with dashed lines.

2. The diversity of perspectives on information economics

One of our first objectives should be to survey the thinking that has been done to date on the economics of information services, and to identify gaps, limitations, and areas of progress. We shall do this only rather sketchily in the following paragraphs, mentioning only some of the highlights of the literature of the subject, and deferring until later sections the discussion of what it has had to say on certain specific topics.

The economics community, whose expertise is obviously greatly needed, has only rarely given any attention to information services. Machlup's famous book,[5] though full of sound philosophy and economic statistics about information for the public and the economy as a whole, covers so broad a territory that scientific and technical information services get lost. On a much smaller scale, there is a thought-provoking 1968 lecture by Marschak,[6] which essays a very broad-brush look at certain issues from an economist's point of view. He approaches the role of information by distinguishing the three functions of inquiring, communicating, and deciding; he points out that the discipline of communication theory telescopes the first and third of these to concentrate on the second, while the discipline of statistical decision theory telescopes for communication link. He also calls attention to the importance of the "problem of optimal assortment" -- the tying together of services aimed at different desiderata into a package which, though not optimum for any single user, may be economically favorable for the community. But from the paucity of economic studies mentioned by Marschak, and their specialized nature, it is natural to infer that not much work relevant to information services has been done.

Fairly recently a few more explicit attempts have been made by economists to deal with some of the problems of information services. A study[7] done for OSIS by a group at Mathematica, Inc., under the direction of W. J. Baumol, has shown how a number of basic principles and methodologies of economics apply to some kinds of information services, discussing such things as externalities, optimal pricing, subsidies, economies of scale, and mathematical modeling. We shall discuss specific points in this report in various places below. In general the treatment is rather abstract, and is admittedly preliminary, not yet incorporating many ideas from the communities of information managers, information scientists, or users.

Also worth noting are some studies by Berg[8] of the economics of scientific journals. These again are formulated in the framework of an idealized economic model; however, some fairly detailed analyses of price elasticity of demand, etc., are given for specific journals.

Though not specifically applied to information services, the work of Becker[9] on people's time as an economic variable is of considerable potential importance for them. We shall use some related concepts in Section III.4. In a more philosophical than economic context, Simon[10] has also stressed the value of people's time, and the importance of conserving it.

When information scientists and information managers approach economic problems, they are apt to be most interested in the costs of their systems and in nondollar measures of their performance. These traits are manifested in many studies (mostly unpublished reports) discussed in the chapters entitled "Design and Evaluation of Information Systems" of the series Annual Review of Information Science and Technology.[11] A large and coherent group of such studies, performed by Westat, Inc. for OSIS, has been brought together and expanded in the book of King and Bryant,[3] which gives an excellent discussion of many aspects of cost analysis and especially performance evaluation for document transfer systems. Though it gives a brief discussion of economic modeling, this book has little to say on ways for determining the coefficients relating benefits to measures of performance. Similar comments could be applied to many other, more specialized, discussions in the information-science literature, some of which we shall cite in the sections to follow.

The user community has contributed to thinking on information economics only in those rare instances where some of its members have temporarily assumed roles as information scientists or information managers; thus their contributions are not sharply distinguishable from the category just discussed. An example that is clearly part of the information-science literature is a book by Morse[12] which shows how operational-research techniques can be employed in library management, though it says little about such central economic issues as the quantification of benefits. Typical of comprehensive user-oriented studies of information problems is the SATCOM Report,[13] which calls attention to the great need for improved understanding of the economics of information services, but contributes no explicit suggestions. More explicit economic statistics, and an attempt at economic modeling and rough quantification of benefits, are contained in a SATCOM task group report on primary journals.[14] Still another type of attack on the economics of information services, based on user surveys, is represented by a study made by a committe of the American Chemical Society,[15] in which time spent by chemists in the use of information resources and time saved by the use of certain services is estimated.

We shall discuss some of the ideas of these latter two studies in items 4 to 6 below.

3. Costs, performance, and analyses using these alone

We are now ready to start discussing the various areas in the table at the top of Fig. 3. From the brief survey of types of literature we have just given in item 2, it is clear that existing studies have treated costs and performance far more adequately than the other items. So it is appropriate for us to begin with a few words about the state of our knowledge of how to treat these two factors, and what can be done with them alone, without having to convert performance into something with a definite dollar value.

As has already been hinted, costs are of many types. Some are transient, and must be amortized over a suitable period: these include developmental costs, initial promotion, capital equipment, etc. For some one-shot activities (e.g., a special conference), all costs are transient. But most information services do have a more or less steady-state operation, and here it is important to distinguish those costs that depend mainly on the supply of information to be processed from those that are roughly proportional to the number of users of the service. In the familiar case of scientific journal publication, these two categories are represented respectively by the editorial and composition (pre-run) and the printing and mailing (run-off) costs.[16] There are also steady-state promotional and marketing activities. More subtle, but often very important, are the hidden costs: their magnitude may be very dependent on where one chooses the boundary of the "universe of concern" in Fig. 2. From a broad social point of view, referees' time and other donated services, authors' time, readers' time, travel time, costs of storage space, etc., should often be taken into account.

As we have noted earlier, the literature on costs of information services is fairly extensive. Annual surveys of it can be found in the chapter of Annual Review of Information Science and Technology entitled "Design and Evaluation of Information Systems";[11] a modest minority of the one or two hundred references cited each year pertain to costing. The book of King and Bryant[3] discusses the subtleties of cost analysis in some detail, especially in Chapter 3. The SATCOM Task Group Report[14] contains much information on both recorded costs and hidden costs for primary journals. Further samples of the literature on costing are provided by papers of Murdock and Liston[4] (retrieval systems), and of Landau[17] (bibliography on costs of document-surrogation systems), and by a recent review by Penner.[18] Though often disregarded, secular trends in costs can be very important. The causes and implications of such trends are discussed in Section V of the Mathematica study.[7]

Probably the most extensively studied of the elements listed in Fig. 3 are those subsumed under the word "performance." Thus, a large part of the book of King and

Bryant,[3] much of Lancaster's book[19] and a large portion of the references in the annual "Design and Evaluation" reviews,[11] have to do with the design and use of measures of performance, usually for document-transfer or retrieval systems. Henderson[20] has given a bibliography of the older literature. Typical examples of what we call performance measures for retrieval systems are speed, completeness, and percentage of retrieved items found relevant. Performance measures for primary journals would include time lag in publication, frequency with which articles are cited, etc. Miscellaneous measures of performance for library operations are described in Morse's book.[12]

There have been many efforts to attack the problem of cost effectiveness using only cost and performance data; the reviews cited[11] mention many such. Of course, what one really wants to do is to maximize the difference between benefits and costs, each expressed in dollar terms, and this can obviously only be done in the general case if the benefits can be quantified. (As King and Bryant[21] make clear, it is the benefit-cost difference, not, as sometimes implied,[4] the ratio, that is important.) But in a restricted domain it may suffice to use an uncalibrated measure of value, such as is provided by performance indices of the type just described. Thus, if one has alternatives to decrease cost while keeping all performance measures unchanged, the choice is obvious;

similarly, an option of improving a performance measure, while keeping costs and all other performance measures constant, again leads to an obvious decision, subject only to the assumption that one knows which direction is "better" for the performance index in question. Morse's book[12] gives examples from the library field. More generally, one can, following the Mathematica Report,[7] eliminate as "inefficient" any alternative for provision of an information service which is "worse" than some other alternative in cost or in one of the several dimensions of performance space, while being "better" in none of these respects. But if one alternative is better than another in one performance measure, worse in another, one will need to be able to compare utility quantitatively for the two measures. In some practical situations one may however be able to do this, or at least may be able to roughly identify, from common sense, a cone of directions in the space within which one is sure the benefit is improved. When different dimensions of performance space can be intercompared, sophisticated optimization studies may sometimes be possible.

Unfortunately, most economic decisions do not lend themselves to this type of analysis, since most contemplated changes in an information service alter both costs and benefits. These must be evaluated in the same units if the optimum trade-off is to be evaluated or even, in many cases, if one is to be confident of moving in the right direction.

4. Quantification of benefits: objective methods

We come thus to the all-important area of calibration of benefits in dollar terms. The neglect of this area in the literature has been due more to the presumed absence of handles for attacking it than to lack of recognition of its importance. For example, in the most recent of the annual reviews Cleverdon[22] states "The major effort in the coming years should be the development of a methodology for determining the value of information-retrieval systems." The difficulties in quantifying value for changes in information services of course are not qualitatively different from those one encounters in valuing other types of innovations in products or services; it is just that most of the usual difficulties are a little more acute. In all cases an innovation changes the way people work, and so changes their productive output. If one is dealing with a service that affects only a particular group of people, one might conceivably conduct an experiment with two matched groups, one with the service in question and one without it, and compare their productivities, which in turn could hopefully be given a dollar value. In practice, this can almost never be done, though we shall presently say a few words about occasional situations where something approaching it would be feasible. In most cases the needs and objectives of different subunits in the user population will be different and will vary with time, so that no matched groups can be found for comparison. Moreover, in the extreme yet very frequent case of a service that operates in parallel with other channels to communicate information among diverse people or groups throughout the nation or the world, such a controlled objective evaluation is obviously impossible. Here one is stymied not only by the diversity of the users, but by the fact that any change in the service in question affects a given user not only directly, but also indirectly, through its influence on his colleagues near and far. Thus no tractable small test population can be isolated.

Information services used in the performance of fairly standardized tasks offer the best opportunity for approximating the objective measure of benefits just mentioned. For example, a computerized handling of medical data on patients might conceivably be evaluated by comparing death rates and convalescence rates for hospitals using it with those for hospitals without it. An especially important class of cases offering a possibility for objective measurement occurs when the principal benefit of an information service consists in reducing the time required for its users to perform some accurately specifiable task which then must do whether the information service is available or not. The reason for the underlined words is that if altering the service can alter the frequency of choosing to perform the task in question, one must then compare the value of these tasks with the value of the other things the users devote their time to. Such a comparison would usually be very difficult, though it would be very useful, as it would enable one to judge whether the workers in question allocate their time optimally.

In view of the rather restrictive requirements that must be met for these objective measures of value to be applied, it is not surprising that no studies have come to our attention which approximate the ideal at all well. The closest approximations we know of are typified by a few studies that use estimates (rather than measurements) of time saved by use of certain information tools, and without checks on how well the proviso underlined in the preceding paragraph is fulfilled. For example, the American Chemical Society study[15] mentioned earlier collected estimates from industrial laboratories regarding time saved by use of various secondary services; the benefits of evaluated data compilations have been given a rough lower bound by estimates of time saved to certain classes of users.[23] Dugger and Klinger[24] have given a sobering comparison of various user estimates of value, including estimates of time saved. Quite noteworthy, incidentally, is an observation by Mueller[25] that engineers often simply mark time while waiting for needed information.

Before leaving this topic we should caution that the conversion of a time saving into dollars, though obvious in a crude sense, has some subtleties; we shall discuss these below in Appendix C.

5. Quantification of benefits: market response

The discussion just concluded has shown that it is rarely feasible to establish an objective dollar equivalent for the benefits conferred by an information service. This being the case, we must fall back on subjective measures, i.e., on judgments by qualified experts that the benefit from an information service is worth so many dollars, or is equal in value to some other quantity which can in turn be given a dollar equivalent. While one is entitled to be suspicious of subjective judgments, their credibility is greatly enhanced if the persons making them are: expert in the field where the service in question is used; motivated, in that their personal goals depend importantly on the soundness of these judgments; and numerous, so that individual idiosyncracies are likely to average out. And there is indeed a measure of value, highly venerated in the science of economics, that is based on subjective judgments with all three of these characteristics: this is the "test of the market-place," in which the value of a product or service is measured by the price its users are willing to pay for it. We shall devote this section and Section III.3 to the market approach; Section II.6 and III.4 will take up another approach, based on subjective judgments with the same three characteristics, that is often even more powerful for application to information services.

Before going further, however, we must call attention to an important question that arises whenever one tries to base an estimate of benefit to a "universe of concern" (cf. Fig. 2) on value judgments made by a variety of people. Each of these people will evaluate the benefit of the service in question to himself (or to the organizational subunit he represents); the benefit to the universe of concern is not necessarily the same as the sum of these self-perceived benefits, though in most cases of interest to us it is usually of the same order. For example, when a scientist who reads a journal article learns facts from it, these benefit not only his own work but that of colleagues to whom he may transmit the information orally, and even colleagues in foreign countries who may be made aware of this information through articles he eventually writes. Yet his decision about paying for a personal subscription to the journal will probably be very little influenced by such considerations. In the other direction, buyers who are subsidized may occasionally purchase services that are not economically justifiable for society. In extreme cases the self-interest of an individual or group may even by strongly antisocial: for example, a would-be hijacker might be willing to pay quite a bit for detailed knowledge about the limitations of metal-detection systems. So if one uses judgments from the user community, one must always ask: what interests do these judgments represent, and how are these related (quantitatively, to the benefit received by the universe of concern one wishes to consider?

To answer the question just posed, one in general needs to set up an economic model that will provide quantitative relations between the individual self-interest benefits and the benefit to the total universe of concern: this step, designated by the words "theory of social value" in Fig. 3, will be discussed in item 7 below. Here and in item 6 we shall discuss only the evaluation, in dollar terms, of the pluralistic self-interest judgments. This may be called "dollar calibration," since it is the crucial step in establishing a connection between the benefits of a service and the world of dollars.

