COMMISSION OF THE EUROPEAN COMMUNITIES

information management

Development and use of models for the prediction of costs for alternative information systems, PT 1 (input model) and PT 2 (output model)

Final report on Project 3, Phase 1

1977

EUR 5693 e

COMMISSION OF THE EUROPEAN COMMUNITIES

information management

Development and use of models for the prediction of costs for alternative information systems, PT 1 (input model) and PT 2 (output model)

Final report on Project 3, Phase 1

PETER VICKERS AND MARTIN ROWAT
Aslib Consultance Service, London

Directorate-General 'Scientific and Technical Information and Information Management' — Luxembourg

1977 EUR 5693 e

Published by the

COMMISSION OF THE EUROPEAN COMMUNITIES
Directorate-General
'Scientific and Technical Information and Information Management'
Bâtiment Jean Monnet - Kirchberg

LUXEMBOURG

Printed in Belgium

A B S T R A C T

The final report consists of two parts:

- the input model (part 1)
- the output model (part 2).

These predictive cost models are able to operate in three dimensions:
- system configuration (i.e. flexibility)
- operating regime (i.e. predicting the cost for any volume of throughput)
- time.

They comprise the mechanical component, the input data and the user interface. The term "mechanical component" refers to the set of mathematical relationships that will determine the cost of each element of an information system, plus the means of performing the necessary calculations. (The financial planning and analysis system PROPHIT II, operating on-line, was used in this study.)

The input model calculates for each operation the staff, materials equipment and services costs as required, prompting the user to consider various systems options where appropriate.

The output model is more complex than the input model since it has to provide for a wider range of system configurations for a variety of different services.

Further research and improvement is needed before system operators could be offered a model into which could be fed details of current operational volumes and costs for a specific system and which the operator could use to determine the effect of changes in methods, staffing, throughput volumes etc.

FINAL REPORT ON PROJECT 3, Phase 1, Pt1 (INPUT MODEL)

CONTENTS

MANAGEMENT SUMMARY

In accordance with the specification for EFAG Project 3, two separate reports have been prepared on the development and testing of cost prediction models for (a) input activities, and (b) output activities of mechanized information systems. The two models are, however, closely related and both reports are summarized here.

Definition of requirements

In designing these models, the first requirements to be considered were the dimensions within which they had to operate. The models should be applicable to most if not all foreseeable system configurations in terms of resources and techniques used, and services provided; they should be able to predict costs for any volume of throughput; and they should be able to predict costs for any reasonable period of future time.

The second requirement was that the models should be easy to use.

Thirdly, the design of the models should not be incompatible with other studies in the present series of EFAG costing projects.

Last but not least, the models should be capable of predicting costs to a satisfactory level of accuracy (which would depend partly on the purpose for which they were used). A factor to be noted here is that, providing reasonable data values are input to the models, the systems they represent could be controlled in such a way as to ensure that the predicted costs were achieved.

General description

The models have three main components:

- the mechanical component
- the input data
- the user interface.

The mechanical component comprises a series of equations that determine the cost of each element of the system. These equations are presented in such a way that the necessary calculations could be performed by hand, but on-line computing facilities were used in developing and testing the models, as described below.

Some of the input data is determined by the model user - such as the configuration of the system and the volume of throughput. The remainder has to be drawn from observation of the behaviour of existing systems, and the accuracy of the models is highly dependent on these values.

When the models are used manually, the user interface can only be rudimentary; little can be done to relieve the drudgery of the repetitive calculations required. With the aid of computer facilities, however the models can be made truly interactive.

The input model

The main sections of the model cover acquisition, selection, cataloguing, indexing, abstracting, translation, and mechanical processing.

The model calculates for each operation the staff, materials, equipment and services costs as required, prompting the user to consider various system options where appropriate. Alternative methods of mechanical processing, such as on- or off-line data preparation, are represented by separate equations. Alternative methods for intellectual operations, such as indexing and abstracting, are dealt with by using unit times appropriate to the quality of work required.

Direct staff costs are calculated on the basis of unit times for each staff activity. These unit times are multiplied by the number of items processed to give the hours required per year. This figure is then divided by the effective number of working hours in a year, taking into account allowances for relaxation, unoccupied time, holidays, etc., to give the number of man-years of effort required for the activity.

Man-years of effort for each activity are multiplied by a salary cost at a level appropriate to the activity. In the model, provision is made for five salary grades. Some types of staff within these grades are regarded as interchangeable (e.g. indexing and abstracting might be done by the same people), and this factor is taken into account in calculating the total numbers of staff required. The numbers thus calculated are rounded up to integer numbers.

Given the total number of staff required for each activity in the system, the model user is invited to determine the number of supervisory and clerical support staff required. It was felt that this decision could not be made in a realistic way by the model.

Computer processing cost calculations are based on unit costs for each operation, or on the estimated percentage occupancy of a computer installation multiplied by a rental charge.

Accommodation costs are calculated for each member of staff. Overheads are added as a percentage of salary costs.

The output model

The output model is inherently more complex than the input model, in that it has to provide for a wider range of system configurations for a variety of different services. It can be linked to the input model, in that the predicted cost of creating a data base can be fed into the output model. Alternatively, the cost of a commercially available data base or data bases can be used.

The output model covers the following services, separately or in combination:

retrospective search (batch processing)

retrospective search (on-line)

SDI

group SDI

secondary publication (alerting service)

secondary publication (abstracts bulletin)

machine-readable services

The model calculates for each operation the staff, equipment, materials and services charges as required for each of the seven output services selected by the model user as part of the design configuration.

Direct staff costs, where applicable, are calculated on the basis of unit times or data values for each activity. The unit times are multiplied by the frequency of the particular activity to give the hours required per year. This figure is then divided by the effective number of working hours in a year, taking into account allowances for relaxation, unoccupied time, holidays etc., to give the number of man-years of effort required for the activity.

Man-years of effort for each activity are multiplied by a salary cost at a level appropriate to the activity. In the model provision is made for five salary grades. Some types of staff within these grades are regarded as interchangeable (e.g. profile formulation for SDI and for group SDI might be done by the same people), and this factor is taken into account in calculating the total numbers of staff required. The numbers thus calculated are rounded up to integer numbers.

Given the total number of staff required for each activity in the system, the model user is invited to determine the number of supervisory and clerical support required.

Accommodation costs are calculated for each member of staff. Overheads are added as a percentage of salary costs.

Costs of materials and printing are calculated as appropriate to each activity. Royalty charges based on volume of usage made of a purchased data-base may be calculated on the basis of charges against numbers of users, frequency of use and/or volume of output produced, according to the conditions obtaining under sales contracts negotiated with individual data-base producers.

Computer processing costs are calculated on the basis of data available for costs of each run (or issue, in the case of secondary publications).

After the model has calculated the direct costs of each service, an apportionment of input and indirect costs is added to give the total cost.

The computerized models

Both models were developed with the aid of the PROPHIT II system, available through the CDC CALL/370 Time Sharing Service.* PROPHIT II is an on-line financial planning and analysis system. When using this facility, the model is expressed as a series of statements (called a definition file) written in a simple user-oriented programming language.

Input can be in the form of a history file (employing data gathered from past experience) or a projection file. With a projection file, data values that will change with time (such as the number of items input, or salary levels) can be generated from an initial value or values by specifying one of a range of projection types (e.g. linear, stepped, compound).

The projection and/or history files are run against the definition file to produce a report covering as many years as required. The effect of changes in data values, methods of projection, or system design options can be explored by means of a WHAT-IF facility.

Data values

For each model, all the variables employed in the equations are defined, and preferred values or ranges of values are presented where appropriate. The reports stress, however, that the model user should be able to apply judgement, based on experience, in selecting values to be used as input to the models.

A significant difference between the input and output models is that while staff costs predominate in the former, computer processing costs are more important in the latter.

The equations for the input model involve 48 variables, although some of these apply only to certain system configurations. The output model, with its range of alternative services, employs 97 variables.

---

* Similar facilities are available from other major timesharing computer services.

Testing the models

Test runs were carried out with both models to ensure that they would operate correctly under a variety of conditions. In the case of the input model, further tests were conducted by simulating known systems.

As required by the project specification, both reports include written specifications for designed experiments to implement the models. The method proposed is to use the models in a retrospective mode, i.e. to make cost predictions for existing systems as of some time in the past, and to compare the results with the actual costs experienced in reality.

Applications of the models

The main application envisaged for these models, in their present form, is at the broad planning level. They can be used to determine the pattern of costs in future years for a proposed new system, and in so doing enable the planner to explore the effect of different system configurations and operating regimes.

They can also be used more generally as a management tool for forecasting manpower requirements, budgets, and unit costs.

The models as presented are highly generalized, and are applicable to most typical system configurations. The methodology that they incorporate could, however, easily be adapted or extended to cover other specialized configurations, or specific applications. For example, they could be developed for application to cooperative networks, or to investigate the effect of changes on existing systems.

CHAPTER I: INTRODUCTION AND TERMS OF REFERENCE

This report is the first of two final reports resulting from the study 'Project 3: Development and use of models for the prediction of costs for alternative information systems'. This report is about the cost prediction model for input activities of mechanized information systems. A companion report deals with the output model.

The project specification is reproduced in Appendix I, but it may be useful to restate here the objectives of the project :

> "To develop models for predicting the costs of various methods of data base creation and provision of information services."

The project as a whole comprises two phases, the first being to develop and test the models, and the second to implement them in an experimental environment. This report is concerned with Phase I, but includes in Appendix 9 a specification for a designed experiment to implement the model.

The nature of the project is such that there can be no detailed statement of the methodology employed. Having studied previous work in this area (see Chapter 2) and determined the requirements of the model (see Chapter 3), we were able to formulate the basic equations and then develop them by an iterative process (see Chapter 4). Some tests were carried out to prove the viability of the model (see Chapter 6). Considerable effort was devoted to research on the data available for input to the model (see Chapter 5).

CHAPTER 2: REVIEW OF PREVIOUS WORK

Little published work has been found that relates to the design of predictive cost models for the input operations of mechanized information systems. Papers that relate to the modelling of output operations are reviewed in the companion report to this one.

Bourne and Ford [1] have reported on the use of a computer-based model designed to simulate the several-year operation of an information system. The model estimates expected operating costs as well as the amount of equipment and personnel required. The use of a model of this kind made it possible to examine the variety of system configurations under various operating regimes. Their paper unfortunately does not describe the model in detail.

The work of Wilkin et al [2] can also be regarded as a modelling exercise, although it was not specifically concerned with computer-based information systems. The aim was to determine the comparative influence of various factors on the time taken to perform a whole range of information system operations. Multiple regression techniques were applied to determine the relative effect of each variable. A model of the form

$$y = b_o + b_1 x_1 + b_2 x_2 \ldots b_a x_a$$

is postulated where y is the dependent variable, x the independent variable(s), and b_o, b_1, b_2 are constants estimated from the data.

CHAPTER 3: DEFINITION OF REQUIREMENTS

3.1 What is a model?

The meaning of the term 'model' is clear enough in the case of an econometric model, or a model of an electric circuit. In the case of a predictive cost model, however, the identity of what we are modelling is less clear. The best definition of our purpose is probably to say that we are trying to model a future situation in which an information system would exist, or in which the operations of an existing system are to be changed, in such a way that its costs can be determined.

3.2 Requirements of the model

3.2.1 Dimensions of operation

The first requirement is that the model should be able to operate in three dimensions :

(1) System configuration;
(2) Operating regime;
(3) Time.

The first of these means simply that the model must be applicable to any foreseeable type of information system in terms of resources used (staff, equipment, or materials) and the type and quality of services provided.

The second implies that the model must be able to predict

costs for any volume of throughput.

The third implies that the model must be able to predict costs for any reasonable period of future time.

In practice, of course, it is probable that in using the model, two of these sets of variables will be held constant, while studying the effect of changing the third one.

3.2.2 Ease of use

The use of a predictive cost model of this kind is inevitably complex. The model provides the mechanism for calculating costs, but the model user must make a series of choices concerning the system configuration, and must select appropriate input data. The model should be designed so that the user is given as much help as possible in making these decisions, and it should also enable the user to determine quickly the effect of changing any of the parameters. We shall return to this aspect in section 3.5.3.

3.2.3 Compatibility

It is obviously desirable that the design of the model should take into account, and where necessary be compatible with, the results of other studies in the present series of EFAG costing projects.

We have attempted to ensure that the classification and definition of cost elements used in the present project are compatible with those proposed in the EFAG Project 2 report. It has to be recognized, however, that the problems of collecting and analysing data from existing operational systems (which were the subject of the

EFAG Project 2*) are fundamentally different from those of forecasting the costs of hypothetical systems. If the EFAG 2 cost-accounting scheme were widely used as a means of collecting data, it would eventually be possible to design a different, and more accurate cost model. But the model we present in this report has to make use of the best data available now.

3.2.4 Accuracy

The accuracy required of the model will to some extent depend on the purpose for which it is applied. In some cases, absolute accuracy will be less important than relative accuracy. For example, if the model is used for comparing the costs of alternative system configurations, it must accurately show the relative effect of these alternatives.

Accuracy of the answers given by the model will depend mainly on the accuracy of the data that is fed into it - a point we shall return to in Section 3.5.2. It also has to be realized that the model can only aim to predict costs dependent on more mechanistic factors, at the same time indicating the extent to which predictions may be distorted by other influences such as quality of management.

On a more encouraging note, it is worth mentioning that the model can to a certain extent provide a self-fulfilling prophecy.

---

* P.H.Vickers .Final report on Project 2: Extension and revision of the cost/accounting scheme to interactive systems of the network.' Aslib Consultancy Service, July 1976.

Providing that the unit times input to the model are sufficient to ensure an acceptable quality of data base, the system it represents could be controlled by its manager in such a way as to ensure that these unit times were realized.

3.3 Applications

The main application envisaged for these models is at the broad planning level. They can be used to determine the pattern of costs in future years for a proposed new system, and in so doing enable the planner to explore the effect of different system configurations and operating regimes.

The models presented in this and the companion report deal with the costs of an individual system. They could only be used for a co-operative network by treating it as a collection of single systems, and combining the results of a series of separate predictions.

3.4 Other design factors

3.4.1 Viewpoint

It is important to recognize that the cost of a system is highly dependent on viewpoint - in other words, we have to decide whose costs we are trying to predict. Should the model be designed to operate at the level of the system, of the organisation which runs it, or of the government of the country concerned? The simplest illustration of this problem is provided by document acquisition costs. A system based on a university, for example, may derive its document input from the university library, and its operating costs would show no outlay for this. Yet the university's budget would show not only the operating cost of the information system, but also

the cost of running the library. We have assumed that the cost models should take into account all local costs relevant to the operation of the information system, and that the costing viewpoint should be that of the parent organisation.

3.4.2 Performance and quality

Ideally, cost predictions should be related to system performance characteristics, listed by Lancaster and Climenson (3) as:

Coverage	Usability	Recall	Precision
Response time	Presentation	User effort	

King and Caldwell (4) have demonstrated the feasibility of designing cost models that relate to levels of performance in terms of recall, precision and some factors affecting user effort. In most situations, however, the practicality of specifying desired levels of recall and precision for a planned system is limited, and we have not attempted to build into our models any direct capability for relating cost to these parameters.

It was considered essential, however, that our models should take into account the system characteristics that can be predetermined and which govern the quality of the services provided. Thus the model parameters include depth of indexing, length of abstract, and print density of output, for example, which affect recall, precision, presentation and user effort.

Response time is a special factor in this context. Our models estimate the staff effort required for each operation, but if response time were a critical design factor, it might be necessary to allow for sub-optimal staff utilization and the values obtained from the model would have to be factored accordingly.

3.4.3 Cost vs economics

It needs to be clearly stated that the models are designed to predict only costs and not overall economics. In other words, they will not take into account the revenue earned by a system to offset its operating costs.

3.5 Components of a predictive cost model

The three main components of a predictive cost model are :

(1) the mechanical component;
(2) the input data;
(3) the user interface.

These components are discussed below.

3.5.1 The mechanical component

The mechanical component comprises a set of mathematical relationships that will determine the cost of each element of an information system. It also implies some means of performing the necessary calculations, such as a slide rule, an electronic calculator, or a computer. In Chapter 4 we shall first present a series of equations which could be used with any calculating device. We shall then show how the same calculations can be performed with the aid of a computer.

3.5.2 The input data

This is the data that must be fed into the model in order that it may calculate the cost of any system. Some of this data is determined by the model user - such as the volume of throughput, and the configuration of the system to be modelled. Much of it, however, has to be drawn from observation of the behaviour of existing systems. The accuracy of the model depends almost entirely on the latter kind of data. The model cannot be better than the data which is available.

The input data for our model are defined, and values are suggested, in Chapter 5.

3.5.3 The user interface

As mentioned in Section 3.2.2, the model should be designed so that it is responsive to the user, and easy to operate. To achieve this, a suitable interface is needed between the user and the mechanism of the model itself.

In the case of a manually-operated model, little can be done to relieve the drudgery of repetitive calculations, and we suspect that use of the model in this mode will be limited. As will be seen from the later parts of this report, however, it is possible to operate the model with aid of a computer, making it truly interactive.

CHAPTER 4: DESCRIPTION OF THE MODEL

In this chapter we shall describe the mechanical component of the model. First we shall explain the function of each part of the model, and present the equations used in sufficient detail for cost predictions to be made manually.

The model is designed to represent what we believe to be the most typical system configurations within the scope of present technology. It does not cover certain ancillary activities, such as microfiche production, but extension of the model to cover such activities would be a simple matter.

Even with the aid of an electronic calculator, manual use of the model can be fairly laborious, and at an early stage in the project it was decided to use computer facilities to develop, test and operate the model. The particular facilities used are described in Section 4.2

The manual and computer-based versions of the model are linked by the line numbers of the computer files. These are shown in parenthesis after each of the parameters used in the equations that follow, and again in Chapter 5, which defines and suggests values for the input data required for the model.

It must be stressed that the computer system merely provides the capability to perform the calculations required by the model, and to prepare cost reports; it does not constitute the actual model.

4.1 The input model

The model calculates for each operation the staff, materials equipment and services costs as required, prompting the user to consider various system options where appropriate.

Direct staff costs are calculated on the basis of unit times for each staff activity. These unit times are essentially 'basic'

times, as defined in BS.3138*, which are multiplied by the number of items processed to give the hours required per year. This figure is then divided by the effective number of working hours in a year, taking into account allowances for relaxation, unoccupied time, holidays, etc., to give the number of man-years of effort required for the activity.

Man-years of effort for each activity are multiplied by a salary cost at a level appropriate to the activity. In the model provision is made for five salary grades. Some types of staff within these grades are regarded as interchangeable (e.g. indexing and abstracting might be done by the same people), and this factor is taken into account in calculating the total numbers of staff required. The numbers thus calculated are rounded up to integer numbers.

Given the total number of staff required for each activity in the system, the model user is invited to determine the number of supervisory and clerical support staff required. It was felt that this decision could not be made in a realistic way by the model.

Accommodation costs are calculated for each member of the staff. Overheads are added as a percentage of salary costs.

Costs of materials, equipment and external services are calculated as appropriate to each activity. Computer processing costs are calculated on the basis of observed unit costs for each activity; or alternatively on the estimated percentage occupancy of a computer installation multiplied by a rental charge.

---

* Glossary of terms used in work study. BS3138 : 1969, London, British Standards Institution, 1969.

The 'manual' model calculates costs for one system configuration in one year of operation. To predice costs for a succession of years with different operating regimes and increasing salaries, equipment rentals, etc, the model user would have to repeat the calculations as many times as necessary.

The equations to be used for each element of the model are presented below.

4.1.1 Acquisition

The model recognizes that the system will acquire a certain number of documents, some or all of which will have to be purchased. Two classes of documents, monographs and journals (i.e. serials) are treated separately. It is also recognized that the number of items input to the computer system may be different from the numbers of documents acquired or purchased. The staff costs of ordering and handling monographs and journals will tend to differ, so different unit times for these operations are called for.

The term 'monographs' is intended to cover all non-serial publications, including books, reports, patents etc. 'Serials' could also include secondary publications, which are sometimes a source of input. If necessary, and if data were available, the equations which follow could be used iteratively with different values for specific types of monograph or serial publications.

The acquisition of input in machine-readable form is covered in the output model, which is appropriate for situations where such input can be used with little or no modification. For a system in which machine-readable input is re-indexed to form, in effect, a new data-base, it would be necessary to adapt the input model equations for this purpose.

The following equations apply to acquisition:

Acquisition effort required in man-years,

$$e_{acq} = \frac{T_b B_a + T_p J_a}{H}$$

Cost of documents purchased,

$$C_{da} = P_b B_p + P_j J_p$$

Total direct cost of acquisition,

$$C_{acq} = \left[\frac{\left[T_b B_a\right] + \left[T_j J_a\right]}{H} \right] S_b + P_b B_p + P_j J_p$$

$$\text{Unit cost per item input} = \frac{C_{acq}}{I_n}$$

where T_b = unit time for ordering monographs (1130)

T_j = unit time for ordering journals (1132)

P_b = average purchase cost of monographs (1160)

P_j = average subscription cost of journals (1162)

B_a = number of monographs acquired (1020)

B_p = number of monographs purchased (1030)

J_a = number of journals acquired (1022)

J_p = number of journals purchased (1032)

S_b = annual salary, Grade B staff (1070)

H = number of hours in a man-year (1040)

I_n = number of document records (for monographs and journal articles) input per year (1010)

4.1.2 Selection

The effort required for selecting items for input is arrived at by multiplying the unit time by the number of items input, and then dividing by the number of man-hours in a year. Thus effort required for selection, $e_{sel} = \frac{T_s I_n}{H}$

Cost of selection,

$$C_{sel} = e_{sel} \; S_c$$

where T_s = unit time for selection (1260)

S_c = annual salary, Grade C staff (1080)

4.1.3 Cataloguing

The effort required for cataloguing is calculated in the same way as that for selection. Thus effort required,

$$e_{cat} = \frac{T_c \; I_n}{H}$$

Cost of cataloguing,

$$C_{cat} = e_{cat} \; S_c$$

where T_c = unit time for cataloguing (1310)

4.1.4 Indexing

The effort required for indexing is calculated by multiplying the unit time by the number of items input, and then dividing by the number of man-hours in a year. Here as in other parts of the model, the model user has to select a unit time appropriate to the quality of indexing, type of document, indexing language, etc. that are to be built into the system.

Thus effort required for indexing,

$$e_{ind} = \frac{T_i \, I_n}{H}$$

Cost of indexing,

$$C_{ind} = e_{ind} \, S_c$$

Where T_i = unit time for indexing (1360)

4.1.5 Abstracting

Abstracts input to the system may be specially prepared, to varying standards of quality and length; or they may be 'author' abstracts (i.e. copied from another source); or a mixture of these types may be used. The model therefore calls for an indication of the proportion of author abstracts to be used, and unit times for each type of abstracting. Effort required for abstracting can then be calculated as follows :-

$$e_{abs} = \frac{I_n \, F_a \, T_a}{100H} + \left[I_n - \frac{I_n \, F_a}{100} \right] \frac{T_w}{H}$$

Cost of abstracting,

$$C_{abs} = e_{abs} \, S_c$$

where F_a = percentage of author abstracts used (1040)

T_a = unit time for preparing author abstracts (1440)

T_w = unit time for preparing written abstracts (1450)

4.1.6 Translation

All or part of the input may be translated from one or more languages. The model user is required to indicate the percentage of input that is to be translated, and the effort required can then be calculated as follows :-

$$e_{tra} = \frac{T_t \; I_n \; F_t}{100H}$$

Cost of translation,

$$C_{tra} = e_{tra} \; S_c$$

Where T_t = unit time for translating an item (1600)

F_t = percentage items translated (1580)

4.1.7 Total cost of intellectual processing of input

At this stage it is possible to calculate the direct cost of all intellectual processing operations, as follows :-

$$C_{ip} = C_{sel} + C_{cat} + C_{ind} + C_{abs} + C_{tra}$$

4.1.8 Mechanical processing of input (data preparation)

In calculating data preparation costs, it is necessary to consider a variety of technical options, and to allow for the use of these separately or in combination. For the purposes of the model, data preparation may be carried out in-house or by a bureau. It may be done on-line or off-line. Off-line methods include the use of punched cards, paper tape, magnetic tape, key-to-disc, or optical character recognition (OCR). For the latter, input is typed on a special typewriter, and then read by an OCR reader which writes the records to magnetic tape.

Data preparation costs may also include the rental of equipment (card punches, terminals, etc.), telecommunications costs (in the case of on-line input), and computer processing costs for input validation, which may be carried out in-house or by a service bureau.

Effort required for off-line data preparation,

$$e_{off} = \left[\frac{F_a L_a}{100} + \left[1 - \frac{F_a}{100}\right]\right]\left[1 + \frac{F_v}{100}\right] \frac{I_n}{K_{off} \; H \times 1000}$$

Effort required for OCR data preparation,

$$e_{ocr} = \left[\frac{F_a L_a}{100} + \left[1 - \frac{F_a}{100}\right] L_w\right] \frac{I_n}{K_{ocr} \; H \times 1000}$$

Effort required for on-line data preparation,

$$e_{on} = \left[\frac{F_a \, L_a}{100} + \left[1 - \frac{F_a}{100} \right] L_w \right] \left[1 + \frac{F_v}{100} \right] \frac{I_n}{K_{on} \, H \times 1000}$$

Bureau data preparation costs can be calculated as follows :-

$$C_{bu} = \left[\frac{F_a \, L_a}{100} + \left[1 - \frac{F_a}{100} \right] L_w \right] \frac{F_c \, I_n \, P_{kb}}{100,000}$$

Conversion costs for writing OCR input to magnetic tape are calculated as follows :-

$$Q = \left[\frac{F_a \, L_a}{100} + \left[1 - \frac{F_a}{100} \right] L_w \right] I_n \, \frac{P_{con}}{1000}$$

Effort required for proof-reading,

$$e_{pr} = \frac{T_p \, I_n}{H}$$

Equipment costs for off-line, OCR and on-line operations can be calculated as follows :-

$G_{off} = E_{off} \, R_k$ where E_{off} represents e_{off} rounded up to the nearest whole number

$G_{ocr} = E_{ocr} \, R_o$ where E_{ocr} represents e_{ocr} rounded up to the nearest whole number

$G_{on} = E_{on} R_t + P_{tc}$ where E_{on} represents e_{on} rounded up to the nearest whole number

Computer processing costs for in-house and bureau operations can be calculated as follows :-

$$M_{in} = \frac{F_{cp} R_c}{100}$$

$$M_{bu} = I_n P_{bu}$$

The total in-house effort required for data preparation (excluding proof-reading) can be expressed as follows :-

$$e_{dp} = \left[1 - \frac{F_c}{100} \right] \left[\left[1 - \frac{F_{on}}{100} \right] e_{off} + \frac{F_{on}}{100} e_{on} \right]$$

e_{ocr} could be substituted for e_{off} in the above equation, in which case the conversion cost Q would need to be included in the final total for the system.

The cost of data preparation can be calculated as follows :-

$$C_{dp} = e_{dp} S_b + C_{bu} + \left[G_{off} \underline{\text{or}} G_{ocr} \right] + G_{on} + \left[M_{in} \underline{\text{or}} M_{bu} \right] + Q + e_{pr} S_c$$

Notation used in the above equations is listed below :-

F_c = percentage of records keyboarded by bureau (1860)

F_{cp} = occupancy of in-house computer expressed as decimal fraction (2630)

F_{on} = percentage of input keyboarded on-line (1950)

F_v = percentage of input records verified (1810)

K_{ocr} = keyboarding rate for OCR input (1730)

K_{off} = keyboarding rate for off-line input (1750)

K_{on} = keyboarding rate for on-line input (1710)

L_a = average length of records with author abstracts, in characters (1680)

L_w = average length of records with written abstracts, in characters (1690)

P_{bu} = cost of computer processing, per record (2600)

P_{con} = cost of reading OCR input, per 1000 characters (2210)

P_{kb} = cost of bureau keyboarding, per 1000 characters (1870)

P_{tc} = communications cost (2060)

R_c = rental of in-house computer (2640)

R_k = rental of keypunch or alternative (2340)

R_o = rental of OCR typewriter (2190)

R_t = rental of terminal (2040)

T_p = unit time for proof reading (2510)

4.1.9 Total effort required

In the model so far, all staff effort has been at Grade B or C. It is assumed that all kinds of keyboard operators would be interchangeable, but that they would not be interchangeable with the staff responsible for acquisitions work. Similarly, it is assumed that staff employed on selection, cataloguing, indexing, abstracting, translation and proof-reading would all be of similar capability and thus interchangeable.

To estimate realistic staff costs, the numbers of staff in each of these three groups need to be rounded up to whole numbers, as follows :-

$E_{b1} = e_{acq}$ rounded up to nearest whole number

$E_{b2} = e_{dp}$ rounded up to nearest whole number

$E_{c1} = e_{sel} + e_{cat} + e_{ind} + e_{abs} + e_{tra} + e_{pr}$

rounded up to nearest whole number

At this point, having determined the numbers of staff needed for each activity, the model user has to decide on the kind of organizational structure that will be required to operate the system, and to estimate the number of supervisory and clerical support staff needed. Supervisory staff might be employed at Grade C, D or E depending on their level in the hierarchy. Clerical support staff are at Grade A. The total numbers and costs of staff can now be calculated as follows :-

total number of staff E_{tot}

$$E_{tot} = E_{b1} + E_{b2} + E_{c1} + E_{c2} + E_{d} + E_{e} + E_{a}$$

where E_{c2} = number of supervisory staff, Grade C (3450)

E_{d} = number of supervisory staff, Grade D (3460)

E_{e} = number of supervisory staff, Grade E (3470)

E_{a} = number of clerical support staff

4.1.10 Accommodation costs

Accommodation costs are calculated on the basis of a space allowance per member of staff, multiplied by a cost per unit of area. Thus accommodation costs,

$$C_{acc} = E_{tot} A_{p} R_{acc}$$

where A_{p} = space required per staff member (3490)

R_{acc} = accommodation cost per unit area (3500)

4.1.11 Total costs

Finally, the total costs can be obtained. This entails multiplying the number of staff E_{b1}, E_{c1}, etc. by the appropriate salaries to convert them to staff costs. Overheads are added as a percentage of staff costs. Materials and equipment costs as determined by the equations above are added into this equation,

$$C_{tot} = \left[1 + F_{ov}\right]\left[E_{b1} S_b + E_{b2} S_b + E_{c1} S_c + E_{c2} S_c + E_d S_d + E_e S_e + E_a S_a\right] + C_{da} + C_{bu} + \left[G_{off} \text{ or } G_{ocr}\right] + G_{on} + \left[M_{in} \text{ or } M_{bu}\right] + Q + C_{acc}$$

where S_d = annual salary, Grade D staff (1090)

S_e = annual salary. Grade E staff (1100)

S_a = annual salary, Grade A staff (1060)

F_{ov} = percentage overhead (3640)

4.2 A computer-based version of the model

The arithmetical operations involved in a cost model of the kind presented in this report are simple, but numerous. A substantial amount of data has to be input, to produce some fairly detailed tabulations and analyses of a future cost situation. At an early stage in the project, it was decided to use computer facilities to run and test the model, and these will now be described. Examples of the output from these trial runs are given in Appendices 6 to 8.

In the course of the work on EFAG Project 2, Mr. D. Barlow of INSPEC brought to our attention the PROPHIT II system available through the CDC CALL/370 Time Sharing Service. PROPHIT II is a financial planning and analysis system, which proved to offer the facilities required for our model at a reasonable cost. This is an on-line system, which greatly facilitated rapid development and refinement of the model. In particular, the ease with which data

values can be adjusted makes it easy to 'tune' the model to give 'reasonable' results.

