International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

The IIASA Health Care Resource Allocation Sub-Model: Mark 1

Gibbs, R.J.

IIASA Research Report
June 1978

THE IIASA HEALTH CARE RESOURCE ALLOCATION SUB-MODEL: MARK 1

R.J. Gibbs

RR-78-8
June 1978

International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

Jeanne Anderer, editor
Martina Segalla, composition
Martin Schobel, graphics

Printed by NOVOGRAPHIC
Maurer-Lange-Gasse 64
1238 Vienna

Preface

The aim of the IIASA Modeling Health Care Systems Task is to build a National Health Care System model and apply it in collaboration with national research centers as an aid to health service planners. The modeling work is proceeding along the lines proposed in earlier papers by Venedictov and Shigan [1] among others. It involves the construction of linked sub-models dealing with population, disease prevalence, resource need, resource supply and resource allocation.

The present paper is concerned with the development of the resource allocation sub-model *DRAM*—disaggregate resource allocation model. It describes the Mark 1 version of the sub-model which simulates the allocation by the Health Care System of a single resource between different types of patients. This version was described briefly in an earlier paper by Gibbs [2] which was written principally for health service planners and other potential *users* of the model. The present paper is written for the scientific or mathematical reader—the model's assumptions are stated formally and the algorithm for solving the model and some methods for estimating the model parameters from empirical data are described in full. It is planned to develop further versions of the sub-model to simulate the allocation of *several* health care resources between patients for whom *alternative modes of treatment* are permitted; this work will be described in future publications. A user's guide to the computer programmes for the Mark 1 version of the sub-model is described in a separate paper [10].

Recent related publications of the IIASA Modeling Health Care Systems Task are listed on the back pages of this Report.

Evgenii N. Shigan
Leader
Health Care Systems Task
May 1978

Summary

Within the context of the IIASA Health Care System model the function of the resource allocation sub-model is to simulate how the HCS allocates limited supplies of resources between competing demands. The principal outputs of the sub-model should be the numbers of patients treated, in different categories, and the modes and standards of treatments they receive. The Mark 1 version of the sub-model is described in this paper. It simulates the allocation of one resource within one mode of treatment but it should be possible to use the approach to develop further versions to cover more general cases. The main assumption of the model is that in allocating its resources the HCS attempts to optimise a utility function whose parameters can be inferred from data on past allocations. Depending upon the type of data that is available different procedures for parameter estimation can be incorporated with the algorithm for solving the model into a computer programme whose main inputs consist solely of empirical data. The programme is fairly small and can readily be installed on most scientific computer installations. The use of the sub-model is illustrated by a hypothetical application using hospital data from England.

Contents

Abstract

Within the context of the IIASA Health Care System model the function of the resource allocation sub-model is to simulate how the HCS allocates limited supplies of resources between competing demands. The principal outputs of the sub-model should be the numbers of patients treated, in different categories, and the modes and standards of treatments they receive. The Mark 1 version of the sub-model is described in this paper. It simulates the allocation of one resource within one mode of treatment but it should be possible to use the approach to develop further versions to cover more general cases. The main assumption of the model is that in allocating its resources the HCS attempts to optimise a utility function whose parameters can be inferred from data on past allocations. Depending upon the type of data that is available different procedures for parameter estimation are required. The procedures for parameter estimation can be incorporated with the algorithm for solving the model into a computer programme whose main inputs consist solely of empirical data. The programme is fairly small and can readily be installed on most scientific computer installations. The use of the sub-model is illustrated by a hypothetical application using hospital data from England.

The IIASA Health Care Resource Allocation Sub-Model: Mark 1

1. THE FUNCTION OF THE RESOURCE ALLOCATION SUB-MODEL

The aim of the IIASA Modeling Health Care Systems Task is to build a National Health Care System Model and apply it in collaboration with national research centers as an aid to health service planners. As described in earlier papers by Venedictov and Shigan [1] and by Gibbs [2] the research plan includes the construction of linked sub-models dealing with population, disease prevalence, resource need, resource supply, and resource allocation. This paper is concerned with the resource allocation sub-model which has been named *DRAM*--disaggregated resource allocation model.

This chapter is concerned with definition of the attributes that are required of DRAM for it to fulfill its role in the overall National Health Care System Model. In Chapter 2 a model formulation is presented which meets some, though not all, of these attributes; the model thus defined is referred to as DRAM Mark 1. The formulation is given in terms of the allocation of hospital beds but this is only an *example* of how the model may be applied; the model is equally applicable to the allocation of other health service resources. An algorithm for running DRAM Mark 1 is described in Chapter 3. There are a number of parameters in the model whose values may be estimated from empirical data. However it is likely that data availability will vary from one country to another. Accordingly, in Chapter 4, three of the most likely cases of data availability are considered and parameter estimation procedures are described for each case. Illustrative model runs for two of the three cases are presented in Chapter 5 using hospital data from England. Finally Chapter 6 suggests how further versions of DRAM might be developed in the future so as to meet all, rather than some, of the required attributes defined below.

The role of the resource allocation sub-model in relation to the other sub-models is shown in Figure 1 and described more fully in Gibbs [2]. Within this schema the function of the resource allocation sub-model is to simulate how the Health Care System (HCS) allocates limited *supplies* of resources between competing *demands*. Accordingly it requires input data on demand and supply.

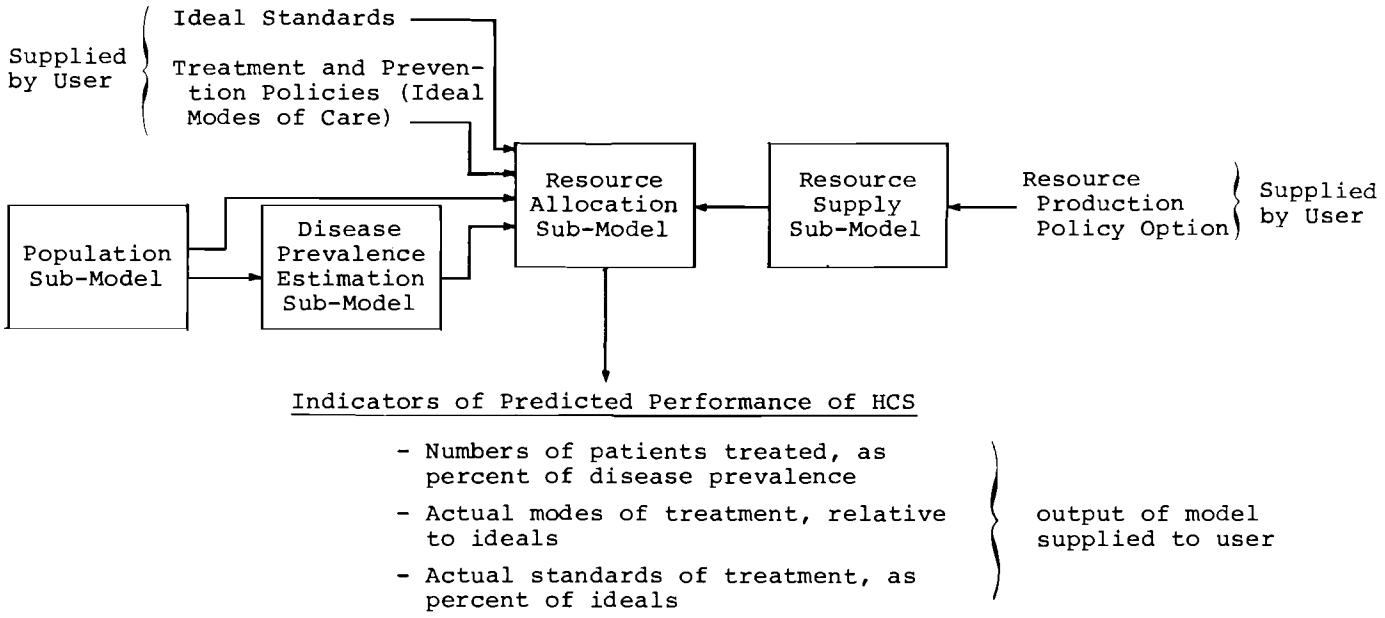

Figure 1. The role of the resource allocation model in the context of the IIASA HCS model.

The demand inputs, which represent ideal or asymptotic demands (to be defined more fully in Chapter 2), are as follows:

- the total number of individuals who could be offered treatment, by category (from the morbidity and population sub-models);

- the policies for treatment (i.e. the feasible modes of treatment for each patient category--in-patient, out-patient, domiciliary, etc.); and

- ideal standards of treatment for each patient category and mode of treatment (e.g. the length of stay in hospital for a given episode of illness).

The supply inputs consist of information on the amounts of resources available for use in the HCS. These inputs can be provided via a resource supply (production) sub-model or, failing this, they can be provided direct by the user in the form of a trial policy option for the provision of HCS resources. An illustration of the latter form of input is given later, in Chapter 5.

It is assumed in the sub-model, as Rousseau [3] among others has observed, that there is never a sufficient supply of resources to saturate all the asymptotic demands for them. Accordingly the sub-model represents the HCS as attempting to achieve an equilibrium between supply and demand by adjustments along three dimensions:

- the numbers of patients of different types who are offered treatment,

- the modes of treatment offered, and

- the standards at which treatment is offered.

Because of the limited supply of resources the performance of the HCS in these three respects falls short of ideal levels:

- a proportion of the morbidity in the population is not treated,

- some patients are not treated in the most desirable mode,

- patients are treated at less than ideal standards.

The degree of short-fall from the ideal levels varies between types of patient and sectors of the HCS according to a set of priorities and preferences which operate in the HCS.