The fundamental assumption of the market approach to dollar calibration is that the prices buyers are willing to pay represent intelligent self-interest judgments. For various reasons this may not always be the case, though usually it is hard to suggest who else could do better. Part of the fallibility arises from the fact that many purchasing decisions are not day-to-day decisions by individual users, but are decisions made only once in a while by administrators, who may maintain only imperfect contact with the actual users of the information services they buy. Green[26] has reported results of a game played with marketing executives who were given decisions to make regarding purchase of market surveys, etc., the rules being such that mathematical methods of decision analysis could be used to determine correct decisions in every case. The decisions of the executives were by no means always optimal. While real-life situations might well lead to grass-roots pressures to revise poor decisions, it is sobering to realize that executives, like everyone else, are fallible.

Even when one accepts the assumption of intelligent self-interest judgments, one must face another difficulty: even the self-interest value to the buyers individually is reflected, not just in the price they actually pay in a given market, but in the prices (perhaps very different from buyer to buyer) which they would be willing to pay if necessary. (The relation to the demand curve of economic theory is detailed in Appendix B.) While market statistics can give some information on this, the information is usually rather sketchy. All too rare is the interesting type of experiment reported by Urbach,[27] in which similar groups of documents from NTIS were marketed at widely different prices, and the demands for the different groups compared. For the case of primary journals, fairly detailed studies of demand curves, etc., have been made, e.g., by Berg,[8] by the Mathematica group,[7] and by the SATCOM Task Group.[14] The first two of these have attempted a rough quantitative accounting for the dependence of demand on other factors than price, namely, characteristics of the journals and of the potential user population. But in all cases one gets only fragmentary information on the value judgments of those buyers to whom the value is greatest. We shall explore the seriousness of this limitation in more detail in Section III.3 below.

6. Quantification of benefits: time investment by users

Though its significance has not been as widely appreciated as that of free-market prices, the time that users of information services choose to spend on them can be equally or even more valuable for purposes of dollar calibration. Each individual user is perpetually making judgments that balance the value he receives from use of an information service against the value of what he might be doing in the same amount of time devoted to one of his other activities. If one can assume that on the average these judgments are sound (note that an analogous assumption was also needed for dollar calibration by market response) then one can estimate the dollar value of the services to the users if one knows the value of their time and the (nonlinear) dependence of their ultimate productivity on the time spent on productive work. As before, one must use some theoretical analysis of the sort to be described in item 7 below, to obtain a correct measure of true social value, or value to a given universe of concern, from the values perceived by individual users.

To our knowledge the only published application of this approach is that given in Section II.4 of a study by the Physics Survey Committee,[28] although some slight use of it was also made in the Report of the SATCOM Task Group;[14] more detailed unpublished studies are, however, available.[27] And although the explicit application to dollar calibration may not have been made, most of the sizable existing literature[30] on time spent in the pursuit of information can readily be applied to this end.

We shall discuss this approach at length in Section III.4 below, and in Appendix D. In general it seems very promising for a large class of cases since the three requirements for reliable dollar calibration from subjective judgments, mentioned at the start of item 5, are often even better fulfilled than for the approach via market response: individual users are in general more <u>expert</u> than anyone else, even their bosses, in assessing the utility of day-to-day use of an information resource; scientists and engineers (unlike some other workers) are usually strongly <u>motivated</u> to optimize their individual productive output; individual users are the most <u>numerous</u> group one could consult. But one condition must be fulfilled if this approach is to work: the investment of users' time must be sufficient to be comparable in value to the other costs of the information service in question. For primary journals[14,28] and for many secondary services[15] this can indeed be verified to be the case; however, it need not always be so. When only negligible time is required from users, their decisions on use of the service will not usually be based in significant measure on the value of this time to them, so the basic assumption of the present approach will fail.

7. Quantification of benefits: theory of social value

As we have mentioned several times already, raw inputs of the type we have been discussing in items 5 and 6 above need a certain amount of processing to convert them into valid measures of equivalent monetary value to the universe of concern, i.e., to society or to a particular organization or group. Thus, one must allow for the fact that supplying an information service to a particular individual in general gives a benefit to society that is significantly different from the self-interest benefit perceived by this individual as he makes his decisions on expenditure of money for the service and of time for the use of it. If the dollar-calibration data have been obtained from market response, one must often make some assumptions about the behavior of the demand curve beyond the range that is empirically accessible. If the dollar calibration is from users' time studies, conversion to social value requires an understanding of the other uses to which a given amount of time can be put, and also of the way in which various channels for information flow are interconnected. (See the more detailed discussion in Section III.4.) In either case, one needs to construct a mathematical-economic model, the various parameters of which can be determined empirically, one of these being the dollar-calibration input previously discussed.

The general problem addressed by such models is tha! computing how the net benefit to a given universe of concern will change if a change is made in one or a few of the controllable parameters of an information system. One must take account of the multiplicity of parameters determining the costs and performance of the system, and the response of users to it, and must allow for the fact that some of the parameters are determined from the others by conditions of economic or sociological equilibration. For example, the benefit delivered to the U.S. by a published data compilation will depend on the number of U.S. users, related to the number of U.S. buyers; the costs to the U.S. as a whole will depend on the latter number and on the sales abroad and the price. Domestic and foreign sales will depend on the price at which itis marketed, according to certain demand curves. To compute the change in net benefit to the U.S. resulting from a change in price one would thus solve some simultaneous equations, after having in some way obtained empirical forms for the functional relationships just mentioned.

As in the case of dollar calibration, the literature on the theory of social value of information services is quite meager; little of the extensive effort economists have lavished on models for the interplay of conventional goods and services has spilled over into the information field. However, two of the studies by economists that we mentioned in item 2 above deserve attention. Berg's study[8] of the scientific journal market considered the interplay of free-market response and possible subsidization, and developed guidelines for the determination of the socially optimum mode of support of journals. The Mathematica study[7] offered a number of general observations on such subjects as externalities, economies of scale, and the relative roles of subsidy and free-market income; it gave especially detailed attention to the question of optimal distribution of a given subsidy budget over a diversity of information services. Machlup's book,[5] though too broad to focus explicitly on the concerns of the present Report, contains many interesting qualitative observations. Among studies originating in the natural-science community, the SATCOM Task Group Report[14] has undertaken a crude modelling of the social value of primary journals, and the information report of the Physics Survey Committee[28] has done a little modelling for information services in general, adapted to a user-time input. We shall discuss these approaches in more detail in Section III.4.

One conclusion, on which there is remarkable unanimity among all the studies cited, is sufficiently important to be mentioned here. Namely, for an information service to be cost-effective from the standpoint of the nation as a whole, it is by no means always necessary for it to be viable as a self-supporting operation when publicly marketed at a standard price. This is because net benefit is usually maximized by marketing at or even below replication cost, while it is characteristic of information services that input or set-up costs are often the largest cost component. Thus some sort of subsidy if often in the public interest.

8. Miscellaneous clues to cost-effectiveness

The preceding sections, based on the framework of Fig. 3, hopefully cover nearly all the possible approaches to economic decisions about steady-state operation of information services. However, the outline is not quite air-tight, and there are occasional avenues of investigation that do not fit simply into it but that may occasionally be helpful. We shall mention here only a few examples that have come to our attention.

Sometimes one can draw useful conclusions by studying the subjective judgments users make regarding relative value or cost to themselves (including the value of their time) involved in the use of alternative information sources, but without attempting an absolute dollar calibration. Data on such judgments can be used, for example, to decide when a given information service will be of no value, because potential users will avoid it in favor of other competing services. This sort of analysis, analogous to the use of performance data to eliminate "inefficient channels" (see item 3 above) has been used by Brookes[31] in a mathematical model for decisions on library operation.

Surveys of user preferences, though often unreliable, can occasionally be very useful. For example, when they reveal near-unanimity, they can be used in the same way as performance measures to eliminate "inefficient channels" when the preferred service is no more costly than the alternatives. Urbach[27] has given a practical example from the NTIS operation. Another possible use of preference studies is to decide which direction of a given performance measure corresponds to increasing utility (e.g., should invited talks at a meeting be longer or shorter?). More frequently, one may wish to use such studies to determine how to weight several different performance measures to get a single (uncalibrated) measure of utility. Design and analysis of surveys for this purpose have been discussed, for example, by Sadacca and Root.[32]

In all such work one must be cautious about treating all users as equivalent. The opinions of those users who make a great deal of use of an information service should be given more weight than those of occasional users. As in all other approaches to economic decisions, one really ought to understand the topology of the many interconnecting information-transfer activities in the user community (see Section III.2 below).

Correlations of productivity of individual scientists or engineers with their amount of use of information resources,[33] though interesting, provide little basis for quantification of benefits, if only because it is difficult to disentangle cause from effect. However, such studies can be useful for thereason mentioned in the preceding paragraph: they can sometimes show, for example, that a time-saving service is more valuable than would be estimated from mean use statistics, if use is predominantly by the most valuable people.

Studies of the frequency with which work previously done by one group is duplicated by another through ignorance of the earlier work can conceivably provide a valuable lower bound to the potential value of an adequate information service. Unfortunately, reliable data of this sort are very hard to come by. Typical of attempts to get such data are some studies by Martyn[34] (industrial chemists and university scientists) and by Brockis and Cole[35] (presumably industrial research projects). These and many similar studies have also obtained data on information received which it would have been valuable to have had earlier; no quantitative measures of <u>how valuable</u> are available, however.

9. <u>Other practical considerations in a changing world</u>

To conclude this section, we should say a few words about some of the items in the second column of Fig. 3, especially stability and motivation. These are important considerations for most economic decisions, and especially so for decisions about the interplay of subsidy and free-market support for information services.

Let us first consider the financial stability of an information service, in the face of possible fluctuations in amount of information to be processed, number of purchasers of the service, or unit costs. The effects of such factors on primary journals have been discussed in detail in various places [36,37,38] and illustrate principles that often apply to other types of services. Three typical conclusions, resulting from quite elementary economic analysis, are: marketing at a fixed price per volume of given size, rather than at a fixed price per year, favors stability with respect to fluctuations in amount of material submitted; support of most of the pre-run costs by page charges provides stability with respect to fluctuations both in material submitted and in demand; stability in the face of fluctuations in unit costs is favored by both these measures.

Motivation is more subtle to analyze. From the national viewpoint, for example, one is interested in motivating the managers of a service to manage efficiently and imaginatively, in such way as to optimize net benefit to the nation. In some situations, for example, the keen use of the for-profit entrepreneur may be the most useful means of recognizing potential markets, lowering costs, etc.; in other cases, as in the common example of unnecessary journals,[39] the profit motive can have very antisocial consequences. The SATCOM Task Group Report[40] contains the most detailed study we know of concerning the influence of various factors in motivation of information managers to serve society well. As the work in this study shows, valid conclusions about motivation usually need to be based on a study of economic models of the type discussed in item 7 above.

The data that market research can provide will often be useful or indispensable in arriving at practical conclusions both on stability and on motivation.

III. DISCUSSION OF THE MORE PROMISING APPROACHES TO QUANTIFICATION OF BENEFITS

From what has just been said in Section II, it should be clear that of all the elements depicted in Fig. 3 as entering into economic decisions, the ones most in need of more data and deepened understanding are the two in the last column encircled by the full line (dollar calibration and theory of social value); we believe that the Science Information Council was thinking specifically of these when it ordered the formation of our Panel. Although stability and motivation are somewhat independent of the items in the last column, and motivation, particularly, is equally challenging and elusive, we have not had time to explore them in detail; this section will therefore be devoted simply to a discussion of the more promising approaches to dollar calibration and theory of social value. We shall start with some considerations common to all the major approaches, and shall then take up the practical aspects of quantifying benefits using inputs of either the market-response or user-time types. Figure 4, which summarizes the import of our survey in Section II, provides a concise reminder of why we consider these to be the most important topics to discuss.

1. Nature of social value

The social value to which scientific and technical information services contribute is that of science and technology generally. Particularly at the