It is not our intention to convey that PROPHIT II is the only or even necessarily the best computer system for running the model. We understand that Time Sharing Ltd, CSS International and Honeywell (in the U.K. alone) all offer financial planning systems that could probably be adapted to the same purpose, and there may be many more. Furthermore, it would not be difficult to write a program to perform the calculations required by the equations in the previous section. To write a complete set of programs giving the same facilities as PROPHIT II would, however, be very costly.

A brief description of the PROPHIT II system is given in Appendix 3, but it may be helpful to outline its main features here.

The model itself is expressed as a series of statements, using a simple user-oriented language, to form a definition file. This can be automatically converted to a plain-language listing which explains the function of each line in the program. This ILLUSTRATE report is shown for the input model in Appendix 4.

The system can also generate an input form of the type shown in Appendix 5. Input can be in the form of a projection file and/or history file. In either case, the first lines (0-12) determine the output format (number of columns, time distribution, report title, etc.) With a projection file, data values that will change with time (such as the number of items input, or salary levels) can be generated from an initial value or values by specify-

ing one of a range of projection types (e.g. linear, stepped, compound). If a history file is provided, containing data from past operations, future values can be calculated to match trends.

The projection and/or history files are run against the definition file to produce a report, an example of which is shown in Appendix 6.

The effect of changes in data values, methods of projection or system design options can be explored by means of a WHAT-IF facility, some examples of which are shown in Appendix 7. The effect of these changes can be displayed more effectively by the use of a sensitivity analysis, which is illustrated in Appendix 8.

An additional feature, which could be useful in performing cost prediction for co-operative systems, is that two or more reports may be combined to produce a single report or tabulation of costs.

It should be noted that the definition file illustrated in Appendix 4 corresponds closely to the manual model presented in the earlier part of this chapter. If it were desired to use this modelling technique to investigate the future costs of an existing, specific system or network, it would be advisable (and cheaper) to prepare a new definition file to suit the problem, rather than use the generalized model we have developed.

CHAPTER 5: INPUT DATA FOR THE MODEL

5.1 Effect of data on model design

In designing the model, the decision as to the kind of data that would be used was a fundamental one. The criterion for this decision was 'from what sources can the most reliable data be obtained'. The possibilities considered were as follows :-

(1) To use global estimates of staff, equipment, materials and indirect costs.

(2) To use published values of overall production unit costs, e.g. cost per item added to a data base.

(3) To use published values of unit costs at the task level, e.g. cost per item indexed.

(4) To use published or estimated unit times (or amounts of effort) for component tasks, to which can be applied appropriate staff, materials and/or equipment cost rates.

The first possibility is often used in real-life situations where a cost estimate based on rule-of-thumb figures is acceptable. It would be too crude for the purposes of our model.

Possibility (2) would still be too crude, and like possibility (3) would entail the use of data gathered in cost surveys of the type

reviewed in the EFAG Project 1 report.* As we know, such data shows excessive scatter and fails to show consistent relationships between cost and the technical characteristics of systems.

Possibility (4), although not offering an ideal solution, appeared to be the best for our purposes. Difficulties are inevitable in dealing with computer processing costs (whatever method is used) but for manual/intellectual tasks, it was considered possible to obtain or estimate unit times of sufficient accuracy.

An important principle that has been adopted concerning data for the model is that the user should be able to apply judgement, based on experience, in selecting values to be used as input to the model. We have endeavoured to strike the right balance between making the model totally prescriptive and the opposite extreme, which would be to make the user provide all his own input data.

5.2 Data definitions and values

In the table which follows, the data elements required for the model are presented in the order in which they are called for in the computerized model (see Appendices 4 and 5), and they are identified by their line numbers. Each element is defined, and preferred values or ranges of values are presented where appropriate. These values have been derived from a variety of sources including

---

* 'Analysis of various cost studies in connection with EURONET'. N.V. System Dynamics SA, February 1976.

computer bureaux, data preparation bureaux and other specialist organizations. In some cases it has been necessary to select, from a mass of published data, values which in our personal experience seem to be the most reasonable. Thus it has not been possible always to quote one specific source for the figures shown.

We would stress that, in applying the model, a user may often have access to data that is more appropriate to a particular situation than the values suggested here. Unit times, salary levels, accommodation costs, computer costs, and overheads are all especially subject to local conditions.

Cost values input to the model can of course be expressed in the currency of the country concerned.

DATA DEFINITIONS AND VALUES

Line No	Data element	Definition
1010	ITEMS INPUT	Number of document records input to the system per year.

To be supplied by user. The volume of throughput for new systems typically increases in the early years of operation, and then levels off. Other growth patterns may however apply in special cases.

1020	MONOGRAPHS ACQUIRED	Number of non-serial documents acquired per year, including those acquired free.

To be supplied by user. Number will tend to increase with time, but will have small effect on total costs.

1022	MONOGRAPHS PURCHASED	Number of non-serial documents purchased per year.

To be supplied by user. Increase in time may need to be controlled to keep within operating budget.

1030	JOURNALS ACQUIRED	Number of serials titles acquired per year, including those acquired free.

To be supplied by user. Number may increase with time, but will have small effect on total costs.

1032	JOURNALS PURCHASED	Number of serials titles purchased (i.e. subscriptions) per year.

To be supplied by user. Increase in time may need to be controlled to keep within operating budget.

1040 MAN-YEAR HOURS — Productive hours worked in a year.

The number of days worked in a year may be calculated as follows :-

days in a year		365
less	weekends	104
	holidays	15 - 25
	sickness (average)	5
	public holidays	7
remainder		224 - 234

At 7 hours per day this would give 1568 - 1638 hours per year, but normal work study practice provides for relaxation and other allowances which reduce these figures by 12½% - 15%. The effective range thus becomes 1333 - 1392. For general use with the model we suggest a figure of 1350.

1060 GRADE A STAFF — Annual salary plus statutory and other related costs, including welfare contributions, government levies, superannuation costs etc.

The model recognizes five staff grades, the salaries for which should represent the average of what may be a wide range. Grade A is intended for clerical support staff. Salary levels for this and other grades will vary considerably from one location or country to another, and therefore should be specified by the user. Increases in salary costs with time will also be dependent on economic conditions in the country concerned.

1070 GRADE B STAFF — Definition as for Grade A staff.

See general notes under Grade A staff. Grade B is intended for senior clerical or sub-professional staff, and in the model is applied to staff responsible for document acquisition procedures and for keyboard operators.

1080 GRADE C STAFF — Definition as for Grade A staff.

See general notes under Grade A staff. Grade C is intended for professional staff and junior supervisors, and in the model is applied to all staff responsible for intellectual processing of input (e.g. indexers, abstractors, translators).

1090 GRADE D STAFF — Definition as for Grade A staff.

See general notes under Grade A staff. Grade D is intended for supervisors and middle management staff.

1100 GRADE E STAFF — Definition as for Grade A staff.

See general notes under Grade A staff. Grade E is intended for senior management responsible for the system.

1130 ORDER UNIT TIME (MONOGRAPHS) — Average time spent in ordering and receipt of non-serial documents, in hours.

This value will vary according to the nature of the system. Published values for university and polytechnic libraries in the UK and USA range from 0.2 to 0.63, but the higher values reflect the more

complex procedures entailed in large academic libraries. Less effort would normally be required for ordering publications as input to a mechanized system. A reasonable value for the purpose of the model would be 0.25.

1132 ORDER UNIT TIME (SERIALS) — Average time taken (in home) per serial for subscription ordering and renewal, control of receipts, and circulation.

Published data are lacking. Value will depend on complexity of procedures used, and average number of issues per title. On the basis of experience, an approximate value would be 1 hour, to cover all handling effort relating to one journal subscription in the course of a year.

1160 UNIT PURCHASE COST (MONOGRAPHS) — Average cost per item purchased.

An excellent source for this data is the annual survey of book and periodical prices published in the Library Association Record, the most recent appearing in the June 1976 issue. This gives average prices by subject class, typical figures for January - April 1976 being :-

Dewey class 500	£8.37
" " 600	£5.93

Price trends are also indicated, non-fiction works in general showing an increase of about 40% since the period July 1974 - June 1975.

1162 UNIT PURCHASE COST (SERIALS) — Average annual subscription cost.

The source mentioned above for monographs is equally applicable for periodicals. Average price for all

periodicals in the field of science and technology, in 1976, is £43.41. This figure represents an increase of 26.7% in comparison with 1975 prices.

1260 SELECT UNIT TIME — Average time taken per document input for scanning and selection, in hours.

Some fairly consistent times for this activity are shown in the OECD cost survey [5], ranging from 0.12 to 0.14 hr. The surveys by Drees [6] and Dubois and Peeters [7] show wider ranges, with values from 0.0003 to 0.40. Obviously this unit time is largely dependant on the number of items selected in relation to the total volume scanned. As a reasonable figure for use with the model, 0.10 is suggested.

1310 CATALOGUE UNIT TIME — Average time taken per document input for descriptive cataloguing, in hrs.

Times of 10 - 16 mins have been reported for polytechnic library systems, and 22 - 29 mins or more for university libraries. It is suspected that the higher values relate to cataloguing of library input rather than to IR systems. The unit time should typically cover scanning the item, writing bibliographic details on an input form, and checking authority files. A value of 0.25 is suggested for use with the model, as a general guide.

1360 INDEX UNIT TIME — Average time taken to index each input document, in hours.

Times reported in surveys of mechanized systems include :-

Vickers	0.12 - 0.98

Drees	0.13 - 3.33
Dubois & Peeters	0.017 - 0.60

A highly significant source for indexing times is however the Aslib Cranfield project [8], in which 10,000 documents were indexed by four different methods (UDC, facet, uniterm, alphabetical). The project team worked to standard times of 12 mins and 8 mins; some documents were also indexed by external collaborators without time control. The retrieval performance achieved by different indexing methods with different measures of input effort was compared. On the basis of this research, it would seem reasonable to use values of 0.13 hr for indexing of adequate depth, and 0.2 hr for a higher quality of indexing.*

1410 AUTHOR ABSTRACTS PERCENTAGE — Percentage of author abstracts used for input.

To be supplied by user. The extent to which author abstracts can be used will depend on several factors such as the type of journals from which input is selected; and the proportion of report literature (which will tend to contain author abstracts) as against the proportion of books (which do not).

1440 AUTHOR ABSTRACTS UNIT TIME — Average time (in hours) taken to prepare an author abstract, including copying and editing.

No reliable published figures have been found, but experience indicates that a time of 0.08 hours would be satisfactory.

1450 WRITTEN ABSTRACTS UNIT TIME — Average time (in hrs) taken to prepare original abstracts.

As in the case of indexing, the unit time to be used

* These values relate to conventional subject indexing. Higher values would apply in the case of special indexing techniques, such as the indexing of chemical compounds, using Wiswesser line notation.

here will depend on the quality of the proposed data base. Of the many published values (see for example refs. 5, 6, 7, 9), the lower values in the range 0.22 - 0.32 seem the most credible. For discussion of the factors affecting abstracting times see Wilkin et al [2] and Wolfe [10]. Suggested values for use with the model are :-

indicative abstract	- 0.25
short informative abstract	- 0.3
long informative abstract	- 0.5

1580 PERCENTAGE ITEMS TRANSLATED — Percentage of input records translated from one language to another.

To be supplied by user. Quantity will depend on proportion of foreign literature covered, and/or volume of input records supplied by collaborating centres in different countries. Note that this proportion may change as system develops.

1600 TRANSLATION UNIT TIME — Average time (in hrs) taken to translate an input record.

No published data found. Time required will relate to the length of the records (mainly the title and abstract), and experience indicates that unit times will be similar to those for abstracting (see 1450 above).

1680 AUTHOR ABSTRACT LENGTH

1690 WRITTEN ABSTRACT LENGTH — Average length, in characters, of complete input records (bibliographic reference, index terms, abstracts, etc.)

Typical record lengths, including abstracts, are :-

Engineering Index	1200
INSPEC	1700
Food Sci. 8 Tech. Abs.	1200
TITUS (ITF)	2000
IRRD	1300

The above are average figures for existing systems, using varying proportions of author abstracts. As a general rule, written abstracts will tend to be rather longer than author abstracts. Where the model user does not wish to lay down different length standards, the same figure could be used for lines 1680 and 1690. Note also that in a system which did not use abstracts, the value to be used could be in the range 100 - 400 chars.

1710 ON-LINE KEYBOARDING RATE

1730 OCR KEYBOARDING RATE

1750 OFF-LINE KEYBOARDING RATE

Keyboarding rate for data preparation, in key-strokes per hour.

Opinions differ as to the extent to which keyboarding rates vary from one technique to another. The options provided for in the model are on-line input, where the keyboard (terminal) is connected directly to the computer; optical character recognition, where input is typed on a special typewriter, and then read by an OCR reader; and off-line input, which includes the use of punched cards, paper tape, magnetic tape, key-to-disc, etc.

The authors' experience indicates that keyboarding rate is more affected by the nature of the work than by the technique used, except that higher rates can be achieved with high-volume work using key-to-disc or other 'pooled processor' methods.

Suggested values are :-

for bibliographic records using upper and lower case characters	6000 - 8000
for work using upper case alpha numeries only	9000 - 12000

Where verification is not used, and records have to be corrected after proof-reading, these values should be factored accordingly. The suggested allowance is a reduction of 10 per cent.

1810 VERIFICATION FACTOR — Percentage of input verified by second keyboarding.

100 per cent verification would entail re-keyboarding all input to check its accuracy - an expensive process. A fairly common practice is to verify only part of each record, such as the author, title, reference, and indexing fields. If verification is to be used (as distinct from proof-reading + correction), the model user has to decide what percentage of the input record is to be so treated. This process is not to be confused with validation, which refers to automatic checking of the content of each part of the record (see computer processing cost elements 2600, 2630, 2640).

1860 CONTRACT KEYBOARDING PERCENTAGE — Percentage of data preparation work carried out by external service bureaux.

In some existing systems, all or part of the data preparation is carried out by commercial service bureaux. The model invites the user to choose what proportion of the work will be so treated. This factor could change with time, if the volume of input were expected to increase, but some limitation

were imposed on the number of staff to be employed.

1870 CONTRACT KEYBOARDING COST — Price changed for data preparation by external service bureaux, per 1000 characters.

Bibliographic records tend to be more complex than much of the routine work handled by service bureaux, so will tend to be charged for at a higher rate. Rates also vary greatly from one location to another. Typical charges in the UK are £1 to £2 per 1000 characters.

Line 1730 above refers to OCR keyboarding done in-house, but it is also possible to get data preparation done in this way through a bureau. Competitive rates are sometimes quoted for keyboarding plus conversion to magnetic tape.

1950 ON-LINE DATA PREPARATION PERCENTAGE — Percentage of data preparation carried out on-line.

If all input is to be keyboarded on-line, the value here will be 100, but in some existing systems, part of the data preparation is done on-line, and the remainder off-line. The model allows for such a practice, and the user would have to indicate what percentage of the work is to be so treated.

2040 TERMINAL RENTAL — Cost of computer terminal, (for input) per year.

If the terminal(s) is to be purchased outright, the cost should be spread over 5 years. Otherwise a rental charge should be shown here. Prices and rental charges vary widely, but typical values in the UK would be :-

teletype	£800 - 1200 purchase cost

teletype	£300 - 360 annual rental
simple VDU	£1000 - 2000 purchase cost
simple VDU	£360 - 600 annual rental

The rental figures shown would be inclusive of maintenance, but up to 20 per cent should be added to figures based on purchase cost, to allow for this.

Rental charges will increase with time, unless covered by a long-term contract.

2060 COMMUNICATIONS COST — Cost per year of communications links between input terminal(s) and computer.

The value to be used here is dependent on system configuration, and many possibilities exist. The terminal(s) may be quite close to the computer, in which case the cost will be negligible; or it might be remotely located. Note that we are concerned here only with costs borne by the system, which will mainly be telephone line costs, plus the cost of equipment. The latter will generally include at least a modem, for which the rental would be £100 - £350 per annum, but in the case of a widely dispersed system might also include multiplexors, concentrators, etc. Telephone charges will be dependent on distance, line occupancy, and line capacity.

For the purpose of a rough estimate, a value could be derived from published telephone line charges. The future availability of EURONET will obviously have an effect on the cost value to be used in this part of the model.

2190 OCR TYPEWRITER RENTAL — Cost per year of renting special typewriter for OCR input.

Suggested value :

purchase cost (to be amortized)	£450
rental	£120 - 132 per year

As with other equipment, costs will increase with time.

2210 OCR CONVERSION COST — Cost of reading and converting OCR typed input to magnetic tape, per 1000 characters.

Few organizations operate their own OCR readers, so this cost will usually relate to a bureau operation. Charges quoted in the UK are about £0.75 per 1000 characters. Cost increase in time should be allowed for in using the model.

2340 KEYBOARD RENTAL — Annual rental (per unit) for data preparation equipment, other than items covered by 2040 and 2190 above.

As for item 2040, an equivalent rental (including maintenance charges) should be used here if equipment is to be purchased rather than rented. Typical values would be :-

card punch	£2000 - 4000 purchase cost
card punch	£480 - 960 annual rental
paper tape punch	£1000 - 5000 purchase cost
paper tape punch	£360 - 1080 annual rental

Rental charges will increase with time, unless covered by a long-term contract.

2510 PROOF-READING UNIT TIME — Average time (in hours) taken to proof-read an input record.

No reliable published data found. On the basis of experience, a value of between 0.03 and 0.08 is suggested.

2600 BUREAU RATE FOR COMPUTER PROCESSING — Cost per record for all computer processing associated with data-base creation.

The model provides for two alternative ways of entering the cost of computer processing. The user can enter a value here in the form of a cost per record, or employ the total rental x occupancy approach of lines 2630 and 2640. Although called 'bureau rate' line 2600 could equally well be used for an in-house situation where charging on a pro-rata basis was preferred. Values will be highly system-dependent, according to the complexity of the computer processing required. The OECD survey [5] quotes several values based on 1972 data, ranging from 0.04 to 2.26 dollars. It is suggested that a reliable value could best be obtained by consulting one or more computer bureaux that should be able to quote realistic figures.

Computer costs will increase with time but generally at a modest rate.

2630 COMPUTER OCCUPANCY — Percentage of computer operating capability used for input processing.

Computer processing costs can be input to the model in the form of computer rental (2640) multiplied by

the percentage of the computer's capacity used for input processing. This method of costing is advocated in the EFAG 2 report 11, but may prove difficult to apply for this part of the model until some values have been collected from existing systems. It could be especially appropriate in the case of systems using dedicated mini-computers.

2640 IN-HOUSE COMPUTER RENTAL — Total cost per year of computer installation.

See notes for 2630 above.

3450 SUPERVISORS - GRADE C
3460 SUPERVISORS - GRADE D
3470 SUPERVISORS - GRADE E
3472 CLERICAL SUPPORT STAFF GRADE A

— Number of staff required in each grade

As explained in Section 4.1.9 the model user is required to designate the numbers of supervisory and clerical support staff required, in the light of the numbers of direct staff calculated by the model. The provision of staff in these grades should allow for system maintenance (including thesaurus maintenance) and development work. The intended levels of seniority of the three supervisory grades are indicated at lines 1080, 1090 and 1100. For a multi-year projection, these numbers may need to be adjusted from one year to another.

3490 SPACE PER STAFF MEMBER — Average working area allowed per staff member.

Standards of accommodation vary from one organi-

zation to another, but the following gives a rough indication of generally accepted space allowances :-

	sq. ft.	sq. metres
senior admin.staff	200-400	18-36
professional staff	100-150	9-14
clerical staff	50-80	4.5 - 7.5
typing staff	40-60	4 - 5.5

The model calls for only one value, which could be estimated on the basis of the mix of staff to be employed.

3500 SPACE RENTAL — Annual cost per sq.foot/ sq. metre (depending on unit used for 3490) of accommodation.

Again, the value to be used here will be location-dependent. It should represent an economic cost including rates, cleaning, etc. Substantial increases in time should be allowed for.

3640 OVERHEAD RATE — Overhead cost expressed as a percentage of salary costs.

To be supplied by user. This factor has to cover all indirect organizational costs other than accommodation, including stationery and other consumable items not specified above.

CHAPTER 6: TESTING THE MODEL

Testing of the computer-based model, during the course of its development, has taken two forms :

(1) test runs with different sets of data values, to ensure that the definition file would operate correctly under a variety of conditions;

(2) simulation of known systems. For these tests, details of a known system and its technical features were fed into the model, and the calculated costs compared with available cost data for the system.

The latter exercises proved invaluable in refining the model, and helped in indicating acceptable limits for certain data values.

The project specification calls for a written specification for a designed experiment to implement the model. The ideal way to check the validity of the model's predictions would, of course, be to design a system; use the model to predict its costs; implement the system; and then compare its costs with the predictions. Unfortunately, such an approach is impractical.

The only practical solution would seem to be to use the model in a retrospective mode, i.e. to make a cost prediction for an existing system as of some time in the past, and compare the results with the actual costs experienced by the system.

In designing any experiment to test the model, three important factors have to be borne in mind. The first is that the model will work best for a user with some knowledge and experience of the environment in which the system will operate. Many of the data values called for will depend on local conditions (e.g. salary rates, computer processing charges, accommodation costs, and overhead rates).

The second factor is one previously referred to in section 3.2.4 - that the model predictions can serve as a self-fulfilling prophecy. In a real-life situation, it should be possible to manage the system in such a way that it would operate within the cost limits predicted by the model. This will not apply if the model is checked against an existing system.

The third factor concerns the accuracy expected of the model. The accuracy required will depend on the purpose for which the model is used. The accuracy achieved will depend on the quality of the data that is fed into the model, coupled with the design of the model itself, which embodies a certain level of approximation. The test we shall describe does not suggest that the model would be deemed to fail, if it did not achieve a specific level of accuracy. The level of accuracy would be measured, and the model judged subjectively.

The specification for the test is given in Appendix 9.

CHAPTER 7: RECOMMENDATIONS

Over and above the test of the model discussed in the previous chapter, we believe that the model could usefully be developed for specific applications. In its present form, it is suitable for making cost predictions at the broad planning level. In the course of the project, interest has been expressed in the use of cost modelling techniques by system operators. Their requirement is for a model into which could be fed details of current operational volumes and costs for a specific system, and which the operator could use to determine the effect of changes in methods, staffing, throughput volumes, etc.

The model would need to be modified to fulfil this role in an effective manner. Since the model would be working on actual cost data of an existing system, it would be possible to dispense with certain features designed to deal with areas of uncertainty. Also the user interface of the model would need to be redesigned with this application in mind.

We therefore recommend that further research on these lines be initiated by the Commission, or by some other interested organization.

APPENDIX 1 - SPECIFICATION OF PROJECT 3

Project 3: Development and use of models for the prediction of costs for alternative information systems

A. Objectives

- To develop models for predicting the costs of various methods of data base creation and provision of information services.

B. Source Material

- The costs of Mechanised Information Systems. - P. Vickers; a study carried out for the OECD Directorate for Scientific Affairs, 1974.

- The costs of Scientific and Technical Information and Documentation Systems. - G. Drees; a study carried out for the CIDST-Brussels Working Party on Pricing, 1974.

- Results of Project 1.

C. Details of project

The study should be carried out in two phases:

- Design and testing of the models;
- Implementation in an experimental environment.

Each of the phases will be broken down into two separate parts. The first part will be concerned with the various methods of data base creation and the second part with the provision of services.

Phase 1, Part 1

- (a) to develop and test a cost prediction model for the input activities of mechanised information systems;

- (b) a written specification for a designed experiment to implement the model in (a) above.

(Phase 2, Part 1)

Phase 1, Part 2

- (a) to develop and test a cost prediction model for the output activities of mechanised information systems, i.e. provision of information services;

- (b) a written specification for a designed experiment to implement the model in (a) above.

(Phase 2, Part 2).

The contractor should produce a separate report for both parts of this study.

APPENDIX 2 - REFERENCES

1. Bourne, C.P. and Ford, D.F. 'Cost analysis and simulation procedures for the evaluation of large information systems'. Amer. Doc., vol. 15, p. 142-9 (April 1964).

2. Wilkin, A.P., Reynolds, R., and Robertson, S.E. 'Standard times for information systems; a method for data collection and analysis.' J. Documentation, vol. 28, p. 131-150(June 1972).

3. Lancaster, F.W. and Climenson, W.D. 'Evaluating the economic efficiency of a document retrieval system.' J. Documentation, vol. 24, p. 16-40 (March 1968).

4. King, D.W. and Caldwell, N.W. 'Study of the cost-effectiveness of retrospective search systems.' ERIC Clearinghouse on Library and Information Sciences, ref. ED 046446 (June 1970).

5. Vickers, P. 'The costs of mechanized Information systems.' Study carried out for the OECD Directorate for Scientific Affairs (1974).

6. Drees, G. 'The costs of scientific and technical information and documentation systems.' Study carried out for the CIDST - Brussels Working Party on Pricing (1974).

7. Dubois, G. & Peeters, E. 'Cost of automatic processing of documentary information.' Report presented at National Conference on Documentation, Brussels (1974).

8. Cleverdon, C.W. 'Aslib Cranfield Research Project: Report on the first stage of an investigation into the comparative efficiency of indexing systems.' College of Aeronautics, Cranfield, September 1960.

9. Bourne, C.P. and others. 'Abstracting and indexing rates and costs: a literature review.' ERIC Clearinghouse on Library and Information Sciences, Review Series No. 5, ref. ED 043 798 (May 1970).

10. Wolfe, J.N. et al. 'The economics of technical information systems.' New York, Praeger Publishers, 1974.

11. Vickers, P.H., 'Final report on Project 2: Extension and revision of the cost/accounting scheme to interactive systems of the network'. Report prepared for the Commission of the European Communities, June 1976.

APPENDIX 3 - DESCRIPTION OF PROPHIT II SYSTEM

Prophit II is a powerful and highly flexible reporting system, designed to assist managers and analysts in planning, analysing, projecting, tracking and controlling business plans and performance. In its simplest form, PROPHIT II can be visualized as a "computerized columnar spread sheet" but it can be adapted to a variety of uses.

In any application, however, two separate files have to be prepared when using the system. One, the Data File, contains the numerical values to be used; the other is the Definition File, which embodies the logic.

A PROPHIT II "definition file" is a line-by-line description of the thinking behind a spread sheet. Definition files are often called "models". Each line of the model will have one line in the definition file, and that line defines :-

1. The number and title of the line.
2. The operation that is to be done (e.g. "Read a line from the data file" or "Add two lines").

A data file contains not only the numbers (values) required by a definition file, but also a 'prologue' containing the title and headings of the report. Each data line is typed with a four-digit line number to correspond to the line number of a "read data" line in the definition file.

There are two types of data files - "history files" and "projection files".

A data line in a history file typically contains the line number and the appropriate number of values. For instance, if five years of data were called for, a line of history file might look like this :-

5000 234, 253, 276, 217, 298

A data line in a projection file includes a "projection type" so that users can enter data in the same way they think of it - in patterns. For instance, a projection type 3 allows a starting number and a growth amount to be entered :

5000 3, 250, 20 would be the equivalent of

5000 250, 270, 290, 310, 330

The 18 projection types cover a range of patterns, from "none" (entering specific data values for each period) to "least-squares projection from historical periods". Projection types applicable to files where no history data is available are shown in Fig. 1.

After a definition and a data file have been typed and saved, PROPHIT II is run. It will ask for a command, and then will perform the "commanded" function. For example, the system will create a report file in a specified format. Most commands ask one or more questions so the operator can specify which files to use. Some, like the PRINT command, ask specific questions that give the operator additional options.

Of particular interest in the application described in this report is the WHAT-IF command, which allows the user to explore the effects of changes in assumptions. Coupled with this is the SENSITIVITY command,

which enables the user to print out the differences arising from a WHAT-IF command. Sensitivity reports can be of the actual differences or percentage differences.

PROJECTION TYPES

(Projections Unrelated to History Data)

1	DETAILED	Value for each projection column. 1000 1, 60,64,68,73,79,85
1.3	DETAILED & LINEAR (Type 1 & 3)	Count of detail values followed by actual values, standard increment of change for rest of projection columns. 1000 1.3, 3,40,-7,59, 18
1.8	DETAILED & COMPOUND (Type 1 & 8)	Count of detail values followed by actual values, percentage value for compounding rest of projection columns. 1000 1.8, 2,65,68, 4.75
2	STEPPED	Value for initial column, to be assigned all succeeding ones until a change occurs. To specify change, enter location (3rd projection column is location 3) and new value. This value is now assigned to columns up to next change. Entry must always end with two zeros (0, 0). 1000 2, 65,3,90,6,110,0,0
3	LINEAR (Increment)	Value for first column, standard increment of change. 1000, 3, 700,25
3.1	LINEAR (Start-end)	Value for first and last columns. Intervening columns projected to increase by a constant increment. 1000 3.1, 100,600
5	CONSTANT	One value assigned all projection columns. 1000 5, 600

PROJECTION TYPES

(Projections Related to History Data)

4	CONTINUE HISTORY	All columns receive value of history column preceding the first projection column. 1000 4
6	AVERAGE OF HISTORY	Average of history assigned to all projection columns. 1000 6
6.1	STRAIGHT LINE	Best-fit curve (least squares) of first degree determined for preceding history and continued through projections. 1000 6.1
6.2	QUADRATIC, PARABOLIC	Same as type 6.1, except second degree curve. 1000 6.2
8	COMPOUND (Single rate)	Enter percent. First column compounded from last history; compounding at this rate continues through rest of columns. 1000 8, 5.25
9	COMPOUND (Separate rates)	Same as type 8 except separate percents for each column. 1000 9,3.5,4,4,4.25,4.5,5
10	CHANGE DEFINITION	Can change line in report definition or insert new line (cannot insert a Type 28 or 40). See manual for full details. 1000 10, 12,1,1,48.5,2010
0	DELETE LINE	Line in report is treated as null, nonprinting (Type 29). 1000 0

Negate history line's data

NOTE: Enter 5.25 percent as 5.25; values, percents, increments can be ± values.

Fig. 1 - Projection types

Finally, within this review of only a few of the PROPHIT II features, it is useful to mention the ILLUSTRATE facility. This produces an explanation in plain language of the logic employed in the definition file.

Examples of the use of the above features are shown in the Appendices which follow.

APPENDIX 4 - STRUCTURE OF COMPUTER-BASED MODEL

The listing which follows was prepared by using the ILLUSTRATE feature of the PROPHIT II system It presents in plain language the operations required by the definition file (DEFIN) for our model. 'READ DATA' lines relate to data required by the model, which are ordinarily supplied from the equivalent line in the projection file (or history file, if used). These data values are defined in Chapter 5 of this report.

Certain elements of the computer-based model may need further explanation :

(1) at line 2930, the computer checks the total direct costs by one method of summation against another, to ensure that no anomalies are present.

(2) after line 2936, there is a section in which numbers of staff for each activity are rounded up to whole numbers. It will be noted that staff grades are identified B1, B2, etc. This is merely a device to separate non-interchangeable staff of any grade.