The type of model that is suitable in this context is one which simulates the way in which the HCS allocates resources by means of a behavioural hypothesis which takes account of these

preferences and priorities. It is argued in an earlier paper [4], which reviewed the literature on HCS resource allocation models, that this behaviour simulation type of model is more appropriate than either the classical econometric or optimisation type. The formulation of a behaviour simulation model, DRAM, is presented in the following chapter. The central behavioural hypothesis is that the HCS allocates its resources so as to maximise a utility function whose parameters can be inferred from observations of past allocations. Like the models of McDonald et al. in the UK [5] and Rousseau in Canada [3], its hypothesis implies that the actors in the HCS are striving to attain some ideal pattern of behaviour within resource constraints.

2. MODEL FORMULATION

The model proposed here, DRAM, is a simplification of the model of McDonald et al. [5]. Of the three main dimensions of the HCS resource allocation process--patient numbers, treatment modes and standards--which were described above and which are included in the McDonald model, the initial, Mark 1, version of DRAM includes only two--patient selection and standard attainment. Thus DRAM Mark 1 can be applied to only one mode of treatment at a time. However one of the advantages of DRAM Mark 1 is that its computing requirements are relatively light (for reasons explained in the next chapter) so that it can be readily implemented on different computers without using elaborate software and so could be relatively easily applied in different countries; (by contrast the McDonald model, in its current form, requires relatively sophisticated software and a large computer in order to solve the non-linear programming formulation). Being more simple this model is also more transparent. Keyfitz [6], among others, has argued persuasively that with a transparent model the user can gain an insight into the workings of the model and is then more likely to have confidence in its results than with a "black box" model. It is planned, as explained in Chapter 6, to develop further versions of DRAM which will both retain some of the computational simplicity of DRAM Mark 1 and include the third dimension of resource allocation--mode selection--which is missing from DRAM Mark 1.

To simplify the exposition, DRAM will be presented here in terms of application to the acute hospital in-patient sector, but this application should be regarded merely as an example. The essence of the model is the concept of the HCS achieving an equilibrium by balancing the desirability of treating more patients of one type against treating more of other types and against the desirability of treating each type patient at a higher average standard. Since this concept is equally valid for many other HCS sectors (e.g. out-patient treatment) and for several resources within each sector (e.g. physicians, beds, nurses), DRAM is offered in the belief that it is widely applicable.

The way in which the HCS achieves such an equilibrium has been extensively researched. One finding, which has been so frequently obtained (e.g. [7,8,9]) that the accumulated evidence for it is by now overwhelming, is that for a wide range of clinical conditions and specialties, both the number of admissions and the average length of in-patient stay are elastic to the supply of beds; that is to say the greater the supply of beds the greater are both the numbers admitted and their length of stay. Furthermore it appears that in *none* of the places studied has the supply of beds reached the level at which in-patient care is given to all individuals who seek it, at the ideal average length of stay.

The model, DRAM, represents how the HCS achieves an equilibrium between numbers of patients and lengths of stay on the one hand and bed supply on the other by means of a hypothesis that the HCS attempts to optimise a utility function. Thus, *if this underlying hypothesis is sound*, DRAM can not merely describe past equilibria, as can classical econometric models, but it can also, unlike classical econometric models, predict how the equilibrium is likely to change in the future as a result of changes in factors such as clinical standards, disease prevalence, and the preferences and priorities operating in the HCS.

The formulation of DRAM, given below, is similar to that of the model of McDonald et al. [5] but the methods for solution and parameter estimation, given in the following chapters, are different.

Definitions

Subscript

i = Patient category (e.g. disease type), $i = 1,2,3,\ldots,N$.

Variables

x_i = Hospital admission rate (cases per million population).

u_i = Average length of stay (days).

Parameters

X_i = Ideal, maximum admission rate for patients needing hospital treatment (to be defined more fully below).

U_i = Ideal average length of stay (to be defined more fully below).

c = Unit cost of a hospital bed-day.

Data

B = Number of hospital bed-days per million population available for occupation, assuming a constant occupancy rate. (The model is applicable for the range $0 < B < \sum_i X_i U_i$.)

Hypothesis

The HCS chooses the x_i, u_i so as to maximise a utility function, Z, where

$$Z = \sum_i g_i(x_i) + \sum_i x_i h_i(u_i) \qquad (1)$$

subject to

$$0 < x_i < X_i \quad , \quad \forall\, i \quad ,$$

$$0 < u_i < U_i \quad , \quad \forall\, i \quad ,$$

and

$$\sum_i x_i u_i = B \quad . \qquad (2)$$

We will now define the components $g_i(x_i)$ and $h_i(u_i)$ of the utility function. To do this we make the following assumptions:

i. $g_i(x_i)$ and $h_i(u_i)$ are monotonically increasing with decreasing gradients.

ii. At the ideal admission rate, X_i, and the ideal length of stay, U_i, the marginal utility of increasing admission rate or length of stay equals the corresponding marginal resource cost. This can be regarded as completing the definitions of the X_i and U_i. An important implication of assumptions (i) and (ii) is that the HCS would not seek allocations for which either $x_i > X_i$ or $u_i > U_i$ since in such cases marginal utility is less than marginal cost.

iii. At the ideal length of stay, U_i, the contribution to utility of treating extra patients of category i is represented by the function $g_i(x_i)$ alone, i.e. $h_i(U_i) = 0$ and $h_i(u_i) < 0$ for $u_i < U_i$. Thus if we consider the marginal utility of treating an additional patient at *less* than the ideal length of stay we see from the additive nature of the utility function defined by equation (1) that there will be a *negative contribution*, $h_i(u_i)$, to be set against a *positive contribution*, $g_i'(x_i)$; the values of the x_i and u_i which maximise total utility under the constraint (2) correspond to a point where these two contributions exactly balance each other--this is one aspect of the way in which we expressed the HCS achieving an equilibrium between competing demands.

iv. The elasticities of admission rate, x_i, and length of stay, u_i, with respect to marginal utility are constant; let us denote the elasticities of the x_i by E_i and of the u_i by F_i. The optimisation model will lead to a solution in which the marginal utility of treating additional patients or of increasing length of stay equals the corresponding opportunity cost. Thus we can also regard the E_i and the F_i as the elasticities of the x_i and the u_i with respect to opportunity cost.

Let us now define the function $g_i(x_i)$. From assumption (iv) we have

$$\frac{d(\log x_i)}{d(\log g_i'(x_i))} = -E_i \quad ;$$

therefore

$$g_i'(x_i) = A_i x_i^{-1/E_i} \quad .$$

To ascribe an appropriate value to A_i, the constant of integration, we invoke assumption (ii). Thus, since the marginal resource cost of treating an extra patient is cU_i, we have

$$g_i'(X_i) = cU_i \quad ,$$

$$A_i = cU_i X_i^{1/E_i} \quad ,$$

$$g_i'(x_i) = cU_i\left(\frac{x_i}{X_i}\right)^{-1/E_i} \quad .$$

Hence, except for a constant of integration which is not relevant to what follows, we have

$$g_i(x_i) = \frac{cU_i X_i}{1 - 1/E_i}\left(\frac{x_i}{X_i}\right)^{1-1/E_i} \quad . \tag{3}$$

We can define the function $h_i(u_i)$ in a similar way. From assumption (iv) we have

$$h_i'(u_i) = B_i u_i^{-1/F_i} \quad .$$

To determine the constant of integration, B_i, we again invoke assumption (ii). Thus, since the direct marginal resource cost per extra day of stay is equal to c, we have

$$h_i'(U_i) = c \quad ,$$

$$B_i = cU_i^{1/F_i} \quad ,$$

$$h_i'(u_i) = \left(\frac{u_i}{U_i}\right)^{-1/F_i} \quad .$$

From (iii) we have $h_i(U_i) = 0$. Thus

$$h_i(u_i) = - \frac{cU_i}{1 - 1/F_i} \left[1 - \left(\frac{u_i}{U_i} \right)^{1-1/F_i} \right] \quad . \qquad (4)$$

The utility function is now fully defined. However we can simplify the expression a little. Firstly note that c appears as a multiplicative constant in each of the $g_i(x_i)$ and $h_i(u_i)$ functions; by choosing to measure utility in the units of the cost of a hospital bed-day we may divide through by c. Secondly it is convenient to replace the elasticity terms by constants α_i and β_i where

$$\alpha_i = - \left(1 - \frac{1}{E_i} \right) \quad ,$$

and

$$\beta_i = - \left(1 - \frac{1}{F_i} \right) \quad .$$

Since the E_i and F_i are expected to lie in the range between zero and unity the α_i and β_i will have positive values. The model formulation is now complete and can be written as follows:

Choose the x_i and u_i to maximise the utility function, Z, where

$$Z = \sum_i g_i(x_i) + \sum_i x_i h_i(u_i) \quad , \qquad (1)$$

and

$$g_i(x_i) = - \frac{X_i U_i}{\alpha_i} \left(\frac{x_i}{X_i} \right)^{-\alpha_i} \quad , \quad \forall\ i \quad , \qquad (5)$$

and

$$h_i(u_i) = \frac{U_i}{\beta_i}\left[1 - \left(\frac{u_i}{U_i}\right)^{-\beta_i}\right] , \quad \forall\, i \ , \tag{6}$$

subject to

$$0 < x_i < X_i \ , \quad \forall\, i \ ,$$

$$0 < u_i < U_i \ , \quad \forall\, i \ ,$$

and

$$\sum_i x_i u_i = B \ . \tag{2}$$

We have now formulated a model whose function is to *simulate* how the HCS allocates a given quantity, B, of hospital bed-days. The user of the model has to supply an input value for B and the model outputs, the x_i and u_i, then constitute a *prediction* of the consequent HCS resource allocation behaviour, conditional upon certain assumptions about the nature of this behaviour. By running the model a number of times with different values for B, the user can examine the consequences of a number of planning options for hospital bed supply. Initially at least, we regard the model parameters--both the ideal allocations, the X_i and U_i, and the priority power parameters, the α_i and β_i--as being properties of the behaviour of the actors in the HCS and *not* under the direct control of the HCS planner. Thus although we have adopted a hypothesis that the behaviour of the actors in the HCS is of an *optimising* nature, we do *not* assume that their objective function necessarily corresponds to any objective function that the HCS planner might have. From the point of view of the HCS planner and the model user, the model is therefore of a *simulation* type even though, as we shall see in the next chapter, *optimisation* techniques are needed to compute the simulation.