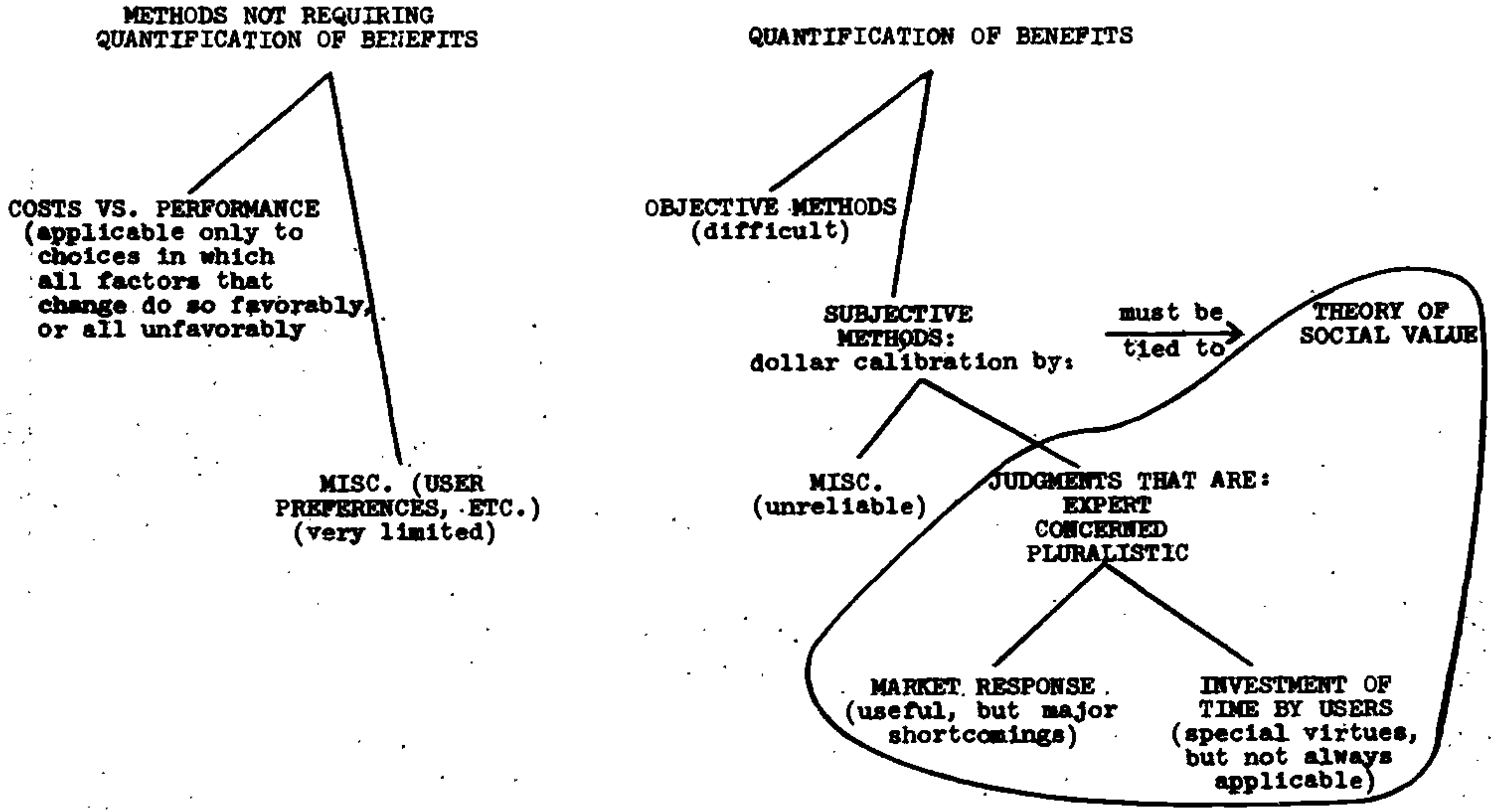

Fig. 4. Avenues for steady-state economic decisions (digest of most of Section II). The areas at the right enclosed by the full line are the ones to which Section III is devoted.

research end of the spectrum, quantification of the value of scientific and technological work is notoriously difficult; the most, therefore, that a theory of value for information services can do will be to "ride piggyback" on whatever judgments may already have been made about the value of the entire scientific and technological enterprise. Specifically, we shall assume that scientific and technological activities now under way do indeed yield more value than they cost, and that the amount of money now being spent on them is somewhere close to optimum. These assumptions may not be correct -- for example, for reasons analogous to those developed in item 2 below, one might suspect that industrial funding of research is suboptimum from the standpoint of the whole society -- but correcting them is a task for those involved with the broadest aspects of the nation's economy, not a task for us. If F dollars of funding give a yield of Y dollars worth of benefits in any enterprise to which these assumptions apply, we may infer that:

(a) if the enterprise is being pursued, $Y>F$. (1)

(b) At the current operating level, $\partial Y/\partial F \approx 1$. (2)

A hypothetical change in an information service will alter the form of the function $Y(F)$. The change in net benefit received by society, due to the change in the information service, can be expressed in either of the alternative ways:

$$\begin{aligned} \delta\,(\text{net benefit}) &= \delta(Y-F) \\ &= -(\delta F)_Y , \\ &= (\delta Y)_F > \left[\left(\frac{\delta Y}{Y}\right)_F\right] \cdot F \end{aligned} \qquad (3)$$

where $(\delta Y)_F$ is the change in Y that would occur if the funding level were held constant, and $(\delta F)_Y$ is the change in funding that would be required to maintain the same output using the altered service.

In Appendix C, where these matters are discussed a little more fully, a further point is brought out which is sufficiently plausible intuitively to be quoted here without proof: namely, that in many cases

$$\delta\,(\text{net benefit}) = \text{a little more than} \quad \frac{\delta_o Y}{Y} F_p , \qquad (4)$$

where $\delta_o Y$ is the change in output which a fixed number of workers in a given enterprise would achieve in consequence of the altered information service, if their funds for supplies, etc., were permitted to rise in proportion to their output, and where F_p is a partially-loaded salary budget for the enterprise, of the order of one and one-half times the bare salary budget.

There is a further characteristic of most scientific and technological enterprises which greatly facilitates the estimation of social value. This is that most of the individuals and organizations involved in the enterprise are attempting to optimize, within what they perceive as their available range of choices, a measure of benefit which is more or less parallel to the benefit of society as a whole, though often not quantitatively equal to it. In other words, one can often write

$$\delta Y \approx \Sigma_i \alpha_i \delta Y_i , \qquad (5)$$

where i runs over the different performing organizations or individuals, δY_i is the benefit which the i-th such tries to optimize (relative to his costs), and the quantities α_i are positive numbers. Thus, universities and their staff members try to optimize their output of good-quality research; industrial laboratories try to optimize the technology for producing useful products; hospitals try to optimize speed and effectiveness of healing; etc. In some areas of human activity this parallelism of individual and societal goals fails completely: for example, in warfare, or in the cigarette industry, α_i can be negative; in schools of astrology, δY_i may be essentially unrelated to δY. While such nonparallelism may sometimes occur in activities affected by scientific and technical information services, such cases are probably a small proportion of the total. When (5) holds, the problem of estimating the change δY in total social benefit reduces to the local problems of estimating the δY_i, and the cooperative problem of estimating the coefficients α_i. The latter are usually greater than unity (though sometimes not by much) because strengthening one element of the scientific-technological enterprise usually provides indirect benefits to other elements.

2. Importance of the topology of information-transfer channels

Useful information is transmitted through a multiplicity of channels, which we have sketched in rather general terms in Fig. 1 above. These operate both in series and in parallel: much of the information delivered through any channel was in turn fed to this channel by some other channel; often the same information can be transmitted independently in two or more channels. Thus a change made in a particular information service affects not only the amount, speed, and quality of information transmitted by this service directly to users, but also affects the amount and utility of information which they receive through other channels which may have been nourished by the service in question. Moreover, the change will cause the relative use of the different channels by the user community to shift. The problem of computing the effect of the change in question on the total flow of information is analogoue to the problem of computing the change in the impedance of an electrical network when one of its elements is altered. Thus a prerequisite to any quantitative understanding of the effects of changes in information services is a knowledge of how the different channels of information flow are interconnected and how much information flows in the different parts of the whole.

There is a modest but significant literature on the relative use of different communication channels by various kinds of scientists.[30] There are also a number of studies of communication among engineers,[41,42] which show markedly different patterns from those prevalent among scientists. Most existing studies fall somewhat short of what one would like, in that they have concentrated mainly on the last link of an information-transfer chain from originator to user, and have not adequately elucidated the way in which different information channels interconnect. Nevertheless, they can still be useful as a guide for the setting up of models for estimating the effect of changes in a given service on total information flow. One rather sketchy study[43] has provided details of this sort for a particular area of physics. This study also made the useful distinction between transmission of scientific information itself (e.g., a formula, a derivation, a datum, a methodology) and transmission of a clue to the existence or whereabouts of such information. Studies by Allen and his collaborators[44] in organizations of a more engineering type have also been very revealing, showing the concentration of information flow through certain individuals called "gatekeepers." It would be most interesting to have further studies that would show how the pattern of interconnection of different communication channels differs from one field to another and in different organizational environments.

To emphasize the practical reality of this issue it may be helpful to cite an example. Every now and then a study appears in the library literature on the question of weeding collections or economizing on subscription and storage costs for periodicals. A common and superficially plausible suggestion is that one should drop subscriptions to journals for which the quotient of frequency of use by bulk or subscription cost is much lower than average.[45] While statistics on frequency of use are obviously vital for intelligent library management, it is risky indeed to take benefit delivered as proportional to number of occasions of use. Certain individuals may pick up important information from journals their colleagues rarely consult, and pass it on orally, etc. In the extreme case of dominance of information flow by "gatekeepers,"[44] the pattern of literature use by the latter, who may constitute no more than a tenth of the professional population, may be more important than that of use by everyone else.

To give a numerical example, in the library of the Bell Laboratories, at Murray Hill, New Jersey, Russian-language periodicals occupy about 5% of the shelf space, but account for less than 1% of periodical use, probably about the 0.5% found by Chen[44] at the MIT library. If these figures were specialized to physics, both numbers would be a little higher, but probably not by much. Yet at a "journal club" where various individuals reported on particularly exciting new discoveries in the current physics, about 12% of several hundred items came from Russian-language journals (before these were translated). Thus the 0.5% use figure, which one might discount even further because of the existence of translation journals, would probably significantly undervalue the Russian-language journals.

3. Benefit analyses based on market response.

The relation of value to buyer response is particularly simple if one can make the simplifying assumptions: first, that the total social benefit delivered by a service is the sum of the benefits delivered through each of the purchasers of the service, and second that the contribution through each purchaser to the total social benefit is the same as the self-interest benefit perceived by that purchaser. In such case Eq. (5) applies to any change in the service, with all $\alpha_i=1$. The benefit realized from providing purchaser i with the service is measured by the maximum price he is willing to pay for it. The "demand curve" of economic theory is simply the curve showing the number of people willing to pay more than any given price; it is customary to plot price vertically and number horizontally, as shown in Fig. 5. It follows from very simple reasoning, detailed in Appendix B, that the sum of the values received by all buyers who purchase a product or service, not available elsewhere and offered at a single price p, is simply the area lying between the demand curve and the price axis and to the left of the vertical line intersecting at p, as shown shaded in Fig. 5(a).

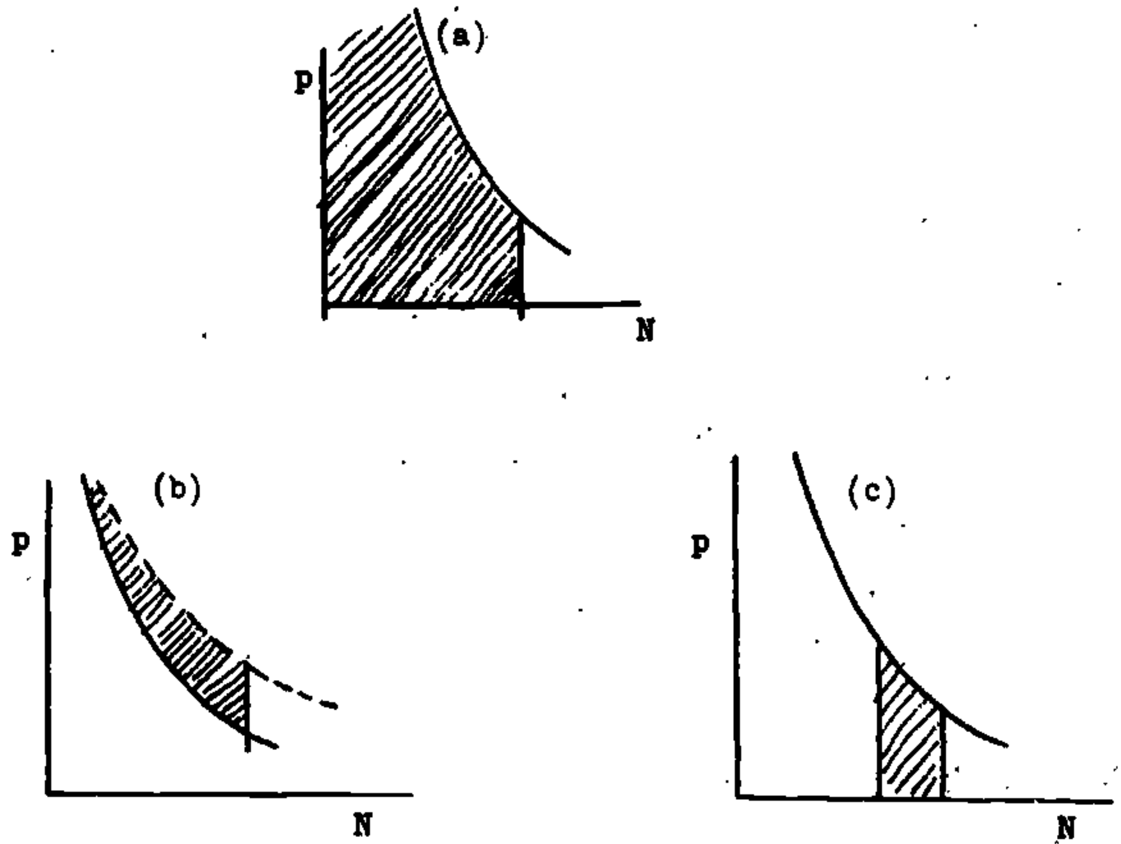

Fig. 5 Demand curve (relation of number N buyers to price p) for a product or service. (a) Shaded area shows total benefit delivered. (b) Change in benefit delivered (shaded area) due to a change in the product resulting in a change of the demand curve from the full to the dashed form. (c) Change in benefit delivered (shaded area) due to a change in price alone.

Aside from the assumptions already made, the weakness of this method of estimating value lies in the fact that it requires knowledge of the shape of the demand curve up to very large values of the price, where empirical market data are almost never available. Any change in the service offered will normally change the entire demand curve, say from the full line to the dashed line in Fig. 5(b). The social value of this change, under the present simplified assumptions, is represented by the shaded area in Fig. 5(b), but this depends on the forms of both curves over an empirically inaccessible range of prices. While one can sometimes approximate the curves by simple analytical forms which allow extrapolation to infinity from the often measurable slopes of the curves in the neighborhood of the present price p, the accuracy of such extrapolations is questionable; for example, the most commonly assumed form, a power law dependence, often gives the nonsensical result of infinite areas (see Appendix B). The situation is considerably better, however, if the change one wishes to evaluate consists simply in a change of the price at which the service is offered, without any change in the nature of the service. In this case, as shown in Fig. 5(c), the change in social value delivered is simply the area of a vertical stripe bounded by the demand curve, and can easily be estimated from the known market response at the existing price.

