DRAM: A MODEL OF HEALTH CARE RESOURCE ALLOCATION

David Hughes
Andrzej Wierzbicki

RR-80-23
May 1980

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
Laxenburg, Austria

Research Reports, which record research conducted at IIASA, are independently reviewed before publication. However, the views and opinions they express are not necessarily those of the Institute or the National Member Organizations that support it.

FOREWORD

The principal aim of health care research at IIASA has been to develop a family of submodels of national health care systems for use by health service planners. The modeling work is proceeding along the lines proposed in the Institute's current Research Plan. It involves the construction of linked submodels dealing with population, disease prevalence, resource need, resource allocation, and resource supply.

This is the second research report on the *d*isaggregated *r*esource *a*llocation sub-*m*odel called DRAM. It describes the extension of the Mark 1 version (RR-78-8) to include the distribution of many resources across different modes of care. The earlier assumption that all available resources must be used has been relaxed, and an extensive analytic treatment suggests various methods for estimating the submodel's parameters. Several case studies that use the model are in progress and reports on these applications will be forthcoming.

This paper is an output of a collaboration between two Areas at IIASA. It describes how a health resource allocation model, developed in the Health Care Systems Task of the Human Settlements and Services Area, may be solved by using optimization techniques studied in the Optimization Task of the System and Decision Sciences Area.

Related publications in Health Care Systems and in Optimization are listed at the end of this report.

ANDREI ROGERS
Chairman
Human Settlements
and Services Area

ANDRZEJ WIERZBICKI
Chairman
System and Decision
Sciences Area

ACKNOWLEDGMENTS

We are indebted to many colleagues at IIASA and elsewhere for their help and advice. We would like to offer special thanks to Rebecca Crow, who prepared the typescript, tables, and diagrams, and to the staff of the South Western Regional Health Authority, who provided the data used in Section 5.

CONTENTS

1 INTRODUCTION

It has been widely observed (Feldstein 1967, Van der Gaag *et al.* 1975, Rousseau 1977) that the demand for health care seems to be insatiable. When more hospitals are opened, more patients are treated, and the hope expressed at the inception of the U.K. National Health Service that increasing supplies of health care would reduce subsequent demands has not been realized there or anywhere. The causes of this phenomenon are various, but it gives rise to the same question in all countries: What health care resources should be made available?

Unfortunately, the principal output of health care systems – health – is almost impossible to define or measure (Cardus and Thrall 1977). Much as we would like to design a health care system that would maximize health, we do not even know how to begin. Instead, we seek to predict how those hospitals and other resources available in the health care system (HCS) will be used. Who gets what?

DRAM (a *d*isaggregated *r*esource *a*llocation *m*odel) is designed to help answer such questions. It is one of the submodels of the HCS model conceived by Venedictov *et al.* (1977), and being developed by a group of scientists from different countries working at the International Institute for Applied Systems Analysis (IIASA). Figure 1 shows the five groups of submodels of the HCS developed so far at IIASA; they are explained in more detail in a recent status report (Shigan *et al.* 1979). This figure represents one part of the complete HCS: the processes by which people fall ill and by which resources are provided and used for their care. DRAM (in the group of resource allocation submodels) represents how the HCS allocates limited supplies of resources among competing demands of morbidity. Specifically, it asks *If a certain mix of health care resources (e.g., hospital beds, nursing care) is available, how will the HCS distribute them among patients?* DRAM does *not* prescribe an optimal allocation of resources.

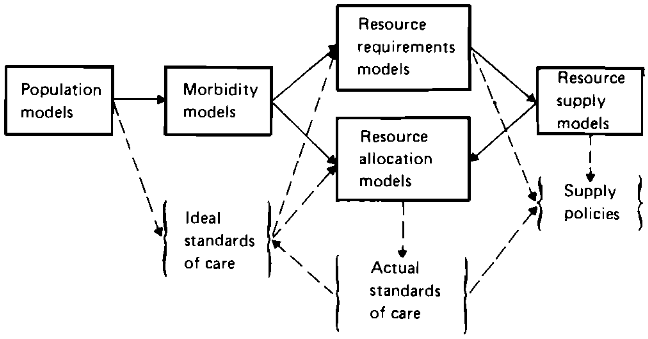

FIGURE 1 The family of HCS submodels constructed at IIASA.

Instead, it *simulates* how the HCS responds when resource availability changes. Even in countries with market economies, there are invariably some planning instruments for controlling the supply of public goods. But even in countries with planned economies, resources cannot be allocated in a rigid, centralized manner. In every country, doctors have clinical responsibility for their patients, and the pattern of care is determined by many local decisions. McDonald *et al.* (1974), Rousseau (1977), and Burton *et al.* (1978) are among those who have modeled this behavior, and DRAM has close links with the first of these models. Other models for health care resource allocation were reviewed by Gibbs (1977) and Nackel *et al.* (1978).

Like many models, DRAM has *accounting* and *behavioral* components. In the accounting in DRAM, different types of resources are distributed among patients

- In different categories (e.g., age, diagnosis)
- In different modes of care (e.g., inpatient, outpatient)
- With different levels of resources per patient (e.g., length of stay in hospital)

and no more resources are allocated than are available. The resources can be determined by a resource supply/production submodel such as the IIASA submodel described in Shigan *et al.* (1979), or they can be set by the user as a trial policy option.

The behavioral assumption in DRAM is that the HCS behaves as if it were maximizing a preference function that increases with the number of patients treated and the resources received by each. Some of the parameters in this function represent demand inputs, like the *ideal levels* at which patients would be treated and would receive resources if no constraints on resource availability existed. These parameters indicate the true "needs" for health care. Other parameters represent the *elasticities* of the actual levels to changes in resource supply, and the balance between need and supply. The relative *costs* of different resources are other parameters used by DRAM to choose between alternative resource mixes. DRAM does not try to include explicitly every behavioral influence that could be active, but to use parameters that can represent the results of all these influences. Because the parameters have meanings outside the model, they can be estimated by methods that do not involve the assumptions underlying DRAM.

Gibbs (1978b) formulated a pilot Mark 1 version of DRAM. This report is the successor, and summarizes progress up to April 1979. Some but not all of the results have appeared in interim IIASA papers, a list of which appears at the end of this report. Much of this report is about the mathematics of the model, and the examples are concerned with hospital services. Our interests, however, are not so restricted. DRAM is designed to model the concept that the HCS balances the desirabilities of more individuals receiving care against higher average levels of care. Such a model should be applicable in many sectors of health care, and perhaps also in other public sectors.

Readers who are uninterested in mathematical details can skip to Section 5 to read about the use of mathematical models in general and to see examples of DRAM. Two examples are presented: one investigates how hospital beds might be used by the HCS, and the other how the balance between inpatient and outpatient care might change. The other parts of the report develop mathematical results that are needed to support such applications. Section 2 solves the simple DRAM and gives three extensions in which certain restrictions applied to the simple model are removed. Not every resource allocation pattern can be simulated by DRAM, so Section 3 investigates its admissible solutions. This is a way of explaining the implications of DRAM's underlying hypothesis. Section 4 presents methods for calibrating DRAM so that it is appropriate for different questions of policy in different regions. The associated computer programs are not described in this report, but Appendix B provides brief details. Section 6 gives a concise summary of the whole report.

2 MODEL FORMULATION AND SOLUTION

The first step in formulating DRAM is to define variables and to make the key assumptions in the model precise. This is done in Section 2.1, and Section 2.2 analyzes a simple version of DRAM in which all the available resources must be used. Three extensions of the model are analyzed in Section 2.3, and Section 2.4 describes a computational method that can be used to solve all four versions of DRAM.

2.1 *Notation and Assumptions*

We use the indices j = patient category ($j = 1, 2, \ldots, J$), k = mode of care ($k = 1, 2, \ldots, K$), and l = resource type ($l = 1, 2, \ldots, L$) in defining the model variables

x_{jk} = numbers of individuals in patient category j who receive resources in mode of care k (per head of population per year)

y_{jkl} = supply of resource type l received by each individual in patient category j in mode of care k

and in writing $\Sigma_j \Sigma_k x_{jk} y_{jkl}$ as the total resources of type l that are allocated (per head of population per year). DRAM seeks to determine x_{jk}, y_{jkl} $\forall j, k, l$, within constraints on total resources, so as to maximize a function

$$U(x, y) = \sum_j \sum_k g_{jk}(x_{jk}) + \sum_j \sum_k \sum_l x_{jk} h_{jkl}(y_{jkl}) \tag{1}$$

where

$$g_{jk}(x) = \frac{\sum_l C_l X_{jk} Y_{jkl}}{\alpha_j} \left[1 - \left(\frac{x}{X_{jk}}\right)^{-\alpha_j}\right] \tag{2}$$

$$h_{jkl}(y) = \frac{C_l Y_{jkl}}{\beta_{jkl}} \left[1 - \left(\frac{y}{Y_{jkl}}\right)^{-\beta_{jkl}}\right] \tag{3}$$

C, X, Y, α, β are model parameters (C denotes $\{C_l, l = 1, 2, \ldots, L\}$ and so on).

The monotonically increasing, concave power functions (2) and (3) follow from general assumptions about aggregate behavior in the HCS. They depict the many agents who control the allocation of health care resources as seeking to attain ideal levels of service (X) and supply (Y), but where the urge to increase the actual levels of service (x) and supply (y) decreases with increasing values of x and y, according to the parameters α and β. The costs of different resources (C) are introduced so that marginal increases in U when ideal levels are achieved ($x = X$, $y = Y$) equal the marginal resource costs. This interpretation is a useful way of introducing meaningful parameters into the model, and Section 4 suggests various ways of estimating X, Y, α, β, C in different applications. For the moment, however, we assume these parameters to be known.

Alternative forms for $U(x, y)$ can be suggested, and some were analyzed by Hughes (1978b). Appendix A presents one of these and shows that minor changes can greatly change both the characteristics of model predictions and the ease of solution. Equations (1)–(3) have convenient analytic properties that make it easy to solve this formulation of the model.

2.2 *The Simple Model*

We seek a solution for x, y that maximizes Eq. (1) subject to the constraints

$$0 \leqslant x_{jk} \leqslant X_{jk} \qquad 0 \leqslant y_{jkl} \leqslant Y_{jkl} \tag{4}$$

In this section, we assume that all available resources of type l, R_l, must be used.

$$F_l(x, y) = R_l - \sum_j \sum_k x_{jk} y_{jkl} = 0 \qquad \forall l \tag{5}$$

With Lagrange multipliers λ_l, $\forall l$, we adjoin the equality constraint, Eq. (5), to the function that is to be maximized, Eq. (1), to give

$$H(x, y, \lambda) = U(x, y) + \sum_l \lambda_l F_l(x, y) \tag{6}$$

When certain convexity and concavity assumptions are satisfied (proved below), the values of x, y, λ that solve the primal problem of $\max_{x,y} \min_\lambda H(x, y, \lambda)$ also solve the dual problem of $\min_\lambda \max_{x,y} H(x, y, \lambda)$. The optimal values $\hat{x}$, $\hat{y}$ are readily found to be

$$\hat{x}_{jk}(\lambda) = X_{jk}\,[\mu_{jk}(\lambda)]^{-1/(\alpha_j+1)} \tag{7}$$

$$\hat{y}_{jkl}(\lambda) = Y_{jkl}(\lambda_l/C_l)^{-1/(\beta_{jkl}+1)} \tag{8}$$

where μ_{jk} is a weighted sum

$$\mu_{jk} = \frac{\sum_l C_l Y_{jkl} \nu_{jkl}}{\sum_l C_l Y_{jkl}} \tag{9}$$

of the terms

$$\nu_{jkl} = [(\beta_{jkl} + 1)(\lambda_l/C_l)^{\beta_{jkl}/(\beta_{jkl}+1)} - 1]/\beta_{jkl} \tag{10}$$

and substituting these values into Eq. (6) yields

$$\begin{aligned} \hat{H}(\lambda) &= H[\hat{x}(\lambda), \hat{y}(\lambda), \lambda] \\ &= \sum_j \sum_k \sum_l \frac{C_l X_{jk} Y_{jkl}}{\alpha_j} \{1 - [\mu_{jk}(\lambda)]^{\alpha_j/(\alpha_j+1)}\} \\ &\quad + \sum_j \sum_k \sum_l \frac{C_l X_{jk} Y_{jkl}}{\beta_{jkl}} \{\mu_{jk}(\lambda)^{-1/(\alpha_j+1)}\}\{1 - (\lambda_l/C_l)^{\beta_{jkl}/(\beta_{jkl}+1)}\} \\ &\quad + \sum_l \lambda_l F_l[\hat{x}(\lambda), \hat{y}(\lambda)] \end{aligned} \tag{11}$$

However, these solutions for x, y are not determined until we find a value $\hat{\lambda}$ that minimizes $\hat{H}(\lambda)$.