3. SOLUTION OF THE MODEL

In this chapter we describe a method for computing the simulated HCS behaviour, in other words a method for determining the values of the x_i and u_i that maximise the utility function defined by equation (1) for a given value of hospital bed supply, B.

Computationally speaking this is clearly an optimisation problem. The method presented below exploits some analytical features of the model DRAM and employs the Lagrange Multiplier technique. The computations can be performed swiftly by a fairly simple Fortran programme which can easily be transferred from one computer installation to another. By contrast the McDonald model [5] employs a sophisticated non-linear mathematical programming algorithm which makes relatively heavy computational demands that few computer installations can satisfy. On the other hand the Mark 1 version of DRAM that is presented here provides a less complete representation of the HCS resource allocation process than the McDonald model. However it is hoped in the future to develop further versions of DRAM that will provide as complete a representation of the resource allocation process as the McDonald model and yet retain the computational advantages of the method described below for DRAM Mark 1.

In the normal way the constrained maximisation can be rewritten as an unconstrained maximisation using the Lagrange Multiplier, λ:

$$\text{Maximise } L = \sum_i g_i(x_i) + \sum_i x_i h_i(u_i) + \lambda\left(B - \sum_i x_i u_i\right) \quad . \tag{7}$$

The optimality conditions are

$$\frac{\partial L}{\partial x_i} = 0 \quad , \qquad \forall\ i \quad , \tag{8}$$

and

$$\frac{\partial L}{\partial u_i} = 0 \quad , \qquad \forall\ i \quad . \tag{9}$$

From (9)

$$x_i h_i'(u_i) - \lambda x_i = 0 \quad ,$$

$$h_i'(u_i) = \lambda \quad , \qquad \text{since } x_i > 0 \quad ,$$

$$u_i = U_i \lambda^{-1/(\beta_i+1)} \quad . \tag{10}$$

From (8)

$$g_i'(x_i) = \lambda u_i - h_i(u_i) \quad ,$$

$$U_i\left(\frac{x_i}{X_i}\right)^{-(\alpha_i+1)} = \lambda U_i \lambda^{-1/(\beta_i+1)} - \frac{U_i}{\beta_i} + \frac{U_i}{\beta_i}\lambda^{\beta_i/(\beta_i+1)} \quad ,$$

$$\left(\frac{x_i}{X_i}\right)^{-(\alpha_i+1)} = \frac{1}{\beta_i}\left[(\beta_i + 1)\lambda^{\beta_i/(\beta_i+1)} - 1\right] \quad ,$$

$$x_i = X_i\left\{\frac{1}{\beta_i}\left[(\beta_i + 1)\lambda^{\beta_i/(\beta_i+1)} - 1\right]\right\}^{-1/(\alpha_i+1)} \quad . \qquad (11)$$

λ is obtained from substituting (10) and (11) in (2), which gives $f(\lambda) = 0$ where

$$f(\lambda) = -B + \sum_i X_i U_i \beta_i^{1/(\alpha_i+1)} \phi_i^{-1/(\alpha_i+1)} \quad , \qquad (12)$$

and

$$\phi_i = (\beta_i + 1)\lambda^{(\alpha_i+\beta_i+1)/(\beta_i+1)} - \lambda^{(\alpha_i+1)/(\beta_i+1)} \quad . \qquad (13)$$

We now have in (10) and (11) analytic expressions for the variables x_i and u_i in which the only unknown is λ. To find λ we merely need to solve $f(\lambda) = 0$. This cannot in general be done analytically, but it is readily amenable to numerical solution by the Newton-Raphson procedure. To prove this we need to obtain an analytic expression for $f'(\lambda)$ and to demonstrate that $f(\lambda)$ and $f'(\lambda)$ have suitable properties in the range of λ which is of interest.

First we recall that we are searching for solutions in the ranges given by

$$0 < x_i < X_i \quad , \qquad 0 < u_i < U_i \quad , \qquad \forall\ i \quad .$$

From inspecting (10) and (11) we see that these ranges imply $\lambda > 1$.

From (12) and (13) we may obtain the following analytic expression for $f'(\lambda)$:

$$f'(\lambda) = - \sum_i \frac{X_i U_i}{(\alpha_i + 1)} \beta_i^{1/(\alpha_i+1)} \phi_i^{-(\alpha_i+2)/(\alpha_i+1)} \theta_i \quad , \quad (14)$$

where

$$\theta_i = \frac{d\phi_i}{d\lambda}$$

$$= (\alpha_i + \beta_i + 1)\lambda^{\alpha_i/(\beta_i+1)} - \frac{\alpha_i + 1}{\beta_i + 1} \lambda^{(\alpha_i-\beta_i)/(\beta_i+1)}$$

$$= \lambda^{\alpha_i/(\beta_i+1)} \left\{ \alpha_i + \beta_i + 1 - \frac{\alpha_i + 1}{\beta_i + 1} \lambda^{-\beta_i/(\beta_i+1)} \right\} \quad . \quad (15)$$

We can now establish that $f(\lambda)$ and $f'(\lambda)$ have the properties required for using the Newton-Raphson process. First we note that $f(\lambda)$ and $f'(\lambda)$ are both continuous in the range $\lambda \geq 1$. Second it can be observed, from (14), that $f'(\lambda)$ is negative throughout this range since for $\lambda \geq 1$ both $\phi_i(\lambda)$ and $\theta_i(\lambda)$ are positive for all i (see (13) and (15)). Third, recalling the fact that the model is applicable for $B < \sum_i X_i U_i$, we see from (12) that $f(1) > 0$. From these three facts it follows that there is only one root, a real one, to $f(\lambda) = 0$ in the range $\lambda \geq 1$, and that this root can be found using the Newton-Raphson process.

Accordingly a small computer programme has been written to solve equation (12) by the Newton-Raphson method; it is described in a separate paper [10]. Computational experience [10] has shown that a good solution can be obtained in a small number of iterations over a wide range of parameters and starting values.

4. PARAMETER ESTIMATION

In order to run the model we require values for the following parameters:

- the X_i and U_i, the ideal admission rates and lengths of stay;
- the α_i and β_i, the power factors of the functions $g_i(x_i)$ and $h_i(u_i)$.

Three possible situations are considered in this paper:

Case 1: Exogenous estimates available for *all* parameters,

Case 2: Exogenous estimates available for *none* of the parameters,

Case 3: Exogenous estimates available for the X_i and U_i but not for the α_i and β_i.

Although we need to consider Case 1 because of its theoretical importance we shall argue that it is unlikely to be relevant in practice. Case 2 describes the practical situation we expect to encounter in those countries where the HCS does *not* have a strong degree of central planning. Case 3 is relevant for those countries which *do* have a strong degree of central planning of the HCS; here we may find that *planning norms* exist which can serve as appropriate values for the X_i and U_i.

This chapter is mainly concerned with describing methods by which parameter values may be estimated for Cases 2 and 3 from certain empirical data on past resource allocations in the HCS. The computations for these methods can be carried out by simple Fortran programmes. For the convenience of the model user the *programmes for parameter estimation* have been incorporated with the *programme for solving the model*, for given parameter values, which was described in the previous chapter. Thus the corporate programmes take the empirical data on past allocations as part of their input set and provide the model solution, the simulation results, as their output; the computed parameter values are, in effect, intermediate quantities within the corporate programme. Some illustrative runs of these programmes are given in the following chapter using empirical data from England. A complete user's guide to the programmes is given in a separate publication [10]. We will now consider the process of parameter estimation for each of the three cases.

Case 1

Exogenous estimates of the X_i may be obtainable from a combination of morbidity estimation and expert opinion on hospitalisation rates, and estimates of the U_i from clinical opinion. Indeed in some countries where there is a strong degree of central planning in the HCS estimates of this type are used within a formal planning process, as described in Gibbs [2].

It is more difficult to see how exogenous estimates of the α_i and β_i might be obtained. It might be possible to estimate the utility functions $g_i(x_i)$ and $h_i(u_i)$ directly by subjective judgements using methods such as those employed by Keeney and Raiffa [11]. However, even if it were possible to obtain estimates of all these parameters by means of subjective judgements, the validity of using them in practice is open to question. The problem is that these subjective judgements define *the utility function of the individual giving the judgements* and there is no reason to suppose that this will correspond to the *utility function of the HCS*. Thus the results of running the model with parameter values of this type will describe a theoretical allocation of bed-days that is optimal from the individual's point of view but this will not, in general, correspond to the resource allocation behaviour of the HCS in practice. Such results would only be relevant for HCS planning if the user of the model had good reason to believe that either (a) the subjective judgements corresponded to the prevailing preferences and priorities in the HCS or (b) that the preferences and priorities implied by the subjective judgements could be implemented in the HCS in place of the prevailing ones.