In practice, the simple assumptions underlying the discussion just given (identity of social and individual value judgments, additivity of these, absence of competing services) are rarely adequate, and corrections to them must be made. Consider first the question of competition. Occasionally it happens that services which from a broad societal viewpoint could be regarded as competing, do not actually compete for user response. Primary journals provide an example: while either of two journals in a given field could provide a publication outlet for authors in this field, buyers who purchase one of the journals and not the other will not have access to the articles in the latter. As has been noted,[39] this fact is responsible for the proliferation of overly expensive journals. But in general things are not so simple. For example, secondary information on the physics literature is provided both by the AIP SPIN tapes and by the INSPEC tapes; the coverage of the latter includes that of the former and is considerably greater, but the former contains citation information which the latter does not. Clearly one would need to know a lot about the market response to both services in order to evaluate from market data the social value of a hypothetical change in one of them.

Consider finally the relation of social value to the individually perceived values of purchasers. Of the various effects that can cause these quantities to differ, probably the most important is the nourishment of informal channels for information transfer by the formal ones, or of various formal channels by each other. Thus, the social value of information delivered by a service to a particular individual lies not only in its effect on the productivity of that individual, but also in its effect on the useful information which he can communicate to his colleagues, either orally or through his writings. The individual will usually not take the latter contribution into account in deciding how much the information service is worth to him. The market estimates of value we have been discussing in the preceding paragraphs therefore need, in many cases, to be multiplied by an amplification factor -- i.e., α_i's > 1 in (5) -- whose evaluation requires knowledge of the pattern of information flow. For some typical cases involving scientific journals, it has been estimated[46] that a factor of the order of two may apply.

4. Benefit analyses based on allocation of time by users.

As we have noted in Section II, a measure of the self-interest value of an information service to a user is provided by the amount of time which he chooses to devote to using it. If users on the average can be assumed to apportion the time they devote to information services optimally, the additional benefit that would be derived from spending one more hour in the use of an information service must have a dollar value (to the user) equal to that of the loss in output resulting from a decrease of one hour in time devoted to productive work. If one makes the tempting assumption that a worker's useful output is the product of a function of the fraction of his time he devotes to productive work by an efficiency factor dependent on how well informed he is, one has the basis for an economic model. Thus, we may represent the value of a worker's output as

$$\text{Value of output} \equiv Y = \varphi(1-t_i-t_o)E(t_i,t_o) \qquad (6)$$

where t_i is the fraction of his time he devotes to use of a particular information service, t_o is the fraction he devotes to other not directly productive activities, including use of other information resources, and E is the efficiency factor. The optimality condition is then, if $\tau = 1-t_i-t_o$,

$$\frac{\partial \log\varphi(\tau)}{\partial\tau} = \frac{\partial \log E}{\partial t_i} \quad . \qquad (7)$$

Since φ is a superlinear function of its argument, whose form can be roughly guessed from common sense or estimated from the experience of people who have had sizable amounts of time siphoned off to administrative or other duties, and since the product φE, or at least a lower bound for it, is known from the expenditure made on the scientific or technological enterprise in question, we can obtain an estimage of $\partial E/\partial t_i$. If we know t_i and can make reasonable assumptions about the functional dependence of E on t_i and on any contemplated innovations, we can then make a quantitative estimate of the effect of these innovations on Y.

Some examples of how this can be done are discussed in Appendix D. One important type of example is that of an innovation capable of saving time for users of a service. As the detailed discussion of the example shows, the value of such an innovation exceeds that of the time saved (if the latter is valued at the normal loaded-salary rate mentioned in Appendix C) because of two types of amplification factor:

(a) A factor of the order of 2 or so due to the fact that only a portion of working time is devoted to directly productive work, and to the fact that the function φ in Eq. (6) increases significantly more rapidly than linearly with its argument.

(b) An additional factor, of the order of $1\frac{1}{2}$ in the example discussed in Appendix D, but likely to vary considerably from case to case, due to the fact, already noted in item 3 above, that supplying one user with information facilitates the work of his colleagues.

Though we feel that this sort of approach through users' time has great possibilities, it needs checking and refinement at several points, and, as we had already noted in Section II.6, it cannot be used for information services that require negligible time from their users. One point requiring further research is the assumption that, at least on the average, individual users allocate an optimal amount of time to each available information service. While this assumption is doubtless no worse than the corresponding assumption of intelligent self-interest that one must make in order to calibrate value from market response, and while one may legitimately argue that judgments of users relating to their everyday work are more likely to be sound, on the average, than anyone else's, there are serious grounds for suspecting that systematic misjudgments in time allocation do indeed occur. For example, there have been a few studies[47] of the use of information resources as a function of the distance one must travel to use them; these have all shown a fall-off with increasing distance rather more rapid than could be explained by the value of the extra time involved. Use therefore cannot have been optimal for all values of the distance; it was probably suboptimal at the larger distances.

Our crude model, based on Eq. (6), has ignored individual variations in the user population, and so has not made use of any information about possible special roles for "gatekeeper" types, etc. Generalizations in this direction would be very desirable. Also, it has assumed that the major conserved quantity involved is simply time; an alternative point of view[48] might be that the important quantity is the individuals' available mental energy. As shown in Appendix D, in the simplest cases one can scale out the dependence of energy consumption on time to give a model that is again of the form (6); however, different kinds of time (e.g., time spent walking to a library versus time spent deciphering an opaque paper) would require different scaling factors. In any event, more sophisticated attention to the multidimensionality of the time variable, for which we have broken out only t_1 and t_o (or, in Appendix D, t_1, t_g, and t_o), would be desirable.

In summary, we feel that intelligent studies of the way users of information resources distribute their time and of the interconnections of the channels through which they receive information provide an extremely promising tool for attaching a quantitative value to hypothetical changes in information services. However, implementation of this approach will require more extensive and more sophisticated studies of these factors than are now available. In Section IV to follow we shall give a few examples of how both this approach and the market approach might be applied to some of the decision questions we listed in the Introduction and in Appendix A.

IV. ILLUSTRATIVE EXAMPLES

In the Introduction, and in more detail in Appendix A, we have mentioned a number of typical examples of economic decisions on various types of information systems, which would be helped if a quantitative measure of benefit were available. We would now like to return to some of these examples -- it would be too time-consuming to consider them all -- in order to show how approaches based on market response or on studies of users' time might conceivably be used in order to obtain the desired quantitative measure of benefit and make the required decision. Our discussion of these examples here will be quite condensed; Appendix E describes our suggestions on them in more detail. However, it is to be stressed that these suggestions are only illustrative: they are intended only to show that it is not difficult to think of plausible ways of biting in on most of the problems. The really optimum procedure for dealing with each could of course only be determined after much more detailed study, and in some cases it might well turn out that our suggested approach has a total flaw; at the very least, better numerical inputs would be required.

We shall consider here four of the thirteen examples mentioned in Appendix A. These will be, respectively, the ones numbered 1, 8, 9, and 11, there:

(1) When should a university or other research organization split off local departmental libraries from a centralized library?

This is a problem for which the user-time approach is especially well suited. The basic input data would be the amount of time that staff members, graduate students, etc., spend in use of library materials, and the amount of additional time they would need to spend for the same use if the library were in a more remote location. If use were optimum in both locations, these data could be used, following the methods sketched in Section III.4 above, to derive a measure of the loss in productive output that would be entailed by shifting from a nearer to a remote library. This loss could be expressed as a product of number of people involved times salary (or equivalent measure, in the case of students) times a loading factor times a sizable amplification factor arising from the superlinear dependence of productivity on time spent in productive work and from the feedback from library use into other channels of communication. However, a major correction would need to be researched and then inserted: all indications are that use of a remote library will fall far more rapidly with increasing distance than can be explained by the effect of increased travel time on an optimal pattern of time allocation; in other words, use falls far below optimum. Carefully designed studies would be needed to quantify this effect, but would probably be feasible.

(8) When is subsidy of a translation journal justified?

Here an approach based on market data could be immediately useful, though it would be helpful to supplement it with specially designed user-time studies. From the general trend of cost-circulation statistics, one could estimate how much the circulation of any given translation journal would increase if its subscription cost were reduced from the value necessary to cover the entire expense of production to a value near the run-off cost. Thus one could estimate the increase in area under the demand curve, as shown in Fig. 5(c). As explained in Subsection 1 above, this area provides a raw measure of the increase in benefit delivered to the user community, a measure that needs to be increased by a factor slightly >1 to take account of the nourishment of other communication channels which benefit the user community as a whole, but may not benefit the individual buyer or purchasing institution. The production cost to society as a whole would rise only slightly (more copies printed). Thus whether to subsidize or not would depend on whether the computed benefit appears to be sizable or only minor, in comparison with considerations in the "practicality" column of Fig. 3.

(9) When should a scientific society adopt a larger page format for its journals?

Here the user-time approach is the obvious one to be used. One would need to make special studies of reading time required for a given level of comprehension using the two page formats. For literature with many equations and figures, the reduction of the need to leaf back and forth to connect these should lead to quite a measurable difference in efficiency, which could well be of the order of the 10% or even 20% difference sometimes found between different typographies for straight text.[49] Multiplying this percentage by the mean total time spent by all users of the journal in question by their mean salary by a loading factor by an amplification factor (due to superlinear dependence of productivity on working time and to interaction of different communication channels, as discussed in Section III.4 above), one could derive a measure of benefit delivered by the proposed change. This could have to be compared with increases, if any, in productive cost, changeover costs, and effects, positive or negative, on library storage costs, etc.

(11) When should the government subsidize the input of citations onto a tape service produced as an adjunct to an A and I publication?

Though market statistics on citation-index services could be very useful if available, it is likely that at the moment a user-time approach would be best. A few data are already available for some populations on the amount of time spent in the use of existing citation indexes. Such data could be augmented and converted, by methods similar to those sketched in Section III.4 above, into a measure of the benefit that would be lost if users who now have access to a citation index were deprived of it, or of the benefit which would be gained if those lacking such an index were supplied it. As this benefit comes out to be quite large (see the sample numbers in Appendix E), and as existing studies (which should, of course, be extended) show that only a minority of those who could benefit from such a service now have access to it, one could probably set lower bounds to the benefit that would result from marketing of an inexpensive service. Obviously, however, some market research would also be desirable, as well as further research on the interaction of information derived from a citation service with other channels of information flow.

V. RECOMMENDATIONS

What should be done in the immediate future? Any organization that supports or provides an information service has an obligation to know something about its social value and to use this knowledge to guide decisions on the expansion, discontinuance, or modification of the service. However, we shall not attempt to detail here the many types of efforts that diverse governmental and non-governmental organizations could usefully make in this direction. Rather, we shall concentrate on those things that OSIS could usefully perform or support for the purpose either of improving the wisdom with which the bulk of its funds are spent, or of providing to COSATI, and hence to other agencies of the government, guidance on the wise expenditure of their funds. Referring again to Fig. 3, we shall further concentrate on recommendations in the circled areas.

As detailed in Sections III and IV, we feel that in many cases there are prospects for significant progress in the quantification of benefits through studies of how users distribute their time and of how useful information diffuses from its originators to its users. So far these prospects have been very inadequately exploited. So we recommend:

(a) Studies should be undertaken, in greater breadth than depth than in the past, of the time scientists, engineers, practitioners, and administrators devote to the use of various kinds of information services, and of the interconnections by which the various channels for the communication of useful information nourish each other. These studies should be undertaken with adequate awareness of, and even in collaboration with, their use in economic models designed to quantify value [see Recommendation (b) below].

In cases where such studies support the hypothesis that users distribute their time optimally for their self-interest, they can provide a foundation for quantitative estimates of value; in cases where nonoptimal use is uncovered, they may suggest ways of improving productive efficiency through user education.

(b) Further work in economic theory should be supported. Several areas are worthy of attention: the general theory of value of information services, taking account of their interactions; specific mathematical models for particular kinds of services; theory of market response. A prerequisite for all such research, however, should be adequate acquaintance of the analysts with the practical characteristics and patterns of use of the information services to which their work is supposed to apply. In particular, they should be given the best possible access to the information developed in pursuit of Recommendations (a) and (c).

Note especially that the wording of these first recommendations calls for integration of economic theory with data collection. Data collection must be planned to provide the inputs that economic models will need, and modelling must be realistic and not outstrip by too far the collection of solid data.

(c) A systematic monitoring should be undertaken of the markets for scientific and technical journals, secondary publications, products of information-analysis centers, and other information services. Data on prices and circulations, and on their changes with time, should be readily available to decision-makers and to researchers in economic theory. Occasional experiments on response over a wide range of prices, like that mentioned earlier,[27] should be stimulated and encouraged. Regularities in the growth of markets should be studied, and in appropriate cases research should be undertaken on the market response to projected services, especially if these are to be subsidized.

(The simplest sort of consideration of the economics of benefits versus costs leads to the conclusion that subsidy of information services is usually socially justifiable to the extent of input costs:[50] if a service is worth providing at all, the net benefit realized from it is optimized if it can be marketed at, or sometimes below, replication cost. However, the service may not be worth providing at all, and one way of finding out if it is worth providing is to know something about the tail of the demand curve. In most cases it is of dubious practical wisdom to offer a service free or to market it at less than incremental cost; in some cases, pricing so as to recover full costs, though not optimum in terms of the strict benefit-cost difference, may be optimum for practical reasons.)

(d) When plausible opportunities present themselves for studying the influence of various possible support mechanisms on the stability of an information service and on the motivations for good performance by those responsible for it, such opportunities should be followed up.