In order to see whether this is possible, we inspect the gradient vector of first derivatives $\hat{H}_\lambda$ evaluated at $x = \hat{x}(\lambda)$, $y = \hat{y}(\lambda)$. After much simplification, this is simply the vector with elements

$$\begin{aligned} \frac{\partial \hat{H}(\lambda)}{\partial \lambda_l} &= F_l[\hat{x}(\lambda), \hat{y}(\lambda)] \\ &= R_l - \sum_j \sum_k X_{jk} Y_{jkl}(\lambda_l/C_l)^{-1/(\beta_{jkl}+1)}[\mu_{jk}(\lambda)]^{-1/(\alpha_j+1)} \end{aligned} \tag{12}$$

The corresponding Hessian matrix of second derivatives $\hat{H}_{\lambda\lambda}$ can be written as the sum of two matrices

$$\hat{H}_{\lambda\lambda} = \left\{ \frac{\partial^2 \hat{H}(\lambda)}{\partial\lambda_l \partial\lambda_m} \right\} = A + B \tag{13}$$

with elements

$$a_{lm} = \frac{1}{C_l} \sum_j \sum_k \frac{X_{jk} Y_{jkl}}{\beta_{jkl} + 1} \left(\frac{\lambda_l}{C_l}\right)^{-(\beta_{jkl}+2)/(\beta_{jkl}+1)} [\mu_{jk}(\lambda)]^{-1/(\alpha_j+1)} \delta_{lm}$$

$$b_{lm} = \sum_j \sum_k \frac{X_{jk} Y_{jkl}}{\alpha_j + 1} \left(\frac{\lambda_l}{C_l}\right)^{-1/(\beta_{jkl}+1)} [\mu_{jk}(\lambda)]^{-(\alpha_j+2)/(\alpha_j+1)} \frac{\partial \mu_{jk}(\lambda)}{\partial \lambda_m}$$

where

$$\frac{\partial \mu_{jk}(\lambda)}{\partial \lambda_m} = \frac{Y_{jkm}}{\sum_m C_m Y_{jkm}} \left(\frac{\lambda_m}{C_m}\right)^{-1/(\beta_{jkm}+1)}$$

and where the Kronecker delta δ_{lm} is 1 when l equals m, and 0 otherwise. A is a diagonal matrix with all elements positive. Therefore, any quadratic form $z'Az$ is always positive, as are all the eigenvalues of A. Equivalently, A is positive definite. The matrix B is symmetric, with typical quadratic forms

$$\begin{aligned} z'Bz &= \sum_{lm} b_{lm} z_l z_m \\ &= \sum_{jk} \left[\frac{X_{jk} [\mu_{jk}(\lambda)]^{-(\alpha_j+2)/(\alpha_j+1)}}{(\alpha_j + 1) \sum_m C_m Y_{jkm}} \right] \left[\sum_l z_l Y_{jkl} (\lambda_l / C_l)^{-1/(\beta_{jkl}+1)} \right]^2 \end{aligned}$$

which are non-negative. Therefore, B is positive semidefinite. It follows that $\hat{H}_{\lambda\lambda}$ is symmetric and positive definite, and this guarantees that $\hat{H}(\lambda)$ is strongly convex. Finally, it can be shown that $\hat{H}(\lambda)$ therefore has a unique minimum for some $\lambda = \hat{\lambda}$.

In order to prove that this solution to the dual problem also solves the primal problem, we consider the matrix of second derivatives of $H(x, y, \lambda)$ with respect to the primal variables $z = (x, y)$, evaluated at $x = \hat{x}(\lambda)$, $y = \hat{y}(\lambda)$. In this partitioned matrix

$$H_{zz} = \begin{bmatrix} H_{xx} & H_{xy} \\ H_{yx} & H_{yy} \end{bmatrix}$$

not only the off-diagonal submatrices, but all the off-diagonal terms are zero. The remaining diagonal elements

$$\frac{\partial^2 g_{jk}[\hat{x}_{jk}(\lambda)]}{\partial x_{jk}^2} = -(\alpha_j + 1) \frac{\sum_l C_l Y_{jkl}}{X_{jk}} [\mu_{jk}(\lambda)]^{(\alpha_j+2)/(\alpha_j+1)}$$

$$\frac{\partial^2 h_{jkl}[\hat{y}_{jkl}(\lambda)]}{\partial y_{jkl}^2} = -(\beta_{jkl} + 1) \frac{C_l}{Y_{jkl}} \left(\frac{\lambda_l}{C_l}\right)^{(\beta_{jkl}+2)/(\beta_{jkl}+1)}$$

are negative, so that $H_{zz}[\hat{x}(\lambda), \hat{y}(\lambda), \lambda]$ is negative definite. This is sufficient to ensure that the solution $[\hat{x}(\hat{\lambda}), \hat{y}(\hat{\lambda})]$ is the saddle point to $H(x, y, \lambda)$, and thus solves both dual and primal problems.

It remains to consider the range of possible solutions for λ. As any λ_l tends to zero, all the elements of $\hat{H}_\lambda$ tend to minus infinity. We deduce therefore that $\hat{\lambda}_l > 0$ for all l. In order for the solutions (7) and (8) to satisfy the constraints (4), we should have $\hat{\lambda}_l > C_l$ for all l. Unfortunately, this cannot be guaranteed, and examples can be found that use all the resources but exceed the ideal standards X, Y. These unrealistic solutions are a deficiency of this simple formulation of DRAM which can be overcome by extending the model.

2.3 *Three Extensions*

In the first extension of the simple model, we remove the constraint on individual resource types (5) and add a constraint on total finance. We seek a solution for x, y that maximizes Eq. (1) subject to constraints (4) and

$$F(x,y) = M - \sum_l C_l \sum_j \sum_k x_{jk} y_{jkl} = 0 \qquad (14)$$

This solution is the optimal allocation under the assumption that finance M should be used to purchase resources that will maximize the returns of Eq. (1). This assumption is not so realistic for our applications, but it gives a model that is easy to solve.

We find that the optimal values $\hat{x}, \hat{y}$ are the same as solutions (7) and (8), but that the Lagrange multiplier λ is now constant across all resource types, $\lambda_1 = \lambda_2 = \ldots = \lambda_L$. The dual function $\hat{H}(\lambda)$ is a function of a single Lagrange multiplier, λ_1 say, and using the earlier results, we can show that it is the sum of a set of strongly convex functions. It is therefore also strongly convex with a unique minimum for some value $\hat{\lambda}_1 > 0$.

In fact, we can demonstrate a stronger result for this version of the model. Because

$$\lambda_1 = 1 \Longrightarrow \frac{\partial \hat{H}(\lambda)}{\partial \lambda_1} = M - \sum_j \sum_k \sum_l C_l X_{jk} Y_{jkl} < 0$$

$$\lambda_1 \to \infty \Longrightarrow \frac{\partial \hat{H}(\lambda)}{\partial \lambda_1} = M > 0$$

and

$$\frac{\partial^2 \hat{H}(\lambda)}{\partial \lambda_1^2} > 0 \qquad 1 \leqslant \lambda_1 < \infty$$

we deduce that there is a unique optimal value $\hat{\lambda}_1 > 1$ that minimizes $\hat{H}(\lambda)$, provided only that the finance available is less than that required to satisfy all demands $M < \Sigma_j \Sigma_k \Sigma_l C_l X_{jk} Y_{jkl}$. In other words, there is always a unique resource mix that will maximize perceived preferences.

In the second extension of the simple model, we replace the equality resource constraint (5) by an inequality constraint

$$F_l(x,y) - r_l = 0; \qquad r_l \geqslant 0 \qquad \forall l \qquad (15)$$

where r_l represents the unused resources of type l, which must always be nonnegative. It is easy to show that there always exists a point (x, y, r) that satisfies constraints (4) and (15) provided that the inequality

$$\sum_j \sum_k X_{jk} Y_{jkl} > R_l > 0 \qquad \forall l \tag{16}$$

is satisfied. When sufficient resources of some type are available to violate Eq. (16), it means that there are more than enough of these resources, and that there is no allocation problem! The resource type in excess can be removed from the model.

It is also possible to show that the model can have no solutions with $\hat{x}_{jk} = 0$, $\hat{y}_{jkl} = 0$, or $\hat{x}_{jk} = X_{jk}$. In other words, these constraints are never active. This is because the first two conditions imply that $U(x, y) = -\infty$, and because the last condition requires $\lambda_l = 1$, $\forall l$, which causes constraint (15) to contradict (16). We conclude then that the only constraints that can be active are the upper constraint on y and the lower constraint on r.

There are now just two possibilities. The first possibility is that $\hat{y}_{jkl} < Y_{jkl}$ for all l. Inspection of the function

$$H(x, y, \lambda) = U(x, y) + \sum_l \lambda_l \left(R_l - r_l - \sum_j \sum_k x_{jk} y_{jkl} \right)$$

shows that it is maximized when r_l is zero for all l. The problem is then identical to that analyzed above, and all the previous results hold true. The second possibility is that $\hat{y}_{jkl} = Y_{jkl}$ for one or more (but not L) resource types l. From Eq. (8), the associated values of (λ_l / C_l) are unity, and the rest of the problem is equivalent to the dual problem specified in Section 2.2, but with the extra constraint

$$\lambda_l \geqslant C_l \qquad \forall l \tag{17}$$

The third extension of the simple model subtracts the costs of the used resources from the preference function

$$U(x, y) = \sum_j \sum_k g_{jk}(x_{jk}) + \sum_j \sum_k \sum_l x_{jk} h_{jkl}(y_{jkl}) - \sum_j \sum_k \sum_l C_l x_{jk} y_{jkl} \tag{18}$$

Other things being equal, the model now tries additionally to maximize the value of unused resources. The optimal values of x, y are similar to solutions (7) and (8)

$$x_{jk}(\lambda) = X_{jk} [\mu_{jk}(\lambda + C)]^{-1/(\alpha_j + 1)} \tag{19}$$

$$y_{jkl}(\lambda) = Y_{jkl} \left(\frac{\lambda_l + C_l}{C_l} \right)^{-1/(\beta_{jkl} + 1)} \tag{20}$$

and we may show that $\hat{H}(\lambda)$ is strictly convex as before, with a unique minimum that now lies in the range $\hat{\lambda}_l > -C_l$ for all l. Should we also wish to replace the equality resource constraint by the inequality constraint (15), the appropriate version of the dual constraint (17) becomes

$$\lambda_l \geqslant 0 \qquad \forall l \tag{21}$$

Note that all three extensions of the simplest model have solutions that are transformations of the simplest solution.

2.4 *Solution Procedure*

So far we have demonstrated only that all the versions of the model discussed above can be solved by solving equivalent dual problems. In each case we have to find $\hat{\lambda}$ so as to minimize $\hat{H}(\lambda)$, sometimes subject to constraints like (21), but with a unique solution always guaranteed. Because we know the gradient vector $\hat{H}_\lambda$ and the Hessian matrix $\hat{H}_{\lambda\lambda}$, we can begin to search for $\hat{\lambda}$ by an iterative technique $\lambda^{i+1} = \lambda^i + td^i$ (the upper index i denotes the iteration number) which finds better approximations $\lambda^i, i = 1, 2, \ldots, N$, to the solution $\hat{\lambda}$, by taking steps with step-size coefficient t, in the Newton direction

$$d^i = -(\hat{H}^i_{\lambda\lambda})^{-1}\hat{H}^i_\lambda \tag{22}$$

Just two refinements are necessary: first, to control the step size, and second, to modify the direction when a constraint like (17) or (21) is applied and encountered.