Accordingly we will examine Case 1 no further and turn our attention to Cases 2 and 3 which describe situations which are likely to be more relevant in practice.

Case 2

Here we assume that in its past resource allocations the HCS has optimised a utility function of the form defined in the previous chapter. We describe a method, based on this assumption, by which values of the parameters of the function--the α_i, β_i, X_i and U_i--can be *inferred* from empirical data on past allocations. With such parameter values we can then use the model to generate predictions of how the HCS would allocate resources in the future for different levels of aggregate resource availability; such predictions are conditional upon the prevailing preferences and priorities in the HCS remaining unchanged. Illustrative examples of using the model in this way are given in the following chapter.

In this section we start by defining a set of empirical resource allocation data. We then derive equations which relate the model parameters to this data; these equations define a set of parameter values which is consistent with the empirical data. Finally we describe an algorithm for solving the equations and producing the required parameter values.

Data for Case 2

Let us consider a geographical region with constituent sub-regions and let us suppose that we can observe, for a given time period, the admission rates, x_i, and the average lengths of stay, u_i, for each sub-region. From these observations we can compute the corresponding quantities for the region as a whole:

$\bar{x}_i$ = regional admission rate for category i;

$\bar{u}_i$ = regional length of stay for category i;

$\bar{B}$ = regional aggregate bed supply;

and these quantities have the natural property

$$\bar{B} = \sum_i \bar{x}_i \bar{u}_i \quad . \tag{16}$$

Let us now define the following elasticities:

γ_i = elasticity of admission rate, x_i, for category i with respect to aggregate bed supply;

η_i = elasticity of average length of stay, u_i, for category i with respect to aggregate bed supply.

Thus, in terms of the model,

$$\gamma_i = \frac{d(\log x_i)}{d(\log B)} \tag{17}$$

and

$$\eta_i = \frac{d(\log u_i)}{d(\log B)} \quad . \tag{18}$$

Now estimates, $\hat{\gamma}_i$ and $\hat{\eta}_i$, of these elasticities may be obtained from cross-section analysis of the sub-region data. For example Feldstein [9] obtained such estimates from English hospital data in 1960 using the following types of regression equation (having previously experimented with other specifications):

$$\log x_i = \hat{\gamma}_i \log B + \text{constant}$$

and

$$\log u_i = \hat{\eta}_i \log B + \text{constant} \quad .$$

Some illustrative elasticity estimates, based on some of Feldstein's results, are shown in Table 1.

Table 1. Elasticities of hospital admissions and lengths of stay with respect to total bed supply* for England, 1960, for certain diseases (adapted from Feldstein [9]).

Disease	Elasticity of:	
	Admissions* per Year	Average Stay
Varicose Veins	0.78	0.62
Haemorrhoids	0.70	0.44
Ischaemic Heart**	1.14	1.08
Pneumonia	0.71	0.23
Bronchitis	1.13	0.05***
Appendicitis	0.05***	0.31

*Per thousand population.
**Excluding acute myocardial infarction.
***The values obtained by Feldstein for these elasticities were negative but not significantly different from zero. Since negative elasticities are somewhat inplausible in this context these results have been interpreted as indicating that the true values for these elasticities are close to zero but positive. For the purposes of the calculations in this paper they have been assigned the value of 0.05.

Case 2: Equations for Parameter Estimation

We now derive equations which relate the values of the model parameters to the empirical data defined above. First we derive equations relating the α_i and β_i to the empirical estimates $\hat{\gamma}_i$ and $\hat{\eta}_i$. From (12) we have

$$f(\lambda) = -B + \sum_i X_i U_i \beta_i^{1/(\alpha_i+1)} \phi_i^{-1/(\alpha_i+1)} \quad . \tag{12}$$

Therefore at the optimum, where $f(\lambda) = 0$, we have

$$B = F(\lambda) \quad ,$$

where

$$F(\lambda) = \sum_i X_i U_i \beta_i^{1/(\alpha_i+1)} \phi_i^{-1/(\alpha_i+1)} \tag{19}$$

and the expression for ϕ_i in terms of λ is given at (13). Now,

$$\frac{dB}{d\lambda} = F'(\lambda) \quad .$$

But we see from (12) and (19) that $F'(\lambda) = f'(\lambda)$. Therefore

$$\frac{dB}{d\lambda} = f'(\lambda) \quad ,$$

$$\frac{d(\log B)}{d\lambda} = \frac{1}{B} \cdot f'(\lambda) \quad ,$$

$$\frac{d\lambda}{d(\log B)} = \frac{B}{f'(\lambda)} \quad . \tag{20}$$

We can now express the γ_i and η_i in terms of the α_i and β_i. From (18)

$$\eta_i = \frac{d(\log u_i)}{d\lambda} \cdot \frac{d\lambda}{d(\log B)} \quad .$$

From (10) and (20)

$$\eta_i = - \frac{1}{(\beta_i + 1)} \cdot \frac{B}{\lambda f'(\lambda)} \quad , \tag{21}$$

and from (17)

$$\gamma_i = \frac{d(\log x_i)}{d\lambda} \cdot \frac{d\lambda}{d(\log B)} \quad .$$

Therefore from (11)

$$\gamma_i = - \frac{B_i \lambda^{-1/(\beta_i+1)}}{(\alpha_i + 1)} \cdot \frac{1}{\left[(\beta_i + 1)\lambda^{\beta_i/(\beta_i+1)} - 1\right]} \cdot \frac{B}{f'(\lambda)}$$

$$= - \frac{B_i}{\alpha_i + 1} \cdot \frac{1}{\left(\beta_i + 1 - \lambda^{-\beta_i/(\beta_i+1)}\right)} \cdot \frac{B}{\lambda f'(\lambda)} \quad . \tag{22}$$

If we have the empirical estimates $\hat{\gamma}_i$ and $\hat{\eta}_i$ of regional elasticities at the current regional average aggregate bed supply, $\bar{B}$, we can use equations (21) and (22) to derive expressions for α_i and β_i. Let

$$C = \frac{-\bar{B}}{\lambda f'(\lambda)} \quad . \tag{23}$$

Then from (21)

$$\beta_i = \frac{C}{\hat{\eta}_i} - 1 \quad , \qquad \forall\ i \quad , \tag{24}$$

and from (22)

$$\alpha_i = \frac{C\beta_i}{\hat{\gamma}_i\left(\beta_i + 1 - \lambda^{-\beta_i/(\beta_i+1)}\right)} - 1 \quad , \qquad \forall\ i \quad . \tag{25}$$

Turning now to the values of parameters X_i and U_i we may derive equations relating them to the observed regional quantities, the $\bar{x}_i$ and $\bar{u}_i$, using the inverses of equations (10) and (11):

$$U_i = \bar{u}_i \lambda^{1/(\beta_i+1)} \quad , \qquad \forall\ i \tag{26}$$

$$X_i = \bar{x}_i \left\{ \frac{1}{\beta_i} \left[(\beta_i + 1)\lambda^{\beta_i/(\beta_i+1)} - 1 \right] \right\}^{1/(\alpha_i+1)} \quad , \qquad \forall\ i \quad . \tag{27}$$

We now have the required equations--(23) through (27)--relating the model parameters which are to be estimated--the α_i, β_i, X_i and U_i--to empirical data--$\bar{B}$ and the $\hat{\gamma}_i$, $\hat{\eta}_i$, $\bar{x}_i$ and $\bar{u}_i$.

Case 2: Algorithm for Parameter Estimation

We now describe a procedure for estimating the model parameters using the equations described above. The procedure is based on the requirement that the parameter values should be such that the consequent behaviour of the model is consistent with the empirical data. In other words it is required that if bed supply is set to the level $\bar{B}$ then the model solution should be given by

$$x_i = \bar{x}_i \quad \text{and} \quad u_i = \bar{u}_i \quad , \qquad \forall\ i \quad , \tag{28}$$

and that the response of the model solution at this point to perturbations in bed supply should be consistent with the empirical elasticity data, i.e. that if the bed supply is perturbed by a small amount δB where $\delta B = o(\bar{B})$ then the perturbations, δx_i and δu_i, in the model solution should satisfy the following:

$$\left.\begin{aligned} \frac{\delta x_i}{\bar{x}_i} &= \gamma_i \frac{\delta B}{\bar{B}} (1 + o(1)) \quad , \quad \forall\, i \quad , \\ \frac{\delta u_i}{\bar{u}_i} &= \eta_i \frac{\delta B}{\bar{B}} (1 + o(1)) \quad , \quad \forall\, i \quad . \end{aligned}\right\} \tag{29}$$

We note that the above mentioned requirement--that the model output should be consistent with the input data--can only be satisfied if the data represent a feasible model solution; this leads to two conditions which the data must satisfy:

$$\sum_i \bar{x}_i \bar{u}_i = \bar{B} \quad , \tag{16}$$

$$\sum_i \bar{x}_i \bar{u}_i (\hat{\gamma}_i + \hat{\eta}_i) = \bar{B} \quad . \tag{30}$$

The derivation of these conditions is given in Appendix 1, although (16) has already been noted in connection with the source of the data on $\bar{x}_i$, $\bar{u}_i$ and $\bar{B}$. (It is possible that because of measuring errors and other reasons the input data will not precisely satisfy (16) and (30); in this case the parameter estimation procedure will be spuriously prevented from converging. Accordingly the computer programme incorporates checks that the data satisfy (16) and (30). In the case that (30) is not satisfied within the necessary margin, a procedure is available in the programme for scaling the input elasticity values by the multiplicative factor $\bar{B}\left[\sum_i x_i u_i (\hat{\gamma}_i + \hat{\eta}_i)\right]^{-1}$ so as to remedy the situation.)