The data and conclusions resulting from any of the work recommended above should be most energetically publicized, to bring them to the attention of decision-makers in societies and research organizations, as well as in government. In addition, measures might be taken to make large organizations, governmental or nongovernmental, more aware than they now are of the possible utility of studies which they might conduct on their own in the area of cost effectiveness and quantification of value. In particular, we urge:

(e) Acting through COSATI, OSIS should keep other agencies of the Federal government informed of the state of development of techniques for quantification of benefits, and should assist them in making economic decisions on their information programs. It should work for the development and acceptance of enlightened guidelines by which agencies can judge when subsidy of an in-house or extramural information service is appropriate.

While the studies envisioned in Recommendations (a), (b), and (d), and part of (c), can for the most part be conducted extramurally, or perhaps by other government agencies stimulated by OSIS, their intelligent coordination will require considerable expertise on the part of the OSIS staff. Such expertise will be especially important for the leadership role called for in Recommendation (e). We therefore make the final recommendation:

(f) OSIS should try systematically to build up the expertise of its staff in the area of economics of information services, with the goal of having as soon as feasible a senior professional of high ability who will devote the major portion of his energies to this field.

ACKNOWLEDGMENTS

Various portions of this Report have benefitted from conversations we have had with a number of our colleagues, including especially W. J. Baumol, J. K. Galt, W. K. Lowry, A. G. Oettinger, and D. Prives.

APPENDIX A. MAJOR TYPES OF INFORMATION ACTIVITIES AND TYPCIAL DECISIONS REQUIRING QUANTITATIVE ESTIMATES OF VALUE

A nice brief summary of the scientific and technical information activities of the Federal government can be found in the periodic NSF publications Federal Funds for Research, Development, and Other Scientific Activities.[51] However, as the caveats in these publications make clear, the expenditures they report are not a complete listing, and the total is only of the order of half the billion dollars a year we quoted in the Introduction. In Fig. 1 of the text and in the discussion below we have broken the activities down into categories of our own making, and have made rough estimates of expenditures from all sources available to us. The categories of Fig. 1 are:

A. Technical libraries. As might be expected from their indispensible role as purveyors of recorded information, libraries loom very large in economic terms. According to the best estimates we have been able to get, the Federal government's expenditures for technical libraries are currently something like $100 million per year; although comprehensive data are hard to obtain, it seems likely that the operation of the scientific and technical component of our nation's libraries involves additional annual expenditures of hundreds of millions of dollars by universities, industries, etc. Some sample decision questions are:

1. When should a university or other research organization split off local departmental libraries from a centralized library?
2. What is the optimum number of subscriptions to an important journal which a library should purchase?
3. How much is it worth to support research and development for better and cheaper microform readers?

B. <u>Systems for selective dissemination of information and computerized services</u>. These are coming increasingly into use in industrial and mission-oriented organizations.[52,53] While we have not seen any breakout of expenditures for such services, we have inferred a lower limit of some tens of millions of dollars a year by the government alone. Especially important are a number of university-centered retrieval systems supported by the NSF and capable of making available to users the content of tapes issued by the producers of large A and I services. Typical questions are:

4. When should a research organization put into operation a selective-dissemination system for its staff?
5. How large a data base should a computerized retrieval system attempt to cover?

C. Information-analysis centers. These are lumped together in Fig. 1 with information service centers (e.g., weather data centers), the dividing line between the two being a little fuzzy. The two together probably account for something like $200 million of Federal expenditures, but probably less than half of this amount is spent on the sort of processing activities envisioned in the Weinberg Report[54] under the term "Specialized information centers." The number of centers of this latter type has, however, been growing steadily; these are included in COSATI's 1968 list of some 111 centers supported by the Federal government;[55] over three-quarters of these were devoted to scientific or technological areas. Much of the most fundamental data analysis in the basic sciences is done in centers supported by the Office of Standard Reference Data in the National Bureau of Standards, which has an annual budget of about $2.4 million. A typical question is:

6. When is it worthwhile to start an information-analysis center in a new field, or to significantly expand an existing center?

D. Journal production. Although no comprehensive compilation exists, available evidence[56,14] suggests that primary publication in the United States of basically new scientific and technical knowledge amounts to over a million kilowords annually. The corresponding expenditure by the publishing organizations must therefore be in the range $50 to $100 million. While direct government support is limited to journals published by government organizations and emergency subventions, such as those for the initiation of translation journals, a sizable and growing proportion of the production cost of regular journals is met by page charges, the majority of which are paid for from research funds supplied by government agencies. Some typical decision questions are:

7. To what extent should the government encourage page-charge financing by providing funds for it in grants and contracts?
8. When is subsidy of a translation journal justified?
9. When should a scientific society adopt a larger page format for its journals?

E. <u>Schemes for distribution for reports and preprints</u>. The National Technical Information Service in 1969 offered some 45,000 reports for distribution.[57] A still larger number of government reports were not made available through this service, but distributed only by the originating agencies. It is probable that a sizable part of the roughly $100 million reported in <u>Federal Funds</u>[51] as spent annually for "publication and distribution" pertains to reports. In a different vein, the notorious Information Exchange Groups operated by the National Institutes of Health sent out in 1966 about one and one-half million copies of preprints. Before the experiment was discontinued, possible expansion to an operation with annual costs in the range $10 to $100 million were being talked about.[51] A typical question might be:

10. When, and on what scale, should preprint exchanges be subsidized?

F. Abstracting and indexing publications. Scores of these are produced directly by the government, most of them rather small but some quite large; the latter are exemplified by Index Medicus, Nuclear Science Abstracts, and STAR. Adding up some rather old data on number of literature items covered by these, we get over half a million for the year 1960; this probably implies well over a million today. There is also sometimes ongoing subsidy of externally produced services, typified by that of International Aerospace Abstracts. OSIS has in the past given considerable developmental support to society-run discipline-oriented systems. The government's expenditure for all these items could well exceed $20 million a year; nongovernmental expenditures are doubtless considerably greater. Some typical decision questions are:

11. When should the government subsidize the input of citations onto a tape service produced as an adjunct to an A and I publication?
12. When is direct or indirect subsidy of an A and I publication justified?

13. Should an abstract journal computerize its annual-index production to enable the index to be produced more promptly?

G. Review grants, etc. In contrast to the extensive work at information-analysis centers, there has been very little investment of resources by government or societies for the purpose of encouraging individuals to produce reviews, treatises, and data compilations. A key question is:

14. Should an agency supporting research divert some of its research funds to a program of grants to authors for the preparation of reviews and compilations?

H. Educational resources. This is a very large area, and one in which the quantification of benefits is even more difficult than for most other types of information services. We shall not attempt to give examples for this area.

I. Meeting support. Most of the support of scientific and technical meetings is supplied by research and development organizations, through payment of registration fees and, most importantly, of travel expenses for attendees. As the total investment of resources has been estimated at something like $8 or $10 million for the field of physics,[28] the total for all fields may well be of the order of $100 million; explicitly

APPENDIX B. DEMAND CURVES, AND THEIR RELATION TO BENEFITS DELIVERED

Figure B1(a) shows a typical demand curve, in which the abscissa is the number of purchasers of a product or service when it is offered at a price represented by the ordinant. We assume for the moment that the price is the same to all buyers, and that the social value delivered to any buyer is, at least on the average, equal to the self-interest value he perceives for the product, in other words, to the maximum price he would be willing to pay for it. The width of the doubly shaded strip represents the number of buyers to whom the value is greater than p_1 but less than the slightly larger value p_2. The area of this strip thus represents -- under our simplifying assumptions -- the value this group of buyers receives from the product (provided the price is below p_1). The total value received by all those who purchase the journal when it is priced at p_o is obtained by adding the contributions from many such vertical strips and is, therefore, the total shaded area:

$$\text{Benefit delivered} = \int_{o}^{N(p_o)} p(N)\varphi N = p_o N(p_o) + \int_{p_o}^{\infty} N(p)\varphi p \qquad (B.1)$$

This is a gross benefit. If the cost of providing the product or service to N buyers has the simple form

$$\text{Cost} = C_o + RN ,$$

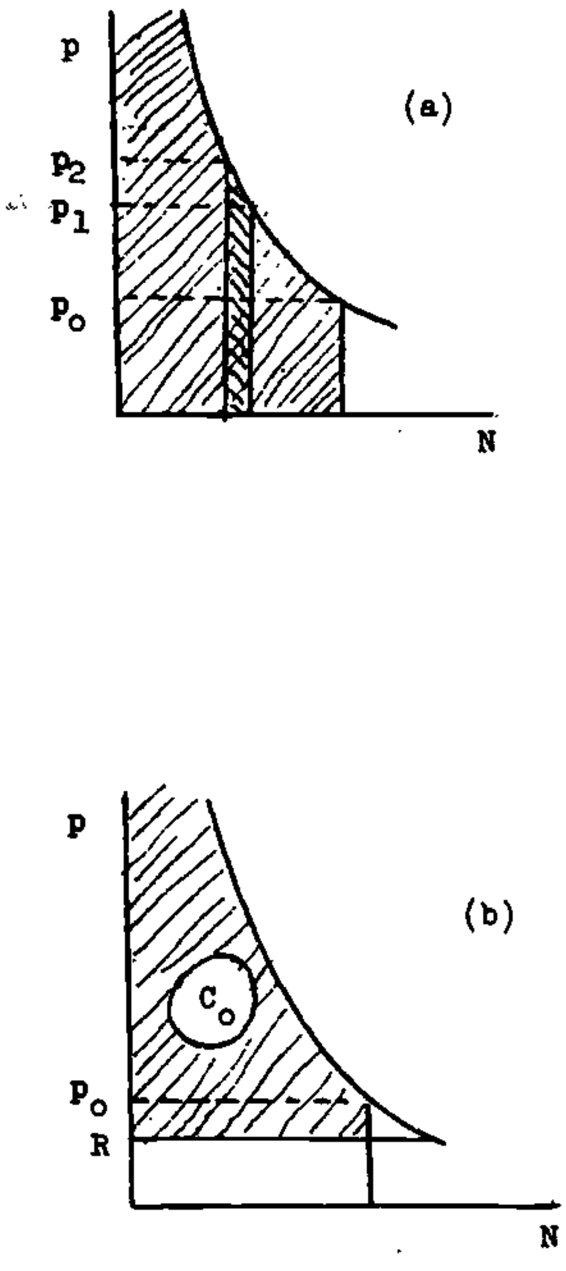

Fig. B.1

where C_o is the "input" or "set-up" cost, and R is the unit replication cost, then the net benefit is given by the shaded area in Fig. B1(b), where the circular hole is chosen to have area C_o.

The logarithmic derivative of the demand curve, $\beta = \partial \ell n N / \partial \ell n p$, is what is called the price elasticity of demand. It is normally negative, and under usual operating conditions has a magnitude that is usually a fraction of unity. If the price elasticity of demand were constant over the whole curve then we should have $N \propto p^{\beta}$. This is, however, impossible, for $\beta > -1$, since it causes $\int^{\infty} N dp$ to diverge.

In practice this simple measure of benefits will need to be corrected both for "externalities" (positive or negative benefits realized by society through providing the product or service to a particular buyer, which do not enter into this buyer's self-interest judgments), and for the interaction of competing but usually not identical products in the marketplace. The former of these has been briefly discussed in the SATCOM Task Group Report,[46] and the Physics Survey Report;[59] the Mathematics Report[7] discusses both, and gives fairly detailed attention to the latter in connection with the optimal distribution of subventions.

APPENDIX C. VALUE OF AN ENTERPRISE AND OF AN IMPROVEMENT IN IT

Let a society, or some organization within it, spend F dollars a year for some enterprise which yields a benefit (yield) judged to be worth Y dollars, i.e., of a value equal to what Y dollars could buy if used for typical other purposes. In general the dependence of Y on F will have the form shown in Fig. C.1: at low F, the enterprise will usually be inefficient; its efficiency will at first rise as F is increased; for large F a region of diminishing returns will be reached, and Y will grow more slowly. If the enterprise

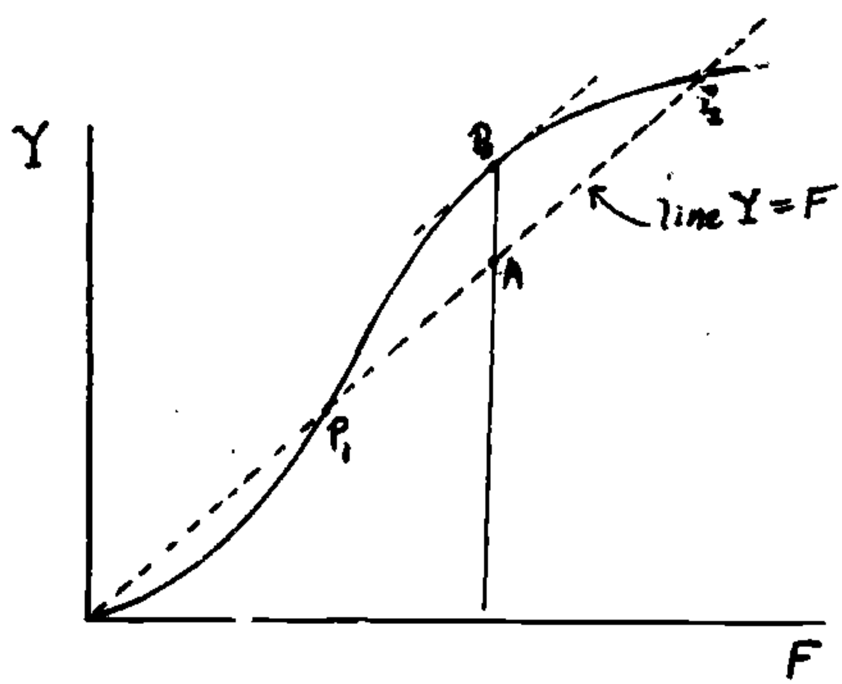

Fig. C.1

consumes only a small fraction of the funds available to the organization, the optimum operating point will be that maximizing the distance above the line Y=F, i.e., the point P_o at which $\partial Y/\partial F = 1$. Operation at any point between P_1 and P_2 will give a net benefit, however. If the full curve never rises above the 45° line, the enterprise is not worth pursuing at all.