In order to control the step-size coefficient, we need only reduce it if a step seems likely to overshoot either the solution or a constraint. Figure 2 depicts an appropriate method that tests for this. To proceed when a constraint like (21) is encountered, we determine the set of resource type indices

$$\bar{L} = \left\{ l : \lambda_l = 0, \frac{\partial \hat{H}(\lambda)}{\partial \lambda_l} > 0 \right\}$$

where the constraint is active, and where $\hat{H}(\lambda)$ can decrease only with negative λ_l. The gradient vector $\hat{H}_\lambda$ and the Hessian matrix $\hat{H}_{\lambda\lambda}$ are then projected onto the space of active constraints by replacing all the elements corresponding to active constraints by zeros. They become the *reduced* gradient vector and Hessian matrix and they determine the Newton direction (22) in the space of inactive constraints $l \notin \bar{L}$, which is complemented by zeros for $l \in \bar{L}$.

Figure 3 shows the complete procedure for determining the optimal $\hat{\lambda}$, and hence the solutions $\hat{x}(\hat{\lambda})$, $\hat{y}(\hat{\lambda})$. A matrix inversion is the only potentially difficult computation. Generally, however, the number of different resource types will be sufficiently small (less than five, say) to prevent problems. (Occasionally in the solution of a badly conditioned problem, a step in the Newton direction will not reduce the function $\hat{H}$ because of numerical errors, and steepest descent $d = -\hat{H}_\lambda$ may be necessary.) Note that there is not too much extra computation introduced by an inequality resource constraint. Most

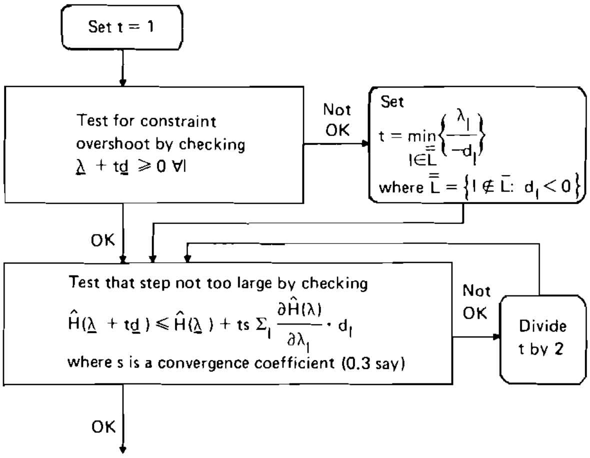

FIGURE 2 Procedure for determining step-size coefficient t.

of the additional refinements are logical rather than computational. All our applications have been solved by a fairly compact computer program, using no special software; Appendix B gives more details about program size and computing efficiency. This program can handle the simple model and all three extensions. In our examples, however, and in most of this report, we refer to the simple model.

3 SOLUTION CHARACTERISTICS

DRAM cannot simulate all patterns of resource allocation that might be observed, and the possibilities for use depend upon the variety of patterns that can be simulated. The analysis given here of admissible solutions to DRAM is restricted to the simplest possible DRAM with one patient category, one treatment mode, and one type of resource, for which all the variants described in Section 2 are identical. Section 3.1 shows how the simplest model can be represented graphically, and gives a fundamental condition on admissible solutions. The results indicate the characteristics of solutions for more complex DRAMs, and suggest ways to fit the model to small numbers of data points. Sections 3.2 and 3.3 derive conditions for fitting two parameters to two data points (Appendix C derives conditions for fitting four parameters to four data points), Section 3.4 derives conditions for fitting two parameters to one data point, and Section 3.5 derives conditions for fitting four parameters to two data points. These results introduce the next section on parameter estimation from many data points.

3.1 *The Simplest DRAM*

For the simplest possible DRAM with $J = K = L = 1$, many elements of the problem can be depicted graphically. First, we can eliminate the Lagrange

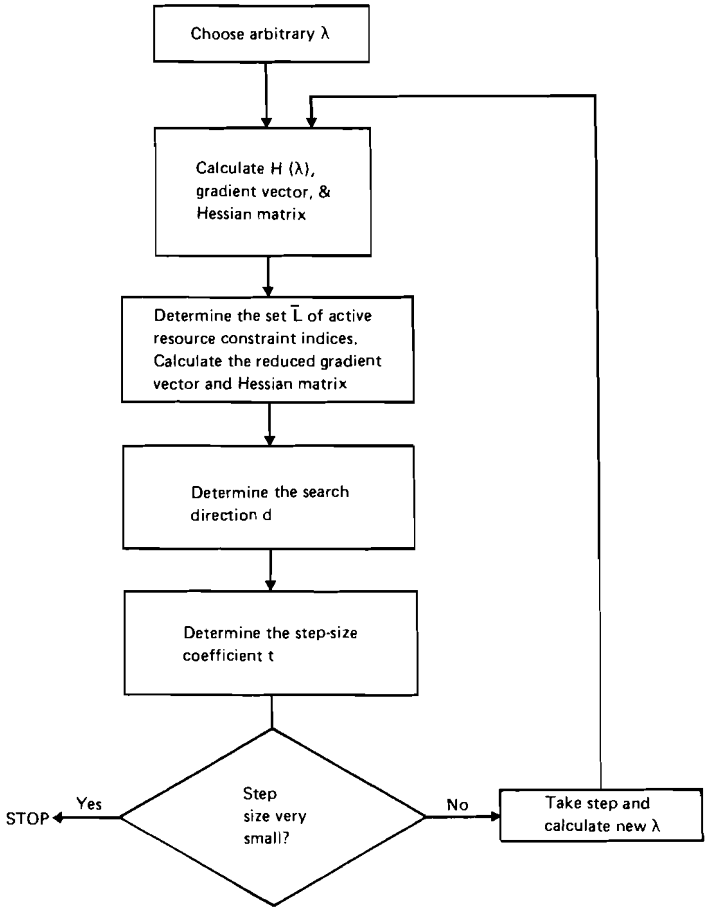

FIGURE 3 Iterative procedure for solving DRAM.

multiplier between Eqs. (19) and (20), to show how the resource level R (which is input to the model) determines the number of individuals treated x, and the supply level y (which are outputs)

$$\left(\frac{R}{XY}\right) = \left(\frac{x}{X}\right)\left[\frac{\beta(x/X)^{-(\alpha+1)} + 1}{\beta + 1}\right]^{-1/\beta} \tag{23}$$

$$\left(\frac{R}{XY}\right) = \left(\frac{y}{Y}\right)\left[\frac{(\beta + 1)(y/Y)^{-\beta} - 1}{\beta}\right]^{-1/(\alpha+1)} \tag{24}$$

It is easy to show that these equations have the shapes shown in Figures 4 and 5. Both curves are concave and monotonically increasing.

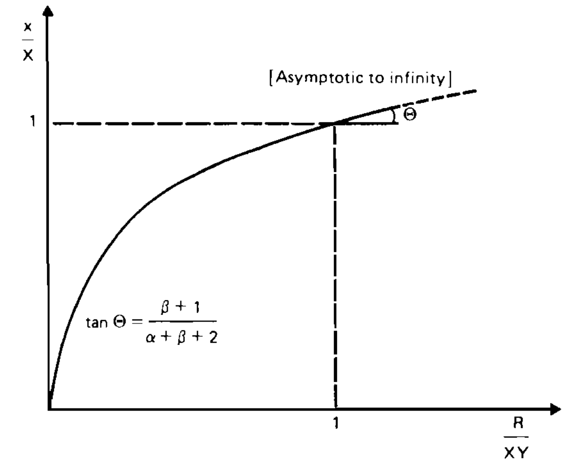

FIGURE 4 (x/X) as a function of (R/XY).

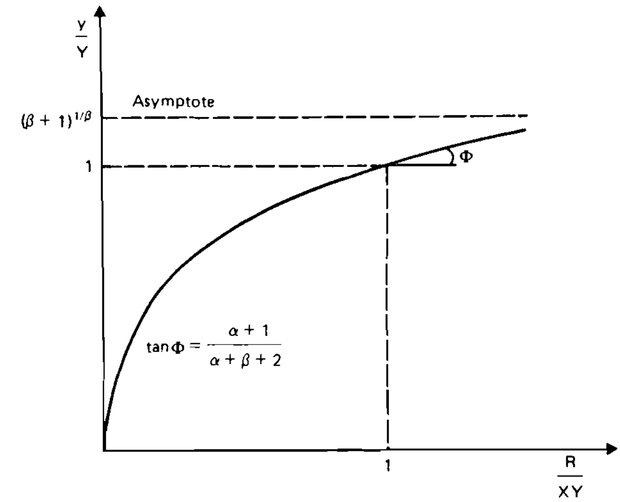

FIGURE 5 (y/Y) as a function of (R/XY).

Alternatively, we can find an equation that relates x and y directly. The result

$$q = \left[\frac{\beta + 1}{\beta p^{-(\alpha+1)} + 1}\right]^{1/\beta} \tag{25}$$

where $p = (x/X)$ and $q = (y/Y)$ is plotted for various ranges of α, β in Figure 6. For $\alpha > \beta - 1$, the curve always has just one point of inflection, and when

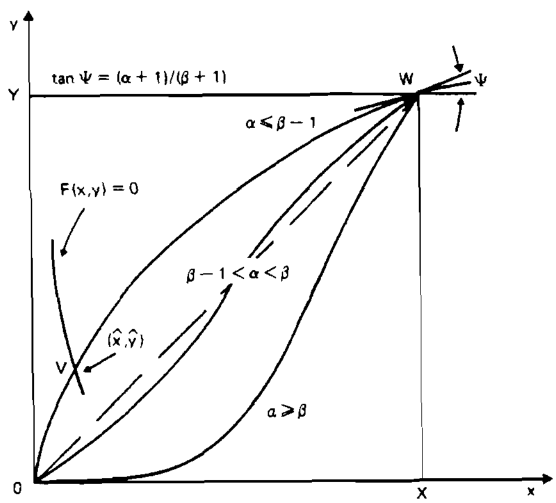

FIGURE 6 Loci of possible solutions on the x–y plane.

$\beta - 1 < \alpha < \beta$, there is just one intersection with the diagonal. From Eq. (25)

$$\frac{\mathrm{d}q}{\mathrm{d}p} = \left[\frac{(\alpha+1)(\beta+1)}{p^{\alpha+2}(\beta p^{-(\alpha+1)}+1)^{\beta+1}}\right]^{1/\beta} \geqslant 0$$

whence we deduce that two data points (p_1, q_1), (p_2, q_2) can be solutions of DRAM only if

$$p_2 > p_1 \Longleftrightarrow q_2 > q_1 \tag{26}$$

This is a fundamental condition on admissible solutions, which we assume for the rest of this section. It means, for example, that the model cannot reproduce increasing available hospital beds and decreasing lengths of stay, simultaneously. (How this condition should be modified when there are two or more resources, perhaps some increasing and others decreasing, is not clear.)

Equation (25) is the locus of solutions of DRAM on the x–y plane. The particular solution for a given resource level is found at the intersection of the locus and the resource hyperbola $F(x, y) = R - xy = 0$, and it is the point on the hyperbola that maximizes the function of Eq. (18). Figure 7 depicts the shape of this function above the x–y plane. We see that

1. $U(X, Y) = 0$, and $x \to 0$ or $y \to 0$ implies $U(x, y) \to -\infty$. Within the constraints $0 < x < X$, $0 < y < Y$, $U(x, y)$ is always negative and concave.
2. $U(x, Y) = g(x)$ and $U(X, y) = h(y)$. Above the point (X, Y) the surface has gradients

$$\left.\frac{\partial U}{\partial x}\right|_{X,Y} = CY \qquad \left.\frac{\partial U}{\partial y}\right|_{X,Y} = C$$

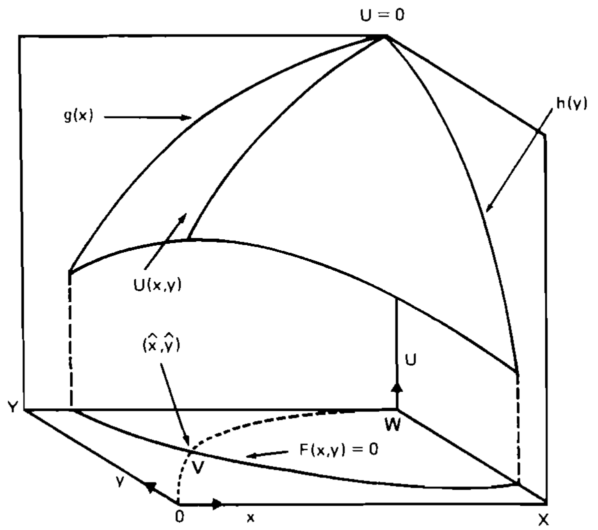

FIGURE 7 The surface $U(x,y)$ above the x–y plane.