DEFINITION FILE 'DEFIN AS OF 7/19/76 PAGE 1

LINE	ACTION
1010 ITEMS INPUT	READ DATA
1020 MONOGS ACQD	READ DATA
1022 JOURNALS ACQD	READ DATA
1030 MONOGS PCHSD	READ DATA
1032 JOURNALS PCHSD	READ DATA
1040 MAN YEAR HRS	READ DATA
1050 GRADE RATES	
1060 GRADE A	READ DATA
1070 GRADE B	READ DATA
1080 GRADE C	READ DATA
1090 GRADE D	READ DATA
1100 GRADE E	READ DATA
1110 ACQUISITION COSTS	
1130 MONOG ORDER TM	READ DATA
1132 JNL ORDER TM	READ DATA
1140 ORDER EFFORT	MONOGS ACQD(1020) X MONOG ORDER TM(1130) / MAN YEAR HRS(1040) + JOURNALS ACQD(1022) X JNL ORDER TM(1132) / MAN YEAR HRS(1040)
1150 SALARY COST	COPY GRADE B(1070)
1160 MONOG UNIT COSTS	READ DATA
1162 JOURNAL UNIT COST	READ DATA
1170 PURCHASE COST	MONOGS PCHSD(1030) X MONOG UNIT COSTS(1160) + JOURNALS PCHSD(1032) X JOURNAL UNIT COST(1162)
1180 LABOUR COST	ORDER EFFORT(1140) X SALARY COST(1150)
1200 TOTAL ACQN COST	-------- + PURCHASE COST(1170) + LABOUR COST(1180)
1220 UNIT ACQN COST	-------- 1 X TOTAL ACQN COST(1200) / ITEMS INPUT(1010)
1230 INTEL OPS	
1240 ITEMS INPUT	COPY ITEMS INPUT(1010)
1260 SEL UNIT TM	READ DATA
1270 SEL EFFORT	SEL UNIT TM(1260) X ITEMS INPUT(1240) / MAN YEAR HRS(1040)
1280 SALARY COST	COPY GRADE C(1080)
1290 SELECTION	SEL EFFORT(1270) X SALARY COST(1280)
1310 CAT UNIT TM	READ DATA
1320 CAT EFFORT	CAT UNIT TM(1310) X ITEMS INPUT(1240) / MAN YEAR HRS(1040)

```
      LINE                    ACTION
---------------------  ----------------------------------------
*
1330 SALARY COST          COPY GRADE C(1080)
1340 CATALOGUING          CAT EFFORT(1320)
                          X SALARY COST(1330)

1360 INDEX UNIT TM        READ DATA
1370 SALARY COST          COPY GRADE C(1080)
1380 INDEX EFFORT         INDEX UNIT TM(1360)  X
                          ITEMS INPUT(1240)  /
                          MAN YEAR HRS(1040)
1390 INDEXING             SALARY COST(1370)
                          X INDEX EFFORT(1380)

1410 AUTH ABSTS PC        READ DATA
1420 NO A ABSTS           1.00000E-02 X
                          ITEMS INPUT(1240)
                          X
                          AUTH ABSTS PC(1410)
1430 NO W ABSTS           ITEMS INPUT(1240)
                          - NO A ABSTS(1420)
1440 A ABSTS UNIT TM      READ DATA
1450 W ABSTS UNIT TM      READ DATA
1460 A ABSTS EFFORT       NO A ABSTS(1420)  X
                          A ABSTS UNIT TM(1440)  /
                          MAN YEAR HRS(1040)
1470 W ABSTS EFFORT       W ABSTS UNIT TM(1450)  X
                          NO W ABSTS(1430)  /
                          MAN YEAR HRS(1040)
1480 SALARY COSTS         COPY GRADE C(1080)
1490 ABSTRACTING          A ABSTS EFFORT(1460)  X
                          SALARY COSTS(1480)  +
                          W ABSTS EFFORT(1470)  X
                          SALARY COSTS(1480)

1580 ITEMS TRANSL PC      READ DATA
1590 NO TRANSLATED        1.00000E-02 X
                          ITEMS INPUT(1240)
                          X
                          ITEMS TRANSL PC(1580)
1600 TRANS UNIT TM        READ DATA
1610 TRANSL EFFORT        TRANS UNIT TM(1600)  X
                          NO TRANSLATED(1590)  /
                          MAN YEAR HRS(1040)
1620 SALARY COSTS         COPY GRADE C(1080)
1630 TRANSLATION          TRANSL EFFORT(1610)
                          X SALARY COSTS(1620)

                          --------
1660 INTEL OPS COSTS      + SELECTION(1290)
                          + CATALOGUTNG(1340)
                          + INDEXING(1390)
                          + ABSTRACTING(1490)
                          + TRANSLATION(1630)
                          --------
1680 A ABSTS LENGTH       READ DATA
```

LINE	ACTION
1690 W ABSTS LENGTH	READ DATA
1700 KEYBOARD RATES	
1710 ON-LINE	READ DATA
1730 OCR	READ DATA
1750 OFF-LINE	READ DATA
1790 OCR OR OTHER	READ DATA
1800 VERIFY YES-NO	READ DATA
1810 VERIFICATION FACT	READ DATA
1820 PROOF IN-HOUSE WK	READ DATA
1840 ITEMS INPUT	COPY ITEMS INPUT(1010)
1860 CONTRACT PC	READ DATA
1870 CONTRACT RATE	READ DATA
1880 NO CONTRACTED	1.00000E-02 X CONTRACT PC(1860) X ITEMS INPUT(1010)
1890 AVGE LENGTH	NO A ABSTS(1420) X A ABSTS LENGTH(1680) / ITEMS INPUT(1010) + NO W ABSTS(1430) X W ABSTS LENGTH(1690) / ITEMS INPUT(1010)
1900 CHARS	NO CONTRACTED(1880) X AVGE LENGTH(1890)
1910 CONTRACT COST	CONTRACT RATE(1870) X CHARS(1900) / 1000
1920 UNIT CNTRT COST	1 X CONTRACT COST(1910) / NO CONTRACTED(1880)
1940 IN-HOUSE PREP	
1950 ON-LINE PC	READ DATA
1960 NO IN-HOUSE	ITEMS INPUT(1010) - NO CONTRACTED(1880)
1970 NO ON-LINE	1.00000E-02 X ON-LINE PC(1950) X NO IN-HOUSE(1960)
1980 ON-LINE RATE	COPY ON-LINE(1710)
1990 ON-LINE EFFORT	NO ON-LINE(1970) X AVGE LENGTH(1890) / ON-LINE RATE(1980) / MAN YEAR HRS(1040)
2000 SALARY COST	COPY GRADE B(1070)
2010 ON-LINE LABOUR	ON-LINE EFFORT(1990) X SALARY COST(2000)
2020 ADDNL MCS	BREAK LEVEL OF ON-LINE EFFORT(1990) INCREMENTS OF 1

LINE	ACTION
2024 TERMINALS	1 + 1 X ADDNL MCS(2020)
2030 TERMINALS	IF ON-LINE PC(1950) GT 0 THEN TERMINALS(2024) ELSE 0
2040 RENTAL RATE	READ DATA
2050 TERMINAL COST	TERMINALS(2030) X RENTAL RATE(2040)
2060 COMS COST	READ DATA
2080 ON-LINE COSTS	+ ON-LINE LABOUR(2010) + TERMINAL COST(2050) + COMS COST(2060)
2100 UNIT ON-LINE COST	1 X ON-LINE COSTS(2080) / NO ON-LINE(1970)
2110 KEYBOARD RATE	COPY OCR(1730)
2120 NO KEYBOARDED	NO IN-HOUSE(1960) - NO ON-LINE(1970)
2130 NO KEYBOARDED	IF OCR OR OTHER(1790) EQ 2 THEN 0 ELSE NO KEYBOARDED(2120)
2140 KEYBOARD EFFORT	NO KEYBOARDED(2130) X AVGE LENGTH(1890) / KEYBOARD RATE(2110) / MAN YEAR HRS(1040)
2150 SALARY COST	COPY GRADE B(1070)
2160 OCR LABOUR	KEYBOARD EFFORT(2140) X SALARY COST(2150)
2170 ADDNL OCR MCS	BREAK LEVEL OF KEYBOARD EFFORT(2140) INCREMENTS OF 1
2174 OCR TYPEWRITERS	1 + 1 X ADDNL OCR MCS(2170)
2180 OCR TYPEWRITERS	IF OCR OR OTHER(1790) EQ 1 THEN OCR TYPEWRITERS(2174) ELSE 0
2190 RENTAL	READ DATA
2200 OCR MC COST	OCR TYPEWRITERS(2180) X RENTAL(2190)
2210 CONVERSION COST	READ DATA
2220 CONVERSION	NO KEYBOARDED(2130) X AVGE LENGTH(1890) / 1000 X CONVERSION COST(2210)
2230 OCR COSTS	+ OCR LABOUR(2160) + OCR MC COST(2200) + CONVERSION(2220)
2240 OCR COSTS	IF OCR OR OTHER(1790) EQ 1 THEN OCR COSTS(2230) ELSE 0

LINE	ACTION
2260 KEYBOARD RATE	COPY OFF-LINE(1750)
2270 NO KEYBOARDED	NO IN-HOUSE(1960) - NO ON-LINE(1970)
2280 NOKEYBOARDED	IF OCR OR OTHER(1790) EQ 1 THEN 0 ELSE NO KEYBOARDED(2270)
2290 KEYBOARD EFFORT	NOKEYBOARDED(2280) X AVGE LENGTH(1890) / KEYBOARD RATE(2260) / MAN YEAR HRS(1040)
2300 SALARY COST	COPY GRADE B(1070)
2310 KEYBOARD LABOUR	KEYBOARD EFFORT(2290) X SALARY COST(2300)
2320 ADDNL KEYBOARD	BREAK LEVEL OF KEYBOARD EFFORT(2290) INCREMENTS OF 1
2324 KEYBOARDS	1 + 1 X ADDNL KEYBOARD(2320)
2330 KEYBOARDS	IF OCR OR OTHER(1790) EQ 2 THEN KEYBOARDS(2324) ELSE 0
2340 RENTAL	READ DATA
2350 KEYBOARDS RENTAL	KEYBOARDS(2330) X RENTAL(2340)
2360 KEYBOARD COSTS	+ KEYBOARD LABOUR(2310) + KEYBOARDS RENTAL(2350)
2370 KEYBOARD COSTS	IF OCR OR OTHER(1790) EQ 2 THEN KEYBOARD COSTS(2360) ELSE 0
2380 NO KEYBOARDED	IF VERIFY YES-NO(1800) EQ 1 THEN NO KEYBOARDED(2270) ELSE 0
2390 VERIFY FACTOR	COPY VERIFICATION FACT(1810)
2400 NO KEYBOARDED	NO KEYBOARDED(2380) X VERIFY FACTOR(2390)
2410 KEYBOARD EFFORT	AVGE LENGTH(1890) X NO KEYBOARDED(2400) / OFF-LINE(1750) / MAN YEAR HRS(1040)
2420 SALARY COST	COPY GRADE B(1070)
2430 VERIFY LABOUR	KEYBOARD EFFORT(2410) X SALARY COST(2420)
2440 ADDNL KEYBDS	BREAK LEVEL OF KEYBOARD EFFORT(2410) INCREMENTS OF 1
2444 KEYBOARDS	1 + 1 X ADDNL KEYBDS(2440)
2450 KEYBOARDS	IF VERIFY YES-NO(1800) EQ 1 THEN KEYBOARDS(2444) ELSE 0
2460 RENTAL	COPY RENTAL(2340)

```
       LINE                    ACTION
---------------------  ----------------------------------------
*
2470 KEYBOARDS RENTAL     KEYBOARDS(2450)
                          X RENTAL(2460)
2480 VERIFY COST          + VERIFY LABOUR(2430)
                          + KEYBOARDS RENTAL(2470)
2490 VERIFY COST          IF VERIFY YES-NO(1800) EQ 1
                          THEN VERIFY COST(2480)
                          ELSE 0

2500 NO IN-HOUSE          COPY NO IN-HOUSE(1960)
2510 PROOF UNIT TM        READ DATA
2520 NO PROOFED           IF PROOF IN-HOUSE WK(1820) EQ 0
                          THEN NO IN-HOUSE(2500)
                          ELSE 0
2530 PROOF EFFORT         NO PROOFED(2520)  X
                          PROOF UNIT TM(2510)  /
                          MAN YEAR HRS(1040)
2540 SALARY COST          COPY GRADE C(1080)
2550 PROOF LABOUR         PROOF EFFORT(2530)
                          X SALARY COST(2540)
2560 PROOF LABOUR         IF PROOF IN-HOUSE WK(1820) EQ 0
                          THEN PROOF LABOUR(2550)
                          ELSE 0

2580 COMPUTER OPS
2590 IN-HOUSE BUREAU      READ DATA
2600 BUREAU RATE          READ DATA
2610 BUREAU COST          ITEMS INPUT(1840)
                          X BUREAU RATE(2600)
2620 BUREAU COST          IF IN-HOUSE BUREAU(2590) EQ 1
                          THEN BUREAU COST(2610)
                          ELSE 0
2630 OCCUPANCY            READ DATA
2640 IN-HOUSE RENTAL      READ DATA
2650 IN-HOUSE COST        IN-HOUSE RENTAL(2640)
                          X OCCUPANCY(2630)
2660 IN-HOUSE COSTS       IF IN-HOUSE BUREAU(2590) EQ 0
                          THEN IN-HOUSE COST(2650)
                          ELSE 0
                          --------
2680 DATA PREP            + CONTRACT COST(1910)
                          + ON-LINE COSTS(2080)
                          + OCR COSTS(2240)
                          + KEYBOARD COSTS(2370)
                          + VERIFY COST(2490)
                          + PROOF LABOUR(2560)
2690 COMPUTER OPS         + BUREAU COST(2620)
                          + IN-HOUSE COSTS(2660)
                          --------
2710 MECHANICAL OPS       + DATA PREP(2680)
                          + COMPUTER OPS(2690)
                          ========

2724 DIRECT INPUT COSTS
2730 LABOUR               + LABOUR COST(1180)
                          + INTEL OPS COSTS(1660)
                          + ON-LINE LABOUR(2010)
                          + OCR LABOUR(2160)
```

LINE	ACTION
2740 LABOUR	+ KEYBOARD LABOUR(2310) + VERIFY LABOUR(2430) + PROOF LABOUR(2550)
2750 LABOUR	+ LABOUR(2730) + LABOUR(2740)
2770 MATERIALS	COPY PURCHASE COST(1170)
2790 TERMINALS	IF ON-LINE PC(1950) GT 0 THEN TERMINAL COST(2050) ELSE 0
2800 LINECOST	IF ON-LINE PC(1950) GT 0 THEN COMS COST(2060) ELSE 0
2810 OCR MCS	IF OCR OR OTHER(1790) EQ 1 THEN OCR MC COST(2200) ELSE 0
2820 KBD RENT	IF OCR OR OTHER(1790) EQ 2 THEN KEYBOARDS RENTAL(2350) ELSE 0
2830 VERIFY KBD	IF VERIFY YES-NO(1800) EQ 1 THEN KEYBOARDS RENTAL(2470) ELSE 0
2840 COMPUTER	COPY IN-HOUSE COSTS(2660)
2850 EQUIPMENT	SUM TERMINALS(2790) THRU COMPUTER(2840)
2870 SERVICES	+ CONTRACT COST(1910) + CONVERSION(2220) + BUREAU COST(2620)
2900 DIRECT INPUT COSTS	-------- + LABOUR(2750) + MATERIALS(2770) + EQUIPMENT(2850) + SERVICES(2870) --------
2920 ALL INPUT OPS	+ TOTAL ACQN COST(1200) + INTEL OPS COSTS(1660) + MECHANICAL OPS(2710)
2930 OUT OF BALANCE	ALL INPUT OPS(2920) - DIRECT INPUT COSTS(2900)
2936 STAFF REQUIRED	
3060 GRADE B1 EFFORT	SUM ITEMS INPUT(1010) THRU OUT OF BALANCE(2930)
3070 EXTRA STAFF	BREAK LEVEL OF GRADE B1 EFFORT(3060) INCREMENTS OF 1
3080 ADD ONE	1 + 1 X EXTRA STAFF(3070)
3090 GRADE B1 STAFF	IF GRADE B1 EFFORT(3060) GT 0 THEN ADD ONE(3080) ELSE 0

LINE	ACTION
3100 GRADE B2 EFFORT	SUM ITEMS INPUT(1010) THRU OUT OF BALANCE(2930)
3110 EXTRA STAFF	BREAK LEVEL OF GRADE B2 EFFORT(3100) INCREMENTS OF 1
3120 ADD ONE	1 + 1 X EXTRA STAFF(3110)
3130 GRADE B2 STAFF	IF GRADE B2 EFFORT(3100) GT 0 THEN ADD ONE(3120) ELSE 0
3140 GRADE B3 EFFORT	SUM ITEMS INPUT(1010) THRU OUT OF BALANCE(2930)
3150 EXTRA STAFF	BREAK LEVEL OF GRADE B3 EFFORT(3140) INCREMENTS OF 1
3160 ADD ONE	1 + 1 X EXTRA STAFF(3150)
3170 GRADE B3 STAFF	IF GRADE B3 EFFORT(3140) GT 0 THEN ADD ONE(3160) ELSE 0
3180 GRADE C1 EFFORT	SUM ITEMS INPUT(1010) THRU OUT OF BALANCE(2930)
3190 EXTRA STAFF	BREAK LEVEL OF GRADE C1 EFFORT(3180) INCREMENTS OF 1
3200 ADD ONE	1 + 1 X EXTRA STAFF(3190)
3210 GRADE C1 STAFF	IF GRADE C1 EFFORT(3180) GT 0 THEN ADD ONE(3200) ELSE 0
3220 GRADE C2 EFFORT	SUM ITEMS INPUT(1010) THRU OUT OF BALANCE(2930)
3230 EXTRA STAFF	BREAK LEVEL OF GRADE C2 EFFORT(3220) INCREMENTS OF 1
3240 ADD ONE	1 + 1 X EXTRA STAFF(3230)
3250 GRADE C2 STAFF	IF GRADE C2 EFFORT(3220) GT 0 THEN GRADE C2 STAFF(3250) ELSE 0
3260 GRADE C3 EFFORT	SUM ITEMS INPUT(1010) THRU OUT OF BALANCE(2930)
3270 EXTRA STAFF	BREAK LEVEL OF GRADE C3 EFFORT(3260) INCREMENTS OF 1
3280 ADD ONE	1 + 1 X EXTRA STAFF(3270)
3290 GRADE C3 STAFF	IF GRADE C3 EFFORT(3260) GT 0 THEN ADD ONE(3280) ELSE 0

```
       LINE                      ACTION
---------------------  ----------------------------------------

3300 GRADE D1 EFFORT     SUM ITEMS INPUT(1010)
                         THRU OUT OF BALANCE(2930)
3310 EXTRA STAFF         BREAK LEVEL OF
                         GRADE D1 EFFORT(3300)
                         INCREMENTS OF 1
3320 ADD ONE             1 +
                          1 X
                         EXTRA STAFF(3310)
3330 GRADE D1 STAFF      IF GRADE D1 EFFORT(3300) GT 0
                         THEN ADD ONE(3320)
                         ELSE 0
3340 GRADE D2 EFFORT     SUM ITEMS INPUT(1010)
                         THRU OUT OF BALANCE(2930)
3350 EXTRASTAFF          BREAK LEVEL OF
                         GRADE D2 EFFORT(3340)
                         INCREMENTS OF 1
3360 ADD ONE             1 +
                          1 X
                         EXTRASTAFF(3350)
3370 GRADE D2 STAFF      IF GRADE D2 EFFORT(3340) GT 0
                         THEN ADD ONE(3360)
                         ELSE 0
3380 GRADE D3 EFFORT     SUM ITEMS INPUT(1010)
                         THRU OUT OF BALANCE(2930)
3390 EXTRA STAFF         BREAK LEVEL OF
                         GRADE D3 EFFORT(3380)
                         INCREMENTS OF 1
3400 ADD ONE             1 +
                          1 X
                         EXTRA STAFF(3390)
3410 GRADE D3 STAFF      IF GRADE D3 EFFORT(3380) GT 0
                         THEN ADD ONE(3400)
                         ELSE 0
                         --------
3430 DIRECT STAFF        SUM GRADE B1 EFFORT(3060)
                         THRU GRADE D3 STAFF(3410)
3450 SUPERVISORS GRADEC  READ DATA
3460 SUPERVISORS GRADED  READ DATA
3470 SUPERVISORS GRADEE  READ DATA
3472 CLERKS GRADE A      READ DATA
                         --------
3480 TOTAL STAFF         SUM DIRECT STAFF(3430)
                         THRU CLERKS GRADE A(3472)
                         --------
3484 OVERHEADS
3486 TOTAL STAFF         COPY TOTAL STAFF(3480)
3490 SPACE PER PERSON    READ DATA
3500 RENTAL              READ DATA
3510 ACCOMODATION COST   TOTAL STAFF(3480)
                         X SPACE PER PERSON(3490)
                         X RENTAL(3500)
3520 SUPERVISORS C COST  SUPERVISORS GRADEC(3450)
                         X GRADE C(1080)
3530 SUPERVISORS D COST  SUPERVISORS GRADED(3460)
                         X GRADE D(1090)
```

```
      LINE                    ACTION
---------------------  ----------------------------------------

3540 SUPERVISORS E COST  SUPERVISORS GRADEE(3470)
                         X GRADE E(1100)
3550 GRADE A STAFF       COPY CLERKS GRADE A(3472)
3560 GRADE B STAFF       SUM GRADE B1 EFFORT(3060)
                         THRU GRADE B3 STAFF(3170)
3570 GRADE C STAFF       SUM GRADE C1 EFFORT(3180)
                         THRU GRADE C3 STAFF(3290)
3580 GRADE D STAFF       SUM GRADE D1 EFFORT(3300)
                         THRU GRADE D3 STAFF(3410)
3590 GRADE A SALARY      GRADE A STAFF(3550)
                         X GRADE A(1060)
3600 GRADE B SALARY      GRADE B STAFF(3560)
                         X GRADE B(1070)
3610 GRADE C SALARY      GRADE C STAFF(3570)
                         X GRADE C(1080)
3620 GRADE D SALARY      GRADE D STAFF(3580)
                         X GRADE D(1090)
3630 ALL SALARIES        SUM SUPERVISORS C COST(3520)
                         THRU GRADE D SALARY(3620)
3640 OVERHEAD RATE       READ DATA
3650 SALARY OVERHEAD     1.00000E-02 X
                         ALL SALARIES(3630)
                         X
                         OVERHEAD RATE(3640)
3660 ACCOMMODATION COST  COPY ACCOMODATION COST(3510)
                         --------
3680 OVERHEAD COSTS      + SALARY OVERHEAD(3650)
                         + ACCOMMODATION COST(3660)
                         --------

3700 PROJ INPUT COSTS

3720 MATERIALS           COPY MATERIALS(2770)
3730 EQUIPMENT           COPY EQUIPMENT(2850)
3740 SERVICES            COPY SERVICES(2870)
3750 DIRECT LABOUR       SUM GRADE A SALARY(3590)
                         THRU GRADE D SALARY(3620)
3760 SUPERVISORY LABOUR  SUM SUPERVISORS C COST(3520)
                         THRU SUPERVISORS E COST(3540)
3770 OVERHEADS           COPY OVERHEAD COSTS(3680)
                         ========
3790 PROJ INPUT COSTS    + MATERIALS(3720)
                         + EQUIPMENT(3730)
                         + SERVICES(3740)
                         + DIRECT LABOUR(3750)
                         + SUPERVISORY LABOUR(3760)
                         + OVERHEADS(3770)
                         ========

3810 DIRECT STAFF USE
3850 GRADE B1            COPY GRADE B1 EFFORT(3060)
3860 GRADE B2            COPY GRADE B2 EFFORT(3100)
3870 GRADE B3            COPY GRADE B3 EFFORT(3140)
3880 GRADE C1            COPY GRADE C1 EFFORT(3180)
```

APPENDIX 5 - INPUT FORM

The computer system automatically generates an input form which can be easily adapted for entering data into the projection or history files.

In the form reproduced on the following pages have been entered the data values (and appropriate projection codes) from which were generated the report shown in Appendix 6.

Certain technical options have to be indicated as follows :

line 1790 - if OCR is to be used, enter 1;
if any other technique, enter 2

line 1800 - if verification is to be used, enter 1;
if not, enter 0

line 1820 - if in-house work is to be proof-read, enter 0;
if not, enter 1

line 2590 - if computer processing to be done in-house, enter 0;
if by bureau enter 1

In the system suggested, 80% of author abstracts are used (line 1410), and 10% of input is translated (line 1580). Half of the data preparation is contracted out, and the remainder is initially done by OCR, but changed after three years to 50% on-line, 50% OCR. The other values shown are self-explanatory.

PROJECTION INPUT DATA FORM (*OPTIONAL ENTRIES)
FROM DEFINITION FILE DEFIN

0 PROJECT 1 5 (FIRST.LAST COLUMNS)

COLUMN DIST.	:*	1	3
START DATE	:*	2	1/1/76
REPORT HEAD 1	:	11	ASLIB INPUT MODEL,
REPORT HEAD 2	:	12	FIVE YEAR PROJECTION;
COLUMN TOTALS	:*	31	.5. .5. .5. .5. .5. .5.0
COLUMN LABELS 1	:*	51	%
	:*	52	:
COLUMN LABELS 2	:*	61	%
	:*	62	:
ITEMS INPUT		1010	3.1, 26000, 32000
		1011	
MONOGS ACQD		1020	3.1, 4000, 6000
		1021	
JOURNALS ACQD		1022	3.1, 2500, 3000
		1023	
MONOGS PCHSD		1030	3.1, 4000, 6000
		1031	
JOURNALS PCHSD		1032	3.1, 2500, 3000
		1033	
MAN YEAR HRS		1040	5, 1350
		1041	
GRADE A		1060	1.8, 1, 2600, 5.0
		1061	
GRADE B		1070	1.8, 1, 3400, 5.0
		1071	
GRADE C		1080	1.8, 1, 4200, 5.0
		1081	

GRADE D	1090	1.8, 1, 5800, 5.0
	1091	
GRADE E	1100	1.8, 1, 7200, 5.0
	1101	
MONOG ORDER TM	1130	5, 0.25
	1131	
JNL ORDER TM	1132	5, 1.0
	1133	
MONOG UNIT COSTS	1160	1.8, 1, 8.37, 28.0
	1161	
JOURNAL UNIT COST	1162	1.8, 1, 43.41, 26.7
	1163	
SEL UNIT TM	1260	5, 0.10
	1261	
CAT UNIT TM	1310	5, 0.25
	1311	
INDEX UNIT TM	1360	5, 0.13
	1361	
AUTH ABSTS PC	1410	5, 80
	1411	
A ABSTS UNIT TM	1440	5, 0.084
	1441	
W ABSTS UNIT TM	1450	5, 0.3
	1451	
ITEMS TRANSL PC	1580	5, 10
	1581	
TRANS UNIT TM	1600	5, 0.3
	1601	
A ABSTS LENGTH	1680	5, 500
	1681	
W ABSTS LENGTH	1690	5, 800

ON-LINE	1710	5, 7200
	1711	
OCR	1730	5, 6000
	1731	
OFF-LINE	1750	5, 6000
	1751	
OCR OR OTHER	1790	5, 1
	1791	
VERIFY YES-NO	1800	5, 0
	1801	
VERIFICATION FACT	1810	5, 0
	1811	
PROOF IN-HOUSE WK	1820	5, 0
	1821	
CONTRACT PC	1860	5, 50
	1861	
CONTRACT RATE	1870	5, 1
	1871	
ON-LINE PC	1950	1, 0, 0, 0, 50, 50
	1951	
RENTAL RATE	2040	1.8, 1, 360, 8.0
	2041	
COMS COST	2060	1.8, 4, 0, 0, 0, 250, 8.0
	2061	
RENTAL	2190	1.8, 1, 120, 8.0
	2191	
CONVERSION COST	2210	1.8, 1, 0.75, 8.0
	2211	
RENTAL	2340	1.8, 1, 480, 8.0
	2341	

PROOF UNTT TM	2510	5,0.03
	2511	
IN-HOUSE BUREAU	2590	5, 0
	2591	
BUREAU RATE	2600	
	2601	
OCCUPANCY	2630	5, 0.1
	2631	
IN-HOUSE RENTAL	2640	1.8, 1, 6000, 10.0
	2641	
SUPERVISORS GRADEC	3450	5, 1
	3451	
SUPERVISORS GRADED	3460	5, 1
	3461	
SUPERVISORS GRADEE	3470	5, 1
	3471	
CLERKS GRADE A	3472	1, 2, 2, 2, 3, 3
	3473	
SPACE PER PERSON	3490	5, 150
	3491	
RENTAL	3500	1.8, 1, 10, 7.0
	3501	
OVERHEAD RATE	3640	5, 75
	3641	

APPENDIX 6 - SUMMARY REPORT

Available data in the projection file is run against the model contained in the definition file and will yield a summary report of all operational costs associated with input activities. A report prepared from the input data shown in Appendix 5 is reproduced on the following pages.

The way in which each line of the report has been calculated can be traced by reference to the ILLUSTRATE listing in Appendix 4. For example, line 1390, showing indexing costs, is seen to be obtained by multiplying line 1370 (Salary cost) by line 1380 (Index effort). Line 1370 is copied from line 1080, which calls for an input value for a Grade C staff salary. The value used for this parameter in producing the report is shown on the input form, in Appendix 5.

ASLIB INPUT MODEL
FIVE YEAR PROJECTION
FOR THE PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 19, 1976

		1 1976	2 1977	3 1978	4 1979	5 1980
1020	MONOGS ACQD	4000	4500	5000	5500	6000
1022	JOURNALS ACQD	2500	2625	2750	2875	3000
1030	MONOGS PCHSD	4000	4500	5000	5500	6000
1032	JOURNALS PCHSD	2500	2625	2750	2875	3000
1170	PURCHASE COST	142005	192587	260202	350380	470403
1180	LABOUR COST	8815	9917	11107	12391	13776
1200	TOTAL ACQN COST	150820	202504	271309	362771	484179
1220	UNIT ACQN COST	5.80	7.36	9.36	11.89	15.13
1240	ITEMS INPUT	26000	27500	29000	30500	32000
1290	SELECTION	8089	8983	9947	10985	12101
1340	CATALOGUING	20222	22458	24867	27461	30252
1390	INDEXING	10516	11678	12931	14280	15731
1490	ABSTRACTING	10289	11427	12653	13972	15392
1630	TRANSLATION	2427	2695	2984	3295	3630
1660	INTEL OPS COSTS	51542	57242	63382	69993	77107
1840	ITEMS INPUT	26000	27500	29000	30500	32000
1880	NO CONTRACTED	13000.0	13750.0	14500.0	15250.0	16000.0
1910	CONTRACT COST	7280.0	7700.0	8120.0	8540.0	8960.0
1970	NO ON-LINE	0.0	0.0	0.0	7625.0	8000.0
2080	ON-LINE COSTS	0.0	0.0	0.0	2432.5	2664.6
2130	NO KEYBOARDED	13000.0	13750.0	14500.0	7625.0	8000.0
2240	OCR COSTS	8635.8	9760.3	11141.0	6260.2	7020.2
2560	PROOF LABOUR	1213.3	1347.5	1492.0	1647.7	1815.1
2580	COMPUTER OPS					
2660	IN-HOUSE COSTS	600.0	660.0	726.0	798.6	878.5
2680	DATA PREP	17129	18808	20753	18880	20460
2690	COMPUTER OPS	600	660	726	799	878
2710	MECHANICAL OPS	17729	19468	21479	19679	21338

		1 1976	2 1977	3 1978	4 1979	5 1980
2724	DIRECT INPUT COSTS					
2750	LABOUR	64626	71900	79738	87836	96889
2770	MATERIALS	142005	192587	260202	350380	470403
2850	EQUIPMENT	720	790	1006	1653	1801
2870	SERVICES	12740	13937	15223	12574	13531
2900	DIRECT INPUT COSTS	220091	279213	356170	452443	582625
2930	OUT OF BALANCE	-0.1	0.1	0.2	0.1	0.1
2936	STAFF REQUIRED					
3090	GRADE B1 STAFF	3.0	3.0	3.0	4.0	4.0
3130	GRADE B2 STAFF	1.0	1.0	2.0	1.0	2.0
3210	GRADE C1 STAFF	13.0	14.0	15.0	15.0	16.0
3430	DIRECT STAFF	17	18	20	20	22
3450	SUPERVISORS GRADEC	1	1	1	1	1
3460	SUPERVISORS GRADED	1	1	1	1	1
3470	SUPERVISORS GRADEE	1	1	1	1	1
3472	CLERKS GRADE A	2	2	2	3	3
3480	TOTAL STAFF	22	23	25	26	28
3484	OVERHEADS					
3486	TOTAL STAFF	22	23	25	26	28
3630	ALL SALARIES	90600	99540	112896	121550	136865
3640	OVERHEAD RATE	75	75	75	75	75
3650	SALARY OVERHEAD	67950	74655	84672	91163	102649
3660	ACCOMMODATION COST	33000	36915	42934	47777	55053
3680	OVERHEAD COSTS	100950	111570	127605	138939	157702
3700	PROJ INPUT COSTS					
3720	MATERIALS	142005	192587	260202	350380	470403
3730	EQUIPMENT	720	790	1006	1653	1801
3740	SERVICES	12740	13937	15223	12574	13531
3750	DIRECT LABOUR	73400	81480	93933	101639	115959
3760	SUPERVISORY LABOUR	17200	18060	18963	19911	20907
3770	OVERHEADS	100950	111570	127605	138939	157702
3790	PROJ INPUT COSTS	347015	418423	516932	625097	780304
3810	DIRECT STAFF USE					
3850	GRADE B1	2.6	2.8	3.0	3.1	3.3
3860	GRADE B2	0.9	1.0	1.0	1.0	1.0
3880	GRADE C1	12.6	13.3	14.0	14.7	15.5

APPENDIX 7 - USE OF THE WHAT-IF FEATURE

The WHAT-IF command makes it possible to examine the effect of changes in input data values, or in the overall cost structure. In the examples which follow, the sequence of prompts from the computer system and the replies given are reproduced. The use can call for a complete revised summary report, or a print-out of specified lines (which is cheaper). The changes investigated all relate to the report shown in Appendix 6.