Now consider a hypothetical change in the enterprise, improving its efficiency or productivity. This will change the function Y(F), say from the full to the dot-dash curve in Fig. C.2.

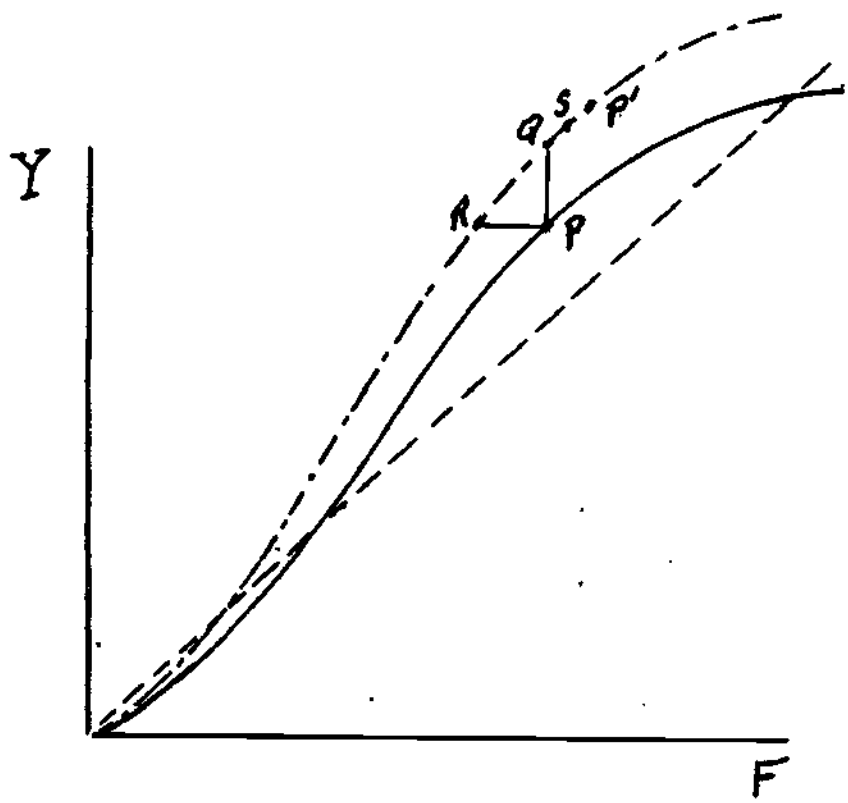

Fig. C.2

Let the optimum operating point (point of unit slope) under the old regime be P, that under the new regime P′. If the change is small, the difference in the heights of P′ and P above the 45° line can be evaluated either as the vertical shift PQ$\equiv (\delta Y)_F$ or as the horizontal shift PR$\equiv -(\delta F)_Y$. Often, however, what is easiest to estimate is not either of these two quantities, but the increase in yield that would result under the new regime if the same professional staff were to be used in the new regime as in the old, with the same standard of support, i.e., the same availability of supplies, secretarial help, etc., even though expenditures for these items might rise in consequence of their working more productively. Let S in the figure be such a hypothetical operating point. (Such an estimate is especially natural if the value of the innovation can be expressed in terms of an equivalent augmention in man-hours worked.) In such case the known quantity is the vertical distance from P to S, and the desired net benefit is this vertical distance minus the horizontal distance from P. to S.

To compute the quantity just mentioned, let us try to divide F into two terms:

$F_w \equiv$ those expenditures that depend mainly on the amount of useful work done, rather than on the number of professionals doing it;

$F_p \equiv$ those expenditures that depend mainly on the number of professionals employed, rather than on their productivity.

Then we have $F = F_w + F_p$, with $F_w \propto Y$, $F_p \propto$ size of professional staff. The horizontal distance from P to S in Fig. C.2 will then be, to first order,

$$\delta F_w \equiv F_w(S) - F_w(P) = F_w \frac{\delta_o Y}{Y},$$

where $\delta_o Y$ is $Y(S) - Y(P)$. Thus the increase in net benefit under the new regime is

$$\delta(Y-F) = \delta_o Y - \delta F_w = \delta_o Y(1 - \frac{F_w}{Y}) > \delta_o Y(\frac{F-F_w}{Y}) = \delta_o Y \frac{F_p}{Y}, \quad \text{(C.1)}$$

since $Y>F$ if the enterprise was worth pursuing in the first place.

Studies of budget items in a typical industrial research and development organization, quoted without details in Ref. 28, have indicated a value for F_p equal to about 1.55 times the bare salary budget for professional staff. We have used this value in the explicit examples of Appendix E. As F_w is normally rather less than F_p, the inequality in (C.1) will not be a very strong one unless Y/F is considerably above unity. But for enterprises where the latter is the case one should try to estimate the value of Y/F in order to improve on the lower limit given by the far right of (C.1).

APPENDIX D. EXAMPLES OF HOW USER-TIME DATA CAN BE CONVERTED TO A MEASURE OF BENEFIT

We shall use the basic assumptions described in Section III.4 of the text, namely, that on the average each user of an information service allots to its use the amount of time which, for fixed characteristics of the environment with which he interacts, maximizes his productive output, and that this productive output has the form of Eq. (6), i.e.,

$$Y = \varphi(1-t_1-t_o)\,E\,(t_1,t_o)\ , \qquad \text{(D.1)}$$

where t_1 is the fraction of his time he devotes to the information service in question, t_o is the fraction of time devoted to other not directly productive activities, and E is an efficiency factor. The function φ, which measures the output of productive work as a function of the time x directly devoted to such work, has the general form shown below:

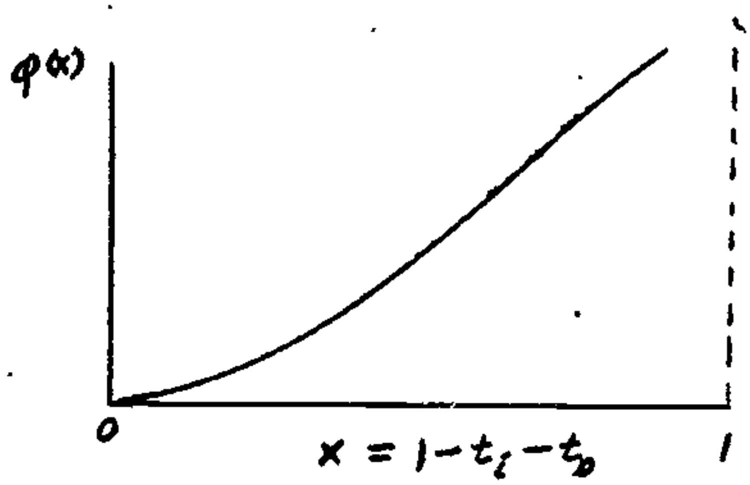

Fig. D.1

For scientific and highly technical work it has a significant upward curvature, because of the need to have many different strands of thought simultaneously in the forefront of consciousness, and because of delays attendant on turn-on and maintenance of apparatus. We have sampled opinions from a number of research physicists, for example, which suggest that the difference in slope between, say, $x = 0.6$ and $x = 0.1$, is probably rather more than a factor of 2. Individual differences may be sizable, however. (We shall return to the concept of φ at the end of this Appendix.)

A correct description of the function E should take account of a number of effects. To begin with, its dependence on t_i is usually very nonlinear, because the user seeks first those pieces of information with the highest anticipated relevance for his work, and then works down to less and less promising items as t_i increases. Moreover, in many cases increasing t_i will produce an increasing redundance of information, the same piece of knowledge being repeatedly encountered in trivially different forms. A further complication has to do with the interaction of different information channels: for example, E can be strongly influenced by the amount of time t_g spent interacting with the "grapevine" of informal contacts; the effectiveness of these contacts, in turn, depends on their interaction with written sources of information; etc. To sort out these various interrelationships one can make use of several types of empirical observations, such as: the relative amounts of time that particular types of users choose to spend in the use of different types of information resources, including the grapevine; the relative frequency of various sequences of information-transfer operations in the acquisition of items of information actually found useful; order-of-magnitude guesses as to the seriousness of the redundancy and diminishing-returns effects. Later below we shall present, mainly for illustrative purposes, some calculations with a rather specific model designed to take acoount of these effects.[29] First, however, we shall illustrate a few particularly important principles by writing a few equations for a deliberately over-simplified model.

Both in our initial over-simplified discussion and in the analysis of more sophisticated models, we shall restrict attention to two specific types of change in an information service:

(i) A change that will enable each user, on the average, to get the same amount of useful information from the service in $100\lambda\%$ less time.

(ii) A change that will enlarge the supply of information from service i from which the user may select items for perusal or study, under the assumption that the additional items, whose number is $100\lambda\%$ of the original number, have the same distribution in quality and appropriateness for the user's needs as those originally available.

In our initial discussion we shall make the simplification of assuming that when either of these types of change is made in a particular information service i, neither the time devoted to use of other information resources, nor the amount of useful information obtained from them, changes. In such case we can just take for E some function of the general shape shown below, and can ignore its dependence on the other times, such as t_g, entering into t_o.

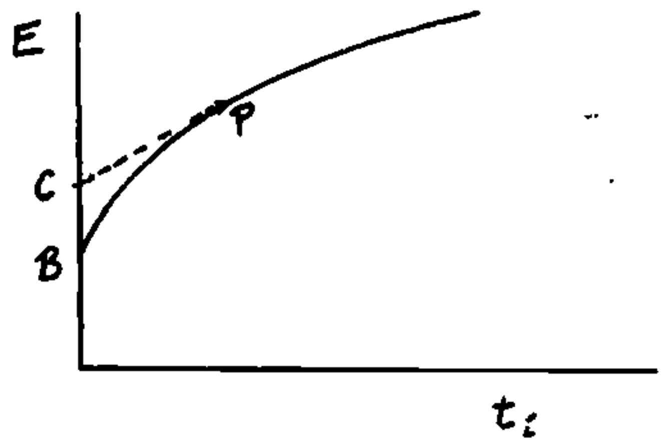

Fig. D.2

The optimality condition, (7) of the text, is thus

$$\frac{\varphi'(1-t_i-t_o)}{\varphi(1-t_i-t_o)} = \frac{E'(t_i)}{E(t_i)} . \qquad (D.2)$$

This already gives us some information about the form of the function E, if t_i and t_o are empirically known and if we can regard φ as approximately known. (There is, of course, an arbitrariness in the normalization of the functions φ and E: either could be multiplied by a constant and the other divided by the same constant; for our purposes we can leave the normalization unspecified.) If we have further information about the relative amounts of useful knowledge acquired through channel i and through all other channels, we can characterize the function E even further, since we will know approximately how far the ordinate of the present operating point P in the figure lies above the intercept B.

Consider first, for this simplified model, the effect of the time-saving measure described in (i) above. Such a measure replaces the function E by a function $\bar{E}$, where

$$\text{Case (i):} \quad \bar{E}(t_i) = E(t_i+\lambda t_i) \ , \qquad (D.3)$$

so that the change in E is, to first order,

$$\delta E \equiv \bar{E}-E \simeq \lambda t_i E'(t_i) \ . \qquad (D.4)$$

From (D.1) and (D.2) we then have

$$\text{Case (i):} \quad \frac{\delta Y}{Y} = \frac{\delta E}{E} = \lambda t_i \frac{\varphi'}{\varphi} \ . \qquad (D.5)$$

Note that we do not need to worry about the change δt_i that will occur in t_i as a result of the innovation, because we have assumed Y to be stationary with respect to changes in

t_1, this being in fact the origin of Eq. (D.2). Now the value of φ'/φ at a typical research-worker's operating point is likely to be in the range 2 to 3: even if one were to ignore the rather appreciable upward curvature of the plot of Fig. D.1, setting $\varphi \propto 1-t_1-t_o$, we would still have $\varphi'/\varphi = (1-t_1-t_o)^{-1}$, which is typically 1.5 to 2 or more. Thus we have our first major conclusion, which as we shall see below remains valid in more sophisticated models: <u>the value of a time-saving innovation is enhanced over the value of the time saved, by an amplification factor of the order of 2 or 3, due simply to the form of the productivity function φ</u>.

Now let us turn to the other type of innovation, called (ii) above, namely, an augmentation of the amount of information available for perusal or study. The assumption enunciated above for this case amounts to replacing E by $\bar{E}$ where

$$\text{Case (ii):} \quad \bar{E}'(t_1) = E'\left(\frac{t_1}{1+\lambda}\right) . \qquad \text{(D.6)}$$

Thus to first order

$$\delta E' \equiv \bar{E}'-E' = -\lambda t_1 E'' , \qquad \text{(D.7)}$$

so that, by integration

$$\delta E = \int_o^{t_1} \delta E' dt_1 = \lambda \left[E(t_1)-E(o) - t_1 E'(t_1)\right] \qquad \text{(D.8)}$$

The resulting fractional increase in output is thus

$$\text{Case (ii):} \quad \frac{\delta Y}{Y} = \frac{\delta E}{E} = \lambda\left[\frac{E-E(o)-t_i E'}{E}\right] = \lambda \frac{\Delta}{E}, \qquad (D.9)$$

where Δ is the distance from the intercept B in Fig. D.2 to the intercept C of the tangent. Unlike (D.5), which depended only on the form of the approximately understood function φ, (D.9) depends sensitively on the details of the function E; to evaluate the benefit from a type (ii) innovation one must therefore conduct studies adequate to map out a rough shape for E.