3. There is always a unique solution $(\hat{x}, \hat{y})$ because constant-U contours are always more concave than constant-F curves.
4. Equation (25) is represented by the line 0VW.

Evidently, it is not always possible to choose parameters X, Y, α, β that will cause the solution locus 0VW to pass through an arbitrary set of data points (x_i, y_i), $i = 1, 2, \ldots, N$. In the rest of this section we investigate the conditions that allow this.

3.2 *Conditions for Fitting X, Y to Two Data Points*

It seems reasonable that at least two points on Figure 6 are needed to specify a solution locus defined by two parameters, although not all such data will be sufficient or consistent. In this section, we assume that α, β are given, together with two data points (x_1, y_1), (x_2, y_2) such that $x_1 > x_2$, $y_1 > y_2$. Can we choose X, Y such that DRAM can reproduce these points? By substituting the two points into Eq. (25), we easily obtain

$$X = x_1 \beta^{-1/(\alpha+1)} \left[\frac{(y_1/y_2)^{\beta} - 1}{(x_1/x_2)^{\alpha+1} - (y_1/y_2)^{\beta}} \right]^{1/(\alpha+1)} \tag{27}$$

$$Y = y_1 (\beta + 1)^{-1/\beta} \left[\frac{(x_1/x_2)^{\alpha+1} - 1}{(x_1/x_2)^{\alpha+1} - (y_1/y_2)^{\beta}} \right]^{1/\beta} \tag{28}$$

The two numerator terms are always positive, and the denominator term is positive if

$$\frac{\beta}{\alpha + 1} < \frac{\ln x_1 - \ln x_2}{\ln y_1 - \ln y_2}$$

This is, therefore, a necessary and sufficient condition for being able to choose X, Y to fit two data points.

3.3 *Conditions for Fitting α, β to Two Data Points*

Alternatively, we can assume that X, Y are given, together with two data points (x_1, y_1), (x_2, y_2) such that $x_1 > x_2, y_1 > y_2$, and ask whether we can choose α, β such that DRAM can reproduce these points. A necessary and sufficient condition for the existence of $\beta > 0$ is easy to find. Writing Eq. (25) as

$$\alpha + 1 = \frac{\zeta(\beta, q)}{-\ln q} \tag{29}$$

where $\zeta(\beta, q) = \ln [(1 + (1/\beta))q^{-\beta} - (1/\beta)]$ we can use the two given data points to eliminate α, giving

$$\zeta(\beta, q_2) - \omega\zeta(\beta, q_1) = 0 \tag{30}$$

where $\omega = (\ln p_2)/(\ln p_1) > 1$. The solution of Eq. (30) is depicted in Figure 8 as the intersection of two curves with known intercepts and asymptotes. There is an intersection for some $\beta > 0$, if $\omega \ln (1 - \ln q_1) > \ln (1 - \ln q_2)$ and $-\omega \ln q_1 < -\ln q_2$, and these conditions can be combined as

$$\tau < \ln y_1 - \ln y_2 < \left(1 + \frac{\tau}{\omega - 1}\right)^{\omega} - \left(1 + \frac{\tau}{\omega - 1}\right) \tag{31}$$

where $\tau = (1 - \omega) \ln q_1$.

A necessary and sufficient condition for the existence of $\alpha > 0$ also comes from Eq. (29). We require that $\zeta(\beta, q_i)/(-\ln p_i) > 1, i = 1, 2$. Unfortunately, it is not easy to remove the dependence on β in this condition. But the two inequalities $\zeta(\beta, q) \geqslant \ln (1 - \ln q)$ and $\zeta(\beta, q) > -\beta \ln q$ lead to two alternative sufficient (but not necessary) conditions

$$q_i < \exp\left(1 - \frac{1}{p_i}\right) \tag{32}$$

$$q_2 < p_2^{1/\beta} \tag{33}$$

where we have used the fact that the second condition is stronger for $i = 2$. We can find a lower bound on β by inspecting the intercepts and asymptotes in Figure 8

$$\ln (1 - \ln q_2) + \beta_{\min}(-\ln q_2) = \omega \ln (1 - \ln q_1)$$

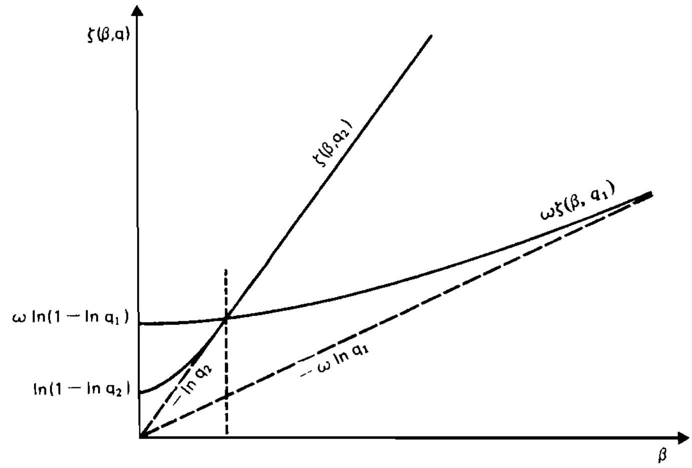

FIGURE 8 Solution of Eq. (30).

When this is used in Eq. (31), the two sufficient conditions become

$$q_i < \exp\left(1 - \frac{1}{p_i}\right)$$

$$\frac{\ln(1 - \ln q_2)}{\ln p_2} - \frac{\ln(1 - \ln q_1)}{\ln p_1} > 1 \tag{34}$$

Empirical evidence suggests that the first condition (32) is less restrictive, at least for small values of (x_1/x_2), and hence closer to being necessary.

These results suggest the following question. Given four points on Figure 6, can we align the solution locus through them all? In other words, given four data points (x_i, y_i), $i = 1, 2, 3, 4$, with $x_1 > x_2 > x_3 > x_4$ and $y_1 > y_2 > y_3 > y_4$, can we choose the four parameters X, Y, α, β such that DRAM can reproduce these points? Sufficient conditions for this, together with an iterative procedure for finding the best fit, are developed in Appendix C. The important conclusion is that even when we have the same number of data points as unknown parameters, and even if the data points satisfy the fundamental condition (26), a perfect fit of the model to the data is not always possible.

3.4 *Conditions for Fitting X, Y or α, β to One Data Point*

In Section 4, we use many data points to estimate pairs of parameters (e.g., X, Y) by combining the estimates suggested by individual data points. We would expect the conditions for fitting two parameters to one data point to be weaker than the conditions derived in Section 3.2 for fitting two parameters to two data points. But is one data point more or less than sufficient to determine two parameters?

In fact, when α, β are given, it is possible to choose an infinite number of pairs X, Y to fit a single data point. Equation (25) shows that for any choice of $X \geqslant x$, there exists some consistent value of $Y \geqslant y$. Similarly, when X, Y are given, it is possible to choose an infinite number of pairs α, β that satisfy Eq. (25) and that therefore fit a single data point. There is, however, a restriction on the minimum possible values of α, β:

$$\alpha_{\min} = \frac{\ln(1 - \ln q)}{-\ln p} - 1, \qquad \zeta(\beta_{\min}, q) = -\ln p$$

which can both be zero, only if $p(1 - \ln q) = 1$.

3.5 *Conditions for Fitting X, Y, α, β to Two Data Points*

Although we do not need the result later, it is interesting to extend and conclude this analysis by asking whether all four parameters can be chosen to fit just two data points (x_i, y_i), $i = 1, 2$; $x_1 > x_2$; $y_1 > y_2$. We analyze this problem in two stages. First, can we choose X, Y so as to satisfy Eq. (31), the necessary condition for the existence of $\beta > 0$? Second, can we also satisfy Eq. (32) or Eq. (34), the sufficient conditions for the existence of $\alpha > 0$?

In order to show that we can always choose X, Y consistent with a $\beta > 0$, we let $\omega \to \infty$ in Eq. (31) giving

$$\tau < \ln(y_1/y_2) < \exp(\tau) - 1 \tag{35}$$

which can always be satisfied by some $\tau > 0$. In practice, ω can be made sufficiently large by setting X close to x_1, and the choice of τ then determines Y.

In order to apply a similar procedure to the sufficient conditions (32) and (34), we write them in the forms

$$\left(\frac{x_1}{x_2}\right)^{\omega/(\omega-1)} - 1 - \frac{\tau}{\omega - 1} < \ln y_1 - \ln y_2$$

$$\ln y_1 - \ln y_2 < \left(1 + \frac{\tau}{\omega - 1}\right)^{\omega} \left(\frac{x_2}{x_1}\right)^{\omega/(\omega-1)} - \left(1 + \frac{\tau}{\omega - 1}\right)$$

where we have set $i = 2$ in Eq. (32). Arguing as earlier that we can choose X to make ω arbitrarily large, we let $\omega \to \infty$ in these equations

$$\left(\frac{x_1}{x_2}\right) - 1 < \ln y_1 - \ln y_2 \tag{36}$$

$$\ln y_1 - \ln y_2 < \left(\frac{x_2}{x_1}\right) \exp(\tau) - 1 \tag{37}$$

Combining Eqs. (35), (36), and (37), we have the sufficient condition

$$\frac{x_1}{x_2} < \min\left\{1 + \ln\left(\frac{y_1}{y_2}\right), \ \left(\frac{y_1}{y_2}\right)\left[1 + \ln\left(\frac{y_1}{y_2}\right)\right]^{-1}\right\} \tag{38}$$

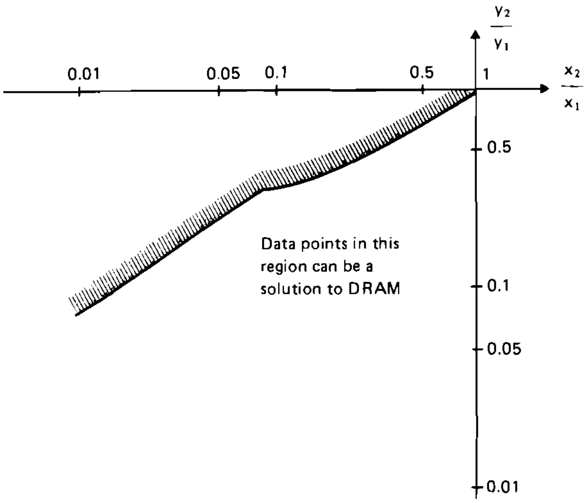

FIGURE 9 Sufficient condition for finding X, Y, α, β consistent with two data points (logarithmic scales).

Figure 9 shows the region of $(x_2/x_1, y_2/y_1)$ in which a consistent choice of the four parameters X, Y, α, β is always possible.

This analysis shows that two arbitrary data points, even when they satisfy condition (26), may not be consistent with any choice of DRAM parameters X, Y, α, β. It suggests, therefore, that simple procedures that estimate parameters from just two data points (Hughes 1978a) may be unsuccessful. For this reason, we turn to more general methods for parameter estimation.

4 ESTIMATION OF PARAMETERS

We turn to the problem of calibrating the model, that is, of estimating parameters for DRAM appropriate for a given region and policy question. Section 4.1 reviews sources of data. Sections 4.2 and 4.3 then describe separate procedures for estimating the two pairs of parameters X, Y and α, β, which are drawn together in Section 4.4. These procedures are quite suitable for small examples and they are illustrated in Section 5. Section 4.5 outlines an alternative approach to parameter estimation that incorporates specific assumptions about the uncertainty of model predictions. It shows that, with certain approximations, the approach is feasible and worth testing. Section 4.6 concludes by briefly mentioning the problems of estimating resource costs.