1. *WHAT-IF the indexing unit time (line 1360) were increased from 0.13 to 0.2 ? Here we have requested to see only the effect on the total input costs (line 3790).*

```
COMMAND? WHAT-IF
WHAT-IF DEFINITION FILE? (T)
REPORT INFILE,OUTFILE? PRINTIN,WHATA
WHATA DOES NOT EXIST BUT IS NOW BEING CREATED
LINE? 1360
TYPE,FIRST, LAST COLUMN? ADD,1,5
ADD FACTOR? 0.07
LINE? 0
REPORT FILE WHATA COMPLETED

COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? SEL
LINES: AFTER LAST 0*
? % 3790,0
SET PAPER,RETURN...
```

ASLIB INPUT MODEL
FIVE YEAR PROJECTION
FOR THE PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 20, 1976

	1 1976	2 1977	3 1978	4 1979	5 1980
3790 PROJ INPUT COSTS	355865	427746	526753	645789	802104

2. *WHAT-IF the system used 50% of author abstracts instead of 80% (line 1410). Again, only the total input costs are requested.*

```
COMMAND? WHAT-IF
WHAT-IF DEFINITION FILE? (T)
REPORT INFILE,OUTFILE? PRINTIN,WHATB
WHATB DOES NOT EXIST BUT IS NOW BEING CREATED
LINE? 1410
TYPE,FIRST, LAST COLUMN? ADD,1,5
ADD FACTOR? -30
LINE? 0
REPORT FILE WHATB COMPLETED

COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? SEL
LINES; AFTER LAST 0*
? % 3790,0
SET PAPER,RETURN...
```

ASLIB INPUT MODEL
FIVE YEAR PROJECTION
FOR THE PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 20, 1976

	1 1976	2 1977	3 1978	4 1979	5 1980
3790 PROJ INPUT COSTS	365482	437968	529200	656535	793379

3. *WHAT-IF all data preparation were done in-house by an off-line method with verification instead of using a mix of OCR and bureau services, and on-line methods in years 4 and 5 (this affects lines 1790,1800, 1810, 1820, 1860, 1950, and 2060). Because the changes are more complex than in the previous examples, several lines of the amended summary report will be output.*

```
COMMAND? WHAT-IF
WHAT-IF DEFINITION FILE? WAAFA
REPORT INFILE,OUTFILE? PRINTIN,WHATC
WHATC DOES NOT EXIST BUT IS NOW BEING CREATED
REPORT FILE WHATC COMPLETED

COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? MRANGE
FIRST,LAST LINES; AFTER LAST 0,0
? % 1840,2720,3790,3791,0,0
SET PAPER,RETURN...
```

ASLIB INPUT MODEL
FIVE YEAR PROJECTION
FOR THE PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 20, 1976

		1 1976	2 1977	3 1978	4 1979	5 1980
1840	ITEMS INPUT	26000	27500	29000	30500	32000
2280	NOKEYBOARDED	26000.0	27500.0	29000.0	30500.0	32000.0
2370	KEYBOARD COSTS	7071.6	7824.2	9195.1	10113.4	11102.1
2490	VERIFY COST	2924.6	3233.4	3566.1	3924.4	4310.2
2580	COMPUTER OPS					
2660	IN-HOUSE COSTS	600.0	660.0	726.0	798.6	878.5
2680	DATA PREP	9996	11058	12261	14038	15412
2690	COMPUTER OPS	600	660	726	799	878
2710	MECHANICAL OPS	10596	11718	13487	14836	16291
3790	PROJ INPUT COSTS	350495	412295	502125	631537	786859

APPENDIX 8 - SENSITIVITY TESTS

The impact upon projected costs of alterations to model parameters can be clearly shown by WHAT-IF reports. But where the recalculated data lines are large or where a minimum change must result before a value is printed, a sensitivity analysis can be performed. In the examples which follow, the results of the WHAT-IF tests in Appendix 7 have been compared with the originally projected figures shown in Appendix 6. Differences are shown here as percentages.

SENSITIVITY REPORT FOR WHAT-IF EXAMPLE 1

```
COMMAND? SENSITIVITY
COMPARATIVE REPORT FILES(2)? PRINTIN,WHIATA
DIFFERENCE OR PERCENTAGE? PERCENTAGE
MINIMUM PERCENT PRINT LEVEL? 1.0
COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? SEL
LINES: AFTER LAST 0*
? % 3790.0
SET PAPER,RETURN...
```

ASLIB INPUT MODEL
SENSITIVITY--PERCENTAGE
FOR THE PERIOD BEGINNING JAN 1,1976
REPORT PREPARED JUL 20,1976

	1.0 1976	2.0 1977	3.0 1978	4.0 1979	5.0 1980
3790 PROJ INPUT COSTS	2.55	2.23	1.90	3.31	2.79

SENSITIVITY REPORT FOR WHAT-IF EXAMPLE 2

```
COMMAND? SENSITIVITY
COMPARATIVE REPORT FILES(2)? PRINTIN,WHAT0
DIFFERENCE OR PERCENTAGE? PERCENTAGE
MINIMUM PERCENT PRINT LEVEL? 1.0
COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? SEL
LINES; AFTER LAST 0*
? % 3790.0
SET PAPER,RETURN...
```

ASLIB INPUT MODEL
SENSITIVITY--PERCENTAGE
FOR THE PERIOD BEGINNING JAN 1,1976
REPORT PREPARED JUL 20,1976

	1.0 1976	2.0 1977	3.0 1978	4.0 1979	5.0 1980
3790 PROJ INPUT COSTS	5.32	4.67	2.37	5.03	1.68

SENSITIVITY REPORT FOR WHAT-IF EXAMPLE 3

```
COMMAND? SENSITIVITY
COMPARATIVE REPORT FILES(2)? PRINTIN,WHATC
DIFFERENCE OR PERCENTAGE? PERCENTAGE
MINIMUM PERCENT PRINT LEVEL? 2.0
COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? MRAN
FIRST,LAST LINES; AFTER LAST 0,0
? % 1840,2720,3790,3791,0,0
SET PAPER.RETURN...
```

ASLIB INPUT MODEL
SENSITIVITY--PERCENTAGE
FOR THE PERIOD BEGINNING JAN 1,1976
REPORT PREPARED JUL 20,1976

		1.0 1976	2.0 1977	3.0 1978	4.0 1979	5.0 1980
1880	NO CONTRACTED	-100.00	-100.00	-100.00	-100.00	-100.00
1910	CONTRACT COST	-100.00	-100.00	-100.00	-100.00	-100.00
1970	NO ON-LINE	**	**	**	-100.00	-100.00
2080	ON-LINE COSTS	**	**	**	-100.00	-100.00
2130	NO KEYBOARDED	-100.00	-100.00	-100.00	-100.00	-100.00
2240	OCR COSTS	-100.00	-100.00	-100.00	-100.00	-100.00
2280	NOKEYBOARDED	**	**	**	**	**
2370	KEYBOARD COSTS	**	**	**	**	**
2490	VERIFY COST	**	**	**	**	**
2560	PROOF LABOUR	-100.00	-100.00	-100.00	-100.00	-100.00
2580	COMPUTER OPS					
2620	BUREAU COST	**	**	**	**	**
2680	DATA PREP	-41.64	-41.21	-38.51	-25.65	-24.67
2710	MECHANICAL OPS	-40.23	-39.81	-37.21	-24.61	-23.66
3790	PROJ INPUT COSTS			-2.86		

NOTE: ** INDICATES DIVISION BY ZERO

If line 3790 were printed out to show differences rather than percentages, the result would be :

		1.0 1976	2.0 1977	3.0 1978	4.0 1979	5.0 1980
3790	PROJ INPUT COSTS	3460	-6128	-14807	6440	6554

APPENDIX 9 - SPECIFICATION FOR EXPERIMENT TO TEST THE INPUT MODEL DEVELOPED IN EFAG PROJECT 3

A. Objectives

To evaluate the predictive cost model for the input activities of mechanized information systems, as developed in Project 3, Phase 1, Part 1.

B. Source material

Final Report on Project 3, Phase 1, Part 1: Development and use of models for the prediction of costs for alternative information systems. Aslib Consultancy Service, July 1976.

C. Details of project

The basic methodology of the test should be to predict the operating costs of a number of existing systems, as from some time in the past, and to check these predictions against operating costs actually recorded. The steps involved would be as follows :-

(1) Select a minimum of three mechanized information systems which create their own data bases. The systems chosen should show as much variation as possible in terms of materials acquired for input (e.g. different mixes of monographs, and serials publications); input record characteristics (e.g. indexing techniques, abstract lengths); data preparation methods; and computer processing techniques.

An essential criterion for selection of candidate systems is that they should have detailed records of their operational activities and costs for at least three years past.

(2) Obtain data on the operating costs of each system for the past three years, as shown in its annual accounts. Data will also be required on the following parameters (for each year of operation), these being the data values that a model user would normally be required to provide :-

- no. of items input per year
- no. of monographs acquired
- no. of monographs purchased
- no. of journals acquired
- no. of journals purchased
- salary scales applicable to the organization responsible for the system
- indexing techniques used
- percentage of author abstracts used (if any)
- percentage of input items translated (if any)
- average length of input records
- data preparation technique(s) used
- percentage of input verified (if any)
- percentage of input keyboarded by external service; bureau (if any)
- percentage of input keyboarded on-line (if any)
- communications cost (if on-line)
- local computer processing costs, and facilities used
- nos. of non-direct staff employed (supervisors and clerical support)
- accommodation cost per unit area

overheads expressed as a percentage of salary costs

(3) Run the model for each system to generate a three-year cost prediction. The projections for data values such as salaries and document purchase costs should be based on known trends for the countries in which the systems are based.

(4) Compare cost predictions for each stage of the model (acquisition, intellectual processing, mechanical processing, etc) with costs recorded for each system in its accounts. The percentage error for each figure should be recorded.

(5) Investigate causes of inaccuracy, modify input values, and re-run model as necessary.

It is recommended that computer facilities be used for running the model. If the PROPHIT II facilities used for development of the model were employed, the necessary program (definition file) could be supplied by Aslib.

FINAL REPORT on PROJECT 3, PHASE 1, PT 2 (OUTPUT MODEL)

CONTENTS

MANAGEMENT SUMMARY

In accordance with the specification for EFAG Project 3, two separate reports have been prepared on the development and testing of cost prediction models for (a) input activities, and (b) output activities of mechanized information systems. The two models are, however, closely related and both reports are summarized here.

Definition of requirements

In designing these models, the first requirements to be considered were the dimensions within which they had to operate. The models should be applicable to most if not all foreseeable system configurations in terms of resources and techniques used, and services provided; they should be able to predict costs for any volume of throughput; and they should be able to predict costs for any reasonable period of future time.

The second requirement was that the models should be easy to use.

Thirdly, the design of the models should not be incompatible with other studies in the present series of EFAG costing projects.

Last but not least, the models should be capable of predicting costs to a satisfactory level of accuracy (which would depend partly on the purpose for which they were used). A factor to be noted here is that, providing reasonable data values are input to the models, the systems they represent could be controlled in such a way as to ensure that the predicted costs were achieved.

General description

The models have three main components:

- the mechanical component
- the input data
- the user interface.

The mechanical component comprises a series of equations that determine the cost of each element of the system. These equations are presented in such a way that the necessary calculations could be performed by hand, but on-line computing facilities were used in developing and testing the models, as described below.

Some of the input data is determined by the model user - such as the configuration of the system and the volume of throughput. The remainder has to be drawn from observation of the behaviour of existing systems, and the accuracy of the models is highly dependent on these values.

When the models are used manually, the user interface can only be rudimentary; little can be done to relieve the drudgery of the repetitive calculations required. With the aid of computer facilities, however the models can be made truly interactive.

The input model

The main sections of the model cover acquisition, selection, cataloguing, indexing, abstracting, translation, and mechanical processing.

The model calculates for each operation the staff, materials, equipment and services costs as required, prompting the user to consider various system options where appropriate. Alternative methods of mechanical processing, such as on- or off-line data preparation, are represented by separate equations. Alternative methods for intellectual operations, such as indexing and abstracting, are dealt with by using unit times appropriate to the quality of work required.

Direct staff costs are calculated on the basis of unit times for each staff activity. These unit times are multiplied by the number of items processed to give the hours required per year. This figure is then divided by the effective number of working hours in a year, taking into account allowances for relaxation, unoccupied time, holidays, etc., to give the number of man-years of effort required for the activity.

Man-years of effort for each activity are multiplied by a salary cost at a level appropriate to the activity. In the model, provision is made for five salary grades. Some types of staff within these grades are regarded as interchangeable (e.g. indexing and abstracting might be done by the same people), and this factor is taken into account in calculating the total numbers of staff required. The numbers thus calculated are rounded up to integer numbers.

Given the total number of staff required for each activity in the system, the model user is invited to determine the number of supervisory and clerical support staff required. It was felt that this decision could not be made in a realistic way by the model.

Computer processing cost calculations are based on unit costs for each operation, or on the estimated percentage occupancy of a computer installation multiplied by a rental charge.

Accommodation costs are calculated for each member of staff. Overheads are added as a percentage of salary costs.

The output model

The output model is inherently more complex than the input model, in that it has to provide for a wider range of system configurations for a variety of different services. It can be linked to the input model, in that the predicted cost of creating a data base can be fed into the output model. Alternatively, the cost of a commercially available data base or data bases can be used.

The output model covers the following services, separately or in combination:

retrospective search (batch processing)

retrospective search (on-line)

SDI

group SDI

secondary publication (alerting service)

secondary publication (abstracts bulletin)

machine-readable services

The model calculates for each operation the staff, equipment, materials and services charges as required for each of the seven output services selected by the model user as part of the design configuration.

Direct staff costs, where applicable, are calculated on the basis of unit times or data values for each activity. The unit times are multiplied by the frequency of the particular activity to give the hours required per year. This figure is then divided by the effective number of working hours in a year, taking into account allowances for relaxation, unoccupied time, holidays etc., to give the number of man-years of effort required for the activity.

Man-years of effort for each activity are multiplied by a salary cost at a level appropriate to the activity. In the model provision is made for five salary grades. Some types of staff within these grades are regarded as interchangeable (e.g. profile formulation for SDI and for group SDI might be done by the same people), and this factor is taken into account in calculating the total numbers of staff required. The numbers thus calculated are rounded up to integer numbers.

Given the total number of staff required for each activity in the system, the model user is invited to determine the number of supervisory and clerical support required.

Accommodation costs are calculated for each member of staff. Overheads are added as a percentage of salary costs.

Costs of materials and printing are calculated as appropriate to each activity. Royalty charges based on volume of usage made of a purchased data-base may be calculated on the basis of charges against numbers of users, frequency of use and/or volume of output produced, according to the conditions obtaining under sales contracts negotiated with individual data-base producers.

Computer processing costs are calculated on the basis of data available for costs of each run (or issue, in the case of secondary publications).

After the model has calculated the direct costs of each service, an apportionment of input and indirect costs is added to give the total cost.

The computerized models

Both models were developed with the aid of the PROPHIT II system, available through the CDC CALL/370 Time Sharing Service.* PROPHIT II is an on-line financial planning and analysis system. When using this facility, the model is expressed as a series of statements (called a definition file) written in a simple user-oriented programming language.

Input can be in the form of a history file (employing data gathered from past experience) or a projection file. With a projection file, data values that will change with time (such as the number of items input, or salary levels) can be generated from an initial value or values by specifying one of a range of projection types (e.g. linear, stepped, compound).

The projection and/or history files are run against the definition file to produce a report covering as many years as required. The effect of changes in data values, methods of projection, or system design options can be explored by means of a WHAT-IF facility.

Data values

For each model, all the variables employed in the equations are defined, and preferred values or ranges of values are presented where appropriate. The reports stress, however, that the model user should be able to apply judgement, based on experience, in selecting values to be used as input to the models.

A significant difference between the input and output models is that while staff costs predominate in the former, computer processing costs are more important in the latter.

The equations for the input model involve 48 variables, although some of these apply only to certain system configurations. The output model, with its range of alternative services, employs 97 variables.

---

* Similar facilities are available from other major timesharing computer services.

Testing the models

Test runs were carried out with both models to ensure that they would operate correctly under a variety of conditions. In the case of the input model, further tests were conducted by simulating known systems.

As required by the project specification, both reports include written specifications for designed experiments to implement the models. The method proposed is to use the models in a retrospective mode, i.e. to make cost predictions for existing systems as of some time in the past, and to compare the results with the actual costs experienced in reality.

Applications of the models

The main application envisaged for these models, in their present form, is at the broad planning level. They can be used to determine the pattern of costs in future years for a proposed new system, and in so doing enable the planner to explore the effect of different system configurations and operating regimes.

They can also be used more generally as a management tool for forecasting manpower requirements, budgets, and unit costs.

The models as presented are highly generalized, and are applicable to most typical system configurations. The methodology that they incorporate could, however, easily be adapted or extended to cover other specialized configurations, or specific applications. For example, they could be developed for application to cooperative networks, or to investigate the effect of changes on existing systems.

CHAPTER 1: INTRODUCTION AND TERMS OF REFERENCE

This report is the second of two final reports resulting from the study 'Project 3: Development and use of models for the prediction of costs for alternative information systems'. The overall objectives of the project, as given in the Project Specification, were as follows:

> "To develop models for predicting the costs of various methods of data base creation and provision of information services".

This report is about the cost prediction model for output activities (i.e. service provision) of mechanized information systems. A companion report* deals with the input model, and also contains a chapter which discusses the definition and application of cost prediction models in general terms. Although this discussion is relevant to both models, it was not considered necessary to reproduce it in this report also. We have also omitted from this report two appendices which can be regarded as common to both reports. One contains the Project Specification, the other a description of the PROPHIT II computer system that has been used to develop both models.

The project as a whole comprises two phases, the first being to develop and test the models, and the second to implement them in an experimental environment. This report is concerned with Phase I, but includes in Appendix 7 a specification for a designed experiment to implement the model.

---

* P.H. Vickers and Martin Rowat. Final report on Project 3, Phase I, Part I: Development and use of models for the prediction of costs for alternative information systems. Aslib Consultancy Service, October 1976.

The nature of the project is such that no detailed statement on methodology is called for. Having studied previous work in this area (see Chapter 2) and determined the requirements of the model in general terms, we were able to formulate the basic equations and develop them by an iterative process (see Chapter 3). Some tests were carried out with notional data values to ensure the viability of the model (Chapter 5). Considerable effort was devoted to research on the data available for input to the model (see Chapter 4).

CHAPTER 2: REVIEW OF PREVIOUS WORK

Relatively little work has been produced on predictive cost models of complete output systems (or indeed of portions of them). Many papers concerned with the description of operational or planned information systems include some elements of cost reporting. These are usually in very broad terms and lack any form of comparability with other reported figures. In particular there are two areas of confusion: the extent to which overheads are incorporated and the pricing policy of the computer unit; and exactly what is included as, e.g. "cost per profile".

Chronologically one of the earliest papers was by King and Caldwell[1] in the study carried out for the American Psychological Association. The study was to explore factors of cost-effectiveness that affect the choice among alternative systems, it is necessary therefore to predict costs for alternative systems. To do this a cost model was produced which suggests that the total cost for any given retrospective search system is composed of:

1. fixed costs associated with each subsystem.
2. variable costs dependent upon the number of items input to the system.
3. variable costs dependent upon the number of searches conducted.

$$\therefore C = C_1 + C_2 X_1 + C_3 X_2$$

A fuller description of the model appears in the OECD survey (Vickers[2]).

Hisinger[3] in 1971 analysed the operating costs of the National Tech-

nological Library of Denmark which ran SDI services from 3 tape series. A broad scale cost equation was developed:

$$C = £\ [1250/N + 14.3 + 0.084 \times X \times H]$$

where C = costs/profile/year

N = total number of profiles

H = total number of references printed out/profile

X = number of lines/reference

Even with this simple model, one can see a similar division into the three cost elements above.

In the OECD survey Vickers[2] produced equations for (a) costs of SDI services and (b) of retrospective retrieval services.

$$(a) \quad C = D + T + \left[U \left[\frac{PR}{100} + M + E \right] + A \right] X$$

where C = total annual operating costs

D = data base cost/year

T = royalties to tape supplier

U = no. of users

P = computer processing costs per record per 100 users

R = no. of records per year

M = profile maintenance costs per year

E = mailing & distribution costs per user per year

A = ancillary costs

X = overheads

(b) $C = D + (S + P.U + A)\ X + N + n.U + t.y$

The variables relate to an online network and the additional ones in this equation are:

S = file storage cost/year
P = computer processing cost/search
U = no. of searches/year
N = telecommunications network costs per year
n = line costs/search
t = terminal cost/year
y = no. of terminals

Although deficient in some respects this still remains one of the more generalised models for this type of system.

Dammers[4] looked at SDI services within the Shell Laboratories at Sittingbourne and has devised a computer based cost model which can be used for simulation purposes. The model incorporates a number of refinements and, interestingly looks at user costs.

The complete equations are as follows:

Current awareness activities model

Summary of equations and parameter values

$C_t = C_1 + C_2 + C_3 + C_4 = C_s + C_4$ = total costs

$C_1 = C_{11} + C_{12} + C_{13} + C_{14} + C_{15}$ = cost of journal acquisitions, etc.

$C_{11} = S = U_1 \cdot P^F$ = cost of journal subscriptions

$C_{12} = U_2 \cdot S$ = cost of binding

$C_{13} = U_3 \cdot S$ = cost of storage

$C_{14} = U_4 \cdot H_{P1} \cdot P$ = professional staff costs

$C_{15} = U_5 \cdot H_C \cdot P$ = clerical costs

$C_2 = C_{21} + C_{22} + C_{23} + C_{24} + C_{25} - C_{26}$ = cost of SDI service

$C_{21} = B$ = cost of data bases

$C_{22} = U_6 \cdot H_M \cdot T_A$ = cost of computer use

$C_{23} = U_7 \cdot H_{P1} \cdot T_A$ = professional staff costs

$C_{24} = U_8 \cdot H_C \cdot T_A$ = clerical staff costs

$C_{25} = U_9 \cdot T_A$ = costs of stationery

$C_{26} = R \cdot T_E$ = recovery from extra-mural users

$R = U_6 \cdot H_M + U_7 \cdot H_{P1} + U_8 \cdot H_C + U_9 + U_{10} \cdot \frac{B}{T_A}$

$C_3 = L_E \cdot V$ = cost of external loans

$L_E = L_S \cdot R_L$ = number of external loans

$L_S = L_O + U_{11} \cdot T_S$ = number of internal loan requests

$R_L = 2400 \cdot p^{-1.2}$ = external loan factor

$V = V_O + U_{12} \cdot H_C$ = unit cost of a loan

$C_4 = C_{41} + C_{42} - C_{43}$ = user cost

$C_{41} = U_{13} \cdot H_{P2} \cdot T_s$ = user cost associated with screening SD1 output

$C_{42} = U_{14} \cdot H_{P2} \cdot L_E$ = user cost arising from non-availability of journals

$C_{43} = G \cdot H_{P2} \cdot T_S$ = SDI cost benefit

$G = U_{15} + U_{16} \cdot P$ = cost benefit factor

$T_S = 25.10^3 \cdot P^{-0.25}$ = locally used search terms

$T_A = T_S + T_E$ = total number of search terms used

$C_X = 1.2\,S + B + 0.06\,T_A - R \cdot T_E + L_E \cdot V_O$ = out-of-pocket expenditure

U_N, T_E, H_X, V_O, L_O, B, F are supplied parameter values which can be varied to produce functions of one variable against others.

Zais[5] in her 1975 thesis provides an economic model of SDI services pricing which is of interest mainly for some of the data derived from her questionnaires. An analysis of the SDI industry indicated an oligopolistic structure and on this basis certain pricing conduct patterns were suggested. These patterns were compared with actual pricing policies. "Evidence is not conclusive that the model applies". The model was purely descriptive and not one that could be used in the way that King, Vickers or Dammers could be.

Cooper[6] is also concerned with user costs and develops equations for the total costs of information retrieval systems. His paper is concerned more with resource allocation than predictive models and related performance to user and system costs. The paper seems more akin to some sections of Flowerdew and Whitehead[7] who are concerned generally with cost-effectiveness and cost-benefit in information science generally. Their paper is concerned with problems more than solutions but provides useful conceptual support to the model builder.

CHAPTER 3: DESCRIPTION OF THE MODEL

In this chapter we shall describe the mechanical component of the model. First we shall explain the function of each part of the model, and present the equations used in sufficient detail for cost predictions to be made manually.

The model is designed to represent what we believe to be the most typical system configurations and services within the scope of present technology. It does not cover certain possible ancillary services, such as microfiche production, but extension of the model to cover such activities would be a simple matter.

Even with the aid of an electronic calculator, manual use of the model can be fairly laborious, and at an early stage in the project it was decided to use computer facilities to develop, test and operate the model. The particular facilities used are described in section 3.2.

The manual and computer-based versions of the model are linked by the line numbers of the computer files. These are shown in parentheses after each of the parameters used in the equations that follow, and again in Chapter 4, which defines and suggests values for the data required for the model.

It must be stressed that the computer system merely provides the capability to perform the calculations required by the model, and to prepare cost reports; it does not constitute the actual model.

3.1 The output model

The output model is inherently more complex than the input model described in the companion report, in that it has to provide for a wider range of system configurations for a variety of different services.

The output model can be linked to the input model, in that the predicted cost of creating a data base can be fed into the output model. Alternatively, the cost of a commercially available data base or data bases can be used, either as the sole input cost or in combination with that of an in-house data base.

The output model covers the following services, separately or in combination:

retrospective search (batch processing)
retrospective search (on-line)
SDI
group SDI*
secondary publication (alerting service)*
secondary publication (abstracts bulletin)
machine-readable services

It was recognized that, in some cases, model users might wish to predict the costs of systems providing, for example, a series of secondary publications in different subject fields rather than a single publication. To accommodate fully such a requirement, the model would have had to be substantially more complex and probably unmanageable. Ways of adapting the model to such a situation will be suggested.

The model calculates for each operation the staff, equipment, materials and services charges as required for each of the seven output services selected by the model user as part of the design configuration.

---

* Explanatory notes on group SDI and alerting publications will be found in sections 3.1.5. and 3.1.6. respectively.

Direct staff costs, where applicable, are calculated on the basis of unit times or data values for each activity. The unit times are essentially 'basic' times, as defined in B.S. 3138*, and are multiplied by the frequency of the particular activity to give the hours required per year. This figure is then divided by the effective number of working hours in a year, taking into account allowances for relaxation, unoccupied time, holidays etc. to give the number of man years of effort required for the activity.

Man-years of effort for each activity are multiplied by a salary cost at a level appropriate to the activity. In the model provision is made for five salary grades. Some types of staff within these grades are regarded as interchangeable (e.g. profile formulation for SDI and for group SDI might be done by the same people), and this factor is taken into account in calculating the total numbers of staff required. The numbers thus calculated are rounded up to integer numbers.

Given the total number of staff required for each activity in the system, the model user is invited to determine the number of supervisory and clerical support staff required. It was felt that this decision could not be made in a realistic way by the model.

Accommodation costs are calculated for each member of the staff. Overheads are added as a percentage of salary costs.

---

* Glossary of terms used in work study. BS 3138 : 1969, London, British Standards Institution, 1969.

Costs of materials and printing are calculated as appropriate to each activity. Royalty charges based on volume of usage made of a purchased data-base may be calculated on the basis of charges against numbers of users, frequency of use and/or volume of output produced, according to the conditions obtaining under sales contracts negotiated with individual data-base producers.

Computer processing costs are calculated on the basis of data available for costs of each run (or issue, in the case of secondary publications). This approach is relatively limited and is discussed in Chapter 4.

After the model has calculated the direct costs of each service, an apportionment of input and indirect costs is added to give the total cost.

The 'manual' model calculates costs for one system configuration, in one year of operation: To predict costs for a succession of years with different operating regimes and increasing salaries, equipment rentals etc., the model user would have to repeat the calculations as many times as necessary.

The detailed working of the model is shown by the equations which follow.

3.1.1 Data-base costs

The model recognizes that the data-base may be purchased from an external source, or be created in-house, or possibly a combination of the two. Provision is therefore made for the inclusion of

projected input costs derived from the predictive model of input costs described in the companion report. Data-bases in machine-readable form may be purchased and may require to be converted into an acceptable format. It may be necessary to strip records from purchased tapes. The following equations apply to data-base costs.