Let us now consider how to solve equations for estimating the model parameters from this data. We have 4N + 1 equations --(23) to (27)--for 4N + 3 unknowns--C, λ, $f'(\lambda)$ and the X_i, U_i, α_i and β_i (where N = the number of patient categories). Thus two additional equations are needed to generate a unique solution. Let us consider two equations which, at first sight, appear to be suitable. The first is obtained from the optimality requirement that $f(\lambda) = 0$ where $f(\lambda)$ is given by (12) and (13); this gives us

$$\sum_i X_i U_i \beta_i^{1/(\alpha_i+1)}$$

$$\cdot \left[(\beta_i + 1)\lambda^{(\alpha_i+\beta_i+1)/(\beta_i+1)} - \lambda^{(\alpha_i+1)/(\beta_i+1)} \right]^{-1/(\alpha_i+1)} = \bar{B} \quad . \tag{31}$$

The second is an expression for $f'(\lambda)$ in terms of λ and the X_i, U_i, α_i and β_i, which can be derived from substituting (13) and (15) in (14).

Unfortunately (31) and (14) add no definition to the equation system. Using equations (26) and (27) it is possible to reduce equation (31) to the data identity given by equation (16). Furthermore it can be shown that equation (14), in conjunction with equations (23) to (27), can be reduced to the data identity given by equation (30). (The proofs of these two results are given in Appendix 1.)

Thus although we have 4N + 3 equations--(23) to (27), (31) and (14)--for 4N + 3 unknowns--C, λ, $f'(\lambda)$ and the X_i, U_i, α_i and β_i--it turns out that only 4N + 1 of them are independent. Thus there are two degrees of freedom in the equation system and an infinite number of solutions. Accordingly the following computational procedure was adopted:

i. Set C and λ at arbitrary initial values.

ii. Using the input data on the $\bar{x}_i$, $\bar{u}_i$, $\hat{\gamma}_i$ and $\hat{n}_i$ and equations (24) to (27), estimate values for the parameters X_i, U_i, α_i and β_i.

iii. With these parameter values the model can now be used to simulate the allocation of any given bed supply B', using equations (10), (11) and solving $f(\lambda) = 0$ for the value $B = B'$ by the Newton-Raphson Method, as for Case 1.

Naturally the values of the parameter estimates obtained by this procedure (in stage (ii)), depend strongly on the arbitrary initial values selected for λ and C in stage (i). However, and this is a most important result, the final outputs of the model--the values of the x_i and u_i obtained in the simulation process in stage (iii)--are *not sensitive* to the initial values of λ and C. For reasons given in Appendix 2 certain bounds on the values of λ and C can be defined *a priori*. If the initial

values of λ and C are restricted to vary within these bounds, computational experience (described in Appendix 2) suggests that the mean absolute variation in the output values (the x_i and u_i) will usually be about 0.1%. Since this is a high level of precision in the field of health services research the computational procedure described above seems to be adequate for practical purposes.

Case 3

We now consider the situation where estimates of the X_i and U_i are supplied exogenously. As mentioned above, for Case 1, they might be obtained during the formal planning process in those countries where there is a strong degree of central planning in the HCS. Given such estimates of the X_i and U_i we now need to estimate only the α_i and β_i; we describe below a method by which this may be done using empirical data on elasticities of the same type as described for Case 2 above. However in this case we encounter some difficulty because we are using two completely different sources of data. The estimates of the X_i and U_i have a *prescriptive* quality since they involve a degree of subjective judgement about what the HCS ought to be doing, either at present or in the *future*; by contrast the elasticity data are *descriptive* of what the HCS *has* done in the past. If we then run the model with such parameter estimates to predict *future* resource allocations we need to assume that the behaviour of the HCS will be consistent with *both* the prescribed ideal allocations *and* the prevailing elasticities. From the behavioural point of view this assumption is tenable if the prescriptive estimates of the X_i and U_i are based on a realistic understanding of the behaviour of the actors in the HCS. It seems reasonable to assume that there are several countries where the HCS has a strong degree of central planning and where the prescriptive quantities, or planning norms, are indeed derived from a careful analysis of HCS behaviour; for example Popov [12] has described how such an analysis is performed within the central planning of the HCS in the USSR. In the belief, therefore, that it is likely to be relevant in several countries we now describe a procedure for estimating the parameters α_i and β_i.

The following procedure was considered initially:

i. Set C and λ to arbitrary initial values.

ii. Using data on $\bar{B}$ and the $\hat{\gamma}_i$ and $\hat{\eta}_i$ and equations (24) and (25) estimate values for the parameters α_i and β_i.

iii. Using the estimates of α_i and β_i from (ii) and the exogenous estimates of the parameters X_i and U_i solve equations (31), (14) and (23) for λ, $f'(\lambda)$ and C.

iv. Repeat (ii) and (iii) until convergence--i.e. until at *the end of stage (ii)* the current values of λ and the α_i and β_i satisfy equation (31) to within a given criterion. (Note that in this case, unlike Case 2, equation (31) cannot be reduced to a data identity because the X_i and U_i are supplied exogenously.) The parameter estimation process is now complete.

v. With the parameter values so obtained the model can be used to simulate the allocation of any given bed supply B', using (10), (11) and solving $f(\lambda) = 0$.

This procedure is less demanding of empirical data on resource allocation than the procedure for Case 2. It requires data on the elasticities $\hat{\gamma}_i$ and $\hat{\eta}_i$ but not on the regional allocations $\bar{x}_i$ and $\bar{u}_i$; thus the latter data can be used to validate the model's performance, as shown in the next chapter.

Unfortunately the parameter estimation part of this procedure does not in general converge because the combination of prospective data on the X_i and U_i and descriptive data for the $\hat{\gamma}_i$ and $\hat{\eta}_i$ is not, in general, consistent with any feasible model solution. To understand why this is so and to remedy the situation we need to examine stage (iii) in more detail. In this stage values of $\bar{B}$ and the X_i, U_i, α_i and β_i are given and the equation $f(\lambda) = 0$ is solved by the Newton-Raphson method. We can regard the computations in this stage as being equivalent to the situation for Mark 1 where the parameters are supplied exogenously and the model is used to simulate the allocations of $\bar{B}$ bed-days; a solution, λ, to the equation $f(\lambda) = 0$ is obtained and simulation results x_i' and u_i' are calculated. However such results do not in general satisfy the condition

$$\sum_i x_i' u_i' (\hat{\gamma}_i + \hat{\eta}_i) = \bar{B}$$

By similar arguments to those given in Appendix 1 it can be shown that unless the elasticity data satisfy this condition there are no values for the parameters α_i and β_i that are consistent with the data; hence the lack of convergence of the procedure described above. The data inconsistency is a direct consequence of the fact we observed above, that for Case 3 the values of the X_i and U_i are supplied from one source, whereas

the elasticity data $\hat{\gamma}_i$ and $\hat{\eta}_i$ are obtained independently from another. Since such mixed data do not, in general, constitute a feasible solution to the model the only way to proceed is to adjust the data until it *does* constitute a feasible solution and then to fit the model parameters to this adjusted data, as with Case 2. We stated above that our discussion of Case 3 is based on the premise that the estimates of the X_i and U_i are derived from a careful *prospective* analysis of HCS behaviour whereas the estimates of elasticities are derived from observations of *past* behaviour. This suggests that it is more reasonable to adjust the latter data rather than the former. Accordingly we propose below a means of scaling the elasticity data so that, together with the data on the X_i and U_i, it is consistent with a feasible model solution. In practice it is probably reasonable to employ such a scaling procedure provided that the consequent alteration of the elasticity data is not very large. If however large alterations were required within this procedure then one would doubt the validity of using the data in this manner.

The elasticity data, the γ_i and η_i, may be scaled by a multiplicative factor, r, given by

$$r = \frac{\bar{B}}{\sum_i x_i' u_i' (\hat{\gamma}_i + \hat{\eta}_i)} \tag{32}$$

and x_i' and u_i' have the meanings described in the preceding paragraph. This scaling can be performed in stage (ii) for *each* iteration after the first. (Note that for Case 2 a similar scaling procedure is required only at the initiation of the procedure.) The same effect can also be achieved, and more conveniently from a computational point of view, by computing the new value of C at the end of stage (iii) by the expression

$$C = \frac{-\sum_i x_i' u_i' (\hat{\gamma}_i + \hat{\eta}_i)}{\lambda f'(\lambda)} \tag{33}$$

rather than

$$C = \frac{-\bar{B}}{\lambda f'(\lambda)} \tag{23}$$

With this modification the procedure is found to converge rapidly.

As with Case 2, the equation system is under-defined but in this case there is only one degree of freedom, whereas with Case 2 there are two. This occurs because, in Case 3, equation (31) cannot be reduced to a data identity since the X_i and U_i are supplied exogenously. Thus for Case 3 we have 2N + 2 equations--(23), (24), (25) and (31)--for 2N + 3 unknowns--C, λ, $f'(\lambda)$ and the α_i and β_i. Computational experience (described in Appendix 2) has shown that the final model outputs, the x_i and u_i, are fairly insensitive to the arbitrary initial values for C and λ. Thus, for the same reasons as for Case 2, the computational procedure suggested above is considered adequate for practical purposes.

5. ILLUSTRATIVE MODEL RUNS

To illustrate how the model can be used we shall examine a hypothetical example of an HCS resource allocation situation--the allocation of acute hospital bed-days in the South Western Region of England in 1968 between patients suffering from six diseases:

- varicose veins;
- haemorrhoids;
- ischaemic heart disease, excluding acute myocardial infarction;
- pneumonia;
- bronchitis; and
- appendicitis.