4.1 *Parameters and Data*

The parameters of the model fall into three groups:

- The *ideal levels* X, Y at which patients would be admitted and receive resources, if there were no constraints on resource availability. Absolute values of these parameters have little meaning, but relative values can be chosen to indicate the relative "needs" for health care.
- The *power parameters* α, β which reflect the elasticities of the actual levels to changes in resource supply. For example, we expect the elasticity of admission rate to bed availability to be less for appendicitis patients than for bronchitis patients, because appendicitis usually requires faster attention.
- The *relative costs* C of different resources. DRAM uses the marginal unit cost of a bed-day, a doctor-hour, and so on, or equivalent parameters, in order to choose between alternative mixes of these resources. We defer discussion of resource cost estimation until Section 4.6.

The level of available resources is not regarded as a model parameter but as an experimental variable. DRAM shows how the levels of satisfied demand vary with changes in resource supply.

There are more data available to estimate X, Y, α, β than there are for many other problems in HCS modeling. The sources include:

- Other models
- Special surveys
- Professional opinions
- Routine statistics

At IIASA, *other models* have been developed for other components of the HCS, and particularly for the estimation of true morbidity from degenerative diseases (Kaihara *et al.* 1977) and infectious diseases (Fujimasa *et al.* 1978). Later at IIASA, these outputs may be useful for setting the ideal rates at which patients in different categories need care. Initially, however, we wish to test and use DRAM independently of other models. Many researchers have performed important and useful *special surveys*. Among others, Newhouse and Phelps (1974) and Feldstein (1967) have estimated both elasticities in hospital care and the costs of acute services, and some of these results were used to calibrate a Mark 1 version of DRAM (Gibbs 1978b). Unfortunately, these results may not be relevant in other regions or countries, or at other times. In an international setting it is necessary to avoid relying on results related to a specific health system.

The *professional opinions* of doctors and health planners can be useful for setting ideal levels of care. Countries where there is a high degree of central planning often set normative figures for ideal hospitalization rates and necessary standards of care, and these can be used in DRAM. However, these are not available in all countries, and probably no professional should be asked to estimate elasticities, in case he supplies his own rather than those of the HCS. This leaves *routine statistics*. Most systems keep regular records on the use and costs of their services, and on how they have allocated resources in the past. If DRAM is a valid model of the HCS, then these figures are typical outputs of the model, which we should be able to use for model calibration.

The aim of DRAM is to model how the HCS reacts to change. Generally, therefore, DRAM's model parameters must be estimated from data that show how an unchanging HCS reacts to external changes, either in space or time. *Cross-sectional* data from subregions of the region of interest may show the HCS operating at different resource levels. So also may *longitudinal* data collected at different times. In both cases, however, the underlying system may be different for the different data. Subregions are often deliberately defined so as to be predominately urban or predominately rural, and we must consider ways of averaging the results across the region. Data collected at different times are highly likely to be affected by historic trends in medicine or management. Ideally, we should model these trends and incorporate the time-varying parameters in a time-dependent model. More probably, we shall use data from a period during which we can assume time variations to be small. The resulting model will still be good for representing those aspects of resource allocation behavior that are independent of time trends. A final and obvious problem is that the available data may be incomplete, either because of recording failures or because the data is insufficiently disaggregated.

Not all of these problems can be overcome simultaneously. But in the next three sections we concentrate on estimation methods that are based on routine statistics about current or past allocation behavior, and that take into account that cross-sectional and longitudinal data may reflect inherent parameter variations. In addition, one of the procedures can be used with incomplete data.

4.2 *Estimation of X, Y*

We consider first the estimation of the ideal service levels X and the ideal supply levels Y, assuming for the moment that the power parameters α, β are known.

Sufficient information to estimate X, Y is given by the current allocation of resources in the region under study. If the current allocation pattern is described by x and y, Eqs. (7) and (8) may be rearranged as

$$X_{jk} = x_{jk}(\mu_{jk})^{1/(\alpha_j+1)} \qquad \forall j, k \tag{39}$$

$$Y_{jkl} = y_{jkl}\,(\lambda_l/C_l)^{\,1/(\beta_{jkl}+1)} \qquad \forall j, k, l \tag{40}$$

which are expressions for X and Y. We have a single equation for each unknown parameter, but as Section 3.4 predicted, we still need some external criterion to determine λ. If we assume that we can define the resources needed to satisfy the ideal levels X_{jk}, Y_{jkl} as some multiple θ_l of the resources used currently

$$\sum_j \sum_k X_{jk} Y_{jkl} = \theta_l \sum_j \sum_k x_{jk} y_{jkl} \qquad \forall l \tag{41}$$

then (39) and (40) can be substituted into (41) to give

$$f_l(\lambda) = 0 \qquad \forall l \tag{42}$$

where

$$f_l(\lambda) = -\theta_l \sum_j \sum_k x_{jk} y_{jkl} + \sum_j \sum_k x_{jk} y_{jkl} (\lambda_l)^{1/(\beta_{jkl}+1)} (\mu_{jk})^{1/(\alpha_j+1)} \tag{43}$$

and where Eq. (42) must be solved for λ. The equations in f are very similar to the equation $\hat{H}_\lambda = 0$ that arises during model solution, and, provided that $\theta_l > 1$, $\forall l$, and that all the terms except λ are known, they may be solved in the same way to give λ. Unfortunately, not all the terms are known. In particular, μ_{jk} is a weighted average involving the terms Y_{jkl}, which are as yet unknown. It is therefore necessary to iterate between solving Eq. (42) for λ, and Eqs. (39) and (40) for X, Y.

This approach suffers from the disadvantage that it only finds values of X, Y that are consistent with the current allocation pattern and the assumed values for α, β. A model with parameters estimated on so little data may have little predictive power. More useful is to estimate X, Y from other data and then to use the current allocation as a test of the model's validity. Other suitable data include cross-sectional and longitudinal data, and given N data points from such sources, we can use Eqs. (39) and (40) to find N estimates of X, Y. The problem remains of how to combine these estimates.

Estimates $X_{jk}(i)$, $Y_{jkl}(i)$ derived for subregions $i = 1, \ldots, N$, may be combined rather easily. If the population of subregion i is $P(i)$, then $X_{jk}(i)P(i)$ is the number of individuals in category j in mode of care k who need treatment in subregion i (per year), and $X_{jk}(i)Y_{jkl}(i)P(i)$ is the number of resources l needed to treat these individuals (per year). These quantities may be summed across the region, and the corresponding *regional* estimates of X, Y are

$$\bar{X}_{jk} = \sum_i X_{jk}(i)P(i) \Big/ \sum_i P(i) \qquad \forall j, k$$

$$\bar{Y}_{jkl} = \sum_i X_{jk}(i) Y_{jkl}(i) P(i) \Big/ \sum_i X_{jk}(i) P(i) \qquad \forall j, k, l$$

This approach (also depicted in Figure 10) is interesting because we do not need to assume that X, Y are constant across the region. The subregional variations are averaged by summing the ideal demands across the region.

Estimates $X_{jk}(i)$, $Y_{jkl}(i)$ derived at different times $i = 1, \ldots, N$ are more difficult to combine. Ideal supply levels Y_{jkl} are probably decreasing with time, and an exponential curve could be fitted to a long sequence of points. The ideal numbers of patients needing care per head of population, $Z_j = \Sigma_k X_{jk}$, $\forall j$, will change because of changes in the age structure and in the morbidity rates. We can correct for the former, but the latter are affected by changes in doctors'

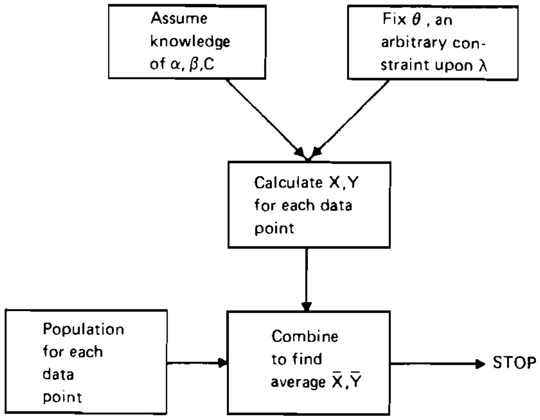

FIGURE 10 Estimation of ideal levels.

preferences between modes of health care. These are reflected in the values of X_{jk}, which could, if necessary, be regarded as experimental variables.

4.3 *Estimation of* α, β

We now consider how to estimate the power parameters α, β, assuming for the moment that the ideal levels X, Y are known. Sufficient information to estimate α, β is given by the current allocation of resources in the region under study. If the current allocation pattern is described by x and y, Eqs. (7) and (8) may be rearranged as

$$\alpha_j = \frac{\ln(\mu_{jk})}{\ln(X_{jk}/x_{jk})} - 1 \qquad \forall j, k \tag{44}$$

$$\beta_{jkl} = \frac{\ln(\lambda_l)}{\ln(Y_{jkl}/y_{jkl})} - 1 \qquad \forall j, k, l \tag{45}$$

which are expressions for α and β. As in Section 4.2, λ must be determined externally. We know, however, that α and β are always positive. This implies then that

$$\lambda_l > \tilde{\lambda}_l = \max_{j,k} \left[\left(\frac{X_{jk}}{x_{jk}}\right), \left(\frac{Y_{jkl}}{y_{jkl}}\right)\right] \qquad \forall l$$

and we can conveniently define λ_l as some (small) multiple $\phi_l > 1$ of the minimum value $\tilde{\lambda}_l$

$$\lambda_l = \phi_l \tilde{\lambda}_l \qquad \forall l \tag{46}$$

A second problem is that Eq. (44) gives K values for each α_j. Generally, these will be different values, but we can overcome this by aggregating the data across modes, and by using Eqs. (44) and (45) for one super mode.

By these means, we may estimate values for the parameters α, β. The model so calibrated will not exactly reproduce the current allocation of resources unless the latter is one of the admissible solutions of DRAM defined in Section 3. However, it will reproduce the actual supply levels y_{jkl}, and the actual numbers of patients in each category $(x_{j1} + x_{j2} + \ldots + x_{jK})$. Whether the estimated elasticities are useful for forward prediction will depend upon whether the current allocation pattern is representative of the HCS's usual behavior. The procedure described above only finds values for α, β that are consistent with this assumption and with the values assumed for X, Y.

A more sophisticated approach is to use more data by estimating *empirical elasticities*. These can then be used to derive the power parameters α, β. Appropriate empirical elasticities for DRAM are γ_{jkl}, the elasticity of the service level x_{jk} to changes in the resource level R_l, and η_{jkml}, the elasticity of the supply level y_{jkm} to changes in the resource level R_l. They can be predicted for given resource levels by DRAM. For example, γ_{jkl} is defined as

$$\gamma_{jkl} = \frac{\partial \ln x_{jk}}{\partial \ln R_l} = \frac{\partial \ln x_{jk}}{\partial \mu_{jk}} \frac{\partial \mu_{jk}}{\partial R_l} R_l$$

We use Eq. (7) to get an expression for $\partial \ln x_{jk}/\partial \mu_{jk}$. Thus,

$$\gamma_{jkl} = \frac{-R_l}{(\alpha_j + 1)\mu_{jk}} \frac{\partial \mu_{jk}}{\partial R_l} \tag{47}$$

Similarly,

$$\eta_{jkml} = \frac{-R_l}{(\beta_{jkl} + 1)\lambda_m} \frac{\partial \lambda_m}{\partial R_l} \tag{48}$$

where

$$\frac{\partial \mu_{jk}}{\partial R_l} = \sum_m \frac{\partial \mu_{jk}}{\partial \lambda_m} \frac{\partial \lambda_m}{\partial R_l}$$

and the derivatives $\partial R_l/\partial \lambda_m = \partial^2 H/\partial \lambda_l \partial \lambda_m$ are given by Eq. (13). Equations (47) and (48) can be written as

$$\alpha_j = \frac{A_{jkl}}{\gamma_{jkl}} - 1 \tag{49}$$

$$\beta_{jkm} = \frac{B_{ml}}{\eta_{jkml}} - 1 \tag{50}$$

where

$$A_{jkl} = \frac{-R_l}{\mu_{jk}} \sum_m \left(\frac{\partial \mu_{jk}}{\partial \lambda_m}\right) \bar{H}_{ml} \tag{51}$$

$$B_{ml} = \frac{-R_l}{\lambda_m} \bar{H}_{ml} \tag{52}$$

and where $\bar{H}_{ml}$ is element ml of the inverted Hessian matrix. However, solution for α, β is still hard. First, this is because A and B are functions of α and β, and iterative solution is necessary. Second, λ must still be chosen externally, and the empirical elasticities must be consistent with the choice of λ, otherwise the procedure may not converge (Gibbs 1978b). Third, there are more empirical elasticities γ, η than there are power parameters α, β. Therefore, unless some of the empirical elasticities are ignored, the parameters will be overspecified. Fourth, the empirical elasticities γ, η, are not directly measurable and are usually the result of some prior data analysis.