Cost of stripping and/or converting records

$$C_{sc} = N_{db} \left[C_s + C_c \right]$$

Cost of data-base

$$C_{db} = C_i + C_t + N_{db} \left[C_s + C_c \right]$$

where C_i = input costs (1140)

C_t = tape purchase costs (1260)

N_{db} = number of items on purchased tapes (1210)

C_s = cost of reading each record and stripping selected records (1300)

C_c = cost of converting each record to system format (1320)

3.1.2 Batch retrospective search

The effort required in formulating search questions is calculated by multiplying the unit time by the number of searches, and then dividing by the number of man-hours in a year. Effort required for search formulation,

$$e_{rs} = \frac{N_{rs} \cdot T_{rs}}{H}$$

Direct cost of retrospective search services in batch mode given by summing the labour costs $e_{rs} \cdot S_c$ together with the costs of mailing, computer processing and royalty payments, as follows:

$$C_{rs} = \frac{N_{rs} \cdot T_{rs}}{H} \cdot S_c + N_{rs} \cdot C_{mrs} + N_{rs} \cdot C_{crs} + N_{rs} \cdot R_{srs} + N_{rs} \cdot I_{rs} \cdot R_{ars}$$

which may be simplified to:

$$C_{rs} = N_{rs} \left[\frac{T_{rs}}{H} \cdot S_c + C_{mrs} + C_{crs} + R_{srs} + I_{rs} \cdot R_{ars} \right]$$

where N_{rs} = number of searches made per year (1360)

T_{rs} = average time to formulate search questions (1370)*

S_c = annual salary of personnel paid on grade C (1080)

C_{mrs} = average cost of mailing one search result (1410)

C_{crs} = average computer cost per search (1430)

R_{srs} = royalty payment per search (1450)

R_{ars} = royalty payment per item retrieved (1460)

I_{rs} = number of items found per search (1470)

3.1.3 On-line retrospective search

Search formulation costs in on-line retrieval mode are normally borne by the user. In certain organizations, however, on-line searches might be performed by in-house staff. In these circum-

---

* Where necessary, this value should also include time spent in screening the search output.

stances the staff effort is the annual number of on-line searches made by in-house staff multiplied by the unit time, divided by the number of hours in a man-year,

$$e_r = \frac{N_r \cdot T_r}{H}$$

where N_r = number of search formulations by in-house staff (1515)

T_r = average time to carry out a search (1520)*

Thus staff costs $= \left[\frac{N_r \cdot T_r}{H} \right] S_c$

Costs of each search is the sum of the average computer cost per search, plusthe royalty payable per search, plus any royalty payable on the items retrieved.

$$= C_{cr} + R_{sr} + I_r \cdot R_{ar}$$

where C_{cr} = computer cost per search (1690)

R_{sr} = royalty payable per search (1610)

R_{ar} = royalty payable per item retrieved (1620)

I_r = items retrieved, per search (1630)

Off-line prints incur a mailing charge

$$= N_{op} \cdot C_{mr}$$

---

* Where necessary, this value should also include time spent in screening the search output.

where N_{op} = number of off-line prints (1570)

C_{mr} = cost of mailing each off-line print-out (1560)

Costs of terminals and communications,

$$= N_t \cdot R_t + C_{coms}$$

where N_t = number of terminals in use (1650)

R_t = annual terminal cost, equivalent-rental (1660)

C_{coms} = telecommunications costs, per year (1680)

Size of data-base available for search influences storage costs directly. Storage costs require multiplication of the data-base size (megabytes), by the annual storage cost per megabyte and by the fraction of each 24 hour period the data-base is available. Use of multiple data-bases simultaneously or sequentially requires repetition on this segment of the calculation. Storage costs are:

$$C_{fs} \cdot N_{db} \cdot A_{fs}$$

where C_{fs} = file storage costs, per megabyte, per year

A_{fs} = fraction of each 24 hour period data-base is available (1710)

N_{db} = size of data-base stored on line (megabytes) (1700)

Thus direct costs

$$C_r = N_{os}\left[C_{cr} + R_{sr} + I_r \cdot R_{ar}\right] + N_{op} \cdot C_{mr} + N_t \cdot R_t$$

$$+ C_{coms} + C_{fs} \cdot N_{db} \cdot A_{fs} + \left[\frac{N_r \cdot T_r}{H} \right] S_c$$

where N_{os} = number of on-line searches per year (1510)

3.1.4 Selective disemmination of information (SDI)

Staff costs are associated with the numbers of profiles added each year, the operational number of profiles and the average times spent in maintaining established or adding new profiles. Staff effort is:

$$e_s = \frac{N_{ps} \cdot T_{ms} + N_{pas} \cdot T_{fs}}{H}$$

where N_{ps} = number of operational profiles (1750)

N_{pas} = number of profiles added each year (1752)

T_{ms} = average time spent in maintaining each operational profile (1774)*

T_{fs} = average time to formulate each new profile (1770)*

Staff costs are thus:

$$e_s \cdot S_c$$

Costs associated with each run are due to computer time, mailing, paper cost and royalties levied on the volume of output. On an annual basis these costs are given by:

---

* Where necessary, these values should also include time spent in screening the SDI output from each profile.

$$N_s \left[C_{cs} \cdot N_{ps} + C_{ms} \cdot N_{ps} + N_{ps} \cdot \frac{I_s \cdot C_{ps}}{I_{ps}} + I_s \cdot R_{as} \right]$$

where C_{cs} = computer costs, per profile, per run (1780)

C_{ms} = average cost of mailing output for each profile per run (1790)

C_{ps} = paper cost, per page (1854)

I_{ps} = average number of items output, per page (1852)

I_s = average number of items output per profile, per run (1850)

N_s = runs per year (1760)

R_{as} = royalty cost per item retrieved (1870)

Additionally royalty charges which might be due on the number of operational profiles is given by:

$$N_{ps} \cdot R_{ps}$$

where R_{ps} = royalty cost per operational profile, per year (1860)

Thus direct costs of SDI

$$C_s = \left[\frac{N_{ps} \cdot T_{ms} + N_{pas} \cdot T_{fs}}{H} \right] S_c + N_s \left[C_{cs} \cdot N_{ps} + C_{ms} \cdot N_{ps} + N_{ps} \cdot \frac{I_s \cdot C_{ps}}{I_{ps}} + I_s \cdot R_{as} \right] + N_{ps} \cdot R_{ps}$$

Material costs (paper)

$$M_s = \frac{N_s \cdot I_s \cdot C_{ps}}{I_{ps}} \cdot N_{ps}$$

3.1.5 Group SDI

By 'group SDI' we mean an SDI service supplied to a group of users with common interests. Typical examples would be the TOPICS standard profiles offered by INSPEC, and the UKCIS MACROPROFILES.

Costs here are calculated in a similar way to the previous case (for SDI). However, materials costs (including reproduction) are likely to be higher since by definition the output from each group profile would normally be sent to a number of users. Thus staff effort:

$$e_g = \frac{N_{pg} \cdot T_{mg} + N_{pag} \cdot T_{fg}}{H}$$

where N_{pag} = number of new group profiles created in one year (2012)

N_{pg} = number of operational group profiles (2010)

T_{fg} = average time to formulate new group profiles (2030)*

T_{mg} = average time in maintaining group profiles (2032)*

Staff costs are:

$$e_g \cdot S_c$$

Costs associated with each run are those due to computer time, mailing, reproduction costs and royalties levied on volume of

---

* Where necessary, these values should also include time spent in screening the SDI output from each profile.

output. On an annual basis these costs are given by:

$$N_g \left[C_{cg} \cdot N_{pg} + C_{mg} \cdot N_{ug} + I_g \cdot N_{pg} \left[R_{ag} + \frac{N_{ug} \cdot C_{rg}}{I_{pg}} \right] \right]$$

where C_{cg} = computer costs, per profile, per run (2040)

C_{mg} = average cost of mailing output, per user, per run (2050)

C_{pg} = paper cost, per page

C_{rg} = reproduction cost, per page (2190)

I_g = average number of items output per group profile, per run (2070)

I_{pg} = average number of items output per page (2072)

N_g = runs per year (2020)

N_{ug} = average number of users (per profile) (2060)

R_{ag} = royalty cost per item retrieved, per profile, per run (2140)

Royalty charges may be levied on the number of group profiles maintained, in which case this cost would be:

$$R_{pg} \cdot N_{pg}$$

where R_{pg} = royalty per group profile (2130)

It follows that direct cost of group SDI is:

$$C_g = \left[\frac{N_{pg} \cdot T_{mg} + N_{pag} \cdot T_{fg}}{H} \right] S_c$$

$$+ N_g \left[C_{cg} \cdot N_{pg} + C_{mg} \cdot N_{ug} + I_g \cdot N_{pg} \left[R_{ag} + \frac{C_{pg}}{I_{pg}} + \frac{N_{ug} \cdot C_{rg}}{I_{pg}} \right] \right]$$

$$+ R_{pg} \cdot N_{pg}$$

Materials and reproduction costs are given by:

$$M_g = \frac{N_g \cdot I_g \cdot N_{pg} \cdot N_{ug} \cdot C_{rg}}{I_{pg}}$$

3.1.6 Alerting publications

By 'alerting publications' are meant secondary publications, usually containing only a minimal record for each item, and intended to provide a current-awareness service. Typical examples would be Chemical Titles and Current Papers in Physics. So far as the model is concerned, the costs of producing such a service are calculated in the same way as the cost of an abstracts bulletin, but it was considered useful to make provision for alerting publications as a separate output from a system.

The equations are written for a single publication, but as mentioned in section 3.1, the model user may wish to predict the costs for a series of separate publications in different subject fields. The overall costs could, of course, be estimated simply by using the cumulated numbers of items and pages in the appropriate equations, but to predict the costs of each publication separately it would be necessary to use the equations iteratively.

The production costs will comprise the costs of editorial effort, computer processing, reproduction, binding, distribution and royalty charges. Calculations must also take into account the proportion of the publication devoted to indexes.

Staff effort will be:

$$e_a = \frac{T_{ea} \cdot N_a}{H}$$

where N_a = number of issues per year (2430)

T_{ea} = time spent in editing each issue (2410)

and staff cost is:

$$e_a \cdot S_c$$

Binding and distribution costs per issue are given by:

$$N_{ca} \left[C_{ba} + C_{ma} \right]$$

where C_{ba} = average cost of binding each copy (2560)

C_{ma} = average cost of mailing each copy (2600)

N_{ca} = number of copies printed of each issue (2540)

Reproduction costs per issue are:

$$N_{ca} \cdot C_{ra} \left[I_{ga} + \frac{I_a}{N_a} + \frac{N_{ia}}{I_{pia}} + \frac{I_a}{N_a} \cdot \frac{1}{I_{aa}} \right]$$

$$= N_{ca} \cdot C_{ra} \left[I_{ea} + \frac{I_a}{N_a} \left[\frac{N_{ia}}{I_{pia}} + \frac{1}{I_{aa}} \right] \right]$$

where C_{ra} = reproduction cost, per page (2530)

I_a = number of items per year (2490)

I_{aa} = number of (alerting) entries per page (2460)

I_{ea} = number of editorial pages per issue (2502)

I_{pia} = number of index entries per page (2480)

N_{ia} = number of index entries per item (2470)

Finally, annual computing and royalty charges must be included. The equation may be simplified and written in the form:

$$C_a = N_a \left\{ \frac{T_{ea} \cdot S_c}{H} + C_{ca} + N_{ca} \left[C_{ba} + C_{ma} + C_{ra} \left[I_{ea} + \frac{I_a}{N_a} \left[\frac{N_{ia}}{I_{pia}} + \frac{1}{I_{aa}} \right] \right] \right] \right\} + R_a$$

where C_{ca} = computer costs per run (issue) (2580)

R_a = royalty charges per year (2618)

Materials costs, including reproduction and binding are given by:

$$M_a = N_a \cdot N_{ca} \cdot C_{ba} + N_a \cdot N_{ca} \cdot C_{ra} \left[I_{ea} + \frac{I_a}{N_a} \left[\frac{N_{ia}}{I_{pia}} + \frac{1}{I_{aa}} \right] \right]$$

3.1.7 Abstracting publications

Cost calculations proceed in the same way as for alerting publications (section 3.1.6) except that data elements would ordinarily assume different values.

If the system were designed to produce a series of publications in separate subject fields, rather than a single publication, it would be necessary to use the equations iteratively, as mentioned under 3.1.6.

Staff effort,

$$e_p = \frac{T_{ep} \cdot N_p}{H}$$

where N_p = number of issues per year (3030)

T_{ep} = time spent in editing each issue of the publication (3010)

and staff cost is:

$$e_p \cdot S_c$$

Binding and distribution costs per issue are:

$$N_{cp} \left[C_{ap} + C_{mp} \right]$$

where C_{bp} = average cost of binding each copy (3160)

C_{mp} = average cost of mailing each copy (3180)

N_{cp} = number of copies per issue (3140)

Reproduction costs per issue are:

$$N_{cp} \cdot C_{rp} \left[I_{ep} + \frac{I_p}{N_p} \left[\frac{N_{ip}}{I_{pip}} + \frac{1}{I_{ap}} \right] \right]$$

where C_{rp} = reproduction cost, per page printed (3130)

I_{ap} = abstracts per page (3060)

I_{ep} = editorial pages per issue (3112)

I_p = number of items per year (3090)

I_{pip} = number of index entries per page (3080)

N_{ip} = number of index items per item (3070)

After adding annual computing and royalty charges the equation may be simplified and written in the form:

$$C_p = N_p \left\{ \frac{T_{ep}}{H} \cdot S_c + C_{cp} + N_{cp} \left[C_{bp} + C_{mp} + C_{rp} \left[I_{ep} + \frac{I_p}{N_p} \left[\frac{N_p}{I_{pip}} + \frac{1}{I_{ap}} \right] \right] \right] \right\} + R_p$$

where C_{cp} = computer costs per issue (3200)

R_p = royalty charges per year (3218)

Materials costs, including binding and reproduction are given by:

$$M_p = N_p \cdot N_{cp} \cdot C_{bp} + N_p \cdot N_{cp} \cdot C_{rp} \left[I_{ep} + \frac{I_p}{N_p} \left[\frac{N_{ip}}{I_{pip}} + \frac{I}{I_{ap}} \right] \right]$$

3.1.8 Machine readable services (Magnetic tapes)

Costs of providing machine readable services can be calculated by multiplying the cost of tape purchase, reproduction and mailing by number of original tapes, frequency and number of copies required.

$$C_{mr} = N_{to} \cdot N_{mr} \cdot N_{tc} \left[C_{tr} + C_{mt} + C_{tp} \right]$$

where C_{mt} = tape mailing cost (3225)

C_{tp} = purchase cost, blank tape (3227)

C_{tr} = reproduction cost, per tape (3223)

N_{mr} = frequency (3224)

N_{tc} = number of copies (of each tape) (3226)

N_{to} = number of original tapes (3222)

3.1.9 Effort required

In the output model so far all staff effort has been at grade C. It is assumed that staff would be interchangeable so far as search or

profile formulation is concerned but that they would not be interchangeable with the staff responsible for editorial work.

To estimate realistic staff costs, the numbers of staff in each of the two groups need to be rounded up to whole numbers, as follows:

$$E_{c1} = e_{rs} + e_{r} + e_{s} + e_{g} \quad \text{rounded up to nearest whole number}$$

$$E_{c2} = e_{a} + e_{p} \quad \text{rounded up to nearest whole number}$$

where E_{c1} = number of direct staff, Grade C1

E_{c2} = number of direct staff, Grade C2

At this point, having determined the numbers of staff needed for each activity, the model user may decide on the kind of organizational structure that will be required to operate the system, and to estimate the number of supervisory and clerical support staff needed. Supervisory staff might be employed at Grade C3, D or E depending upon their level in the hierarchy. Clerical support staff are at Grade A. The total numbers and costs of staff can now be calculated as follows:

total number of staff E_{tot}

$$E_{tot} = E_{c1} + E_{c2} + E_{c3} + E_d + E_e + E_a$$

where E_{c3} = number of supervisory staff, Grade C (4310)

E_d = number of supervisory staff, Grade D (4320)

E_e = number of supervisory staff, Grade E (4330)

E_a = number of clerical support staff (4332)

3.1.10 Accommodation costs

Accommodation costs are calculated on the basis of a space allowance for each member of staff, multiplied by a cost per unit of area. The accommodation costs:

$$C_{acc} = E_{tot} \cdot A_p \cdot R_{acc}$$

where A_p = space required per staff member (4390)

R_{acc} = accommodation cost per unit area (4400)

3.1.11 Total salary costs

Calculating total salary costs entails multiplying the numbers of staff E_{c1}, E_{c2} etc. by the appropriate salaries to convert them to staff costs.

Thus direct staff costs:

$$C_{direct} = S_a \cdot E_a + S_c \cdot E_{c1} + S_c \cdot E_{c2}$$

and $C_{grade\ A} = S_a \cdot E_a$

$$C_{grade\ C1} = S_c \cdot E_{c1}$$

$$C_{grade\ C2} = S_c \cdot E_{c2}$$

Supervisory staff costs:

$$C_{super} = S_c \cdot E_{c3} + S_d \cdot E_d + S_e \cdot E_e$$

while total staff costs:

$$C_{staff} = C_{direct} + C_{super}$$

where S_c = annual salary, Grade C staff (1080)

S_d = annual salary, Grade D staff (1090)

S_e = annual salary, Grade E staff (1100)

S_a = annual salary, Grade A staff (1060)

Overheads are calculated as a percentage of staff costs and accommodation costs are added to the salary overhead.

$$C_o = F_{ov} \cdot C_{staff} + C_{acc}$$

where C_o = total overhead cost

F_{ov} = percentage overhead (4540)

3.1.12 Total costs of each service

The direct output cost elements calculated so far may now be summed to find the total direct output costs.*

$$C_{op} = C_{rs} + C_r + C_s + C_g + C_a + C_p + C_{mr}$$

At this stage it is necessary to apportion all other costs between the output services to be provided. For overhead, supervisory labour costs, clerical labour and data-base costs this is done in proportion to the direct costs shown above. It will be recalled (from section 3.1.9) that staff utilisation is summed for interchangeable grades and rounded up to whole numbers of people to be employed at that grade. Equitable apportionment of direct labour costs results if this is done in proportion to the actual labour effort expended by each output service. Thus in Table 1 equations are given for each service showing the proportion of these various costs assignable in each case.

---

* including salary costs of labour effort actually utilised.

SERVICE	Portion of direct salary assignable	Portion of clerical salary assignable	Portion of supervisory cost assignable	Portion of overheads assignable	Portion of input costs assignable
Batch retrosearch	$C_{\text{grade c1}} \frac{e_{rs}}{e_{c1}}$	$C_{\text{grade A}} \frac{C_{rs}}{C_{op}}$	$C_{\text{super}} \frac{C_{rs}}{C_{op}}$	$C_{ov} \frac{C_{rs}}{C_{op}}$	$C_{db} \frac{C_{rs}}{C_{op}}$
On-line retrosearch	$C_{\text{grade c1}} \frac{e_{r}}{e_{c1}}$	$C_{\text{grade A}} \frac{C_{r}}{C_{op}}$	$C_{\text{super}} \frac{C_{r}}{C_{op}}$	$C_{ov} \frac{C_{r}}{C_{op}}$	$C_{db} \frac{C_{r}}{C_{op}}$
SDI	$C_{\text{grade c1}} \frac{e_{s}}{e_{c1}}$	$C_{\text{grade A}} \frac{C_{s}}{C_{op}}$	$C_{\text{super}} \frac{C_{s}}{C_{op}}$	$C_{ov} \frac{C_{s}}{C_{op}}$	$C_{db} \frac{C_{s}}{C_{op}}$
Group SDI	$C_{\text{grade c1}} \frac{e_{g}}{e_{c1}}$	$C_{\text{grade A}} \frac{C_{g}}{C_{op}}$	$C_{\text{super}} \frac{C_{g}}{C_{op}}$	$C_{ov} \frac{C_{g}}{C_{op}}$	$C_{db} \frac{C_{g}}{C_{op}}$
Alerts	$C_{\text{grade c2}} \frac{e_{a}}{e_{c2}}$	$C_{\text{grade A}} \frac{C_{a}}{C_{op}}$	$C_{\text{super}} \frac{C_{a}}{C_{op}}$	$C_{ov} \frac{C_{a}}{C_{op}}$	$C_{db} \frac{C_{a}}{C_{op}}$
Abstracts	$C_{\text{grade c2}} \frac{e_{p}}{e_{c2}}$	$C_{\text{grade A}} \frac{C_{p}}{C_{op}}$	$C_{\text{super}} \frac{C_{p}}{C_{op}}$	$C_{ov} \frac{C_{p}}{C_{op}}$	$C_{db} \frac{C_{p}}{C_{op}}$
Machine readable services	-	$C_{\text{grade A}} \frac{C_{mr}}{C_{op}}$	$C_{\text{super}} \frac{C_{mr}}{C_{op}}$	$C_{ov} \frac{C_{mr}}{C_{op}}$	$C_{db} \frac{C_{mr}}{C_{op}}$

TABLE 1 - APPORTIONMENT OF DIRECT SALARY COSTS AND INDIRECT COST

The values e_{c1} and e_{c2} referred to in Table 1 are composed as follows:

$$e_{c1} = e_{rs} + e_r + e_s + e_g$$

$$e_{c2} = e_a + e_p$$

Where a service does not exist, the costs assignable will have a zero value.

Total costs for each service can be shown to be

for batch retrosearch:

$$C_{rs.\,tot} = N_{rs} \left[C_{mrs} + C_{crs} + R_{srs} + I_{rs} \cdot R_{ars} \right]$$

$$+ C_{rs}/C_{op} \left[C_{\text{grade A}} + C_{super} + C_{ov} + C_{db} \right]$$

$$+ C_{\text{grade C1}} \cdot \frac{e_{rs}}{e_{c1}}$$

for on-line retrosearch

$$C_{r\,.\,tot} = N_{os} \left[C_{cr} + R_{sr} + I_r \cdot R_{ar} \right]$$

$$+ Cr/C_p \left[C_{\text{grade A}} + C_{super} + C_{ov} + C_{db} \right]$$

$$+ C_{\text{grade C1}} \cdot \frac{e_r}{e_{c1}}$$

$$+ N_{op} \cdot C_{mr} + N_t \cdot R_t + C_{coms} + C_{fs} \cdot N_{db} \cdot A_{fs}$$

for SDI

$$C_{s.\,tot} = N_s \left[C_{cs}\ N_{ps} + C_{ms} \cdot N_{ps} + \frac{N_{ps} \cdot I_s \cdot C_{ps}}{I_{ps}} + I_s \cdot R_{as} \right] + C_s/C_{op} \left[C_{grade\ A} + C_{super} + C_{ov} + C_{db} \right] + C_{grade\ C1} \cdot \frac{e_s}{e_{c1}} + N_{ps} \cdot R_{ps}$$

for group SDI

$$C_{g.\,tot} = N_g \left[C_{cg} \cdot N_{pg} + C_{mg} \cdot N_{ug} + I_g \cdot N_{pg} \left[R_{ag} + \frac{C_{pg}}{I_{pg}} + \frac{N_{ug} \cdot C_{rg}}{I_{pg}} \right] \right] + C_g/C_{op} \left[C_{grade\ A} + C_{super} + C_{ov} + C_{db} \right] + C_{grade\ c1} \cdot \frac{e_g}{e_{c1}} + R_{pg} \cdot N_{pg}$$

for alerting publication

$$C_{a.\,tot} = N_a \left\{ C_{ca} + N_{ca} \left[C_{ba} + C_{ma} + C_{ra} \left[I_{ea} + \frac{I_a}{N_a} \left[\frac{N_{ia}}{I_{pia}} + \frac{I}{I_{aa}} \right] \right] \right] \right\} + C_a / C_{op} \left\{ C_{grade\ A} + C_{super} + C_{ov} + C_{db} \right\} + C_{grade\ C2} \cdot \frac{e_a}{e_{c2}} + R_a$$

for abstracting publications

$$C_{p.tot} = N_p \left\{ C_{cp} + N_{cp} \left[C_{bp} + C_{mp} + C_{rp} \left[I_{ep} + \frac{I_p}{N_p} \left[\frac{N_p}{I_{pip}} + \frac{I}{I_{ap}} \right] \right] \right] \right\} + C_p / C_{op} \left[C_{grade\ A} + C_{super} + C_{ov} + C_{db} \right] + C_{grade\ C2} \cdot \frac{e_p}{e_{c2}} + R_p$$

for machine readable services

$$C_{mr.tot} = N_{to} \cdot N_{mr} \cdot N_{tc} \left[C_{tr} + C_{mt} + C_{tp} \right]$$

$$+ \frac{C_{mr}}{C_{op}} \left[C_{grade\ A} + C_{super} + C_{ov} + C_{db} \right]$$

Thus total costs for all output operations are:

$$C_{tot} = C_{rs.tot} + C_{r.tot} + C_{s.tot} + C_{g.tot} + C_{a.tot} +$$

$$C_{p.tot} + C_{mr.tot}$$

3.2 A computer-based version of the model

The arithmetical operations involved in a cost model of the kind presented in this report are simple, but numerous. A substantial amount of data has to be input, to produce some fairly detailed tabulations and analyses of a future cost situation. At an early stage in the project, it was decided to use computer facilities to run and test the model, and these will now be described. Examples of the output from these trial runs are given in Appendices 4 to 6.

In the course of the work on EFAG Project 2, Mr. D. Barlow of INSPEC brought to our attention the PROPHIT II system available through the CDC CALL/370 Time Sharing Service. PROPHIT II is a financial planning and analysis system, which proved to offer the facilities required for our model at a reasonable cost. This is an on-line system, which greatly facilitated rapid development and refinement of the model. In particular, the ease with which data values can be adjusted makes it easy to 'tune' the model to give 'reasonable' results.

It is not our intention to convey that PROPHIT II is the only or even necessarily the best computer system for running the model. We understand that Time Sharing Ltd, CSS International and Honeywell (in the U.K. alone) all offer financial planning systems that could probably be adapted to the same purpose, and there may be many more. Furthermore, it would not be difficult to write a program to perform the calculations required by the equations in the previous section. To write a complete set of programs giving the same facilities as PROPHIT II would, however, be very costly.

With PROPHIT II, the model itself is expressed as a series of statements, using a simple user-oriented language, to form a definition file. This can be automatically converted to a plain-language listing which explains the function of each line in the program. This ILLUSTRATE report is shown for the output model in Appendix 2.

The system can also generate an input form of the type shown in Appendix 3. Input can be in the form of a projection file and/or history file. In either case, the first lines (0-12) determine the output format (number of columns, time distribution, report title, etc.) With a projection file, data values that will change with time (such as the number of items input, or salary levels) can be generated from an initial value or values by specifying one of a range of projection types (e.g. linear, stepped, compound). If a history file is provided, containing data from past operations, future values can be calculated to match trends.

The projection and/or history files are run against the definition file to produce a report, an example of which is shown in Appendix 4.

The effect of changes in data values, methods of projection or system design options can be explored by means of a WHAT-IF facility, some examples of which are shown in Appendix 5. The effect of these changes can be displayed more effectively by the use of a sensitivity analysis, which is illustrated in Appendix 6.

It should be noted that the definition file illustrated in Appendix 2 corresponds closely to the manual model presented in the earlier part of this chapter. If it were necessary to use this modelling technique to investigate the future costs of an existing system or network, it would be advisable (and cheaper) to prepare a new definition file to suit the problem, rather than use the generalized model we have developed.

CHAPTER 4: DATA FOR THE MODEL

4.1 Effect of data on model design

We have explained in the introduction to Chapter 5 of the companion report the relationship between the design of the models and the kinds of data available. It would seem unnecessary to repeat that introduction here, but it is worth emphasizing that the model's predictions cannot be better than the data allows.

As in the case of the input model, we regard it as an important principle that the model user should be able to apply judgment, based on experience, in selecting values to be used in the output model. We have endeavoured to strike the right balance between making the model totally prescriptive and the opposite extreme, which would be to make the user provide all his own data.

4.2 Data definitions and values

In the table which follows, the data elements required for the model are presented in the order in which they are called for in the computerized model (see Appendices 2 and 3), and they are identified by their line numbers. Each element is defined, and preferred values or ranges of values are presented where appropriate. These values have been derived from a variety of sources, including computer bureaux and other specialist organizations. In some cases it has been necessary to select, from a mass of published data, values which in our personal experience seem to be the most reasonable. Thus it has not been possible always to quote one specific source for the figures shown.

Cost values input to the model can, of course, be expressed in the currency of the country concerned.

DATA DEFINITIONS AND VALUES

Line No	Data element	Definition
1040	MAN-YEAR HOURS	Productive hours worked in a year.

The number of days worked in a year may be calculated as follows :-

days in a year		365
less	weekends	104
	holidays	15 - 25
	sickness (average)	5
	public holidays	7
remainder		224 - 234

At 7 hours per day this would give 1568 - 1638 hours per year, but normal work study practice provides for relaxation and other allowances which reduce these figures by $12\frac{1}{2}\%$ - 15%. The effective range thus becomes 1333 - 1392. For general use with the model we suggest a figure of 1350.

1060	GRADE A STAFF	Annual salary plus statutory and other related costs, including welfare contributions, government levies, superannuation costs etc.

The model recognizes five staff grades, the salaries for which should represent the average of what may be a wide range. Grade A is intended for clerical support staff. Salary levels for this and other grades will vary considerably from one location or country to another, and therefore should be specified by the user. Increases in salary costs with time will also be dependent on economic conditions in the country concerned.

1070 GRADE B STAFF — Definition as for Grade A staff.

See general notes under Grade A staff. Grade B is intended for senior clerical or sub-professional staff, and in the model is applied to staff responsible for document acquisition procedures and for keyboard operators.

1080 GRADE C STAFF — Definition as for Grade A staff.

See general notes under Grade A staff. Grade C is intended for professional staff and junior supervisors, and in the model is applied to all staff responsible for intellectual processing of input (e.g. indexers, abstractors, translators).

1090 GRADE D STAFF — Definition as for Grade A staff.

See general notes under Grade A staff. Grade D is intended for supervisors and middle management staff.

1100 GRADE E STAFF — Definition as for Grade A staff.

See general notes under Grade A staff. Grade E is intended for senior management responsible for the system.

1140 INPUT PREPARATION COST — Total annual cost of input prepared in-house.

This value will be the known cost for an existing system, or a predicted cost which might be calculated by the input model (line 3790). It should include all appropriate direct and indirect costs.

1200	RECORDS INPUT	Number of items input per year to the system, using an in-house data base.

This value will be known for an existing data base, or may be an estimate provided by the user. It corresponds to line 1010 in the input model. If all input is in the form of purchased data bases, this value will be zero.

1210	RECORDS PURCHASED	Number of records contained in purchased* data base(s) per year.

To be supplied by user. If all input is prepared in-house, this value will be zero.

1220	RECORDS STRIPPED	Number of records extracted from purchased data base(s) per year.

This caters for a situation where selected records are extracted from a purchased data base, on the basis of subject content, source journals, etc. This value can only be supplied by the user from knowledge of the data base concerned. If the entire contents of the purchased data base are input, the value should be set at zero.

1260	PURCHASE COST OF DATA BASE(S)	Total annual expenditure on machine-readable data-bases (exclusive of royalty charges).

These costs vary widely from one data-base to another, and can be found in a number of published sources (see refs 8 - 12).

---

* 'Purchased' here implies acquired from an external source, and may be taken to include 'acquired by exchange' or even 'acquired at zero cost'.

1300 STRIPPING COST

Cost of computer processing associated with selecting records from a purchased data base, expressed as cost per item read.

No published data found. The cost of the stripping operation would be similar to the cost of performing a search on the same data base, and this could be divided by the number of records on the tape (line 1210) to give the value required. Alternatively, an estimate could be obtained from the computer department which is to carry out the work.

1320 CONVERSION COST

Cost of computer processing associated with format conversion of a purchased data base, expressed as a unit cost per record.

Little published data available. Suggested range of values would be £10 - £20 per megabyte. Alternatively, an estimate could be obtained from the computer department which is to carry out the work.

1360 NUMBER OF SEARCHES (BATCH)

Number of retrospective searches carried out per year, in batch mode.

To be supplied by user.

1370 SEARCH FORMULATION UNIT TIME (BATCH)

Average time (in hours) spent in formulating each search statement for batch processing, and checking output.

Published values vary widely. Unit time will depend on system characteristics, and especially whether controlled or uncontrolled vocabulary is used. An approximate value, based on experience, would be 3 hours.

1410 MAILING COST — Cost of sending search output to enquirer, per search.

This cost will include postage and packing. The staff effort entailed should be taken into account when estimating requirements for clerical support staff (see line 4332).