Our discussion so far has neglected two effects, one, at least, of which can be fairly important, in the direction of enhancing the benefit of improvements in information services. The two effects are:

(a) Different information channels can nourish each other. Improvement in a particular information service i affects the efficiency E of a given individual not only via the direct route we have been considering (increasing the amount of information he can get for a given expenditure t_i of his own time), but also by increasing the amount or value of the information he receives through other channels (since the individuals with whom he interacts in these other channels will themselves be better informed as a result of the improvement in channel 1).

(b) The time distribution that individuals choose to make between use of the literature, use of the grapevine, and other activities will not in general be such as to optimize the total social yield from all individuals together, even though each individual undertakes to optimize his own output. Specifically, an individual will not redistribute his own time between these various types of activities if such redistribution will not improve his own output, even though it might improve the output of others. (This statement may be a little extreme, but there certainly is an effect of this sort.)

In a study[29] which will not be described in detail here, calculations have been made for a few quantitative models which embody allowance for these two effects, and for the fact that information received through different channels may be redundant. For a considerable range of certain parameters entering these models, the following statements seem to hold true: the benefit received from a time-saving innovation in a formal information channel i is enhanced by a further amplification factor, beyond that given in (D.5), due to nourishment of informal by formal channels; if comparable amounts of information are received through formal and informal channels, this enhancement factor is typically the order of 1.5, and it increases slowly with increase in the relative amount received from the informal channels. The effects of the redistribution of users' time produced by the innovation seem in nearly all cases to be minor.

In the more sophisticated models the effect of an augmentation of available information, case (ii), continues to be rather more model-sensitive than in case (i). The effects of mutual nourishment of different channels, and of their mutual redundancy, may either increase or decrease the benefit in comparison with that given by (D.6), depending on the assumed values of various parameters entering the model. Here again, however, the effects of readjustment in the distribution of users' time are usually minor.

In closing we shall indicate, again for an over-simplified model, why it is that mutual nourishment of different information channels produces an amplification of benefit, as mentioned in the underlined sentence above. Suppose information is received through two channels, "literature" (fraction t_ℓ of time) and "grapevine" (fraction t_g), and suppose there is no redundancy, so that

$$E(t_\ell, t_g) = e_\ell(t_\ell) + e_g(t_g) \ . \qquad \text{(D.10)}$$

If the individuals with whom the typical worker interacts in the grapevine are like himself, and get information from similar sources, it will be reasonable to suppose that

$$e_g(t_g) = f(t_g)E \ , \qquad \text{(D.11)}$$

where f is some increasing function which tends to saturate for large t_g, and that each individual will try to maximize his own e_g by choosing an appropriate t_g in $f(t_g)$ treating the E of (D.11) as a constant [this is the source of the effect (b) noted above]. From (D.10) and (D.11) we have

$$E = \frac{e_\ell(t_\ell)}{1-f(t_g)} \ , \qquad \text{(D.12)}$$

and from (D.1),

$$Y = \frac{\varphi(1-t_\ell-t_g)\, e_\ell(t_\ell)}{1-f} \ . \qquad \text{(D.13)}$$

Thus if one ignores changes in t_g and t_ℓ,

$$\frac{\delta Y}{Y} = \frac{\delta e_\ell}{e_\ell} \ . \qquad \text{(D.14)}$$

Since the optimality condition re t_ℓ is

$$\frac{e_\ell'}{e_\ell+e_g} = \frac{\varphi'}{\varphi} , \qquad (D.15)$$

we have for $i=\ell$ in case (i), by (D.4) with e_ℓ replacing E,

$$\frac{\delta Y}{Y} = \lambda t_\ell \left(\frac{e_\ell+e_g}{e_\ell}\right) \frac{\varphi'}{\varphi} , \qquad (D.16)$$

which differs from (D.5) by the "mutual-nourishment" amplification factor $(e_\ell+e_g)/e_\ell$. What this means is that an innovation saving time in use of the literature makes everyone not only get more information from the literature, but also deliver more information via the grapevine. As noted above, allowance for redundancy decreases the amplification effect: typically, if $e_\ell \approx e_g$, the factor $(e_\ell+e_g)/e_\ell \approx 2$ in (D.16) becomes about 1.5 in the more sophisticated models.

The importance of the function $\varphi(x)$ of Fig. D.1 for this analysis suggests that it would be worthwhile to give some attention to techniques for estimating it, and to ways of generalization that has been suggested is based on the conception that it is intellectual energy, rather than time, that is the scarce commodity, and that time spent in different types of activities uses up this energy at different rates. This one might replace the $\varphi(1-t_1-t_o)$ in (D.1) by $\varphi(1-\beta_1 t_1-\beta_o t_o)$, where β_1 is the ratio of the energy demand per unit time while using information resource i to that while doing typical productive work, and β_o is a similar average ratio for the remaining activities. With

$$1-\beta_i t_i-\beta_o t_o = \beta_i \left[1-t_i-\frac{\beta_o}{\beta_i} t_o + \left(\frac{1}{\beta_i}-1\right)\right]$$

we could write

$$\varphi(1-\beta_i t_i-\beta_o t_o) = \chi_i(1-t_i-\frac{\beta_o}{\beta_i} t_o) , \qquad (D.17)$$

where χ_i is some function of its argument. Our analysis through (D.9) could now be repeated with no change except replacement on φ by χ_i. But if rates of different channels were to be compares, as in the equations (D.10)-(D.16) for "literature" plus "grapevine" channels, one would have to take account of the dependence of χ_i on the channel i. Unfortunately, while time spent in doing various types of things can be directly measured, expenditures of intellectual energy cannot.

APPENDIX E. POSSIBLE PROCEDURES FOR BENEFIT QUANTIFICATION IN TYPICAL SPECIFIC CASES

The brief discussions below, for several of the items on the list of Appendix A, are intended to show how one might go about getting a quantitative value measure. Any numbers quoted are only illustrative, and make no pretense to authority or correctness; the important point is rather that procedures can be identified by which roughly correct numbers could be obtained with a little work. Moreover, it must be remembered that the measure of value we are discussing here is only one of several factors that enter into the decision on any item: one must also consider such things as costs, economic stability, and organizational or political practicality. The numbers on the items correspond to those of Appendix A.

1. Replacement of a centralized library by local libraries.

Various studies[47] have shown that the use made of any information resource decreases rapidly with increasing distance from the user or inconvenience to him. In part this decrease is logical, since increase in time required will change the optimum division of the user's time between use of this resource and other activities (see the study[29] briefly summarized in Appendix D, for an attempt at mathematical evaluation of this effect). But the observed decrease is so rapid that one must conclude that most users at some point fail to optimize their time distribution. This introduces a worrisome element into the process of making dollar calibration from users' time studies; further investigations are needed to clarify the nature of the misjudgments involved. For the present discussion we shall assume that such studies would show that use of a resource such as a library becomes more nearly optimum the less the delay and inconvenience involved, so that increasing delay, etc., leads to sub-optimum use.

Let us consider, then, that users who have a departmental library in their own building make optimal use of it; the fraction t_L of their time which they devote to such use can be measured in similar institutions where departmental libraries already exist. For our illustrative calculation, we shall assume $t_L = 0.007$, a figure that might be reasonable for physicists, though too low for chemists. The difference in commuting time between use of a central library and use of the departmental library can be measured: suppose that in a given case it amounts to 10 minutes per round trip. The average duration of a library visit can also be measured: let us assume it to be 50 minutes in the present exercise. Then if the users were to optimize their distribution of time in both cases, the shift from the one library system to the other might be treated as a 20% change in the time required to get a given amount of information from the library-visit channel [like "case (i)" of Appendix D, but for a model isolating library visits from other types of literature use, including phoned requests to the library, etc.]. As in Appendix D, the effect of this 20% change will be equivalent to the value of this time times an amplification factor which can be at least roughly computed from obtainable inputs, and which we shall take to be 3. Thus for a physics research group, with a loading factor 1.55 applied to bare salaries to take account of nonproductivity-related expenses incurred through employment of professionals (see Appendix C),

$$\begin{pmatrix}\text{Excess value received}\\ \text{from departmental}\\ \text{library per year}\end{pmatrix} \approx 0.007 \times 0.2 \times 3 \times 1.55 \times \begin{pmatrix}\text{total}\\ \text{staff}\\ \text{salaries}\end{pmatrix}$$

$$= .0065 \times \begin{pmatrix}\text{staff}\\ \text{size}\end{pmatrix} \times \begin{pmatrix}\text{mean}\\ \text{salary}\end{pmatrix}$$

For a staff of 60 this would be about $6000 per year.

But this is on the assumption that use of the more remote library would be optimal, and as noted above, it probably falls far below optimal. Values of t_L for groups with remote libraries can be obtained: suppose the value of t_L were 0.002. All indications are that the function describing the dependence on t_L of the utility of the information received from library use is a rapidly saturating one, similar to the $E(t_1)$ of Fig. D.2 in Appendix D. Thus:

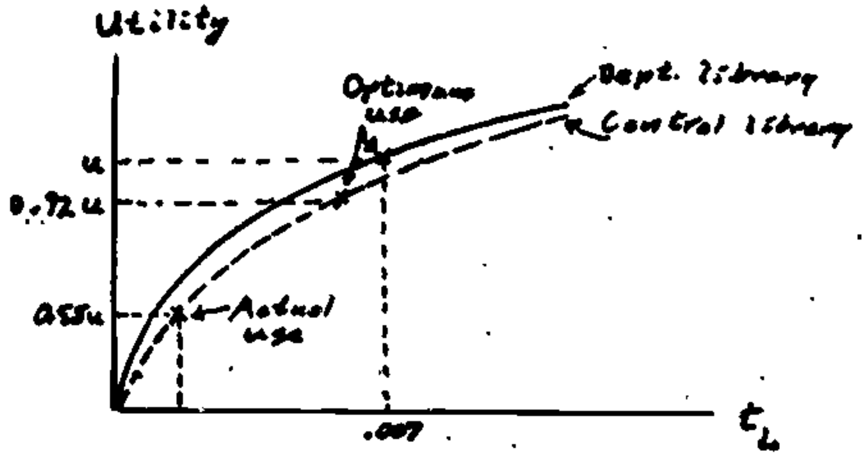

Fig. E.1.

The optimum-use point for the central library would be the point at which the slope of the dashed curve is the same as that of the full curve at $t_L = .007$ and might correspond to a t_L only slightly smaller. For a plausible set of parameters (appropriate parameters should of course be sought in a special study) this new optimum point might have 92% as much utility as for the departmental library, as shown. But the actual use $t_L = 0.002$ might have only 55% as much utility. In such case the dollar value difference between the two library systems would be about five times the $6000/year figure quoted above.

The $30 K figure would probably exceed the total annual acquisition cost of a physics library, and certainly would far exceed the differential cost of acquisitions that would need to be duplicated if departmental libraries were set up. The $6 K figure, on the other hand, might be of the same order as the latter. Putting in space and staff expenses, etc., one could probably derive reasonable criteria for when (in terms of size of staff, etc.) individual chemistry, physics, mathematics, etc. libraries should be set up, when libraries for several departments should be consolidated, etc.

Our simple discussion here has not taken any account of students, a very important consideration in university affairs; they could, of course, be put into the picture with similar techniques.

8. Subsidy of translation journals.

This could benefit society in two ways: by increasing the availability of existing translation journals, and by bringing new ones into existence. Here we shall discuss only the former channel. As in most cases, dollar calibration for this problem can be obtained either from market data or from user-time studies; as neither approach can be made very accurate, it would be best to do one's best with both approaches and compare the results.

At the moment the needed data are more nearly available from the market approach. Let us take as an example a particular translation journal, Soviet Physics JETP, published by the American Institute of Physics. According to the SATCOM study,[14] the 1968 circulation of this journal was 1611 and its subscription cost, per equivalent kiloword of text, 3.4¢. From the general trend of cost-circulation statistics for many physics journals (Ref.14, Fig. 7) it can be estimated that if the journal could be marketed at near run-off cost (about 0.3-0.5¢ per equivalent kiloword, according to Fig. 11 of

Ref. 14) it would, in the quasi-steady state, have about 2800 subscribers. The increase in area under the demand curve -- see Fig. B1 (b) above -- would be roughly that of the triangle of height 1200 and width 3¢/kiloword, or $18 per kiloword. This, then, is an estimate of the benefit society in general (including foreign countries) would receive if support of prerun costs by subsidy were to make it possible to market this journal at run-off cost. This is a minor, though appreciable, fraction of the total prerun cost (presumably about $45 per kiloword, the expense of translation being compensated by the saving in photo-copying equations, etc.), i.e., of the amount of subsidy that would be involved. Thus society would be getting a 63/45 return on its subsidy money.

As the $18 figure just estimated could be in error by a factor of two in either direction because of the uncertainty in the elasticity of the demand curve, and because the market approach has other shortcomings, studies aimed at dollar calibration via users' time would be worth performing. Reducing the subscription price to the run-off level would have two kinds of effect: it would make the journal available in institutions where it is not now available, and it would make it more easily available in other institutions via sub-libraries, individual subscriptions, etc. In the former case the effect would be like that of "case (ii)" of Appendix D and the contribution to the value could be estimated by the methods described there, if market data on the number of new institutions subscribing could be obtained. In the latter (probably more important) case the effect would be roughly one of decreasing the time involved in use of the translation journal, and could be estimated if data could be collected on the amounts of time spent in use of this journal (as compared, say, with a more widely available one like Physical Review) in institutions having it in many sub-libraries and in those having it only in a central library; the mathematical techniques would be similar to those of Appendix D and Ref. 29.