Some of these difficulties can be avoided by incorporating the prior data analysis within the solution of Eqs. (49)–(52). For example, estimates $\hat{\gamma}$, $\hat{\eta}$ are found by assuming that some N known data points $x_{jk}(i), y_{jkl}(i), R_l(i), i = 1, \ldots, N$, satisfy the linear models

$$\ln x_{jk}(i) = a_{jk}^x + \sum_l \gamma_{jkl} \ln [R_l(i)] + \epsilon_{jk}^x(i) \tag{53}$$

$$\ln y_{jkm}(i) = a_{jkm}^y + \sum_l \eta_{jkml} \ln [R_l(i)] + \epsilon_{jkm}^y(i) \tag{54}$$

in which a^x, a^y are unknown constants, and ϵ^x, ϵ^y are random, uncorrelated error terms with zero means. If we eliminate γ, η by combining Eqs. (49), (50), (53), and (54) to give

$$\ln x_{jk}(i) = a_{jk}^x + \left(\frac{1}{\alpha_j + 1}\right) \sum_l A_{jkl} \ln [R_l(i)] + \epsilon_{jk}^x(i) \qquad \forall\, j, k, i \tag{55}$$

$$\ln y_{jkm}(i) = a_{jkm}^y + \left(\frac{1}{\beta_{jkm} + 1}\right) \sum_l B_{ml} \ln [R_l(i)] + \epsilon_{jkm}^y(i) \qquad \forall\, j, k, m, i \tag{56}$$

we can use the following iterative scheme in order to estimate α and β.

1. Fix λ arbitrarily for some resource level R, perhaps by using Eq. (46) on one of the data points.
2. Assume some initial estimates of α, β (e.g., unity).
3. Derive μ from Eqs. (9) and (10), $\hat{H}_{\lambda\lambda}$ from Eq. (13), and A, B from Eqs. (51) and (52).
4. Find the best least-squares estimators of $(\alpha_j + 1)^{-1}$, $(\beta_{jkm} + 1)^{-1}$ in Eqs. (55) and (56).
5. Hence, estimate α, β and repeat from step 3.

This procedure (also depicted in Figure 11) is likely to be lengthy because it incorporates regression estimation at each iteration. Nor can we ensure the positive estimates of α, β that are necessary for convergence. On the other hand, it has the advantage that more of the original data can be used directly. If a full data set

$$\{x_{jk}(i), y_{jkl}(i), R_l(i);\quad i = 1, \ldots, N, j = 1, \ldots, J$$
$$k = 1, \ldots, K, l = 1, \ldots, L\}$$

is available, KN equations are available to estimate each α_j, and perhaps not all of the $x_{jk}(i)$ need be known. Fewer equations (just N) are available to estimate each β_{jkl}, and it may be necessary to introduce some further simplifying

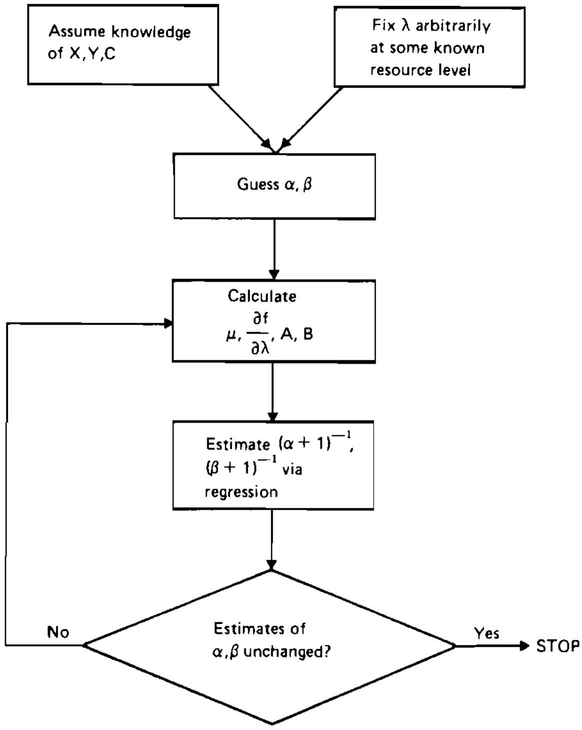

FIGURE 11 Estimation of power parameters.

assumptions such as $\beta_{j_1kl} = \beta_{j_2kl}$, $\forall j_1, j_2 \in \{1, \ldots, J\}$, in order to obtain reliable estimates. A second advantage of this procedure is that it is not necessary to modify any of the input data to make them consistent with the model. A third advantage is that the parameter estimated in each regression has an estimated standard error associated with it. These errors provide a measure of the reliability of α, β.

Perhaps the main assumption in the above analysis is that the underlying elasticities are constant across the set of data points. Because there is little information about how elasticities are likely to vary in time or space, we have not attempted to model this variation here. But Appendix D shows that in a certain sense, the procedure described above gives *unbiased* estimates. This is a reassuring result, and the estimates can be further tested to see if the model so calibrated can reproduce the current allocation of resources.

4.4 *Estimation of α, β and X, Y*

In the most general case, neither of the parameter pairs X, Y or α, β is known, and we require estimates of both. In this circumstance, the two procedures described above may be used together in the following scheme.

1. With some arbitrary initial estimates of α, β, use the methods of Section 4.2 to estimate X, Y.
2. With these estimates of X, Y, use the methods of Section 4.3 to estimate α, β.
3. Repeat from step 1.

The analysis in Section 3 showed that not even all small data sets can be consistent with DRAM, so that convergence of this scheme cannot be guaranteed. For this reason, although we have implemented on the computer the procedures for estimating both X, Y and α, β, we prefer not to link these programs together, but rather to use them alternately to obtain consistent pairs of estimates. (Note however that when neither parameter pair is given exogenously, the same data cannot be used to estimate both pairs of parameters.)

The parameter estimation procedures described above involve the choice of additional constraint variables such as ϕ and θ. Fortunately, however, this is not a problem. Although different values of ϕ, θ lead to different values for α, β, X, Y, each set of parameter values will reproduce with similar accuracy the data points used for estimation. Provided that predictive runs of the model do not involve resource levels very different from those used in estimation, the results will be relatively insensitive to ϕ, θ. Section 5 illustrates how these procedures were used to estimate model parameters in two examples.

4.5 *An Alternative Approach*

We now describe an alternative approach to parameter estimation that takes into account that DRAM's predictions are subject to uncertainty, and that incorporates this uncertainty mathematically. It is not fully implemented or tested, but the preliminary analysis given below is encouraging.

We consider how to use historical resource allocations $x(i)$, $y(i)$, $i = 1, \ldots, N$ in order to estimate the model parameter set $P = \{X, Y, \alpha, \beta\}$. As mentioned in Section 4.1, these are not the only data available. Nor does P include all the parameters: we have omitted the resource costs C because they seem to be more naturally estimated from external studies of financial or related statistics. Nevertheless, procedures to estimate these parameters from these data would be useful.

If reality conformed exactly to DRAM, we would expect the historical allocations $x(i)$, $y(i)$ to be exactly those $\hat{x}(i)$, $\hat{y}(i)$ prescribed by DRAM for the historical resource levels. These solutions are the result of (constrained) maximization over x and y of a function $U(x, y, P, C, R)$ that depends also upon the parameters P, the costs C, and the resource levels R. This function is known, and is presumably also maximized by choosing the correct parameters

$$\max_{\hat{P}\text{ given past }\{x, y, R, C\}} U(x, y, P, R, C)$$

because with wrong parameters, it would be maximized by different values of x, y.

However, DRAM is only a model of reality. The historical allocations are related to the model predictions by equations like $x(i) = \hat{x}(i) + \xi_1(i)$ and $y(i) = \hat{y}(i) + \xi_2(i)$ where $\xi_1(i)$, $\xi_2(i)$ are stochastic processes with statistics S that need to be specified. Such a specification would be quite complicated. The probability distributions involved in S depend upon the reasons why the assumptions in DRAM are not perfect, the reasons that influence actual decisions, and the reasons that give rise to inaccurate data. But if such a specification were possible, the parameter set P could be estimated through

$$\max_{P} \underset{\substack{\text{with respect to} \\ \hat{x},\, \hat{y} \text{ given } x,\, y,\, S}}{\text{conditional expectation}} U(\hat{x}, \hat{y}, P, R, C) \tag{57}$$

where

$$U(\hat{x}, \hat{y}, P, R, C) = \max_{\substack{x, y \text{ given} \\ \text{past } \{R, C\}}} U(x, y, P, R, C) \tag{58}$$

Such a calculation would also be quite complicated, however, because the integral involved in the conditional expectation is unlikely to be analytic. In short, the ideal estimation procedure is extremely difficult both to formulate and solve. It does, however, suggest a more practical approach.

If the function $U(\hat{x}, \hat{y}, P, R, C)$ in Eq. (58) were twice differentiable in $\hat{x}$, $\hat{y}$, it could be expanded as a Taylor series about the point x, y, with terms in the prediction errors $(\hat{x} - x)$, $(\hat{y} - y)$. If, in addition, S were such that EXPECTATION $\xi_1(i) =$ EXPECTATION $\xi_2(i) = 0$, term-by-term expansion of this series would eliminate all first-order terms, causing the dominant terms of the series to be the squares and cross-products of the prediction errors. Whereas this is hardly a feasible way to solve (57), it suggests the idea of formulating the parameter estimation problem as the minimization of a function of the squared prediction errors

$$\min_{P} J(P) \tag{59}$$

where

$$J(P) = \frac{1}{2} \sum_{ijk} \rho^x_{ijk} [\hat{x}_{jk}(i) - x_{jk}(i)]^2 + \frac{1}{2} \sum_{ijkl} \rho^y_{ijkl} [\hat{y}_{jkl}(i) - y_{jkl}(i)]^2 \tag{60}$$

in which

$\hat{x}(i)$, $\hat{y}(i)$ are the optimal model allocations for assumed P and known past $R(i), C(i), i = 1, \ldots, N$

$x(i)$, $y(i)$ are the observed historical resource allocations for known past $R(i), i = 1, \ldots, N$

ρ^x_{ijk}, ρ^y_{ijkl} are weighting coefficients to be specified later.