1430 COMPUTER PROCESSING COST (BATCH) — Computer processing cost per search (batch processing).

This can only be an approximate value, unless typical search processing costs are known for the particular system under consideration. The scatter of observed values is evident from a number of published surveys (refs 13 to 16). For most purposes, a value could be selected from the range £2 - 5 per search.

1450 ROYALTY COST (BATCH SEARCHES) — Royalty charges payable per search, when using purchased data base(s).

The structure of royalty charges varies according to the data base used, and details of these charges can be obtained from several published sources (refs 8 - 12). The model provides for royalty charges based on the number of searches and/or on the number of references retrieved (see line 1460). If charges are not levied on a per search basis, this value should be set at zero.

1460 ROYALTY COST (ABSTRACTS) — Royalty charges payable per reference retrieved, when using purchased data base(s).

See notes for line 1450.

1470 ITEMS RETRIEVED — Average number of items retrieved per search.

To be supplied by user. This value is only required where a positive value is input for item 1460. Guidance on typical values could be obtained from other systems using the same data bases.

1510 NUMBER OF SEARCHES (ON LINE) — Number of on-line retrospective searches carried out per year.

To be supplied by user.

1515 SEARCH FORMULATIONS — Number of on-line search formulations carried out in-house, per year.

This value is only required where all or some of the on-line searches are carried out by staff employed by the organization responsible for the system, on behalf of the end users. A typical example is the present version of the TITUS system, of the Institut Textile de France. The value would have to be supplied by the user, and may be the same as line 1510.

1520 SEARCH FORMULATION UNIT TIME (ON-LINE) — Average time (in hours) spent in formulating search statement, checking output, and operating terminal, per search.

Published values vary widely. There is no evidence that this value will differ significantly from the formulation unit time for batch searches (line 1370) to achieve an equivalent result. Thus the same approximate value of 3 hours could be used here. This value is only required where all or some of the searches are performed in-house, as explained for line 1515.

1560	MAILING COST	Cost of sending search output to enquirer, per search.

This value may be required for the cases explained under line 1515, i.e. where an on-line search is performed in-house on behalf of an external user; but as well as this case, external users requesting off-line prints (line 1570) will result in a cost being incurred in delivering the result. The value will comprise postage and packing.

1570	SEARCHES REQUIRING OFF-LINE PRINT-OUTS	Number of on-line searches for which off-line print-out of results are required.

To be supplied by user. This value relates to external users who call for hard-copy print-out of the results of searches they have performed on-line.

1590	COMPUTER PROCESSING COST (ON-LINE)	Computer processing cost per search (on-line).

Few reliable published figures available. Some values may be found in published surveys (see refs 13 to 16), otherwise it is suggested that a value be used in the range £5 - 10.

1610	ROYALTY COST (ON-LINE SEARCHES)	Royalty charge payable per search, when using purchased data base(s)

See notes for line 1450.

1620	ROYALTY COST (ABSTRACTS)	Royalty charge payable per reference retrieved, when using purchased data base(s).

See notes for line 1460. This value is however less likely to be applicable to on-line searches.

1630 ITEMS RETRIEVED — Average number of items retrieved per search.

See notes for line 1470. This value is only required if line 1620 applies.

1650 NUMBER OF TERMINALS — Number of terminals supported by the system for on-line searching.

This value will normally be required only for a system of the type referred to at line 1515. The model user may, however, include provision here for terminals required within the system for testing or monitoring.

1660 TERMINAL RENTAL — Cost of computer terminal per year.

If the terminal(s) is to be purchased outright, the cost should be spread over 5 years. Otherwise, a rental charge should be shown here. Prices and rental charges vary widely, but typical values in the U.K. would be:

teletype	£800 - 1200 purchase cost
teletype	£300 - 360 annual rental
simple VDU	£1000 - 2000 purchase cost
simple VDU	£360 - 600 annual rental

The rental figures shown would be inclusive of maintenance, but up to 20 per cent should be added to figures based on purchase cost, to allow for this.

Rental charges will increase with time, unless covered by a long-term contract.

1680 LINE RENTAL — Annual cost of telecommunications lines between terminals and computer.

These costs will generally be borne by the users rather than the system, in which case this value is not required. But where all or part of these costs are borne by the system, it will be necessary to estimate them separately, according to the system configuration. For information, the rental charges for a private line (2400 baud) in the U.K., range from £19 (0-0.2 km) to £3890 (> 480 km).

1690 FILE STORAGE COSTS — Annual on-line file storage costs, per megabyte.

Two different charging methods for file storage costs have to be considered. Some large systems have disc drives or other equipment dedicated to their own use, and pay a rental for the equipment. Smaller systems may use bureau facilities, where they may be charged according to the amount of storage occupied, and the time for which the files are made accessible. Present costs for the first case seem to be of the order of £125 per megabyte per year. With improvements in technology, file storage costs are decreasing steadily.

1700 SIZE OF DATA-BASE STORED — Size of on-line file storage, in megabytes.

This value may not be the same as the input file size (i.e. no. of records x average no. of characters per record), as a result of file inversion or compression.

1710 FILE ACCESSIBILITY — Fraction of total time for which files are available for searching.

In the case of dedicated file storage equipment (see line 1690) this value will normally be unity. But in the case of facilities paid for according to the access time required, a fractional value should be used, e.g. if the data base is made accessible on-line for 8 hours out of every 24, the value would be 0.33.

1750 NUMBER OF OPERATIONAL SDI PROFILES — Number of SDI profiles serviced per year.

To be supplied by user.

1752 NUMBER OF PROFILES ADDED — Number of SDI profiles added per year.

To be supplied by user.

1760 SDI COMPUTER RUNS — Number of SDI computer runs per year.

To be supplied by user, according to frequency of service. Frequencies of commercially available tapes can be found in the directories referred to earlier (refs 8 - 12).

1770 SDI PROFILE FORMULATION UNIT TIME — Average time (in hours) spent in formulating a new SDI profile and checking output.

Values will vary according to system characteristics (see line 1370), and are unlikely to be very different from those for line 1370. Thus an approximate value of 3 hours could be used.

1774 SDI PROFILE MAINTENANCE UNIT TIME

Average time (in hours) spent in maintaining and updating an existing SDI profile, including output screening.

No reliable published data available. Suggest use of similar value to line 1370.

1780 COMPUTER PROCESSING COST

Average computer processing cost per profile per run/issue.

Wide variations in published data. Values calculated from OECD survey (ref 2) range from $0.42 to $11.75. An average of the middle-range figures would be $4, equivalent to £2, which is in line with experience.

1790 MAILING COST (SDI)

Cost of sending SDI output to users, per despatch.

See notes for line 1410.

1850 SDI ITEMS OUTPUT

Average number of items output per SDI profile per run.

To be supplied by user. Value will normally be in range 10 - 50, but highly dependent on frequency and size of files searched.

1852 ITEMS PER PAGE

Average number of items per print-out page.

Dependent on form of record output. Typical values would be:

abstracts	6 per page
citations and descriptors	10 per page

	citations only	12 per page.

1854 PAGE COST — Cost per page of output stationery.

This value is only applicable to systems using special output stationery for SDI, such as preprinted card stock, where the cost can be substantial. Cost of normal computer output paper would be absorbed in computer processing costs. No published data available.

1860 ROYALTY COST (SDI PROFILES) — Royalty charges payable per SDI profile, per year.

See notes for line 1450.

1870 ROYALTY COST (OUTPUT) — Royalty charges payable per item output.

See notes for line 1450.

2010 NUMBER OF OPERATIONAL GROUP PROFILES — Number of group SDI profiles provided.

To be supplied by user.

2012 NUMBER OF GROUP PROFILES ADDED — Number of new group profiles added per year.

To be supplied by user.

2020 GROUP SDI COMPUTER RUNS — Number of group SDI computer runs per year.

To be supplied by user, according to frequency of issue of tapes. Frequencies of commercially available tapes can be found in published directories (refs 8 to 12).

2030 GROUP PROFILE FORMULATION UNIT TIME — Average time (in hours) spent in formulating a new group SDI profile, and checking output.

Can be assumed to be equivalent to line 1770.

2032 GROUP PROFILE MAINTENANCE UNIT TIME — Average time (in hours) spent in maintaining and updating existing group profiles, including output screening.

See notes for line 1774.

2040 COMPUTER PROCESSING COST — Average computer processing cost per group profile per run/issue.

Likely that this cost will be of the order of that for operations for individual SDI. See notes for line 1780.

2050 MAILING COST (GROUP SDI) — Cost of sending group SDI output to recipients, per despatch.

Suggest that due to more general nature of a group SDI profile the output produced each run is likely to be more bulky. Nevertheless likely that postage costs will be similar to those previously suggested (see notes for line 1790).

2060 RECIPIENTS — Average number of recipients per group profile.

To be supplied by user. There will be a wide variation in numbers of recipients over all group profiles produced by any particular system, but for calculation purposes the average value is suggested.

2070	GROUP SDI ITEMS OUTPUT	Average number of items output per group SDI profile per run.

To be supplied by user. Suggest values will be a little higher than for individual SDI (see line 1850) since profiles represent broader interests of a group.

2072	ITEMS PER PAGE	Average number of items per print-out page.

See notes for line 1852.

2130	ROYALTY COST (GROUP SDI PROFILES)	Royalty charges payable per group SDI profile, per year, when using purchased data base(s).

See notes for line 1450.

2140	ROYALTY COST (OUTPUT)	Royalty charges per item output.

See notes for line 1450.

2190	PHOTOCOPY/REPRODUCTION COST	Average cost per original page of reproducing output for subsequent distribution.

Cost of reproduction may vary with volume, but essentially is for low-volume runs. Suggested values for A.4 page size are:

Xerox	-	£0.02/page/copy
Printing	-	£0.035 - 0.053/page/copy

Lines 2410 - 2618 are data elements for calculating the costs of "alerting" services. An explanation of how this service differs from other secondary

publications is given in section 3.1.6.

2410 EDIT TIME (ALERTING) — Editorial time per issue (hours).

Supplied by user. Given input is carefully done and software utilised inserts appropriate page and/or section headings then this value should be low.

Typical values 1 - 3 hours/issue.

2430 ISSUES PER YEAR — Number of issues (runs) per year.

To be supplied by user. Frequency may depend upon frequency of receipt of tapes or upon marketing decisions based upon the literature size and number of users.

2460 ALERTS PER PAGE — Number of entries output per page of print out.

To be supplied by user. Dependent upon the precise form of record format and record length.

Typical values : Bio Research Index 42.7 items/page
BEI 33 items/page

2470 INDEX PAGES RATIO (ALERTS) — Number of pages of indexes per page of 'alert' entries.

To be supplied by user. This ratio will depend primarily upon length of each input entry and the type of indexes prepared. An examination of several services produced these values:

Bio Research Index	0.37
Chemical Abstracts	0.21
Computer Control Abstracts	0.05
Excerpta Medica	0.64

2490 ITEMS PER YEAR — Number of items announced per year.

This value could be the same as either line 1200 or line 1210 (or their sum) depending upon the system configuration being modelled. This value could be greater (if this segment of the model represents several overlapping bulletins), or less (if selective alerting services are proposed).

2502 EDITORIAL PAGES (ALERTS) — Number of editorial pages inserted per issue.

To be supplied by user. There should be some correlation between the value for this line and that for line 2410. Standard introductory pages which appear in each issue should be included here. The number of pages will be determined primarily by economic factors and it is suggested will be in the range of 1 - 5.

2530 REPRODUCTION COST — Average cost of reproducing (printing) each page of output, per copy, including paper costs.

This value will vary with volume and size of the bulletin. It is likely that greater volume will be required than for group SDI (line 2190), and the method used will influence costs. It is likely that fairly modest print runs will be required so that lithographic methods and small jobbing firms may be economically used.

Tucker[16] has suggested values on a table which is partially reproduced below.

Number of copies	Number of A4 pages	
	5	50
1	5p	5p
2	5p	4p
5	4p	4p
10	4p	4p
20	3p	3p
50	1.8p	1.8p
100	1.6p	1.4p
1000	0.4p	0.35p

See also line 2540.

2540 COPIES PER ISSUE — Number of copies prepared of each issue.

To be supplied by user. The value used here will partially determine the values applicable to line 2530.

2560 BINDING — Cost of collating and binding each finished copy of alerting bulletin.

The method of binding will largely determine costs here. Selection of method will in turn depend upon function of the bulletin (durability etc.), desired quality of production, physical size (thickness) and numbers. Typical values from a small printing house are :

		£	Unit Cost £
collating	1000 sheets	2.50	-
stapling	1000 sets top left	7.00	0.007
stapling	1000 2 side or saddle	8.00	0.008
trimming	1000 sheets 1 side	.15	-
trimming	1000 sheets 3 sides	.20	-
comb. binding	100 books to 5/16"	12.00	0.12
comb. binding	100 books to 1"	35.00	0.35
perfect binding	100 books	12.00	0.12
(including trimming)	1000 books	90.00	0.09

2580 COMPUTER PROCESSING COST

Average computer processing cost per run (issue).

No published data available. The costs will vary according to sophistication of software and peripherals used. Should cover all computer processing to convert data base records to output format and to produce indexes. User should design the desired service and then get a quotation from a bureau. Tucker[16] includes some useful data of limited application, since it is based on limited experience with an IBM 360/50 installation and is primarily for catalogue production.

e.g. Fortnightly print-out procedures with weekly update (values as £).

Additions per week	Cumulated number of entries					
	1,000	10,000	50,000	100,000	200,000	500,000
10	1.2	8.5	41	81	200	410
100	3.4	12	44	85	210	410
1000	37	44	76	120	240	440
2000	73	80	110	160	280	480
4000	140	150	180	220	350	550

2600 MAILING COST (ALERTS)

Cost of sending alerting bulletins to recipients, per copy.

To be supplied by user. Where alerting bulletins are produced and distributed in-house only, then this cost may not apply. In this case set this value to zero. See also notes on line 1410.

2618 ROYALTY CHARGES (ALERTS)

Annual cost of royalty charges when using purchased data base(s).

Where publications are produced from purchased data bases for in-house use additional royalty charges

might not be payable. If distribution is made outside the purchasers premises then royalty charges will become due and would probably be negotiated on an annual basis. See also notes for line 1450.

3010 EDIT TIME (SECONDARY PUBLICATIONS) — Editorial time spent per issue (hours).

To be supplied by user. It is likely that these secondary publications are intended for widespread distribution, and that more time will be necessary to ensure maintenance of high standards, liaison with printers, composing editorial comment etc. See also notes at line 2410.

3030 ISSUES PER YEAR — Number of issues (runs) per year.

See notes on line 2430.

3060 ITEMS PER PAGE — Number of entries output per page of print-out.

To be supplied by user. Dependent upon the precise length and format of each record. Typical values obtained from an examination of several published services are :

Chemical Abstracts	12.3	items/page
Excerpta Medica	5	
ERIC	6.8	
INIS	10.6	

See also notes on line 1852.

3070 INDEX PAGES RATIO (SECONDARY PUBLICATIONS) — Number of pages of indexes per page of abstracts.

To be supplied by user. See notes on line 2470.

3090 ITEMS PER YEAR — Number of items announced per year.

To be supplied by user. See notes on line 2490.

3112 EDITORIAL PAGES (SECONDARY PUBLICATIONS) — Number of editorial pages inserted per issue.

To be supplied by user. See also notes on lines 3010, 2502.

3130 REPRODUCTION/PRINTING COST — Cost per page of original, per copy made.

Tucker[16] supplies some data from which is will be seen that with extended print runs unit page costs do not have a linear relationship. It may be necessary to adjust this value therefore as the numbers of copies/issue changes (line 3140). The values suggested are for extended "instant print" charges. It is likely that costs will be substantially higher if a traditional printing house is used.

Number of copies	Number of pages 5	50	500
50	1.8p	1.8p	1.6p
100	1.6p	1.4p	1.3p
1000	0.4p	0.35p	0.35p

3140 NUMBER OF COPIES (SECONDARY PUBLICATIONS) — Number of copies printed of each issue, equivalent to the number of subscribers plus copies for in-house use.

To be supplied by user. See also notes on line 3130.

3160 BINDING — Cost of collating and binding each finished

copy of the secondary publication.

See notes on line 2560.

3180 MAILING COST (SECONDARY PUBLICATIONS) — Cost of despatching secondary publications to recipients, per copy.

To be supplied by user. Likely to be fairly bulky and be despatched at printed paper rates. Cost includes postage and packing. The staff effort entailed should be taken into account when estimating requirements for clerical support staff. (See line 4332).

3200 COMPUTER PROCESSING COSTS — Average computer processing costs per run (issue).

Little published data available, see notes on line 2580. The costs here are likely to be higher since a quality output will probably be required necessitating additional computer effort, for computer typesetting, for example.

3218 ROYALTY CHARGES (SECONDARY PUBLICATIONS) — Annual costs of royalty charges when using purchased data base(s).

See notes on line 2618.

3222 NUMBER OF ORIGINAL TAPES — Number of tapes occupied by machine readable data-base, each issue.

To be supplied by user. Number of tapes will depend upon number of entries, packing density, tape format, and will be related to frequency of

issue and total (annual size of data-base(s) offered.

3223	REPRODUCTION COST (TAPES)	Computer processing charges for duplicating one tape.

Little published data available. Seems to be of the order of £1/tape.

3224	FREQUENCY	Number of occasions tapes are issued, each year.

To be supplied by user.

3225	MAILING COST (TAPES)	Cost of packing and posting one magnetic tape, inland, or of best alternative method overseas.

To be supplied by user.

3226	NUMBER OF COPIES (TAPES)	Number of copies required of each (original) tape. Likely to equal number of subscribers.

To be supplied by user.

3227	PURCHASE COST (TAPES)	Purchase price of a blank tape.

Cost of each blank tape. Current prices appear to be of the order of £5.

4310	SUPERVISORS GRADE C	Number of staff required in each grade
4320	SUPERVISORS GRADE D	
4330	SUPERVISORS GRADE E	
4332	CLERICAL SUPPORT STAFF GRADE A	

As explained in Section 3.1 the model user is required to designate the numbers of supervisory and clerical support staff required, in the light of the numbers of direct staff calculated by the model. The provision of staff in these grades should allow for system maintenance (including thesaurus maintenance) and development work. The intended levels of seniority of the three supervisory grades are indicated at lines 1080, 1090 and 1100. For a multi-year projection, these numbers may need to be adjusted from one year to another.

4390 SPACE PER STAFF MEMBER — Average working area allowed per staff member.

Standards of accommodation vary from one organisation to another, but the following gives a rough indication of generally accepted space allowances :

	sq. ft	sq. metres
senior admin. staff	200 - 400	18 - 36
professional staff	100 - 150	9 - 14
clerical staff	50 - 80	4.5 - 7.5
typing staff	40 - 60	4 - 5.5

The model calls for only one value, which could be estimated on the basis of the mix of staff to be employed.

4400 SPACE RENTAL — Annual cost per sq.foot/ sq.metre (depending on unit used for 4390) of accommodation.

Again, the value to be used here will be location-dependent. It should represent an economic cost including rates, cleaning, etc. Substantial increases in time should be allowed for.

4540 OVERHEAD RATE — Overhead cost expressed as a percentage of salary costs.

To be supplied by user. This factor has to cover all indirect organizational costs other than accommodation.

CHAPTER 5: TESTING THE MODEL

Testing of the computer-based model, during the course of its development, has mainly taken the form of test runs with different sets of data values, to ensure that the definition file would operate correctly under a variety of conditions.

The project specification calls for a written specification for a designed experiment to implement the model. The ideal way to check the validity of the model's predictions would, of course, be to design a system; use the model to predict its costs; implement the system; and then compare its costs with the predictions. Unfortunately, such an approach is impractical.

The only practical solution would seem to be to use the model in a retrospective mode, i.e. to make a cost prediction for an existing system as of some time in the past, and compare the results with the actual costs experienced by the system.

In designing any experiment to test the model, three important factors have to be borne in mind. The first is that the model will work best for a user with some knowledge and experience of the environment in which the system will operate. Many of the data values called for will depend on local conditions (e.g. salary rates, computer processing charges, accommodation costs, and overhead rates).

The second factor is that the model predictions can serve as a self-fulfilling prophecy. In a real-life situation, it should be possible to manage the system in such a way that it would operate within the cost limits

predicted by the model. This will not apply if the model is checked against an existing system.

The third factor concerns the accuracy expected of the model. The accuracy required will depend on the purpose for which the model is used. The accuracy achieved will depend on the quality of the data that is fed into the model, coupled with the design of the model itself, which embodies a certain level of approximation. The test we shall describe does not suggest that the model would be deemed to fail, if it did not achieve a specific level of accuracy. The level of accuracy would be measured, and the model judged subjectively.

The specification for the test is given in Appendix 7.

CHAPTER 6: RECOMMENDATIONS

Over and above the test of the model discussed in the previous chapter, we believe that the model could usefully be developed for specific applications. In its present form, it is suitable for making cost predictions at the broad planning level. In the course of the project, interest has been expressed in the use of cost modelling techniques by system operators. Their requirement is for a model into which could be fed details of current operational volumes and costs for a specific system, and which the operator could use to determine the effect of changes in methods, staffing, throughput volumes, etc.

The model would need to be modified to fulfil this role in an effective manner. Since the model would be working on actual cost data of an existing system, it would be possible to dispense with certain features designed to deal with areas of uncertainty. Also the user interface of the model would need to be redesigned with this application in mind.

We therefore recommend that further research on these lines be initiated by the Commission, or by some other interested organization.

APPENDIX 1 - REFERENCES

1. King, D.W. and Caldwell, N.W. 'Study of the cost-effectiveness of retrospective search systems'. American Psychological Association, 1970 (Available as ED 046446).

2. Vickers, P. 'The costs of mechanized information systems'. Organisation for Economic Co-operation and Development, 1974.

3. Hisinger, B. 'A multi-disciplinary and computerized SDI service for industry and research - practical experience and costs'. Inform. Stor. Retr., vol 7, p. 175 - 187 (Nov. 1971).

4. Dammers, H. 'Economic evaluation of current awareness systems'. In Eurim: a European conference on research into the management of information services and libraries (Aslib, 1974), p. 108 - 124.

5. Zais, H.W. 'The pricing of information : a model for selective dissemination of information services'. University of California, Berkeley. Lawrence Radiation Laboratory, 1975. (LBL - 4257).

6. Cooper, M.D. 'A cost model for evaluating information retrieval systems'. J. Amer. Soc. Inf. Sc., vol. 23, p. 306 - 312.

7. Flowerdew, A.D.J. and Whitehead, C.M.E. 'Cost-effectiveness and cost/benefit analysis in information science'. London School of Economics and Political Science, 1974. (OSTI Report 5206).

8. Pratt, G. (ed), 'Data bases in Europe'. A directory to machine-readable data bases and data banks in Europe. Aslib EUSIDIC, European User Series 1. Aslib (1975).

9. Thomas, Angela (comp). 'LUCIS guide to computer-based information services'. Central Information Services, Interim Library Resources Co-ordinating Committee. (December 1975).

10. Finer, R.F., 'A guide to selected computer-based information services'. Aslib (1972).

11. British Library Research and Development Department, 'Inventory of bibliographic data bases produced in the U.K.' Report No. 5256 (January 1976).

12. Williams, Martha E. and Rouse, Sandra H. (eds)., 'Computer-readable bibliographic data bases : a directory and data source book'. ASIS (1976).

13. Allaire, Elisabeth., 'Enquete sur les couts des systemes de documentation automatique en France'. Bureau National de l'Inform-Scientifique et Technique. (Decembre 1975).

14. Drees, G., 'The costs of scientific and technical information and documentation systems'. Study carried out for the CIDST - Brussels Working Party on Pricing (1974).

15. Dubois, G. & Peeters, E., 'Cost of automatic processing of documentary information'. Report presented at National Conference on Documentation, Brussels (1974).

16. Tucker, C.J. 'A comparison of the production costs of different physical forms of catalogue output.' Program, vol.18, p.59-74 (April 1974)

APPENDIX 2 - STRUCTURE OF COMPUTER-BASED MODEL

The listing which follows was prepared by using the ILLUSTRATE feature of the PROPHIT II system. It presents in plain language the operations required by the definition file (DEFOPAA) for our model. 'READ DATA' lines relate to data required by the model, which are ordinarily supplied from the equivalent line in the projection file (or history file, if used). These data values are defined in Chapter 4 of this report.

One element of the computer-based model may need further explanation. After line 4000, there is a section in which numbers of staff for each activity are rounded up to whole numbers. It will be noted that staff grades are identified C1, C2. This is merely a device to separate noninterchangeable staff at any grade.

DEFINITION FILE 'DEFOPCO AS OF 7/28/76

LINE	ACTION
1040 MAN YEAR HOURS	READ DATA
1060 GRADE A	READ DATA
1070 GRADE B	READ DATA
1080 GRADE C	READ DATA
1090 GRADE D	READ DATA
1100 GRADE E	READ DATA
1140 INPUT PREP COST	READ DATA
1200 RECORDS INPUT	READ DATA
1210 RECORDS PCHSD	READ DATA
1220 RECORDS STRPD	READ DATA
1230 RECORDS USED	IF RECORDS STRPD(1220) GT 0 THEN RECORDS STRPD(1220) ELSE RECORDS PCHSD(1210)
1240 RECORDS IN D/B	+ RECORDS INPUT(1200) + RECORDS USED(1230)
1260 PURCHASE COST	READ DATA
1300 STRIPPING	READ DATA
1310 STRIPPING	RECORDS PCHSD(1210) X STRIPPING(1300)
1320 CONVERSION	READ DATA
1330 CONVERSION	RECORDS PCHSD(1210) X CONVERSION(1320)
1340 DATA BASE CST	+ INPUT PREP COST(1140) + PURCHASE COST(1260) + STRIPPING(1310) + CONVERSION(1330)
1350 RETRO BATCH	
1360 SEARCHES	READ DATA
1370 FORMULATE TM	READ DATA
1380 STAFF COST	COPY GRADE C(1080)
1390 EFFORT	SEARCHES(1360) X FORMULATE TM(1370) / MAN YEAR HOURS(1040)
1400 LABOUR	STAFF COST(1380) X EFFORT(1390)
1410 MAILING	READ DATA
1420 MAILING	MAILING(1410) X SEARCHES(1360)
1430 COMPUTER COST	READ DATA
1440 COMPUTER	COMPUTER COST(1430) X SEARCHES(1360)
1450 ROYALTY SEARCHES	READ DATA
1460 ROYALTY ABSTRACT	READ DATA
1470 ITEMS RETRIEVED	READ DATA
1480 ROYALTIES	SEARCHES(1360) X ROYALTY SEARCHES(1450) + ITEMS RETRIEVED(1470) X ROYALTY ABSTRACT(1460) X SEARCHES(1360)
1490 DIRECT COSTS	+ LABOUR(1400) + MAILING(1420) + COMPUTER(1440) + ROYALTIES(1480)
1500 ONLINE RETRO	

LINE	ACTION
1510 SEARCHES	READ DATA
1515 FORMULATIONS	READ DATA
1520 FORMULATE TM	READ DATA
1530 STAFF COST	COPY GRADE C(1080)
1540 EFFORT	FORMULATE TM(1520) X FORMULATIONS(1515) / MAN YEAR HOURS(1040)
1550 LABOUR	STAFF COST(1530) X EFFORT(1540)
1560 MAILING	READ DATA
1570 OFF-LINE PRINTS	READ DATA
1580 MAILING	MAILING(1560) X OFF-LINE PRINTS(1570)
1590 COMPUTER COST	READ DATA
1600 COMPUTER	COMPUTER COST(1590) X SEARCHES(1510)
1610 ROYALTY SEARCHES	READ DATA
1620 ROYALTY ABSTRACT	READ DATA
1630 ITEMS RETRIEVED	READ DATA
1640 ROYALTIES	ROYALTY SEARCHES(1610) X SEARCHES(1510) + ROYALTY ABSTRACT(1620) X ITEMS RETRIEVED(1630) X SEARCHES(1510)
1650 TERMINALS	READ DATA
1660 RENTAL	READ DATA
1670 TERMINAL COST	TERMINALS(1650) X RENTAL(1660)
1680 LINE RENTAL	READ DATA
1690 FILE STORAGE	READ DATA
1700 STORED D/B SIZE	READ DATA
1710 ACCESS	READ DATA
1720 STORAGE	STORED D/B SIZE(1700) X FILE STORAGE(1690) X ACCESS(1710)
1730 DIRECT COSTS	+ LABOUR(1550) + MAILING(1580) + ROYALTIES(1640) + TERMINAL COST(1670) + LINE RENTAL(1680) + STORAGE(1720)
1740 SDI	
1750 OPERNL PROFILES	READ DATA
1752 PROFILES ADDED	READ DATA
1760 RUNS PER YEAR	READ DATA
1770 FORMULATE TM	READ DATA
1774 MAINTAIN TM	READ DATA
1780 COMPUTER	READ DATA
1790 MAILING	READ DATA
1800 MAILING	MAILING(1790) X OPERNL PROFILES(1750) X RUNS PER YEAR(1760)
1810 PROFILE EFFORT	OPERNL PROFILES(1750) X MAINTAIN TM(1774) / MAN YEAR HOURS(1040) + PROFILES ADDED(1752) X FORMULATE TM(1770) / MAN YEAR HOURS(1040)

LINE	ACTION
1820 STAFF COST	COPY GRADE C(1080)
1830 LABOUR	PROFILE EFFORT(1810)
	X STAFF COST(1820)
1840 COMPUTER	RUNS PER YEAR(1760)
	X OPERNL PROFILES(1750)
	X COMPUTER(1780)
1850 ITEMS OUTPUT	READ DATA
1852 ITEMS PER PAGE	READ DATA
1854 PAGE COST	READ DATA
1856 PAPER	ITEMS OUTPUT(1850) X
	RUNS PER YEAR(1760) X
	OPERNL PROFILES(1750) /
	ITEMS PER PAGE(1852) X
	PAGE COST(1854)
1860 ROYALTY PROFILE	READ DATA
1870 ROYALTY ABSTRACT	READ DATA
1880 ROYALTIES	OPERNL PROFILES(1750) X
	ROYALTY PROFILE(1860) +
	ROYALTY ABSTRACT(1870) X
	OPERNL PROFILES(1750) X
	RUNS PER YEAR(1760) X
	ITEMS OUTPUT(1850)
1890 DIRECT COSTS	+ LABOUR(1830)
	+ COMPUTER(1840)
	+ ROYALTIES(1880)
	+ MAILING(1800)
2000 GROUP SOI	
2010 GROUP PROFILES	READ DATA
2012 NEW PROFILES	READ DATA
2020 RUNS PER YEAR	READ DATA
2030 FORMULATE TM	READ DATA
2032 MAINTAIN TM	READ DATA
2040 COMPUTER	READ DATA
2050 MAILING	READ DATA
2060 USERS	READ DATA
2070 ITEMS OUTPUT	READ DATA
2072 ITEMS PER PAGE	READ DATA
2080 MAILING	USERS(2060)
	X MAILING(2050)
	X RUNS PER YEAR(2020)
2090 STAFF COST	COPY GRADE C(1080)
2100 EFFORT	GROUP PROFILES(2010) X
	MAINTAIN TM(2032) /
	MAN YEAR HOURS(1040) +
	NEW PROFILES(2012) X
	FORMULATE TM(2030) /
	MAN YEAR HOURS(1040)
2110 LABOUR	EFFORT(2100)
	X STAFF COST(2090)
2120 COMPUTER	COMPUTER(2040)
	X RUNS PER YEAR(2020)
	X GROUP PROFILES(2010)
2130 ROYALTY PROFILE	READ DATA
2140 ROYALTY ABSTRACT	READ DATA
2150 ROYALTIES	ROYALTY PROFILE(2130) X
	GROUP PROFILES(2010) +
	USERS(2060) X