9. Adoption of a larger page format by a journal.

This issue is a particularly simple one for which to get a measure of value. With two groups of subjects one could compare reading time required for a given level of comprehension with, say, a 500-word journal page and a 1300-word page. For literature with many equations and figures, the reduction of the need to leaf back and forth to connect these should lead to quite a measurable difference in efficiency. While we do not know how much difference in speed of comprehension would be found in such a study, the fact that differences in typography are known to affect reading speed for non-technical material by 10% or even 20%[49] suggests that a 10% figure might well be found for the present case. If so, the value could be computed by the methods of Appendix D, "case (i)", i.e.,

$$\text{Value} = 0.1 \begin{pmatrix}\text{Mean time spent}\\ \text{by all users in}\\ \text{reading given}\\ \text{journal}\end{pmatrix} \times \begin{pmatrix}\text{Mean}\\ \text{salary}\end{pmatrix} \times 1.55 \times \begin{pmatrix}\text{Amplifi-}\\ \text{cation}\\ \text{factor}\\ \text{of 3 or so}\end{pmatrix}$$

where the factor 1.55 takes account of loading as described above and in Appendix C, and the last factor is due to the increase of efficiency with increasing time worked and to feedback effects, as discussed above and in Appendix D.

11. Support of citation indexes.

Let us consider the special case of a citation index for chemistry, supposed in existence but very costly. Should such an index be operated on a self-supporting basis by sale of subscriptions, or should it be socially subsidized, so that it can be marketed at something approaching run-off cost? Here again one could in principle use either a demand-curve or a user-time approach; however, statistics on the demand curves for citation indexes are almost non-existent, so our remarks here will be concerned only with the user-time approach.

According to some statistics collected by a committee of the American Chemical Society,[15] industrial research chemists who have access to the Science Citation Index spend an average of about 0.2 hours per week using it. [One can question the methodology used in collecting these statistics -- a more controlled study is needed if such figures are to be used for an important application. Also, the data are doubtless not representative of steady-state user judgments, as SCI is still new to many potential users.] But only 20% of the chemists queried had such access. From the nature of the sample it seems likely that most of the remaining 80% would have found SCI comparably useful if they had had it; statistics on chemical publications, patents, etc., from different kinds of institutions could probably be collected which would give an estimate, or at least a decent lower limit, for the effective number of chemists who do not now have access to SCI but who could benefit comparably from it. To combine with such data one would need only to know the dollar value of the loss, per chemist, that the average present user of SCI would suffer if he were deprived of it.

In a crude model, one could ignore the "grapevine" effects discussed in Appendix D, and simply look at this average chemist as an individual. The benefit he receives from spending a fraction t_i of his time with SCI will depend on t_i according to a saturating curve of the form of Fig. D.2 of Appendix D, i.e.,

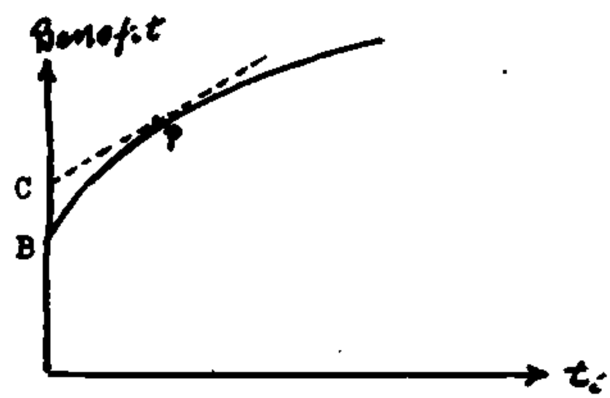

Fig. E.2

If he uses good judgment, he will pick his operating point P to be such that the slope of this curve represents the marginal value of his time when devoted to other uses. Since t_i is very small ($\approx 0.2/40 = .005$ in the present example), this marginal value can be taken as unaffected by t_i; in such case the net loss to him if he is deprived of SCI will be the distance BC. To get a numerical value for this one must know something about the curvature of the curve. A rough guess would set BC ≈ the vertical distance from C to P; this would be the case if the curve were parabolic with a slope at A much greater than that at P. Better guides to the shape of the curve could be obtained from studies of search yields vs. time (see for example Ref. 60), time spent with SCI when it is inconveniently available, etc. (Another estimate, though probably not a very reliable one, can be obtained from the estimate in Ref. 15 that the average user of SCI saved 0.4 hours/week by such use; this means BC = twice the vertical distance from C to P.)

With the above guess BC ≈ vertical CP, the value .005 for t_1, and the slope CP ≈ 2.0 × 1.55 × $15K/man yr. as estimated from the discussion of the φ function in Appendix D, one would get a value of about $230 per chemist year, resulting from supplying him with a citation index. One would need to combine such a figure, suitably sharpened, with an estimate of the number of chemists in the country, not now able to use SCI, whose use patterns might be comparable with those of the ACS sample. One should also check the assumption that nearly all of these would indeed have access to a citation index marketed at near run-off cost.

It is noteworthy that the benefit figure just estimated is so large that, even if it is an overestimate by a factor three, any company with more than 30 or 35 research-level chemists should purchase SCI at its current price of about $2500 a year. As the great majority of the chemists in the ACS study were employed by companies with over $100 million annual sales, while only 20% of them had access to SCI, a serious hysteresis in market response is indicated.

REFERENCES

1. For example, a 1966 Task Force Report on National Systems for Scientific and Technical Information, prepared by OST, states that "... the Federal Government invests about $400 million per year directly in activities designed to make scientific and technical information easily available. About another $600 million in Federal funds is spent indirectly for this purpose -- mostly by contractors and scientists who spend Federal funds to purchase needed data."
2. A detailed study, not yet available, has just been prepared for COSATI on Federal expenditures for scientific and technical information.
3. D. W. King and E. C. Bryant, The Evaluation of Information Services and Products (Information Resources Press, Washington 1971).
4. J. W. Murdock and M. Liston, Jr., Am. Doc. 18, 197 (1967).
5. F. Machlup, The Production and Distribution of Knowledge in the United States (Princeton University Press 1962).
6. J. Marschak, Am. Econ. Rev. 58, Papers & Proc. Sec. p. 1 (1968).
7. A Cost Benefit Approach to Evaluation of Alternative Information Procedures (Report to OSIS on Contract NSF-C606, Mathematica, Inc. 1971).
8. S. V. Berg, Structure, Behavior, and Performance in the Scientific Journal Market (thesis, Yale 1970); J. ASIS 23, 23 (1972).

9. G. S. Becker, Econ. J. 75, 493 (1965); Economic Theory (Knopf. 1971).

10. H. A. Simon, Chapter 2 of Computers, Communications, and the Public Interest (M. Greenberger, Ed., Johns Hopkins University Press 1971).

11. Annual Review of Information Science and Technology (C. Cuadra, Ed., Encyclopedia Brittannica, annual): chapters entitled "Design and Evaluation of Information Systems", by D. W. King (V. 3, 1968), M. M. Henderson (V. 4, 1969), F. W. Lancaster and C. J. Gillespie (V. 5, 1970), C. W. Cleverdon (V. 6, 1971).

12. P. M. Morse, Library Effectiveness (M.I.T. Press 1969).

13. Scientific and Technical Communication a Pressing National Problem and Recommendations for its Solution (Committee on Scientific and Technical Communication, NAS-NAE 1969).

14. Report of the Task Group on the Economics of Primary Publication (Committee on Scientific and Technical Communication, NAS-NAE 1970).

15. N. B. Hannay et al., Cost Effectiveness of Information Systems (report of a committee of the American Chemical Society, 1969); see also Chem. Eng. News 47, July 28, 1969, p. 45.

16. See Ref. 14, Appendix Secs. IIIA 3-5 and Fig. 8 for a more detailed discussion of these categories.

17. H. B. Landau, Am. Doc. 20, 302 (1969).

18. R. J. Penner, J. ASIS 21, 67 (1970).

19. F. W. Lancaster, Information Retrieval Systems; Characteristics, Testing, and Evaluation (Wiley 1968).

20. M. M. Henderson, Evaluation of Information Systems: a Selected Bibliography with Information Abstracts (NBS Technical Note 297, 1967).

21. Ref. 3, Chapter 9.

22. Cleverdon, Ref. 11, p. 43.

23. S. Rossmassler, private communication.

24. E. Dugger and R. F. Klinger, "User Evaluation of Information Services," in Information Retrieval (A. B. Tonick, Ed., International Information 1967), p. 135.

25. M. W. Mueller, "Time, Cost, and Value Factors in Information Retrieval" (IBM Information Retrieval Systems Conference, Poughkeepsie 1959); see also Ref. 19, p. 178.

26. P. E. Green, Experiments on the Value of Information in Simulated Marketing Environments (Allyn and Bacon 1967).

27. P. F. Urbach, in Information Retrieval: the User's Viewpoint as an Aid to Design (A. B. Tonick, Ed., International Information, Inc. 1967) p. 7.

28. Report of the Panel on the Dissemination and Utilization of the Information of Physics (Physics Survey Committee, NAS 1972, to be published in Vol. II of their Report); see also the summary ibid., Vol. I, Chapter 13.

29. C. Herring, unpublished work 1972.

30. In his book, Techniques of Information Retrieval (Butterworths 1970), especially Ch. V., Appendix. B. C. Vickery has surveyed many studies of users and their apportionment of their time. A brief summary of some studies is also given in Ref. 14, Appendix Section III C. Current work is summarized in the articles entitled "Information Needs and Uses" in Annual Review of Information Science and Technology (C. Cuadra, Ed., Encyclopedia Brittannica, Chicago, annual), by H. Menzel (V. 1, 1966), S. and M. Herner (V. 2, 1967), W. J. Paisley (V. 3, 1968), T. J. Allen (V. 4, 1969), B.-A. Lipetz (V. 5, 1970), and D. Crane (V. 6, 1971). Particularly useful among the older studies are those of the Case Institute group, An Operational Research Study of the Dissemination and Use of Recorded Scientific Information. A Study Under NSF G-8434 (Report PB 171503, available from NTIS); see also M. W. Martin, Jr., IRE Trans, Eng. Management EM9, 66 (1962), M. W. Martin and R. L. Ackoff, Management Sci. 9, 322 (1963). Further valuable data can be found in Ref. 15 and in Ref. 28, Sections II.1 and II.2.

31. B. C. Brookes, Inform. Storage & Retrieval 6, 127 (1970).

32. R. Sadacca and R. T. Root, Human Factors 10, 5 (1968).

33. See for example R. von Zelst and W. A. Kerr, J. Abn. Soc. Psych. 46, 470 (1951); R. E. Maizell, Am. Doc. 11, 9 (1960).

34. J. Martyn, "Literature Searching by Research Scientists" (ASLIB Research Dept., London 1964); New Scientist 21, 338 (1964).

35. F. J. Brockis and P. F. Cole, Chemistry in Britain 3, 421 (1967).

36. H. A. Barton, Physics Today 16, No. 6, p. 45 (1963).

37. H. W. Koch, Report ID 69-5, American Institute of Physics (1969).

38. Ref. 14, Appendix Section IV A.5.

39. Ref. 14, pp. 165-6.

40. Ref. 14, Appendix Sections IV A.2 to IV A.4.

41. D. G. Marquis and T. J. Allen, Am. Psychologist 21, 1052 (1966); T. J. Allen, Ind. Management Rev. 8, 87 (1966).

42. R. S. Rosenbloom and F. W. Wolek, Technology, Information, and Organization: Information Transfer in Industrial R&D (Report prepared for OSIS, Grant GN-305, Harvard University Graduate School of Business Administration 1967).

43. Ref. 28, Figs. 7 and 8.

44. T. J. Allen, Managing the Flow of Scientific and Technological Information (thesis M.I.T. 1966); T. J. Allen and S. I. Cohen, Admin. Sci. Quarterly 14, 12 (1969); T. J. Allen, in

communication Among Scientists and Engineers (C. E. Nelson and D. K. Pollock, Eds., Heath Lexington Books 1970), p. 191.

45. C. C. Chen, J. ASIS 23, 254 (1972).

46. Ref. 14, Appendix Section IV A.6.

47. V. Rosenberg, Info. Storage and Retrieval 3, 119 (1967); P. G. Gerstberger and T. J. Allen, J. Appl. Psych. 52, 272 (1968); T. J. Allen, in Communication among Scientists and Engineers (C. E. Nelson and D. K. Pollock, Eds., Heath Lexington Books 1970) p. 191.

48. Suggested in a private conversation by J. K. Galt.

49. M. A. Tinker, Legibility of Print (Iowa State University Press 1963).

50. Ref. 14, Appendix Sections II and IV A.1; Ref. 7, pp. I.12, I.15, III.3.

51. Federal Funds for Research, Development, and Other Scientific Activities (annual, National Science Foundation).

52. J. H. Connor. Lib. Quarterly 37, 373 (1967).

53. H. East, ASLIB Proc. 20, 482 (1968).

54. Science, Government, and Information (Report of the President's Science Advisory Committee, GPO 1963).

55. Directory of Federally Supported Information Analysis Centers (COSATI 1968).

56. T. H. Campbell and J. Edmisten, Characteristics of Scientific Journals (Herner and Co., 1965).

57. Recommendations for Improving the Dissemination of Federal Scientific and Technical Information. Report of the COSATI Task Group on Dissemination of Information (COSATI 1970).

58. P. Abelson, Editorial, Science, Nov. 11, 1966.

59. Ref. 28, Section II.4.

60. C. C. Spencer, Am. Doc. 18, 87 (1967).