To obtain the required input quantities we shall use hospital data for the 15 Hospital Regions of England for 1968 and the elasticity estimates of Feldstein shown in Table 1. Illustrative runs for Cases 2 and 3 are presented below.

Case 2

The required input consists of data on the $\bar{x}_i$, $\bar{u}_i$, $\hat{\gamma}_i$ and $\hat{\eta}_i$ and $\bar{B}$. For these illustrative runs the values used for the $\bar{x}_i$, $\bar{u}_i$ and $\bar{B}$ are taken from data [13] on the actual use of hospital beds in the South Western Region in 1968. The elasticity data are derived from the results of Feldstein, Table 1. The full list of input data is shown below in Table 2.

Table 2. Input data for illustrative runs of DRAM for Case 2.

Patient Category i	Mean Admission Rate $\bar{x}_i$	Mean Length of Stay $\bar{u}_i$	Elasticities $\hat{\gamma}_i$	$\hat{\eta}_i$
1. Varicose Veins	6.3	11.3	0.78	0.62
2. Haemorrhoids	4.1	13.1	0.70	0.44
3. Ischaemic Heart Disease	4.6	40.2	1.14	1.08
4. Pneumonia	12.3	14.7	0.71	0.23
5. Bronchitis	11.8	27.4	1.13	0.05
6. Appendicitis	24.8	11.3	0.05	0.31
All Categories	63.9	17.1	-	-
Total bed-days used = $\bar{B} \equiv \sum_i \bar{x}_i \bar{u}_i$ = 1094.2				

Two illustrative runs for Case 2 are described below. In the first run the bed supply input, figure B, is set at a level, 800 bed-days, which is considerably below the level $\bar{B}$ (1094.2 bed-days per million population), which existed in the South Western Region in 1968. In the second run the figure for bed supply is set at a higher level, 1200 bed-days. The results of the two runs are displayed below in Table 3.

Table 3. Output from two illustrative runs of DRAM for Case 2.

Patient Category i	Run 1: B = 800 bed-days/million		Run 2: B = 1200 bed-days/million	
	Admission Rate x_i	Av. Length of Stay u_i	Admission Rate x_i	Av. Length of Stay u_i
1. Varicose Veins	5.0	9.4	6.7	11.9
2. Haemorrhoids	3.3	11.5	4.3	13.6
3. Ischaemic Heart Disease	3.3	29.2	5.1	43.9
4. Pneumonia	10.0	13.7	13.0	15.0
5. Bronchitis	8.4	27.0	12.9	27.5
6. Appendicitis	24.4	10.3	24.9	11.6
All Categories	54.6	14.7	67.0	17.9

Naturally both the admission rate and the length of stay for each category increase from Run 1 to Run 2 as a consequence of the increase in bed supply. However the amount of increase varies considerably, in accordance with the values of the corresponding elasticities. For example, the admission rate for appendicitis increases very little, from 24.4 in Run 1 to 24.9 in Run 2 (see Table 3); this is a direct consequence of the low value of the elasticity for appendicitis admissions, 0.05 (see Table 2). By contrast the admission rate for bronchitis changes a great deal, from 8.4 in Run 1 to 12.9 in Run 2, because of the relatively high value, 1.13, of the elasticity of bronchitis admissions. For similar reasons the length of stay for bronchitis increases very little between runs but the length of stay for ischaemic heart disease increases a great deal.

Case 3

The required input consists of data on the X_i, U_i, $\hat{\gamma}_i$ and $\hat{\eta}_i$ and $\bar{B}$. The data for the elasticities was taken as before from the Feldstein results shown in Table 1. The value used for $\bar{B}$ is set, as before, at the level, 1094.2, of actual usage in the South Western Region in 1968. In a real application data on the X_i and U_i would be obtainable from morbidity estimates and clinical opinion, as described in the previous chapter. Since such data were not available for this exercise, proxy measures were used. These were obtained using data [13] for the 15 regions of England and Wales in 1968; for each individual parameter the highest figure from the 15 regions was selected. For example the largest figure for pneumonia admissions per million population is 12.8, from the North West Metropolitan Region, and this figure was used for the parameter value, X_i, for pneumonia admissions. A full list of the input data is shown in Table 4 below.

Table 4. Input data for illustrative runs of DRAM for Case 3.

Patient Category i	Ideal Admission Rate X_i	Ideal Average Length of Stay U_i	Elasticities $\hat{\gamma}_i$	$\hat{\eta}_i$
1. Varicose Veins	12.8	15.0	0.78	0.62
2. Haemorrhoids	7.7	13.1	0.70	0.44
3. Ischaemic Heart Disease	10.4	52.1	1.14	1.08
4. Pneumonia	21.0	19.7	0.71	0.23
5. Bronchitis	21.3	34.2	1.13	0.05
6. Appendicitis	24.8	10.1	0.05	0.31
All Categories	98.0	22.8	-	-
Bed-day supply for which elasticity estimates apply = $\bar{B}$ = 1094.2				

With this data two illustrative runs of the model were performed with bed supply, B, set at 800 and 1200 respectively, as before. The results are displayed in Table 5 below. It will be seen that they are somewhat similar to those of the illustrative runs for Case 2. The similarity arises for two reasons:

i. the elasticity data is the same in both cases, and

ii. the observed mean regional allocation (the $\bar{x}_i$ and $\bar{u}_i$), used for the Case 2 runs, are, within fairly small margins, consistent with the data on ideal allocations (the X_i and U_i) used for the Case 3 runs and with the elasticity data. In other words the proxy estimates of the ideal allocations used here are, reassuringly, reasonably consistent with the ideals *implied* by the actual regional allocation data used for Case 2.

Thus if DRAM Mark 3 is run with the bed supply B set equal to the value, 1094.2, used for $\bar{B}$ in the runs for Case 2 then the outputs from this run are approximately equal to the corresponding figures for mean regional allocations used for the input data for the runs for Case 2. This can be seen in Table 6 where the results of this run are displayed.

Table 5. Output from two illustrative runs for Case 3.

Patient Category i	Run 1: B = 800 bed-days/million		Run 2: B = 1200 bed-days/million	
	Admission Rate x_i	Av. Length of Stay u_i	Admission Rate x_i	Av. Length of Stay u_i
1. Varicose Veins	6.4	9.1	8.5	11.4
2. Haemorrhoids	4.1	9.0	5.4	10.6
3. Ischaemic Heart Disease	3.6	20.7	5.4	30.9
4. Pneumonia	11.3	16.2	14.7	17.6
5. Bronchitis	8.1	32.8	12.3	33.3
6. Appendicitis	23.7	7.7	24.2	8.7
All Categories	57.2	14.0	70.4	17.0

Table 6. Output from run for Case 3 compared with data used for Case 2 runs.

Patient Category i	Output from Case 3 run for B = 1094.2 bed-days/million		Data used for Case 2 (actual regional allocations for bed usage = 1094.2)	
	Admission Rate x_i	Av. Length of Stay u_i	Admission Rate $\bar{x}_i$	Av. Length of Stay $\bar{u}_i$
1. Varicose Veins	7.9	10.9	6.3	11.3
2. Haemorrhoids	5.1	10.2	4.1	13.1
3. Ischaemic Heart Disease	4.9	28.3	4.6	40.2
4. Pneumonia	13.9	17.3	12.3	14.7
5. Bronchitis	11.2	33.2	11.8	27.4
6. Appendicitis	24.1	8.5	24.8	11.3
All Categories	67.1	16.3	63.9	17.1

For Case 3 it is possible to compare the output quantities, x_i and u_i, with the corresponding ideal quantities X_i and U_i, supplied exogenously. As can be observed from Tables 4 and 5, the output quantities which vary little between runs because of their low elasticities are closer to the corresponding ideals than the quantities for which the elasticities are large. Thus in both sets of results shown in Table 5 the admission rates for appendicitis are close to the ideal level, 24.8, whereas those for bronchitis are relatively far from the ideal level, 21.3. Thus we can interpret a low elasticity, γ_i, as implying high priority for admission and vice versa. This situation is illustrated in Figure 2 which shows admission rate as a function of bed supply for appendicitis and bronchitis. Even if bed supply is low a high proportion of appendicitis cases are admitted whereas for bronchitis, with a lower implied priority, the admission rate is low and only rises as the bed supply is increased.

The response of average length of stay to changing bed supply can be interpreted in a similar way. For example in the results of both of the runs for Case 3, shown in Table 5, the length of stay for bronchitis is close to the ideal level, 34.2, quoted in Table 4; this is a consequence of the low elasticity, 0.05. By contrast the result for the length of stay for varicose veins increases considerably from Run 1 to Run 2, as a result of its relatively high elasticity, 0.62. This situation is illustrated in Figure 3; in regions with a low bed supply the length of stay for varicose veins will be much shorter than in regions with high supply whereas the length of stay for bronchitis is much less affected by bed supply. It is possible to interpret

this situation, in the English context, as follows. In the case of varicose veins it is often possible to discharge patients after a stay as short as two days [14]. There is evidence [13] from the English hospital service that this practice is more common in regions where the bed supply is low than where it is high, thus bringing the *average* length of stay down in the former. By contrast there is less possibility for such practice in the case of bronchitis because of the risk of relapse if the patient is discharged early.