LINE	ACTION
2170 ITEMS PER PAGE	COPY ITEMS PER PAGE(2072)
2180 PAGES OUTPUT	ITEMS OUTPUT(2070) /
	ITEMS PER PAGE(2072) X
	GROUP PROFILES(2010) X
	RUNS PER YEAR(2020) X
	USERS(2060)
2190 REPRO COST	READ DATA
2200 REPRODUCTION	PAGES OUTPUT(2180)
	X REPRO COST(2190)
2300 DIRECT COSTS	+ MAILING(2080)
	+ LABOUR(2110)
	+ COMPUTER(2120)
	+ ROYALTIES(2150)
	+ REPRODUCTION(2200)
2400 ALERT PUBS	
2410 EDIT TM	READ DATA
2420 SALARY COST	COPY GRADE C(1080)
2430 ISSUES PER YEAR	READ DATA
2440 EFFORT	EDIT TM(2410) X
	ISSUES PER YEAR(2430) /
	MAN YEAR HOURS(1040)
2450 LABOUR	SALARY COST(2420)
	X EFFORT(2440)
2460 ALERTS PER PAGE	READ DATA
2470 INDEX PAGE RATIO	READ DATA
2490 ITEMS PER YEAR	READ DATA
2492 ITEMS PER ISSUE	1 X
	ITEMS PER YEAR(2490)
	/
	ISSUES PER YEAR(2430)
2500 PAGES ALERTS	ITEMS PER ISSUE(2492) /
	ALERTS PER PAGE(2460)
2502 EDITORTAL PAGES	READ DATA
2510 PAGES INDEX	PAGES ALERTS(2500)
	X INDEX PAGE RATIO(2470)
2520 PAGES/TSSUE	+ PAGES ALERTS(2500)
	+ PAGES INDEX(2510)
	+ EDITORIAL PAGES(2502)
2530 REPRODUCTION	READ DATA
2540 COPIES PER ISSUE	READ DATA
2550 REPRODUCTION	PAGES/ISSUE(2520) X
	COPIES PER ISSUE(2540) X
	REPRODUCTION(2530) X
	ISSUES PER YEAR(2430)
2560 BINDTNG	READ DATA
2570 BINDING	COPIES PER ISSUE(2540)
	X ISSUES PER YEAR(2430)
	X BINDING(2560)
2580 COMPUTER COST	READ DATA
2590 COMPUTER	COMPUTER COST(2580)
	X ISSUES PER YEAR(2430)
2600 MAILTNG	READ DATA
2610 MAILING	ISSUES PER YEAR(2430)
	X COPIES PER ISSUE(2540)
	X MAILING(2600)

LINE	ACTION
2618 ROYALTIES	READ DATA
2620 DIRECT COSTS	+ LABOUR(2450)
	+ REPRODUCTION(2550)
	+ BINDING(2570)
	+ COMPUTER(2590)
	+ MAILING(2610)
	+ ROYALTIES(2618)
3000 ABSTRACT PUBS	
3010 EDIT TM	READ DATA
3020 SALARY COST	COPY GRADE C(1080)
3030 ISSUES PER YEAR	READ DATA
3040 EFFORT	EDIT TM(3010) X
	ISSUES PER YEAR(3030) /
	MAN YEAR HOURS(1040)
3050 LABOUR	SALARY COST(3020)
	X EFFORT(3040)
3060 ITEMS PER PAGE	READ DATA
3070 INDEX PAGE RATIO	READ DATA
3090 ITEMS PER YEAR	READ DATA
3092 ITEMS PER ISSUE	1 X
	ITEMS PER YEAR(3090)
	/
	ISSUES PER YEAR(3030)
3100 PAGES ABSTRACTS	ITEMS PER ISSUE(3092) /
	ITEMS PER PAGE(3060)
3110 PAGES INDEX	PAGES ABSTRACTS(3100)
	X INDEX PAGE RATIO(3070)
3112 EDITORIAL PAGES	READ DATA
3120 PAGES/ISSUE	+ PAGES ABSTRACTS(3100)
	+ PAGES INDEX(3110)
	+ EDITORIAL PAGES(3112)
3130 REPRODUCTION	READ DATA
3140 COPIES PER ISSUE	READ DATA
3150 REPRODUCTION	PAGES/ISSUE(3120) X
	COPIES PER ISSUE(3140) X
	ISSUES PER YEAR(3030) X
	REPRODUCTION(3130)
3160 BINDING	READ DATA
3170 BINDING	ISSUES PER YEAR(3030)
	X COPIES PER ISSUE(3140)
	X BINDING(3160)
3180 MAILING	READ DATA
3190 MAILING	MAILING(3180)
	X COPIES PER ISSUE(3140)
	X ISSUES PER YEAR(3030)
3200 COMPUTER	READ DATA
3210 COMPUTER	COMPUTER(3200)
	X ISSUES PER YEAR(3030)
3218 ROYALTIES	READ DATA
3220 DIRECT COSTS	+ LABOUR(3050)
	+ REPRODUCTION(3150)
	+ BINDING(3170)
	+ MAILING(3190)
	+ COMPUTER(3210)
	+ ROYALTIES(3218)
3221 M/R SERVICES	

LINE	ACTION
3222 ORIGINAL TAPES	READ DATA
3223 REPRO COST	READ DATA
3224 FREQUENCY	READ DATA
3225 MAILING	READ DATA
3226 NO OF COPIES	READ DATA
3227 TAPE PURCHASE	READ DATA
3228 UNIT COST	+ REPRO COST(3223) + MAILING(3225) + TAPE PURCHASE(3227)
3229 M/R COST	UNIT COST(3228) X ORIGINAL TAPES(3222) X FREQUENCY(3224) X NO OF COPIES(3226)
3230 DATABASE COSTS	COPY DATA BASE CST(1340)
3240 DIRECT O/P COSTS	SUM DIRECT COSTS(1490) THRU M/R COST(3229)
3250 RETRO D/B	DIRECT COSTS(1490) / DIRECT O/P COSTS(3240) X DATABASE COSTS(3230)
3260 ONLINE RETRO D/B	DIRECT COSTS(1730) / DIRECT O/P COSTS(3240) X DATABASE COSTS(3230)
3270 SDI D/B	DIRECT COSTS(1890) / DIRECT O/P COSTS(3240) X DATABASE COSTS(3230)
3280 GROUP SDI D/B	DIRECT COSTS(2300) / DIRECT O/P COSTS(3240) X DATABASE COSTS(3230)
3290 ALERTS D/B	DIRECT COSTS(2620) / DIRECT O/P COSTS(3240) X DATABASE COSTS(3230)
3300 ABSTRACTS D/B	DIRECT COSTS(3220) / DIRECT O/P COSTS(3240) X DATABASE COSTS(3230)
3302 M/R SERVICES D/B	M/R COST(3229) / DIRECT O/P COSTS(3240) X DATABASE COSTS(3230)
3320 DATABASE COST	
3330 RECORDS INPUT	COPY RECORDS INPUT(1200)
3340 RECORDS USED	COPY RECORDS USED(1230)
3350 RECORDS IN D/B	COPY RECORDS IN D/B(1240)
3360 PURCHASE COST	COPY PURCHASE COST(1260)
3370 INPUT PREP COST	COPY INPUT PREP COST(1140)
3390 STRIPPING	COPY STRIPPING(1310)
3400 CONVERSION	COPY CONVERSION(1330)

3420 DATABASE COSTS	COPY DATA BASE CST(1340)

3460 OUTPUT SERVICES	
3480 RETRO BATCH	+ DIRECT COSTS(1490) + RETRO D/B(3250)
3490 SEARCHES	IF RETRO BATCH(3480) GT 0 THEN SEARCHES(1360) ELSE 0

LINE	ACTION
3510 ONLINE RETRO	+ DIRECT COSTS(1730) + ONLINE RETRO D/B(3260)
3520 SEARCHES	IF ONLINE RETRO(3510) GT 0 THEN SEARCHES(1510) ELSE 0
3540 SDI	+ DIRECT COSTS(1890) + SDI D/B(3270)
3550 PROFILES	IF SDI(3540) GT 0 THEN OPERNL PROFILES(1750) ELSE 0
3560 RUNS/YEAR	IF SDI(3540) GT 0 THEN RUNS PER YEAR(1760) ELSE 0
3580 GROUP SDI	+ DIRECT COSTS(2300) + GROUP SDI D/B(3280)
3590 GROUP PROFILES	IF GROUP SDI(3580) GT 0 THEN GROUP PROFILES(2010) ELSE 0
3600 USERS/PROFILE	IF GROUP SDI(3580) GT 0 THEN USERS(2060) ELSE 0
3610 RUNS/YEAR	IF GROUP SDI(3580) GT 0 THEN RUNS PER YEAR(2020) ELSE 0
3630 ALERT PUBS	+ DIRECT COSTS(2620) + ALERTS D/B(3290)
3640 ITEMS/YEAR	COPY ITEMS PER YEAR(2490)
3650 COPIES/YEAR	ISSUES PER YEAR(2430) X COPIES PER ISSUE(2540)
3670 ABSTRACT PUB	+ DIRECT COSTS(3220) + ABSTRACTS D/B(3300)
3680 ITEMS/YEAR	COPY ITEMS PER YEAR(3090)
3690 COPIES/YEAR	COPIES PER ISSUE(3140) X ISSUES PER YEAR(3030)
3710 M/RSERVICES	+ M/R COST(3229) + M/R SERVICES D/B(3302)
3720 SUBSCRIBERS	COPY NO OF COPIES 3226)
4000 STAFF REQUIRED	
4050 GRADE C1 EFFORT	SUM MAN YEAR HOURS(1040) THRU LINESKIP(3700)
4060 EXTRA STAFF	BREAK LEVEL OF GRADE C1 EFFORT(4050) INCREMENTS OF 1
4070 ADD ONE	1 + 1 X EXTRA STAFF(4060)
4080 GRADE C1 STAFF	IF GRADE C1 EFFORT(4050) GT 0 THEN ADD ONE(4070) ELSE 0

```
       LINE                      ACTION
-----------------------  ------------------------------------
4090 GRADE C2 EFFORT     SUM MAN YEAR HOURS(1040)
                         THRU LINESKIP(3700)
4100 EXTRA STAFF         BREAK LEVEL OF
                         GRADE C2 EFFORT(4090)
                         INCREMENTS OF 1
4110 ADD ONE             1 +
                          1 X
                         EXTRA STAFF(4100)
4120 GRADE C2 STAFF      IF GRADE C2 EFFORT(4090) GT 0
                         THEN ADD ONE(4110)
                         ELSE 0
                         --------
4300 DIRECT STAFF        SUM GRADE C1 STAFF(4080)
                         THRU GRADE C2 STAFF(4120)
4310 SUPERVISORS GRADEC  READ DATA
4320 SUPERVISORS GRADED  READ DATA
4330 SUPERVISORS GRADEE  READ DATA
4332 CLERKS GRADE A      READ DATA
                         --------
4340 TOTAL STAFF         SUM DIRECT STAFF(4300)
                         THRU CLERKS GRADE A(4332)
                         --------

4370 OVERHEADS
4380 TOTAL STAFF         COPY TOTAL STAFF(4340)
4390 SPACE PER PERSON    READ DATA
4400 RENTAL              READ DATA
4410 ACCOMMODATION       TOTAL STAFF(4380)
                         X SPACE PER PERSON(4390)
                         X RENTAL(4400)
4420 SUPERVISORS C COST  SUPERVISORS GRADEC(4310)
                         X GRADE C(1080)
4430 SUPERVISORS D COST  SUPERVISORS GRADED(4320)
                         X GRADE D(1090)
4440 SUPERVISORS E COST  SUPERVISORS GRADEE(4330)
                         X GRADE E(1100)
4460 GRADE A SALARY      CLERKS GRADE A(4332)
                         X GRADE A(1060)
4510 GRADE C1 SALARY     GRADE C1 STAFF(4080)
                         X GRADE C(1080)
4520 GRADE C2 SALARY     GRADE C2 STAFF(4120)
                         X GRADE C(1080)
4530 ALL SALARIES        SUM SUPERVISORS C COST(4420)
                         THRU GRADE C2 SALARY(4520)
4540 OVERHEAD RATE       READ DATA
4550 SALARY OVERHEAD     0.01 X
                         ALL SALARIES(4530)
                         X
                         OVERHEAD RATE(4540)
4560 ACCOMMODATION       COPY ACCOMMODATION(4410)
                         --------
4580 OVERHEADS           + SALARY OVERHEAD(4550)
                         + ACCOMMODATION(4560)
```

LINE	ACTION

5000 DIRECT SALARIES	SUM GRADE A SALARY(4460) THRU GRADE C2 SALARY(4520)
5010 DIRECT O/P COSTS	COPY DIRECT O/P COSTS(3240)
5020 RETRO LABOUR	GRADE A SALARY(4460) X DIRECT COSTS(1490) / DIRECT O/P COSTS(5010) + GRADE C1 SALARY(4510) X EFFORT(1390) / GRADE C1 EFFORT(4050)
5030 ON-LINE LABOUR	GRADE A SALARY(4460) X DIRECT COSTS(1730) / DIRECT O/P COSTS(5010) + GRADE C1 SALARY(4510) X EFFORT(1540) / GRADE C1 EFFORT(4050)
5040 SDI LABOUR	GRADE A SALARY(4460) X DIRECT COSTS(1890) / DIRECT O/P COSTS(5010) + GRADE C1 SALARY(4510) X PROFILE EFFORT(1810) / GRADE C1 EFFORT(4050)
5050 GROUP LABOUR	GRADE A SALARY(4460) X DIRECT COSTS(2300) / DIRECT O/P COSTS(5010) + GRADE C1 SALARY(4510) X EFFORT(2100) / GRADE C1 EFFORT(4050)
5060 ALERT LABOUR	GRADE A SALARY(4460) X DIRECT COSTS(2620) / DIRECT O/P COSTS(5010) + GRADE C2 SALARY(4520) X EFFORT(2440) / GRADE C2 EFFORT(4090)
5070 ABSTS LABOUR	GRADE A SALARY(4460) X DIRECT COSTS(3220) / DIRECT O/P COSTS(5010) + GRADE C2 SALARY(4520) X EFFORT(3040) / GRADE C2 EFFORT(4090)
5080 M/R LABOUR	GRADE A SALARY(4460) X M/R COST(3229) / DIRECT O/P COSTS(5010)
5090 SUPERVISORY LAB	SUM SUPERVISORS C COST(4420) THRU SUPERVISORS E COST(4440)
5100 RETRO SUPER	DIRECT COSTS(1490) / DIRECT O/P COSTS(5010) X SUPERVISORY LAB(5090)
5110 ONLINE SUPER	DIRECT COSTS(1730) / DIRECT O/P COSTS(5010) X SUPERVISORY LAB(5090)
5120 SDI SUPER	DIRECT COSTS(1890) / DIRECT O/P COSTS(5010) X SUPERVISORY LAB(5090)
5130 GROUP SUPER	DIRECT COSTS(2300) / DIRECT O/P COSTS(5010) X SUPERVISORY LAB(5090)

LINE	ACTION
5140 ALERT SUPER	DIRECT COSTS(2620) /
	DIRECT O/P COSTS(5010) X
	SUPERVISORY LAB(5090)
5150 ABSTRACTS SUPER	DIRECT COSTS(3220) /
	DIRECT O/P COSTS(5010) X
	SUPERVISORY LAB(5090)
5160 M/R SUPER	M/R COST(3229) /
	DIRECT O/P COSTS(5010) X
	SUPERVISORY LAB(5090)
5170 OVERHEADS	COPY OVERHEADS(4580)
5180 RETRO	DIRECT COSTS(1490) /
	DIRECT O/P COSTS(5010) X
	OVERHEADS(5170)
5190 ONLINE	DIRECT COSTS(1730) /
	DIRECT O/P COSTS(5010) X
	OVERHEADS(5170)
5200 SDI	DIRECT COSTS(1890) /
	DIRECT O/P COSTS(5010) X
	OVERHEADS(5170)
5210 GROUP	DIRECT COSTS(2300) /
	DIRECT O/P COSTS(5010) X
	OVERHEADS(5170)
5220 ALERTS	DIRECT COSTS(2620) /
	DIRECT O/P COSTS(5010) X
	OVERHEADS(5170)
5230 ABSTRACTS	DIRECT COSTS(3220) /
	DIRECT O/P COSTS(5010) X
	OVERHEADS(5170)
5240 M/R	M/R COST(3229) /
	DIRECT O/P COSTS(5010) X
	OVERHEADS(5170)
5250 PROJ OUTPUT COST	
5270 RETRO SEARCH	
5280 STAFF	COPY RETRO LABOUR(5020)
5290 SUPERVSTON	COPY RETRO SUPER(5100)
5292 ROYALTIES	COPY ROYALTIES(1480)
5302 MAILING	COPY MAILING(1420)
5310 EQUIPMENT	COPY COMPUTER(1440)
5320 INPUT	COPY RETRO D/B(3250)
5330 OVERHEADS	COPY RETRO(5180)

5334 RETRO SEARCH	SUM STAFF(5280)
	THRU OVERHEADS(5330)

5350 ONLINE SEARCH	
5360 STAFF	COPY ON-LINE LABOUR(5030)
5370 SUPERVISION	COPY ONLINE SUPER(5110)
5372 ROYALTIES	COPY ROYALTIES(1640)
5382 MAILING	COPY MAILING(1580)
5390 EQUIPMENT	+ COMPUTER(1600)
	+ TERMINAL COST(1670)
	+ LINE RENTAL(1680)
	+ STORAGE(1720)

LINE	ACTION
5400 INPUT	COPY ONLINE RETRO D/B(3260)
5410 OVERHEADS	COPY ONLINE(5190)

5414 ON LINE SEARCH	SUM ONLINE SEARCH(5350)
	THRU OVERHEADS(5410)

5440 SDI	
5450 STAFF	COPY SDI LABOUR(5040)
5460 SUPERVISION	COPY SDI SUPER(5120)
5462 ROYALTIES	COPY ROYALTIES(1880)
5470 MATERIALS	COPY PAPER(1856)
5472 MAILING	COPY MAILING(1800)
5480 EQUIPMENT	COPY COMPUTER(1840)
5490 INPUT	COPY SDI D/B(3270)
5500 OVERHEADS	COPY SDI(5200)

5504 SDI	SUM STAFF(5450)
	THRU OVERHEADS(5500)

5530 GROUP SDI	
5540 STAFF	COPY GROUP LABOUR(5050)
5550 SUPERVISION	COPY GROUP SUPER(5130)
5552 ROYALTIES	COPY ROYALTIES(2150)
5560 MATERIALS	+ REPRODUCTION(2200)
5562 MAILING	COPY MAILING(2080)
5570 EQUIPMENT	COPY COMPUTER(2120)
5580 INPUT	COPY GROUP SDI D/B(3280)
5590 OVERHEADS	COPY GROUP(5210)

5604 GROUP SDI	SUM STAFF(5540)
	THRU OVERHEADS(5590)

5620 ALERTING JOURNAL	
5630 STAFF	COPY ALERT LABOUR(5060)
5640 SUPERVISION	COPY ALERT SUPER(5140)
5642 ROYALTIES	COPY ROYALTIES(2618)
5650 MATERIALS	+ REPRODUCTION(2550)
	+ BINDING(2570)
5652 MAILING	COPY MAILING(2610)
5660 EQUIPMENT	COPY COMPUTER(2590)
5670 INPUT	COPY ALERTS D/B(3290)
5680 OVERHEADS	COPY ALERTS(5220)

5694 ALERTING JOURNAL	SUM STAFF(5630)
	THRU OVERHEADS(5680)

5710 ABSTRACTS JOURNAL	
5720 STAFF	COPY ABSTS LABOUR(5070)
5730 SUPERVISION	COPY ABSTRACTS SUPER(5150)
5732 ROYALTIES	COPY ROYALTIES(3218)
5740 MATERIALS	+ REPRODUCTION(3150)
	+ BINDING(3170)

LINE	ACTION
5742 MAILING	COPY MAILING(3190)
5750 EQUIPMENT	COPY COMPUTER(3210)
5760 INPUT	COPY ABSTRACTS D/B(3300)
5770 OVERHEADS	COPY ABSTRACTS(5230)

5774 ABSTRACTS JOURNAL	SUM STAFF(5720)
	THRU OVERHEADS(5770)

5800 M/R SERVICES	
5801 STAFF	COPY M/R LABOUR(5080)
5802 SUPERVISION	COPY M/R SUPER(5160)
5803 MATERIALS	ORIGINAL TAPES(3222)
	X FREQUENCY(3224)
	X NO OF COPIES(3226)
	X TAPE PURCHASE(3227)
5804 MAILING	ORIGINAL TAPES(3222)
	X FREQUENCY(3224)
	X NO OF COPIES(3226)
	X MAILING(3225)
5805 EQUIPMENT	REPRO COST(3223)
	X ORIGINAL TAPES(3222)
	X FREQUENCY(3224)
	X NO OF COPIES(3226)
5806 INPUT	COPY M/R SERVICES D/B(3302)
5807 OVERHEADS	COPY M/R(5240)

5809 M/R SERVICES	SUM STAFF(5801)
	THRU OVERHEADS(5807)
	========
5811 PROJ OUTPUT COSTS	+ RETRO SEARCH(5334)
	+ ON LINE SEARCH(5414)
	+ SDI(5504)
	+ GROUP SDI(5604)
5812 PROJ OUTPUT COSTS	+ ALERTING JOURNAL(5694)
	+ ABSTRACTS JOURNAL(5774)
	+ M/R SERVICES(5809)
	+ PROJ OUTPUT COSTS(5811)
	========
5820 PROJ OUTPUT COST	
5830 STAFF	COPY DIRECT SALARIES(5000)
5840 SUPERVISION	COPY SUPERVISORY LAB(5090)
5842 ROYALTIES	+ ROYALTIES(5552)
	+ ROYALTIES(5642)
	+ ROYALTIES(5732)
	+ ROYALTIES(5292)
	+ ROYALTIES(5372)
	+ ROYALTIES(5462)
5850 MATERIALS	+ MATERIALS(5470)
	+ MATERIALS(5560)
	+ MATERIALS(5650)
	+ MATERIALS(5740)
	+ MATERIALS(5803)

LINE	ACTION
5852 MAILING	+ MAILING(5302)
	+ MAILING(5382)
	+ MAILING(5472)
	+ MAILING(5562)
5853 MAILING	+ MAILING(5852)
	+ MAILING(5652)
	+ MAILING(5742)
	+ MAILING(5804)
5860 EQUIPMENT	+ EQUIPMENT(5310)
	+ EQUIPMENT(5390)
	+ EQUIPMENT(5480)
	+ EQUIPMENT(5570)
5861 EQUIPMENT	+ EQUIPMENT(5860)
	+ EQUIPMENT(5660)
	+ EQUIPMENT(5750)
	+ EQUIPMENT(5805)
5870 INPUT	COPY DATABASE COSTS(3230)
5880 OVERHEADS	COPY OVERHEADS(5170)
	========
5900 PROJ OUTPUT COSTS	SUM STAFF(5830)
	THRU OVERHEADS(5880)
6800 DIRECT STAFF USE	
6870 GRADE C1	COPY GRADE C1 EFFORT(4050)
6880 GRADE C2	COPY GRADE C2 EFFORT(4090)

APPENDIX 3 - INPUT FORMS

The computer system automatically generates an input form which can be easily adapted for entering data into the projection or history files.

In the two forms reproduced on the following pages have been entered the data values (and appropriate projection codes) from which were generated the summary reports shown in Appendix 4.

In the first system suggested, the data base is created in-house and the parameters describing this are taken from the input system proposed in the companion report (lines 1140,1200). The output services offered are :

SDI	*Fortnightly*	*In-house*
Group SDI	*Fortnightly*	
Alerting bulletin	*Fortnightly*	
Abstracting journal	*Monthly*	*Published*
Magnetic tapes	*Monthly*	

The second system proposed purchases its entire data base in the form of compatible magnetic tapes and initially offers two output services :

On-line retrosearch	*(on entire data base)*
SDI	*(weekly)*

PROJECTION INPUT DATA FORM (*OPTIONAL ENTRIES)
FROM DEFINITION FILE DEFOPAA

0 PROJECT 1 5 (FIRST,LAST COLUMNS)

COLUMN DIST. :* 1 3

START DATE :* 2 1/1/76

REPORT HEAD 1 : 11 ASLIB OUTPUT MODEL, SYSTEM A

REPORT HEAD 2 : 12 FIVE YEAR PROJECTION

COLUMN TOTALS :* 31 .5, .5, .5, .5, .5, .5,0

COLUMN LABELS 1 :* 51 %

:* 52 :

COLUMN LABELS 2 :* 61 %

:* 62 :

MAN YEAR HOURS	1040	5,1350
	1041	
GRADE A	1060	1.8, 1, 2600, 5.0
	1061	
GRADE B	1070	1.8, 1, 3400, 5.0
	1071	
GRADE C	1080	1.8, 1, 4200, 5.0
	1081	
GRADE D	1090	1.8, 1, 5800, 5.0
	1091	
GRADE E	1100	1.8, 1, 7200, 5.0
	1101	
INPUT PREP COST	1140	1,347015, 418423 516923 625097 780304
	1141	
RECORDS INPUT	1200	1,26000 27500, 29000, 30500 32000
	1201	

OPERNL PROFILES	1750	3.1, 150, 320
	1751	
PROFILES ADDED	1752	1, 0, 50, 50, 40, 30
	1753	
RUNS PER YEAR	1760	5, 26
	1761	
FORMULATE TM	1770	5, 0.5
	1771	
MAINTAIN TM	1774	5, 0.5
	1775	
COMPUTER	1780	1.8, 1, 2.0, 7.0
	1781	
MAILING	1790	5, 0
	1791	
ITEMS OUTPUT	1850	5, 35
	1851	
ITEMS PER PAGE	1852	5, 12
	1853	
PAGE COST	1854	5, 0
	1855	
ROYALTY PROFILE	1860	5, 0
	1861	
ROYALTY ABSTRACT	1870	5, 0
	1871	
GROUP PROFILES	2010	3.1, 150, 200
	2011	
NEW PROFILES	2012	5, 10
	2013	
RUNS PER YEAR	2020	5, 26
	2021	

FORMULATE TM	2030	5,0.5
	2031	
MAINTAIN TM	2032	5,0.5
	2033	
COMPUTER	2040	1.8,1,2.0,7.0
	2041	
MAILING	2050	5,0
	2051	
USERS	2060	5,9
	2061	
ITEMS OUTPUT	2070	5,80
	2071	
ITEMS PER PAGE	2072	5,12
	2073	
ROYALTY PROFILE	2130	5,0
	2131	
ROYALTY ABSTRACT	2140	5,0
	2141	
REPRO COST	2190	18,1,0.02,8.0
	2191	
EDIT TM	2410	5,1.0
	2411	
ISSUES PER YEAR	2430	5,26
	2431	
ALERTS PER PAGE	2460	5,30
	2461	
INDEX PAGE RATIO	2470	5,0.2
	2471	
ITEMS PER YEAR	2490	1,26000,27500,29000,30500,32000
	2491	
EDITORIAL PAGES	2502	5,3

REPRODUCTION	2530	1.8, 1, 0.065, 8.0
	2531	
COPIES PER ISSUE	2540	3.1, 500, 800
	2541	
BINDING	2560	1.8, 1, 0.12, 8.0
	2561	
COMPUTER COST	2580	3.1, 80, 150
	2581	
MAILING	2600	1.8, 1, 0.35, 12.0
	2601	
ROYALTIES	2618	5, 0
	2619	
EDIT TM	3010	5, 14
	3011	
ISSUES PER YEAR	3030	5, 12
	3031	
ITEMS PER PAGE	3060	5, 10
	3061	
INDEX PAGE RATIO	3070	5, 0.31
	3071	
ITEMS PER YEAR	3090	3.1, 26000, 32000
	3091	
EDITORTAL PAGES	3112	5, 7
	3113	
REPRODUCTION	3130	1.8, 1, 0.09, 8.0
	3131	
COPIES PER ISSUE	3140	3.1, 250, 790
	3141	
BINDING	3160	1.8, 1, 0.25, 9.0
	3161	
MAILING	3180	1.8, 1, 1.50, 12.0

COMPUTER	3200	3.1, 110, 190
	3201	
ROYALTIES	3218	5, 0
	3219	
ORIGTNAL TAPES	3222	5, 2
REPRO COST	3223	1.8, 1, 1.0, 5.0
FREQUENCY	3224	5, 12
MAILING	3225	1.8, 1, 1.85, 12.0
NO OF COPIES	3226	3.1, 50, 130
TAPE PURCHASE	3227	1.8, 1, 5.75, 5.0
SUPERVISORS GRADEC	4310	
	4311	
SUPERVISORS GRADED	4320	5, 1
	4321	
SUPERVISORS GRADEE	4330	
	4331	
CLERKS GRADE A	4332	5, 3
	4333	
SPACE PER PERSON	4390	5, 150
	4391	
RENTAL	4400	1.8, 1, 10, 7.0
	4401	
OVERHEAD RATE	4540	5, 75
	4541	

PROJECTION INPUT DATA FORM (*OPTIONAL ENTRIES)
FROM DEFINITION FILE DEFOPAA

0 PROJECT 1,5 (FIRST,LAST COLUMNS)

COLUMN DIST. :* 1 3

START DATE :* 2 1/1/76

REPORT HEAD 1 : 11 ASLIB OUTPUT MODEL, SYSTEM B :

REPORT HEAD 2 : 12 FIVE YEAR PROJECTION :

COLUMN TOTALS :* 31 .5, .5, .5, .5, .5, .5,0

COLUMN LABELS 1 :* 51 %

:* 52 :

COLUMN LABELS 2 :* 61 %

:* 62 :

MAN YEAR HOURS 1040 5,1350

1041

GRADE A 1060 1.8,1,2600,5.0

1061

GRADE B 1070 1.8,1,3400,5.0

1071

GRADE C 1080 1.8,1,4200,5.0

1081

GRADE D 1090 1.8,1,5800,5.0

1091

GRADE E 1100 1.8,1,7200,5.0

1101

RECORDS PCHSD 1210 1.8,1,36000,3.0

1211

RECORDS STRPD 1220 5,0

1221

PURCHASE COST	1260	1.8, 1,2000 , 10.0
	1261	
STRIPPING	1300	5,0
	1301	
CONVERSION	1320	5,0
	1321	
SEARCHES	1510	3.1,200,700
	1511	
FORMULATIONS	1515	5,0
	1516	
FORMULATE TM	1520	5,0
	1521	
MAILING	1560	1.8,1,0.50,8.0
	1561	
OFF-LINE PRINTS	1570	3.1,66,233
	1571	
COMPUTER COST	1590	1.8,1,8.0,10.0
	1591	
ROYALTY SEARCHES	1610	3.1,1,0.30
	1611	
ROYALTY ABSTRACT	1620	1.8,1,0.054,10.0
	1621	
ITEMS RETRIEVED	1630	5,45
	1631	

TERMINALS	1650	5,0
	1651	
RENTAL	1660	5,0
	1661	
LINE RENTAL	1680	5,0
	1681	
FILE STORAGE	1690	1.8,1,125,—16.0
	1691	
STORED D/B SIZE	1700	3.1,9,12.5
	1701	
ACCESS	1710	5,0.33
	1711	
OPERNL PROFILES	1750	3.1,40,.85
	1751	
PROFILES ADDED	1752	1,9,9,9,9,9
	1753	
RUNS PER YEAR	1760	5,52
	1761	
FORMULATE TM	1770	5,0.5
	1771	
MAINTAIN TM	1774	5,0.5
	1775	
COMPUTER	1780	1.8,1,2.0,8.0
	1781	
MAILING	1790	1.8,1,0.35,8.0
	1791	
ITEMS OUTPUT	1850	5,10
	1851	
ITEMS PER PAGE	1852	5,10
	1853	

PAGE COST	1854	5,0
	1855	
ROYALTY PROFILE	1860	1.8,1,1.5,8.0
	1861	
ROYALTY ABSTRACT	1870	5,0
	1871	
SUPERVISORS GRADEC	4310	
	4311	
SUPERVISORS GRADED	4320	
	4321	
SUPERVISORS GRADEE	4330	5,1
	4331	
CLERKS GRADE A	4332	5,1
	4333	
SPACE PER PERSON	4390	5,150
	4391	
RENTAL	4400	1.8,1,10,7.0
	4401	
OVERHEAD RATE	4540	5,75
	4541	

APPENDIX 4 - SUMMARY REPORTS

Available data in each of the projection files is run in turn against the model contained in the definition file and will yield summary reports of all operational costs associated with output services. Reports prepared from the data shown in Appendix 3 are reproduced on the following pages.