DRAM's most useful feature is that the solutions $\hat{x}, \hat{y}$ are analytic functions of the parameters P. This means that we can calculate the gradient vector and Hessian matrix of $J(P)$, opening the way for powerful techniques for solving (59). The gradient vector is

$$\frac{\partial J(P)}{\partial P} = \sum_{ijk} \rho^x_{ijk}[\hat{x}_{jk}(i) - x_{jk}(i)] \frac{\partial \hat{x}_{jk}(i)}{\partial P} + \sum_{ijkl} \rho^y_{ijkl}[\hat{y}_{jkl}(i) - y_{jkl}(i)] \frac{\partial \hat{y}_{jkl}(i)}{\partial P} \quad (61)$$

and the Hessian matrix is

$$\frac{\partial^2 J(P)}{\partial P' \partial P} = \sum_{ijk} \rho^x_{ijk} \left\{ [\hat{x}_{jk}(i) - x_{jk}(i)] \frac{\partial^2 \hat{x}_{jk}(i)}{\partial P' \partial P} + \frac{\partial \hat{x}_{jk}(i)}{\partial P'} \frac{\partial \hat{x}_{jk}(i)}{\partial P} \right\}$$

$$+ \sum_{ijkl} \rho^y_{ijkl} \left\{ [\hat{y}_{jkl}(i) - y_{jkl}(i)] \frac{\partial^2 \hat{y}_{jkl}(i)}{\partial P' \partial P} + \frac{\partial \hat{y}_{jkl}(i)}{\partial P'} \frac{\partial \hat{y}_{jkl}(i)}{\partial P} \right\} \quad (62)$$

$$\simeq \sum_{ijk} \rho^x_{ijk} \frac{\partial \hat{x}_{jk}(i)}{\partial P'} \frac{\partial \hat{x}_{jk}(i)}{\partial P} + \sum_{ijkl} \rho^y_{ijkl} \frac{\partial \hat{y}_{jkl}(i)}{\partial P'} \frac{\partial \hat{y}_{jkl}(i)}{\partial P} \quad (63)$$

if the prediction errors are small. Expressions for the elements in the sensitivity derivative vectors $\partial \hat{x}_{jk}(i)/\partial P$ and $\partial \hat{y}_{jkl}(i)/\partial P$ are evaluated and listed in Appendix E.

The dimension of these vectors, and also of the Hessian matrix, is the same as the number of parameters $(2JKL + JK + J)$ in the parameter set P. Each element in the Hessian matrix is the sum of the $N(JK + JKL)$ terms enumerated in Eq. (63). Renumbering these terms as $m = 1, 2, \ldots, N(JK + JKL)$, we obtain the simpler form

$$\frac{\partial J(P)}{\partial P' \partial P} = \sum_m \rho_m v_m v'_m \quad (64)$$

where ρ are scalars

$$\rho_1 = \rho^x_{111}, \qquad \rho_2 = \rho^x_{112}, \ldots$$

$$\rho_{NJK+1} = \rho^y_{1111}, \qquad \rho_{NJK+2} = \rho^y_{1112}, \ldots$$

and v are vectors

$$v_1 = \frac{\partial \hat{x}_{11}(1)}{\partial P'}, \qquad v_2 = \frac{\partial \hat{x}_{12}(1)}{\partial P'}, \ldots$$

$$v_{NJK+1} = \frac{\partial \hat{y}_{111}(1)}{\partial P'}, \quad v_{NJK+2} = \frac{\partial \hat{y}_{112}(1)}{\partial P'}, \ldots$$

By arguments similar to those in Section 2.2, a matrix such as Eq. (64) is always positive semidefinite, which is useful for search procedures to solve (59).

However, the Hessian matrix will not be positive definite, and such searches will fail, unless the vectors v_m are linearly independent and span the parameter space. Just $2JKL + JK + J$ parameters X_{jk}, Y_{jkl}, α_j, β_{jkl}, $\forall j, k, l$, have to be estimated, and each data point x_{jk}, y_{jkl}, $\forall j, k, l$, provides $JK + JKL$ degrees of freedom that are subject to L resource constraints. Therefore, the number of data points N needed to identify P must satisfy $N(JK + JKL - L) \geqslant 2JKL + JK + J$. When $J = K = L = 1$, N must be four or more, but when $J = K = 3$ and $L = 2$, N can be as small as 2, although more data than this would be needed to achieve reasonable confidence in the estimated parameters.

An attempt to choose parameters P that will minimize $J(P)$ may also fail if the problem is badly conditioned, and specifically if the eigenvalues of

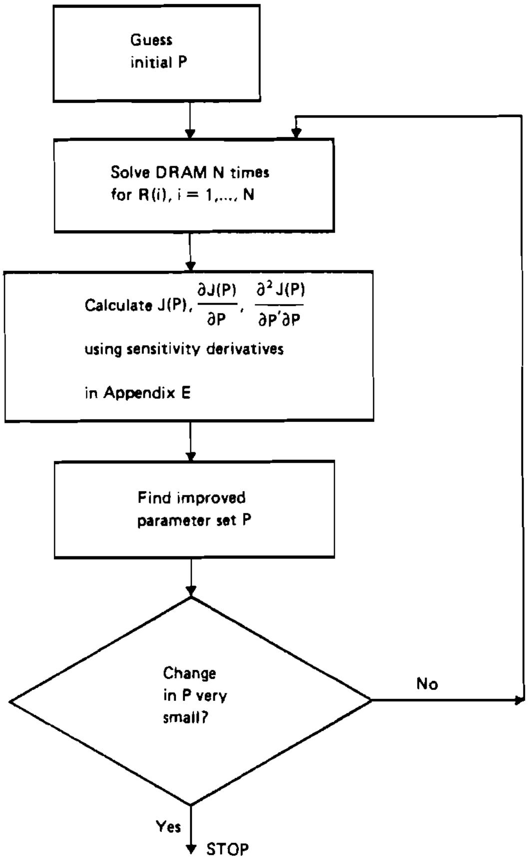

FIGURE 12 Proposed alternative parameter estimation procedure.

$\partial^2 J(P)/\partial P'\partial P$ are very dissimilar. We can control this by appropriate choice of the weights ρ in the definition of $J(P)$ in Eq. (62). Setting $\rho_m = \|v_m\|^{-2}$ is equivalent to normalizing all the vectors in Eq. (64) to unit length. If the vectors are additionally orthogonal, all the eigenvalues would be equal. When they are not orthogonal, the eigenvalues are approximately equal.

Figure 12 shows a way of using these results to estimate the parameter set P by solving (59) according to an iterative procedure. It uses initially some guesses about P to derive the function $J(P)$ in Eq. (60), and then the gradient vector (61) and perhaps the Hessian matrix (62) to find a new parameter set that is closer to the solution of (59).

What computation is involved in this procedure? At each step in the iteration, DRAM must be solved N times to give the model predictions $x(i)$, $y(i)$ corresponding to each of the observed data points $x(i)$, $y(i)$. Probably this procedure is most useful when large amounts of data are available (N at least greater than 20). This means more than 20 DRAM solutions per step, and probably at least 200 DRAM solutions for convergence. But with the typical model solution times reported in Appendix B, this is not too many, especially when each solution also gives the gradient vector and Hessian matrix of $J(P)$.

Potentially, the storage requirements could be excessive. Fortunately, however, all three terms $J(P)$, $\partial J(P)/\partial P$, $\partial^2 J(P)/\partial P'\partial P$ are formed by summation, and the individual terms can be calculated and added sequentially. Appendix E shows that many of the sensitivity derivatives are identically zero, and the remaining derivatives can be computed in logical and space-saving order. The Hessian is symmetrical, permitting further saving. For Example 2 in Section 5, where $J = 7$, $K = 2$, and $L = 2$, the number of locations needed to store these three functions is $1 + (2JKL + JK + J) + \frac{1}{2}(2JKL + JK + J)(2JKL + JK + J + 1) = 1 + 77 + 3003 = 3081$, which is quite reasonable. It remains only to specify how a new parameter set is determined. This problem is similar to that of finding improved estimates of λ in Section 2.4, and similar or more sophisticated gradient methods can easily be devised.

4.6 *Estimation of C*

We now discuss how to estimate the unit resource costs C needed in the model. These parameters are defined rather carefully. Specifically, C_l is the marginal cost of using one more resource of type l, when all needs for health care are met. Strictly speaking, these costs are not money costs but opportunity costs. They reflect the benefit in some alternative that is foregone through buying the extra resource. How then can they be estimated? Often, we have financial data that we can use directly, but when these are unavailable or inappropriate, how can equivalent model parameters be inferred?

Two assumptions will enable us to estimate the costs C from financial data, when these are available. The first assumption is that in long-term planning, opportunity costs are approximately measured by money costs. Given sufficient time, every option is an alternative, and all resources are substitutable. The second assumption is that marginal costs are approximately measured by average costs. The cost function of an individual hospital or medical school is certainly nonlinear, with marginal costs being generally less than average costs. But when many such hospitals or medical schools are operating in a single region, the aggregate cost function may be approximately linear, as shown in Figure 13. In these circumstances, the average costs recorded in historical accounts approximate the marginal costs at some hypothetical resource level.

However, not all countries compare alternative plans in terms of financial feasibility. In the Soviet Union, for example, planning seeks mainly to reconcile the real outputs between producers while satisfying aims such as full employment and constant growth. For application of the model in these countries, it is not necessary to estimate resource costs, but only some parameters that have an equivalent function in the model. The purpose of the C parameters is to

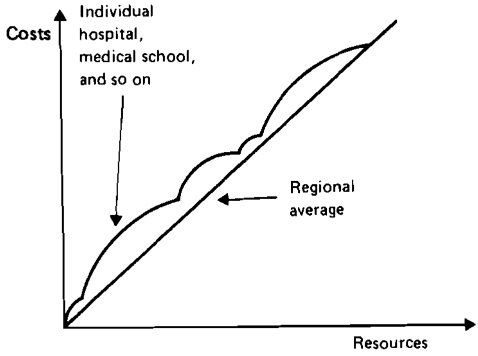

FIGURE 13 A linear regional cost function.

reflect the relative value of different resources; or, conversely, their relative scarcity or the relative difficulty of providing different resources. In a society with uniform and constant growth, resources increase in proportion to their current levels, and these ratios may be adequate first estimates of the C parameters. When different growths are expected in different parts of the HCS, the ratios may be adjusted accordingly, or more detailed analysis may reveal the shadow prices of each constrained resource.

The principal unsolved problem is that of resource definition. The costs of a hospital bed could be the capital cost of creating it, or the cost of maintaining the patient in it with food, heat, and laundry. The cost of a doctor could include his training, his accommodation, or just his salary. The choices made at this stage actually *define* the resources for the purposes of the model, and they depend mainly upon which alternatives are interesting to the users of the model. Finally, of course, we desire to estimate C at some future time instead of at the present. A full treatment of this issue would need and could use more sophisticated predictive models.

5 EXAMPLES

This section contains no mathematics; instead, it illustrates how DRAM can be used. Section 5.1 has some general comments about the use of mathematical models, and Sections 5.2 and 5.3 contain two examples. The first is used mainly to compare different methods of parameter estimation. The second uses the full structure available in DRAM (categories, modes, and resources) in order to investigate questions of resource balance.

5.1 *Application of DRAM*

A mathematical model represents some common mechanism or process. The process that DRAM represents is the distribution of scarce resources within a large and complex system. The value of having such a model is that changes proposed for the system can be tried out first on the model to see what effects they are likely to have. This helps in debates about which changes are best.

Three points need to be noted. First, DRAM is not a model of a complete HCS. Rather, it is a model of resource allocation in such systems. Second, DRAM is not a model of health resource allocation in, for example, Austria. Rather, it is a model for all regions (nations or districts) where its hypotheses are justified. Third, DRAM is not a model with certain specific data needs. Rather, it is a tool that can be calibrated for different problems, large and small.

A large problem might concern the use of all health resources throughout a country. To apply DRAM here (we have not attempted it), we would need a detailed study of the appropriate patient categories and resource groupings. Because DRAM uses generalized variables (e.g., resources) which are not restricted in number or type, as many categories as desired, of whichever type, could be used. A lot of data about past allocation patterns would need to be collected and related to other sources in order to estimate parameters, and methods such as those proposed in Section 4.5 would be useful. Such a model, linked with other models for population, morbidity, and education, would be a tool similar in scope to a large-scale economic model.

A small problem might concern the age distribution of hospital patients. In such an exploratory application, not all the dimensions available in DRAM would be needed. Just three patient categories (young, middle-aged, and old), one resource (beds) and one mode (hospital) might be enough. But if subsequent work suggests that lack of convalescent care is affecting discharges from hospital, then no new structure would be needed to extend the analysis to include this extra mode of care. It might be interesting to use the model with alternative age groupings to see if the results are sensitive to this. Because DRAM is easy to solve, many runs are possible at small expense.