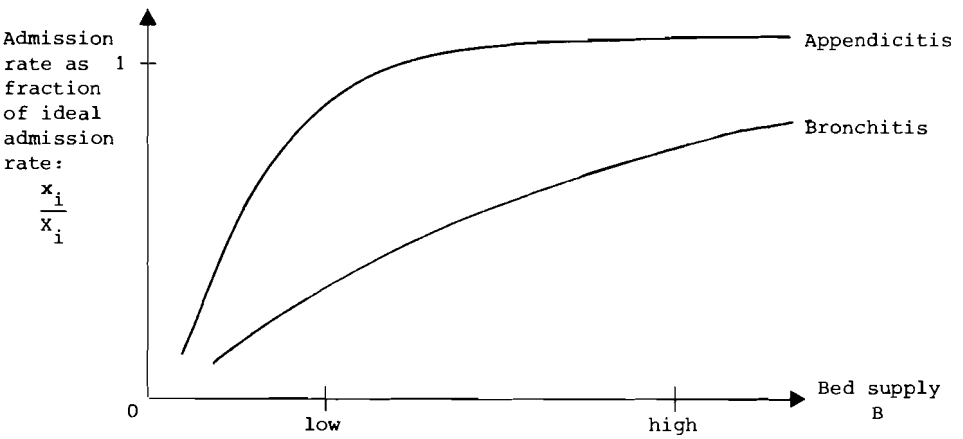

Figure 2. Admission rate as a function of bed supply.

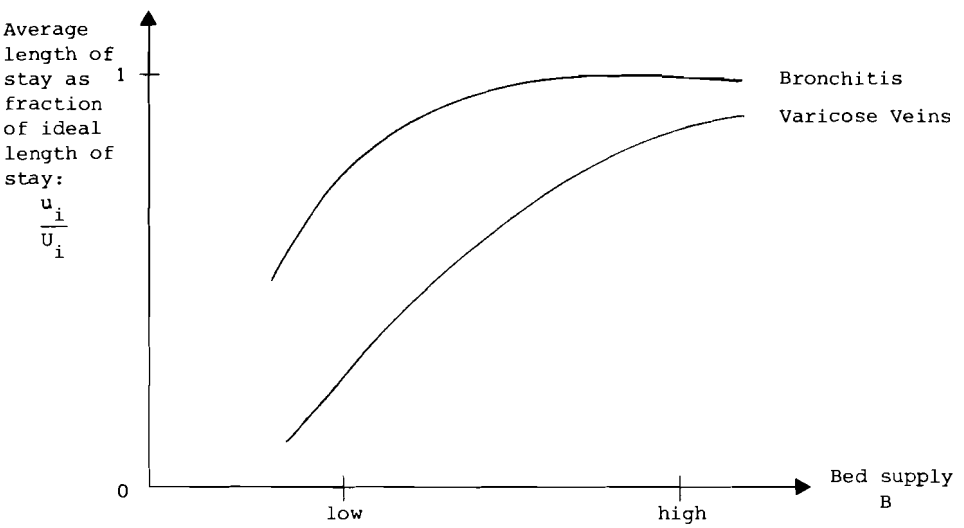

Figure 3. Length of stay as a function of bed supply.

6. FUTURE DEVELOPMENT OF DRAM

DRAM Mark 1, the version of the model described in this paper, represents the allocation of one HCS resource (such as hospital beds) within one mode of treatment (such as in-patient treatment). In the future it is planned to develop more general versions of the model, in particular:

- a Mark 2 version, to represent the allocation of *several* resources within *one* treatment mode; and then
- a Mark 3 version, to include substitution between alternative treatment modes.

The main preoccupation in this development work will be to try to retain the computational convenience of the methods described for DRAM Mark 1.

With DRAM Mark 2 the user would be able to explore a wider range of planning issues than with DRAM Mark 1; rather than merely study the consequences of changing the supply of one resource he would also be able to study the consequences of changing the *mix* of resources within a service. For example the hospital in-patient service could be represented as a mix of different resources--beds, physicians, nurses, laboratories, X-ray equipment, etc.--rather than as a single composite resource, bed-days, as we have had to do for DRAM Mark 1. DRAM Mark 2 would not only be more useful from a planner's point of view, but it would also be a more accurate model from the scientific point of view since it would represent both how different types of patient make different demands on each resource and how some resources have a greater effect on admissions and length of stay than others. For example Feldstein [9] and Prevett [15] have shown that lengths of stay are much more elastic to the availability of physicians than to the availability of nurses.

Computationally it should not prove too difficult to develop DRAM Mark 2. The change in formulation is fairly simple. An additional subscript, k, to represent resource type is introduced. The single constraint (2) in DRAM Mark 1 is replaced by a set of constraints

$$\sum_i x_i u_{ik} = B_k \quad , \qquad \forall\, k \quad ,$$

and terms of the form $\sum_i \sum_k x_i h_{ik}(u_{ik})$ replace the term $\sum_i x_i h_i(u_i)$ in the objective function (1). The problem can be solved as before, using Lagrange Multipliers. However in this case there are a set of multipliers, λ_k (one for each resource), rather than a single multiplier, λ, to be determined.

The scope and usefulness of DRAM Mark 3 would be even greater since it would allow the user to examine the balance between alternative modes of treatment (e.g. in-patient and out-patient treatment) as well as the mix of resources within a mode. The difficulty here will be to retain sufficient simplicity in the formulation so as to allow efficient solution by Lagrange Multipliers (and so avoid being forced to use large and highly specialised computer programmes) while at the same time capturing the essence of the problem of the balance between alternative modes of treatment. Only future study will reveal whether this difficulty can be overcome.

ACKNOWLEDGEMENTS

I am indebted to many colleagues at IIASA for the help I received in the work described in this paper, in particular to Sasha Klementiev, who worked jointly with me on developing the methods and computer programmes for parameter estimation, and to Martina Segalla, who patiently and accurately converted my execrable handwriting into elegant typescript.

REFERENCES

[1] Venedictov, D.D., and E.N. Shigan, *The IIASA Health Care System Model*, paper presented at IIASA Conference on Modeling Health Care Systems, International Institute for Applied Systems Analysis, Laxenburg, Austria, November, 1977.

[2] Gibbs, R.J., *A Disaggregated Health Care Resource Allocation Model*, RM-78-1, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1978.

[3] Rousseau, J.-M., The Need for an Equilibrium Model for Health Care System Planning, in E.N. Shigan and R. Gibbs, eds., *Modeling Health Care Systems - Proceedings of a IIASA Workshop*, CP-77-8, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1977.

[4] Gibbs, R.J., *Health Care Resource Allocation Models - A Critical Review*, RM-77-53, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1977.

[5] McDonald, A.G., G.C. Cuddeford, and E.M.L. Beale, Mathematical Models of the Balance of Care, *British Medical Bulletin*, 30, 3 (1974), 262-270.

[6] Keyfitz, N., *Understanding World Models*, RM-77-18, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1977.

[7] Roemer, M.I., and M. Shain, *Hospital Utilization Under Insurance*, Hospital Monograph Series, No. 6, American Hospital Association, Chicago, 1959.

[8] Harris, D.H., Effect of Population and Health Care Environment on Hospital Utilisation, *Health Services Research*, 10, 229 (1975).

[9] Feldstein, M.S., *Economic Analysis for Health Services Efficiency*, North-Holland, Amsterdam, 1967.

[10] Gibbs, R.J., *Computer Programs for the IIASA Health Care Resource Allocation Sub-Model, Mark 1--A User's Guide*, internal paper, International Institute for Applied Systems Analysis, Laxenburg, Austria (forthcoming).

[11] Keeney, R.L., and H. Raiffa, *Decision Analysis with Multiple Competing Objectives*, Wiley, New York, 1976.

[12] Popov, G.A., *Principles of Health Care Planning in the USSR*, World Health Organization, Geneva, 1971.

[13] Department of Health and Social Security, *Report on Hospital In-Patient Enquiry for the Year 1968*, Her Majesty's Stationery Office, London, 1972.

[14] Doran, F.S.A., M. White, and M. Drury, The Scope and Safety of Short Stay Surgery in Treatment of Groin Herniae and Varicose Veins--Report on 705 Cases, *British Journal of Surgery*, 59 (1972), 333.

[15] Prevett, C., *Hospital Efficiency*, MSc Thesis, Brunel University, London, 1971.

Appendix 1

Restrictions on the Data Used in Case 2 and the Consequent Degrees of Freedom in the Parameter Estimation Process

In Case 2 empirical input data are supplied for $\bar{B}$ and the $\bar{x}_i$, $\bar{u}_i$, $\hat{\gamma}_i$ and $\hat{\eta}_i$. From these, estimates are made of the parameters--the X_i, U_i, α_i and β_i--by the procedure described in Chapter 4. This procedure is based on the proposition that the model solution for the value $B = \bar{B}$ should be consistent with the input data. Thus the data must itself be a feasible model solution. This gives rise to two restrictions on the data, as shown below. It is then shown that these restrictions give rise to two degrees of freedom in the parameter estimation procedure.