The way in which each line of the reports has been calculated can be traced by reference to the ILLUSTRATE listing in Appendix 2. For example, line 3540, showing SDI costs, is seen to be the sum of lines 1890 and 3270. Line 1890 (Direct costs) is the sum of lines 1830, 1840, 1880 and 1800. Line 1830 (Labour) calls for multiplication of line 1810 (Profile effort) by line 1820 (Staff cost). Line 1820 is copied from line 1080, which calls for an input value for a Grade C staff salary. The value used for this parameter, in producing the report, is shown on the input form in Appendix 3. The same process of tracing back can be applied to any part of the reports which follow.

ASLIB OUPUT MODEL SYSTEM A
FIVE YEAR PROJECTION
FOR THF PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 27, 1976

		1 1976	2 1977	3 1978	4 1979	5 1980
3320	DATABASE COST					
3330	RECORDS INPUT	26000.0	27500.0	29000.0	30500.0	32000.0
3350	RECORDS IN D/B	26000.0	27500.0	29000.0	30500.0	32000.0
3370	INPUT PREP COST	*347015.0	418423.0	516923.0	625097.0	780304.0
3420	DATABASE COSTS	*347015.0	418423.0	516923.0	625097.0	780304.0
3460	OUTPUT SERVICES					
3540	SDI	25315.4	29819.9	35594.3	41875.9	50009.0
3550	PROFILES	150.0	192.5	235.0	277.5	320.0
3560	RUNS/YEAR	26.0	26.0	26.0	26.0	26.0
3580	GROUP SDI	40112.4	39732.5	42047.1	45134.4	50109.9
3590	GROUP PROFILES	150.0	162.5	175.0	187.5	200.0
3600	USERS/PROFILE	9.0	9.0	9.0	9.0	9.0
3610	RUNS/YEAR	26.0	26.0	26.0	26.0	26.0
3630	ALERT PUBS	*140565.9	155633.9	181600.5	212743.4	255729.1
3640	ITEMS/YEAR	26000.0	27500.0	29000.0	30500.0	32000.0
3650	COPIES/YEAR	13000.0	14950.0	16900.0	18850.0	20800.0
3670	ABSTRACT PUB	*269805.6	400269.0	560321.3	745077.5	984131.1
3680	ITEMS/YEAR	26000.0	27500.0	29000.0	30500.0	32000.0
3690	COPIES/YEAR	3000.0	4620.0	6240.0	7860.0	9480.0
3710	M/RSERVICES	32521.3	41314.1	51837.5	63082.1	77173.6
3720	SUBSCRIBERS	50.0	70.0	90.0	110.0	130.0
4000	STAFF REQUIRED					
4080	GRADE C1 STAFF	1.0	1.0	1.0	1.0	1.0
4120	GRADE C2 STAFF	1.0	1.0	1.0	1.0	1.0
4300	DIRECT STAFF	2	2	2	2	2
4320	SUPERVISORS GRADED	1	1	1	1	1
4332	CLERKS GRADE A	3	3	3	3	3
4340	TOTAL STAFF	6	6	6	6	6
4370	OVERHEADS					
4380	TOTAL STAFF	6	6	6	6	6
4530	ALL SALARIES	22000	23100	24255	25468	26741
4540	OVERHEAD RATE	75	75	75	75	75
4550	SALARY OVERHEAD	16500	17325	18191	19101	20056
4560	ACCOMMODATION	9000	9630	10304	11025	11797
4580	OVERHEADS	25500	26955	28495	30126	31853

		1 1976	2 1977	3 1978	4 1979	5 1980
5250	PROJ OUTPUT COST					
5270	RETRO SEARCH					
5334	RETRO SEARCH	0	0	0	0	0
5350	ONLINE SEARCH					
5414	ON LINE SEARCH	0	0	0	0	0
5440	SDI					
5450	STAFF	2420.7	2943.2	3159.1	3338.7	3525.3
5460	SUPERVISION	288.9	272.4	261.2	253.8	248.8
5480	EQUIPMENT	7800.0	10710.7	13990.7	17677.3	21811.6
5490	INPUT	17282.0	18713.1	21114.9	23626.9	27535.6
5500	OVERHEADS	1269.9	1205.5	1164.0	1138.7	1124.0
5504	SDI	29062	33845	39690	46035	54245
5530	GROUP SDI					
5540	STAFF	2783.3	2321.1	2237.6	2232.4	2249.7
5550	SUPERVISION	457.7	362.9	308.5	273.5	249.3
5560	MATERIALS	4680.0	5475.6	6368.5	7369.3	8489.4
5570	EQUIPMENT	7800.0	9041.5	10418.6	11944.1	13632.2
5580	INPUT	27383.5	24933.6	24942.8	25465.3	27591.2
5590	OVERHEADS	2012.2	1606.2	1375.0	1227.3	1126.3
5604	GROUP SDI	45117	43741	45651	48512	53338
5620	ALERTING JOURNAL					
5630	STAFF	2719.8	2502.7	2412.7	2385.5	2395.0
5640	SUPERVISION	1603.9	1421.5	1332.6	1289.3	1272.2
5650	MATERIALS	37895.0	49487.4	63374.5	79903.6	99467.7
5652	MAILING	4550.0	5860.4	7419.8	9269.0	11455.2
5660	EQUIPMENT	2080.0	2535.0	2990.0	3445.0	3900.0
5670	INPUT *	95960.1	97666.1	107727.1	120032.2	140807.9
5680	OVERHEADS	7051.5	6291.7	5938.4	5784.9	5747.9
5694	ALERTING JOURNAL	151860	165765	191195	222109	265046
5710	ABSTRACTS JOURNAL					
5720	STAFF	7777.2	8735.5	9539.5	10282.7	11004.9
5730	SUPERVISION	3078.5	3655.9	4111.7	4515.3	4895.8
5740	MATERIALS *	79274.9	139214.8	213816.1	305487.5	416964.5
5742	MAILING	4500.0	7761.6	11741.2	16564.1	22375.4
5750	EQUIPMENT	1320.0	1560.0	1800.0	2040.0	2280.0
5760	INPUT *	184188.1	251183.9	332387.8	420380.9	541875.9
5770	OVERHEADS	13534.8	16181.4	18322.8	20260.0	22120.0
5774	ABSTRACTS JOURNAL	293673	428293	591719	779531	1021516

		1 1976	2 1977	3 1978	4 1979	5 1980
5800	M/R SERVICES					
5801	STAFF	499.0	507.5	511.6	514.1	516.3
5802	SUPERVISION	371.1	377.3	380.4	382.3	383.9
5803	MATERIALS	6900.0	10143.0	13693.0	17572.7	21806.1
5804	MAILING	2220.0	3481.0	5012.6	6861.7	9082.3
5805	EQUIPMENT	1200.0	1764.0	2381.4	3056.1	3792.4
5806	INPUT	22201.3	25926.2	30750.5	35591.6	42492.8
5807	OVERHEADS	1631.4	1670.2	1695.1	1715.3	1734.6
5809	M/R SERVICES	35023	43869	54425	65694	79808
5812	PROJ OUTPUT COSTS	554735	715513	922679	1161880	1473953
5820	PROJ OUTPUT COST					
5830	STAFF	16200	17010	17860	18753	19691
5840	SUPERVISION	5800	6090	6394	6714	7050
5842	ROYALTIES	0	0	0	0	0
5850	MATERIALS	128750	204321	297252	410333	546728
5853	MAILING	11270	17103	24174	32695	42913
5861	EQUIPMENT	20200	25611	31581	38163	45416
5870	INPUT	347015	418423	516923	625097	780304
5880	OVERHEADS	25500	26955	28495	30126	31853
5900	PROJ OUTPUT COSTS	554735	715513	922679	1161880	1473953
6800	DIRECT STAFF USE					
6870	GRADE C1	0.1	0.2	0.2	0.2	0.2
6880	GRADE C2	0.1	0.1	0.1	0.1	0.1

ASLIB OUTPUT MODELSYSTEM B
FIVE YEAR PROJECTION
FOR THE PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 27, 1976

		1 1976	2 1977	3 1978	4 1979	5 1980
3320	DATABASE COST					
3340	RECORDS USED	36000.0	37080.0	38192.4	39338.1	40518.2
3350	RECORDS IN D/B	36000.0	37080.0	38192.4	39338.1	40518.2
3360	PURCHASE COST	2000.0	2200.0	2420.0	2662.0	2928.2
3420	DATABASE COSTS	2000.0	2200.0	2420.0	2662.0	2928.2
3460	OUTPUT SERVICES					
3510	ONLINE RETRO	1446.9	1935.9	2453.0	3013.9	3634.1
3520	SEARCHES	200.0	325.0	450.0	575.0	700.0
3540	SDI	6667.6	8746.5	11123.1	13826.8	16892.3
3550	PROFILES	40.0	51.3	62.5	73.8	85.0
3560	RUNS/YEAR	52.0	52.0	52.0	52.0	52.0
4000	STAFF REQUIRED					
4080	GRADE C1 STAFF	1.0	1.0	1.0	1.0	1.0
4300	DIRECT STAFF	1	1	1	1	1
4330	SUPERVISORS GRADEE	1	1	1	1	1
4332	CLERKS GRADE A	1	1	1	1	1
4340	TOTAL STAFF	3	3	3	3	3
4370	OVERHEADS					
4380	TOTAL STAFF	3	3	3	3	3
4530	ALL SALARIES	14000	14700	15435	16207	17017
4540	OVERHEAD RATE	75	75	75	75	75
4550	SALARY OVERHEAD	10500	11025	11576	12155	12763
4560	ACCOMMODATION	4500	4815	5152	5513	5899
4580	OVERHEADS	15000	15840	16728	17668	18661
5250	PROJ OUTPUT COST					
5270	RETRO SEARCH					
5334	RETRO SEARCH	0	0	0	0	0
5350	ONLINE SEARCH					
5360	STAFF	463.6	494.7	517.9	538.7	559.5
5370	SUPERVISION	1283.8	1370.0	1434.3	1491.7	1549.4

		1 1976	2 1977	3 1978	4 1979	5 1980
5372	ROYALTIES	686.0	1136.8	1615.6	2132.9	2700.4
5382	MAILING	33.0	58.2	87.2	120.5	158.5
5390	EQUIPMENT	1971.2	3202.2	4668.9	6406.8	8455.6
5400	INPUT	356.6	398.7	437.3	476.4	518.4
5410	OVERHEADS	2674.6	2870.6	3022.5	3162.0	3303.9
5414	ON LINE SEARCH	7469	9531	11784	14329	17246
5440	SDI					
5450	STAFF	6336.4	6645.3	6979.1	7333.2	7705.9
5460	SUPERVISION	5916.2	6189.9	6503.7	6843.2	7202.2
5462	ROYALTIES	60.0	83.0	109.3	139.4	173.5
5472	MAILING	728.0	1007.4	1326.8	1690.8	2104.7
5480	EQUIPMENT	4160.0	5756.4	7581.6	9662.0	12026.7
5490	INPUT	1643.4	1801.3	1982.7	2185.6	2409.8
5500	OVERHEADS	12325.4	12969.4	13705.7	14505.7	15357.5
5504	SDI	31169	34453	38189	42360	46980
5530	GROUP SDI					
5604	GROUP SDI	0	0	0	0	0
5620	ALERTING JOURNAL					
5694	ALERTING JOURNAL	0	0	0	0	0
5710	ABSTRACTS JOURNAL					
5774	ABSTRACTS JOURNAL	0	0	0	0	0
5800	M/R SERVICES					
5809	M/R SERVICES	0	0	0	0	0
5812	PROJ OUTPUT COSTS	38638	43984	49973	56689	64226
5820	PROJ OUTPUT COST					
5830	STAFF	6800	7140	7497	7872	8265
5840	SUPERVISION	7200	7560	7938	8335	8752
5842	ROYALTIES	746	1220	1725	2272	2874
5850	MATERIALS	0	0	0	0	0
5853	MAILING	761	1066	1414	1811	2263
5861	EQUIPMENT	6131	8959	12250	16069	20482
5870	INPUT	2000	2200	2420	2662	2928
5880	OVERHEADS	15000	15840	16728	17668	18661
5900	PROJ OUTPUT COSTS	38638	43984	49973	56689	64226

APPENDIX 5 - USE OF THE WHAT-IF FEATURE

The WHAT-IF command makes it possible to examine the effect of changes in input data values, or in the overall cost structure. In the examples which follow, the sequence of prompts from the computer system and the replies given are reproduced. The user can call for a complete revised summary report, or a print-out of specified lines (which is cheaper). The changes investigated relate to the two reports shown in Appendix 4.

SYSTEM A

1. *WHAT-IF the SDI service (lines 1750-1870) were abandoned. Here we have requested to see the effect on the summary of overall costs (lines 5820-5900).*

```
COMMAND? WHAT-IF
WHAT-IF DEFINITION FILE? ALTA
REPORT INFILE,OUTFILE? OUTPRINA,WHOPA
WHOPA DOES NOT EXIST BUT IS NOW BEING CREATED
REPORT FILE WHOPA COMPLETED

COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? RAN
FIRST,LAST LINES*? 5820,5900
SET PAPER,RETURN...
```

ASLTB OUPUT MODEL SYSTEM A
FIVE YEAR PROJECTION
FOR THE PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 27, 1976

		1 1976	2 1977	3 1978	4 1979	5 1980
5820	PROJ OUTPUT COST					
5830	STAFF	16200	17010	17860	18753	19691
5840	SUPERVISION	5800	6090	6394	6714	7050
5842	ROYALTIES	0	0	0	0	0
5850	MATERIALS	128750	204321	297252	410333	546728
5853	MAILING	11270	17103	24174	32695	42913
5861	EQUIPMENT	12400	14900	17590	20485	23605
5870	INPUT	347015	418423	516923	625097	780304
5880	OVERHEADS	25500	26955	28495	30126	31853
		=======	=======	=======	=======	=======
5900	PROJ OUTPUT COSTS	546935	704802	908689	1144203	1452142

2. ***WHAT-IF print-size were reduced in the alerting and abstracting journals (lines 2460 and 3060) allowing more items to be printed on each page. Here we wish to know the effect on overall projected costs (line 5900).***

```
COMMAND? WHAT-IF
WHAT-IF DEFINITION FILE? (T)
REPORT INFILE,OUTFILE? OUTPRINA,WHOPB
WHOPB DOES NOT EXIST BUT IS NOW BEING CREATED
LINE? 2460
TYPE,FIRST, LAST COLUMN? REP,1,5
ENTER DATA( 5 ITEMS)
? % 42,42,42,42,42
LINE? 3060
TYPE,FIRST, LAST COLUMN? REP,1,5
ENTER DATA( 5 ITEMS)
? % 13,13,13,13,13
LINE? 0
REPORT FILE WHOPB COMPLETED

COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? SEL
LINES; AFTER LAST 0*
? % 5900,0
SET PAPER,RETURN...
```

ASLIB OUPUT MODEL SYSTEM A
FIVE YEAR PROJECTION
FOR THE PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 27, 1976

	1 1976	2 1977	3 1978	4 1979	5 1980
5900 PROJ OUTPUT COSTS	527393	671716	858490	1072718	1354505

3. ***WHAT-IF the number of items appearing in the abstracts journal were substantially decreased due to a change in coverage (line 3090). Here we wish to see the detailed effects on costs for this service (lines 5710-5774) as well as on overall costs.***

```
COMMAND? WHAT-IF
WHAT-IF DEFINITION FILE? (T)
REPORT INFILE,OUTFILE? OUTPRINA,WHOPC
WHOPC DOES NOT EXIST BUT IS NOW BEING CREATED
LINE? 3090
TYPE,FIRST, LAST COLUMN? 7,1,5
PERCENT ADDED? -17.0
LINE? 0
REPORT FILE WHOPC COMPLETED

COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? MRAN
FIRST,LAST LINES; AFTER LAST 0,0
? % 5710,5776,5900,6790,0,0
SET PAPER,RETURN...
```

ASLIB OUPUT MODEL SYSTEM A
FIVE YEAR PROJECTION
FOR THE PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 27, 1976

		1 1976	2 1977	3 1978	4 1979	5 1980
5710	ABSTRACTS JOURNAL					
5720	STAFF	7455.6	8402.7	9200.4	9937.8	10653.2
5730	SUPERVISION	2839.4	3408.4	3859.6	4258.8	4634.3
5740	MATERIALS	* 66246.9	116296.7	178562.0	255047.7	348030.8
5742	MAILING	4500.0	7761.6	11741.2	16564.1	22375.4
5750	EQUIPMENT	1320.0	1560.0	1800.0	2040.0	2280.0
5760	INPUT	*169881.9	234181.6	312008.3	396499.4	512935.8
5770	OVERHEADS	12483.6	15086.1	17199.4	19109.0	20938.6
5774	ABSTRACTS JOURNAL	264727	386697	534371	703457	921848
5900	PROJ OUTPUT COSTS	541707	692595	887425	1111441	1405019

SYSTEM B

1. ***WHAT-IF advanced technology permitted a substantial reduction in storage costs to be made (line 1690). Here we wish to see the effects upon the cost of the on-line retro search service and again upon overall costs.***

```
COMMAND? WHAT-IF
WHAT-IF DEFINITION FILE? (T)
REPORT INFILE,OUTFILE? OUTPRINC,WHOPD
WHOPD DOES NOT EXIST BUT IS NOW BEING CREATED
LINE? 1690
TYPE,FIRST, LAST COLUMN? 7,1,5
PERCENT ADDED? -20
LINE? 0
REPORT FILE WHOPD COMPLETED

COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? MRAN
FIRST,LAST LINES; AFTER LAST 0,0
? % 5350,5416,5900,6790,0,0
SET PAPER,RETURN...
```

ASLIB OUTPUT MODELSYSTEM B
FIVE YEAR PROJECTION
FOR THE PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 27, 1976

		1 1976	2 1977	3 1978	4 1979	5 1980
5350	ONLINE SEARCH					
5360	STAFF	437.3	476.6	504.7	528.7	551.9
5370	SUPERVISION	1211.1	1319.7	1397.6	1464.1	1528.3
5372	ROYALTIES	686.0	1136.8	1615.6	2132.9	2700.4
5382	MAILING	33.0	58.2	87.2	120.5	158.5
5390	EQUIPMENT	1897.0	3133.7	4606.3	6350.0	8404.3
5400	INPUT	336.4	384.0	426.1	467.6	511.4
5410	OVERHEADS	2523.1	2765.1	2945.2	3103.6	3258.9
5414	ON LINE SEARCH	7124	9274	11583	14167	17114
5900	PROJ OUTPUT COSTS	38564	43916	49910	56632	64175

2. ***WHAT-IF the number of searches made each year were increased by 34 per cent (line 1510). In this case we wish to see the effect upon staff requirement (lines 4000 to 4120) and overall costs (line 5900).***

```
COMMAND? WHAT-IF
WHAT-IF DEFINITION FILE? (T)
REPORT INFILE,OUTFILE? OUTPRINC,WHOPE
WHOPE DOES NOT EXIST BUT IS NOW BEING CREATED
LINE? 1510
TYPE,FIRST. LAST COLUMN? 7,1,5
PERCENT ADDED? 34.0
LINE? 0
REPORT FILE WHOPE COMPLETED

COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? MRAN
FIRST,LAST LINES; AFTER LAST 0,0
? % 4000,4120,5900,6790,0,0
SET PAPER,RETURN...
```

ASLIB OUTPUT MODELSYSTEM B
FIVE YEAR PROJECTION
FOR THE PERIOD BEGINNING JAN 1, 1976
REPORT PREPARED JUL 27, 1976

	1 1976	2 1977	3 1978	4 1979	5 1980
4000 STAFF REQUIRED					
4080 GRADE C1 STAFF	1.0	1.0	1.0	1.0	1.0
5900 PROJ OUTPUT COSTS	39415	45343	52003	59496	67932

APPENDIX 6 - SENSITIVITY TESTS

The impact upon projected costs of alterations to model parameters can be clearly shown by WHAT-IF reports. But where the recalculated data lines are large in number or where a minimum change must result before a value is printed a sensitivity analysis can be performed. In the examples which follow the results of the WHAT-IF changes in Appendix 5 have been compared with the originally projected figures shown in Appendix 4. Differences are shown here as percentages.

SENSITIVITY REPORT FOR WHAT-IF EXAMPLE 1 ON SYSTEM A

```
COMMAND? SENSIVITY
COMPARATIVE REPORT FILES(2)? OUTPRINA,WHOPA
DIFFERENCE OR PERCENTAGE? PERCENT
MINIMUM PERCENT PRINT LEVEL? 1.0
COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? RAN
FIRST,LAST LINES*? 5820,5900
SET PAPER,RETURN...
```

ASLIB OUPUT MODEL SYSTEM A
SENSITTVITY--PERCENTAGE
FOR THE PERTOD BEGINNING JAN 1,1976
REPORT PREPARED JUL 27,1976

	1.0 1976	2.0 1977	3.0 1978	4.0 1979	5.0 1980
5820 PROJ OUTPUT COST					
5842 ROYALTIES	**	**	**	**	**
5861 EQUIPMENT	-38.61	-41.82	-44.30	-46.32	-48.03
5900 PROJ OUTPUT COSTS	-1.41	-1.50	-1.52	-1.52	-1.48

NOTE: ** INDICATES DIVISTON BY ZERO

SENSITIVITY REPORT FOR WHAT-IF EXAMPLE 2 ON SYSTEM A

```
COMMAND? SENSITIVITY
COMPARATIVE REPORT FILES(2)? OUTPRINA,WHOPB
DIFFERENCE OR PERCENTAGE? PERCENT
MINIMUM PERCENT PRINT LEVEL? 4.0
COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? SEL
LINES: AFTER LAST 0*
? % 5900,0
SET PAPER,RETURN...
```

ASLIB OUPUT MODEL SYSTEM A
SENSITIVITY--PERCENTAGE
FOR THE PERIOD BEGINNING JAN 1,1976
REPORT PREPARED JUL 27,1976

	1.0 1976	2.0 1977	3.0 1978	4.0 1979	5.0 1980
5900 PROJ OUTPUT COSTS	-4.93	-6.12	-6.96	-7.67	-8.10

SENSITIVITY REPORT FOR WHAT-IF EXAMPLE 3 ON SYSTEM A.

```
COMMAND? SENSITIVITY
COMPARATIVE REPORT FILES(2)? OUTPRINA,WHOPC
DIFFERENCE OR PERCENTAGE? PERCENT
MINIMUM PERCENT PRINT LEVEL? 7.5
COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? MRAN
FIRST,LAST LINES; AFTER LAST 0,0
? % 5710,5776,5900,6790,0,0
SET PAPER.RETURN...
```

ASLIB OUPUT MODEL SYSTEM A
SENSITIVITY--PERCENTAGE
FOR THE PERIOD BEGINNING JAN 1,1976
REPORT PREPARED JUL 27,1976

	1.0 1976	2.0 1977	3.0 1978	4.0 1979	5.0 1980
5710 ABSTRACTS JOURNAL					
5730 SUPERVISTON	-7.77				
5732 ROYALTIES	**	**	**	**	**
5740 MATERIALS	-16.43	-16.46	-16.49	-16.51	-16.53
5760 INPUT	-7.77				
5770 OVERHEADS	-7.77				
5774 ABSTRACTS JOURNAL	-9.86	-9.71	-9.69	-9.76	-9.76

NOTE: ** INDICATES DIVISION BY ZERO

SENSITIVITY REPORT FOR WHAT-IF EXAMPLE 1 ON SYSTEM B

```
COMMAND? SENSITIVITY
COMPARATIVE REPORT FILES(2)? OUTPRINC,WHOPD
DIFFERENCE OR PERCENTAGE? DIFFERENCE
MINIMUM DIFFERENCE PRINT LEVEL? 100
COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? MRAN
FIRST,LAST LINES; AFTER LAST 0,0
? % 5350,5416,5900,6790,0,0
SET PAPER,RETURN...
```

ASI TB OUTPUT MODELSYSTEM B
SENSITIVITY--DIFFERENCE
FOR THE PERIOD BEGINNING JAN 1,1976
REPORT PREPARED JUL 27,1976

	1.0	2.0	3.0	4.0	5.0
	1976	1977	1978	1979	1980
5350 ONLINE SEARCH					
5410 OVERHEADS	-151.5	-105.5			
5414 ON LINE SEARCH	-344	-257	-201	-161	-132

SENSITIVITY REPORT FOR WHAT-IF EXAMPLE 2 ON SYSTEM B

```
COMMAND? SENSITIVITY
COMPARATIVE REPORT FILES(2)? OUTPRINC,WHOPE
DIFFERENCE OR PERCENTAGE? PERCENTAGE
MINIMUM PERCENT PRINT LEVEL? 5.0
COLUMNS? ALL
TOTAL COLUMNS? NO
LINES? MRAN
FIRST,LAST LINES; AFTER LAST 0.0
? % 4000,4120,5900,6790,0,0
SET PAPER,RETURN...
```

ASI IB OUTPUT MODEL SYSTEM B
SENSITIVITY--PERCENTAGE
FOR THE PERIOD BEGINNING JAN 1,1976
REPORT PREPARED JUL 27,1976

	1.0 1976	2.0 1977	3.0 1978	4.0 1979	5.0 1980
4000 STAFF REQUIRED					
4120 GRADE C2 STAFF	**	**	**	**	**
5900 PROJ OUTPUT COSTS					5.27

NOTE: ** INDICATES DIVISION BY ZERO

APPENDIX 7 - SPECIFICATION FOR EXPERIMENT TO TEST THE OUTPUT MODEL DEVELOPED IN EFAG PROJECT 3

A. Objectives

To evaluate the predictive cost model for the output activities of mechanized information systems, as developed in Project 3, Phase 1, Part 1.

B. Source material

Final Report on Project 3, Phase1,Part 2: Development and use of models for the prediction of costs for alternative information systems. Aslib Consultancy Service, July 1976.

C. Details of project

The basic methodology of the test should be to predict the operating costs of a number of existing systems, as from some time in the past, and to check these predictions against operating costs actually recorded. The steps involved would be as follows :

(1) Select a minimum of three mechanized information system which provide one or a range of output services, using a data base(s) created in-house or purchased from an external source. The systems chosen should if possible be representative of the most common types of system, in terms of the mix of services provided.

An essential criterion for selection of candidate systems is that they should have detailed records of their operational activities and costs for at least three years past.

(2) Obtain data on the operating costs of each system for the past three years, as shown in its annual accounts. Data will also be required on the annual volume of throughput of each service, its operational characters, and all other parameters that would normally be determined by the model user.

(3) Run the model for each system to generate a three-year cost prediction. The projections for data values such as salaries and equipment rentals costs should be based on known trends for the countries in which the systems are based.

(4) Compare cost predictions for each service, with costs recorded for each system in its accounts. The percentage error for each figure should be recorded.

(5) Investigate causes of inaccuracy, modify input values, and re-run model as necessary.

It is recommended that computer facilities be used for running the model. If the PROPHIT II facilities used for development of the model were employed, the necessary program (definition file) could be supplied by Aslib.

SALES OFFICES

BELGIQUE – BELGIË

Moniteur belge – Belgisch Staatsblad

Rue de Louvain 40-42 – Leuvenseweg 40-42
1000 Bruxelles – 1000 Brussel – Tél. 512 00 26
CCP 000-2005502-27 – Postrekening 000-2005502-27

Sous-dépôt – Agentschap

Librairie européenne – Europese Boekhandel
Rue de la Loi 244 – Wetstraat 244
1040 Bruxelles – 1040 Brussel

DANMARK

J.H. Schultz – Boghandel

Møntergade 19
1116 København K – Tel. 14 11 95
Girokonto 1195

BR DEUTSCHLAND

Verlag Bundesanzeiger

Breite Straße
Postfach 108 006
5000 Köln 1 – Tel. (0221) 21 03 48
(Fernschreiber: Anzeiger Bonn 08 882 595)
Postscheckkonto 834 00 Köln

FRANCE

Service de vente en France des publications des Communautés européennes – Journal officiel

26, rue Desaix – 75 732 Paris Cedex 15
Tél. (1) 578 61 39 – CCP Paris 23-96

IRELAND

Stationery Office

Beggar's Bush – Dublin 4
Tel. 68 84 33

ITALIA

Libreria dello Stato

Piazza G. Verdi 10
00198 Roma – Tel. (6) 85 08
Telex 62008
CCP 1/2640

Agenzia di Roma

00 187 Roma - Via XX Settembre
(Palazzo Ministero del tesoro)

GRAND-DUCHÉ DE LUXEMBOURG

Office des publications officielles des Communautés européennes

5, rue du Commerce
Boîte postale 1003 – Luxembourg
Tél. 49 00 81 – CCP 191-90
Compte courant bancaire : BIL 8-109/6003/300

NEDERLAND

Staatsdrukkerij- en uitgeverijbedrijf

Christoffel Plantijnstraat, 's Gravenhage
Tel. (070) 81 45 11 – Postgiro 42 53 00

UNITED KINGDOM

H.M. Stationery Office

P.O. Box 569
London SE1 9NH – Tel. 01 - 928 6977, ext. 365
National Giro Account 582-1002

UNITED STATES OF AMERICA

European Community Information Service

2100 M Street N.W. – Suite 707
Washington, D.C. 20 037 – Tel. (202) 872 8350

SCHWEIZ–SUISSE–SVIZZERA

Librairie Payot

6, rue Grenus – 1211 Genève
CCP 12-236 Genève - Tel. 31 89 50

SVERIGE

Librairie C.E. Fritze

2, Fredsgatan – Stockholm 16
Post Giro 193, Bank Giro 73/4015

ESPAÑA

Librería Mundi-Prensa

Castelló 37 – Madrid 1
Tel. 275 46 55

OTHER COUNTRIES

Office for Official publications of the European Communities

5, rue du Commerce
Boîte postale 1003 – Luxembourg
Tél. 49 00 81 – CCP 191-90
Compte courant bancaire : BIL 8-109/6003/300

NOTICE TO THE READER

All scientific and technical reports published by the Commission of the European Communities are announced in the monthly periodical '**euro-abstracts**'. For subscription (1 year : FB 1 200) please write to the address below.

FB 720 DKr 112,50 DM 46,15 FF 96 Lit 16 750 Fl 48,30 £ 11 US $ 19,25

OFFICE FOR OFFICIAL PUBLICATIONS OF THE EUROPEAN COMMUNITIES
Boîte postale 1003 - Luxembourg

7533