For a given bed supply, B, the model produces a solution, the x_i and u_i, to satisfy the constraint

$$\sum_i x_i u_i = B \quad . \tag{2}$$

Thus the first restriction on the input data for the model is

$$\sum_i \bar{x}_i \bar{u}_i = \bar{B} \quad . \tag{16}$$

Let us now consider how the model solution changes in response to changes in the value of B. As before, let γ_i and η_i be the elasticities of the x_i and u_i with respect to B. From (2) we have

$$\sum_i \left(x_i \frac{du_i}{dB} + u_i \frac{dx_i}{dB} \right) = 1 \quad . \tag{32}$$

From the definition of γ_i and η_i (see (17) and (18)), we can show that

$$\gamma_i = \frac{B}{x_i} \frac{dx_i}{dB}$$

and

$$\eta_i = \frac{B}{u_i} \frac{du_i}{dB} \quad .$$

Substituting these in (32) we have

$$\sum_i \left(\frac{x_i u_i \eta_i}{B} + \frac{x_i u_i \gamma_i}{B} \right) = 1 \quad ;$$

therefore

$$\sum_i x_i u_i (\gamma_i + \eta_i) = B \quad . \tag{33}$$

Thus there is a second restriction on the input data

$$\sum_i \bar{x}_i \bar{u}_i (\hat{\gamma}_i + \hat{\eta}_i) = \bar{B} \quad . \tag{28}$$

We will now show how these two restrictions on the data cause two degrees of freedom in the parameter estimation procedure. One of the equations used in the procedure is

$$\sum_i X_i U_i \beta_i^{1/(\alpha_i+1)} \times \left[(\beta_i + 1) \lambda^{(\alpha_i+\beta_i+1)/(\beta_i+1)} - \lambda^{(\alpha_i+1)/(\beta_i+1)} \right]^{-1/(\alpha_i+1)} = \bar{B} \quad . \tag{31}$$

By rearranging equations (26) and (27) we have

$$U_i \lambda^{-1/(\beta_i+1)} = \bar{u}_i \tag{34}$$

and

$$X_i \left\{ \frac{1}{\beta_i} \left[(\beta_i + 1) \lambda^{\beta_i/(\beta_i+1)} - 1 \right] \right\}^{-1/(\alpha_i+1)} = \bar{x}_i \quad . \tag{35}$$

By substituting these equation (31) reduces to the data identity represented by (16); this creates one degree of freedom in the parameter estimation process.

Next consider another equation used in the procedure

$$f'(\lambda) = -\sum_i \frac{X_i U_i}{(\alpha_i + 1)} \beta_i^{1/(\alpha_i+1)} \phi_i^{-(\alpha_i+2)/(\alpha_i+1)} \theta_i \quad , \tag{14}$$

where ϕ_i and θ_i are given by (13) and (15). Substituting (34) and (35) in (14) we have

$$f'(\lambda) = -\sum_i \bar{x}_i \bar{u}_i \phi_i^{-1} \theta_i (\alpha_i + 1)^{-1} \quad . \tag{36}$$

From (13) we have

$$\phi_i = \lambda^{(\alpha_i+\beta_i+1)/(\beta_i+1)} \left(\beta_i + 1 - \lambda^{-\beta_i/(\beta_i+1)} \right) \quad .$$

We recall, see (15), that $\theta_i = \frac{d\phi}{d\lambda} = \phi'$. Thus we can write

$$\phi_i^{-1} \theta_i = \phi_i^{-1} \phi_i'$$

$$= \frac{d \log \phi_i}{d\lambda}$$

$$= \frac{\alpha_i+\beta_i+1}{\beta_i+1}\,\lambda^{-1} + \frac{\beta_i}{\beta_i+1}\,\lambda^{-(2\beta_i+1)/(\beta_i+1)}$$

$$\times \left[\beta_i+1-\lambda^{-\beta_i/(\beta_i+1)}\right]^{-1}$$

$$= \frac{\alpha_i+1}{\lambda}\left\{\frac{1}{\beta_i+1} + \frac{\beta_i}{(\alpha_i+1)\left(\beta_i+1-\lambda^{-\beta_i/(\beta_i+1)}\right)}\right\} .$$

Therefore from (23), (24) and (25)

$$\phi_i^{-1}\theta_i = \frac{\alpha_i + 1}{\lambda}\left\{-\frac{\lambda f'(\lambda)\hat{\eta}_i}{\bar{B}} - \frac{\lambda f'(\lambda)\hat{\gamma}_i}{\bar{B}}\right\} .$$

Substituting this in (36) we have

$$f'(\lambda) = f'(\lambda)\sum_i \bar{x}_i\bar{u}_i(\hat{\gamma}_i + \hat{\eta}_i)\bar{B}^{-1} .$$

$$\sum_i \bar{x}_i\bar{u}_i(\hat{\gamma}_i + \hat{\eta}_i) = \bar{B} .$$

Thus equation (14) reduces to the data identity given by (28) and a second degree of freedom is created in the parameter estimation process for Case 2.

Appendix 2

Suitable Initial Values in the Parameter Estimation Process

The parameter estimation process for Cases 2 and 3 are described in Chapter 4. For each Case the stage (i) consists of setting arbitrary initial values for λ and C. In this Appendix suitable ranges for these initial values are derived. Within such ranges the final outputs of the model are shown to be insensitive to the initial values selected.

First of all, it can be shown, by reference to equations (24) and (25), that the smallest value, C_{min}, of C which guarantees positive values for each of the α_i and β_i in stage (ii) of the parameter estimation process is given by

$$C_{min} = \underset{i}{\text{Max}} \left(\hat{\gamma}_i + \hat{\eta}_i \right) \quad . \tag{42}$$

In order to set an upper bound for C we need to consider plausible upper bounds for the α_i and β_i.

We recall, from Chapter 2, that the α_i and β_i are related to the elasticities E_i and F_i of admissions and length of stay with respect to marginal opportunity cost (and, equivalently, marginal utility)

$$E_i = \frac{1}{\alpha_i + 1} \quad , \qquad F_i = \frac{1}{\beta_i + 1} \quad .$$

Substituting these in (24) and (25) we have

$$\hat{\eta}_i = CF_i$$

and

$$\hat{\gamma}_i = CE_i\beta_i \left(\beta_i + 1 - \lambda^{-\beta_i/(\beta_i+1)} \right)^{-1}$$

$$\approx CE_i \qquad \text{for a wide range of values of } \beta_i > 0 \text{ and } \lambda > 1.$$

On *a priori* grounds it is reasonable that the hypothetical elasticities E_i and F_i with respect to opportunity cost should be similar in magnitude to the empirical elasticities $\hat{\gamma}_i$ and $\hat{\eta}_i$ with respect to bed supply. This implies a value of C that is within an order of magnitude of unity.

Using the test problem described in Chapter 5 some sensitivity analysis was performed. The initial value of C was set first at the minimum value, C_{min}, given by (42), which was 2.22, and second at the value 10.0. It was found that the sensitivity of the final model outputs, the x_i and u_i, to this change was about 0.1% for Case 2 and 1% for Case 3. It was concluded that any value of C between C_{min} and 10.0 would be suitable for most problems but the value, C_{min}, given by (43), being closest to unity, would be most suitable. Since the largest of the elasticities $\hat{\gamma}_i$ and $\hat{\eta}_i$ are rarely much in excess of unity, the application of (42) will typically elicit a starting value for C that is of the order of two.

From this definition of a suitable initial value for C we may deduce suitable values for λ. In Case 2 we assume that we have no information about the ideal quantities given by the parameters X_i and U_i. We wish to use the model to explore the consequences of setting the bed supply B to a number of different values. Let us suppose that the largest conceivable value of interest is given by

$$B = 2\bar{B} \quad , \tag{43}$$

i.e. a bed supply of double the current regional value. It is also necessary in solving the model that the parameters X_i and U_i should satisfy

$$B \leq \sum_i X_i U_i \quad . \tag{44}$$

The following equality satisfies both conditions:

$$\sum_i X_i U_i = 2\bar{B} \quad . \tag{45}$$

Computational experience with Case 2 suggests that the estimates produced for the X_i and U_i depend on the starting values for λ and C in a manner given, approximately, by

$$\sum_i X_i U_i \approx \bar{B}\lambda^{1/C} \quad .$$

Thus, from (45) we have

$$\lambda^{1/C} \approx 2 \quad ;$$

therefore

$$\lambda \approx 2^C \quad . \tag{46}$$

It is suggested that equation (46) be used for assigning a suitable initial value for λ. Since the most suitable value for C is of the order of two, the application of (46) will result in a starting value for λ that is of the order of four.

Sensitivity analysis using the test problem described in Chapter 5 has shown that if the starting value of λ is changed from 2.0 to 10.0 the sensitivity of the final model outputs, the x_i and u_i, is of the order of 0.1% for Case 2 and only 0.01% for Case 3. Thus the procedures described above for setting the initial values of C and λ are considered satisfactory for practical purposes.

Papers of the Modeling Health Care Systems Study

Kiselev, A., *A Systems Approach to Health Care*, RM-75-31, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1975.

Venedictov, D.D., Modeling of Health Care Systems, in *IIASA Conference '76*, Vol. 2, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1976.

Fleissner, P., *Comparing Health Care Systems by Socio-Economic Accounting*, RM-76-19, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1976.

Klementiev, A.A., *A Computer Method for Projecting a Population's Sex-Age Structure*, RM-76-36, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1976.

Klementiev, A.A., *Mathematical Approach to Developing a Simulation Model of a Health Care System*, RM-76-65, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1976.

Kaihara, S., et al., *An Approach to Building a Universal Health Care Model: Morbidity Model of Degenerative Diseases*, RM-77-06, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1977.

Shigan, E.N., *Alternative Analysis of Different Methods for Estimating Prevalence Rate*, RM-77-40, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1977.

Klementiev, A.A., *On the Estimation of Morbidity*, RM-77-43, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1977.

Fleissner, P., and A. Klementiev, *Health Care System Models: A Review*, RM-77-49, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1977.

Gibbs, R.J., *Health Care Resource Allocation Models - A Critical Review*, RM-77-53, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1977.

Gibbs, R.J., *A Disaggregated Health Care Resource Allocation Model*, RM-78-1, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1978.

Kaihara, S., et al., *Analysis and Future Estimation of Medical Demands Using a Health Care Simulation Model: A Case Study of Japan*, RM-78-3, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1978.

Fujimasa, I., S. Kaihara, and K. Atsumi, *A Morbidity Submodel of Infectious Diseases*, RM-78-10, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1978.