What sort of problems are amenable to investigation with DRAM? The most obvious ones are questions about the consequences of changing levels of resources. When all resources are increasing, DRAM probably has little to say about who gets what, beyond what could be deduced directly from the empirical elasticities of demand to supply. But when some resources are increasing (e.g., numbers of doctors), and others are decreasing (e.g., hospital beds), either by design or through natural trends, such simple deductions become difficult. DRAM recognizes that there will be substitution between resources, and can show where the balance will lie.

Slightly different questions arise when resource levels are constant but the behavior of the HCS is changing. Morbidity levels (the X parameters in DRAM) change with population age structure. Ideal standards of care (the Y parameters in DRAM) change as alternative forms of care become popular. These sorts of assumptions lead to model runs that predict what will happen in the future to a single sector (e.g., care of children), if no change is made in the present HCS.

More unorthodox applications are possible. DRAM is deliberately designed with parameters that can be interpreted outside the model. When ideal standards of care (Y) have been proposed by professional consensus, DRAM is useful for seeing how nearly they can be achieved when resources are scarce. But this approach can also be reversed. The parameter estimation procedures reveal what ideal standards are implied by current behavior, and how these compare with professionally set standards. Such procedures can also be used to estimate the levels of potential demands for care, and thereby make a comparison of underlying morbidity. Effectively, the model is inverted in order to predict inputs from outputs. The examples that follow, however, are rather more straightforward.

5.2 *Example 1: Hospital Beds*

If more hospital beds are provided to increase the numbers of short-stay patients, might the result just be the same number of long-stay patients staying still longer? Because hospital beds are an expensive form of care, this is an important question. To illustrate how DRAM can be used to study it, consider the distribution of acute hospital bed-days between patients suffering from six diseases: varicose veins, hemorrhoids, ischemic heart disease (excluding acute myocardial infarction), pneumonia, bronchitis, and appendicitis. Table 1 gives the numbers of patients admitted to hospitals in England in 1968 and 1973 with these diseases, and their average lengths of stay (Department of Health and Social Security 1972, 1977a). Together, these patients use only about 8 percent of all hospital beds (excluding maternity beds), but an extension of this example to include the remainder, either as a group or individually, would not be difficult. We notice that during these 5 years, the number of bed-days used for these diseases has fallen by about 28 percent. Furthermore, admissions and lengths of stay in each disease category have nearly all fallen. Is it possible to calibrate a model of these changes?

Gibbs (1977, 1978a, b) did this using the empirical elasticities estimated by Feldstein (1967) from 1960 data, and exogenous 1968 estimates of the ideal levels X, Y. The corresponding model parameters, summarized in Table 2, were used to reproduce the 1968 allocations in one region of England (the South Western Regional Health Authority – SWRHA), and to investigate the effects of changing the number of beds available there by 20 percent. The analysis was repeated with X, Y chosen to reproduce regional admission and supply levels.

We have repeated this exercise, applying the parameter estimation methods described in Section 4 to the actual admissions and lengths of stay in the 14 health regions of England in 1968 and 1973 (Department of Health and Social Security 1972, 1977a). Table 3 gives the parameters estimated by using the 1968 figures to estimate α, β and the 1973 figures to estimate X, Y recursively

TABLE 1 Allocation of hospital bed-days in England.

	1968		1973	
	Admissions per 10,000 people	Average stay (days)	Admissions per 10,000 people	Average stay (days)
Varicose veins	9.8	12.0	7.6	10.1
Hemorrhoids	5.6	10.1	4.7	7.8
Ischemic heart	6.5	39.8	8.5	24.9
Pneumonia	14.2	25.4	14.0	18.0
Bronchitis	14.1	25.6	10.8	23.1
Appendicitis	20.4	9.1	17.5	7.9
Total bed-days per 10,000 people	1,340.1		964.8	

SOURCE Department of Health and Social Security (1972, 1977a).

TABLE 2 First set of model parameters for Example 1.

	Empirical elasticities[a]		Model parameters			
	γ	η	α^b	β	X^b	Y
Varicose veins	0.78	0.62	1.64	3.03	12.8	15.4
Hemorrhoids	0.70	0.44	2.11	4.68	7.7	13.1
Ischemic heart	1.14	1.08	0.54	1.31	10.4	52.1
Pneumonia	0.71	0.23	2.28	9.87	21.0	19.7
Bronchitis	1.13	−0.23	1.14	49.00	21.3	34.2
Appendicitis	−0.16	0.31	44.40	7.06	24.8	10.1

[a] Feldstein (1967, p. 219).
[b] α was estimated from γ, η with arbitrary constant $c = 25$, and X was chosen exogenously (Gibbs 1978b).

as described in Section 4.4. For this example, we have assumed that the parameters are constant over time, but we could have incorporated exogenous information to correct for this. We could also have corrected for the effects of changing age structure, but they were small. At some points in the iteration towards the results of Table 3, negative elasticities were estimated, but their associated standard errors were so large that they could reasonably be changed to small positive numbers. If professional opinions about ideal admission rates or lengths of stay had been available to us, we could have incorporated them also within this scheme.

Gibbs (1978b) used data from the SWRHA for model testing, and we have done the same. In 1973, only 633 bed-days per 10,000 people were used for the six diseases and Table 4 shows how they were distributed. Making the

TABLE 3 Second set of model parameters for Example 1.

	Empirical elasticities[a]		Model parameters[b]			
	γ	η	α[c]	β	X	Y
Varicose veins	0.54	0.43	1.68 (0.7)	3.27 (0.5)	31.6	30.9
Hemorrhoids	0.34	0.31	3.63 (0.5)	5.00 (−0.9)	11.2	17.0
Ischemic heart	0.66	0.93	0.50 (0.7)	1.00 (−12)	71.0	247.5
Pneumonia	0.66	0.18	1.57 (0.8)	9.44 (0.5)	75.2	28.1
Bronchitis	0.90	0.04	1.04 (0.8)	50.00 (−4)	102.7	24.9
Appendicitis	0.04	0.14	40.00 (0.3)	12.75 (0.6)	19.5	11.1

[a]Derived from α, β, X, Y and R = 1340.1 bed-days per 10,000 people.
[b]Estimated from 1968 and 1973 allocations across 14 English regions (Department of Health and Social Security 1972, 1977a) with arbitrary constants $\phi = 5$, $\theta = 20$.
[c]Confidence coefficients (in parentheses) are defined as 1 — (estimated standard error ÷ estimated value).

TABLE 4 Allocation of hospital bed-days[a] in 1973 in the South Western Region of England (Example 1).

	Actual[b]		Predicted by model using Table 2 parameters		Predicted by model using Table 3 parameters	
	Admissions per 10,000 people	Average stay (days)	Admissions per 10,000 people	Average stay (days)	Admissions per 10,000 people	Average stay (days)
Varicose veins	6.1	14.4	5.5	8.1	6.1	8.7
Hemorrhoids	4.2	7.7	3.7	8.3	4.1	6.9
Ischemic heart	5.3	17.4	3.0	16.9	6.5	16.5
Pneumonia	11.4	14.4	9.9	15.5	10.7	16.7
Bronchitis	9.9	16.8	6.5	32.5	7.5	22.4
Appendicitis	15.4	7.8	23.5	7.3	17.2	7.5

[a] 663 bed-days available per 10,000 people in 1973.
[b] From Department of Health and Social Security (1977a).

assumption that the model parameters that we have estimated from English data are appropriate to SWRHA, we can use the model to make predictions of this distribution, also shown on Table 4. The parameters from Table 3 give slightly better predictions than those from Table 2; the average error is about

14 percent. Note also that predictions from two sets of model parameters indicate the sensitivities of the model outputs to changes in model parameters. (Appendix E shows that expressions for these sensitivities can also be derived explicitly.) If these parameters are judged acceptable, the model can be used with different bed supply levels to predict the effects of an increase or a decrease in the number of beds. It is important to note that such predictions have little value unless the model is adequately calibrated. It is for this reason that parameter sets estimated from different sources are valuable.

The two sets of model parameters in Tables 2 and 3 vary because of different data and because of different values used for the arbitrary constants. (It is not easy to choose equivalent values when both procedures are solved iteratively.) Nevertheless, they show very similar variations across diseases. Appendicitis is clearly represented as a disease where most patients go to the hospital (high α), and bronchitis appears as a disease afflicting many patients (high X) for whom hospital care is not essential (low α). The empirical elasticities in Table 3 are values derived via the model, Eqs. (47) and (48), using the 1968 English resource level. This calculation incorporates DRAM's behavioral assumption (see Section 1). Because Feldstein's estimates given in Table 2 do not incorporate this assumption, the reasonable agreement between them suggests that DRAM's assumptions are valid, and supports the previous results.

5.3 *Example 2: The Balance of Inpatient and Outpatient Care*

If hospital beds are decreased and medical staff are increased, will more or fewer patients receive treatment and how will the balance of inpatient and outpatient care be affected? This is a question facing health managers in England and elsewhere, and DRAM can be used to help answer it.

Table 5 shows how beds and doctors were used in the SWRHA in England for 1977 in the seven largest acute hospital specialties: general surgery, general medicine, obstetrics and gynecology, trauma and orthopedic (T & O) surgery, ear, nose, and throat (ENT), pediatrics, and ophthalmology (Department of Health and Social Security 1977b). In this example, the patient categories are the seven specialties, the two modes of care are inpatient and outpatient, and the two resources are beds and doctors. Therefore, this example uses all the structure available in DRAM, although it has the simplifying feature that one of the resources (beds) is used in only one mode of care (inpatient).

Because the problem is more complicated than the previous one, formulating a suitable DRAM model is more difficult. For example, hospital specialties are not as precisely defined as disease categories, and the division of doctor's time between inpatients and outpatients is not directly measurable. The first is not so important if the definitions are reasonably consistent across the region. If the definitions are consistent but not universal, comparisons beyond SWRHA may be suspect. The second difficulty can be overcome by subtracting from each consultant's working year (measured in half days), the number of

TABLE 5 Beds and doctors in the South Western Regional Health Authority in 1977 (Example 2).

	Relevant catchment population (thousands)	Admissions per 1,000 people		Average hospital stay (days)[a]	Half-day consultant sessions per admission[b]	
		Inpatient	Outpatient		Inpatient[c]	Outpatient
General surgery[d]	3,035.4	20.9	19.0	7.87	0.170	0.153
General medicine[e]	3,035.4	14.8	10.5	10.18	0.183	0.345
Obstetrics and gynecology	1,563.8[f]	39.5	37.1	5.78	0.072	0.139
T & O surgery	3,035.4	9.1	22.4	13.60	0.252	0.121
ENT	3,035.4	4.4	11.1	4.39	0.346	0.128
Pediatrics	641.8[g]	29.7	17.7	6.28	0.266	0.362
Ophthalmology	3,035.4	2.8	10.3	6.59	0.427	0.214

SOURCE Department of Health and Social Security (1977b).

[a] 892 bed-days available per 1,000 people in 1977.

[b] Assuming each full-time consultant works the equivalent of 450 half-day sessions per year; 46 half-day consultant sessions available per 1,000 people in 1977.

[c] Derived by subtracting actual outpatient sessions from total number of sessions.

[d] Includes urology.

[e] Includes cardiology.

[f] Excludes males.

[g] Excludes people more than 15 years old.

TABLE 6 Estimated model parameters for Example 2.

	α_j	X_{j1} (Inpatient)	X_{j2} (Outpatient)	β_{j11} (IP, beds)	β_{j12} (IP, doctors)	β_{j22} (OP, doctors)	Y_{j11} (IP, beds)	Y_{j12} (IP, doctors)	Y_{j22} (OP, doctors)
General surgery[a]	10.0	26.3	22.2	10.8	6.1	1.0	10.5	0.34	0.46
	(−0.4)			(0.7)	(0.7)	(−3.9)			
General medicine[b]	0.01	217.7	83.3	10.7	2.7	11.2	13.3	0.42	0.41
	(0.5)			(0.3)	(0.7)	(0.8)			
Obstetrics and gynecology	16.5	44.8	38.7	10.3	1.5	0.001	7.7	0.22	1.32
	(0.2)			(0.6)	(0.8)	(0.6)			
T & O surgery	10.0	10.8	26.5	1.0	12.7	10.0	58.5	0.37	0.15
	(−1.4)			(−1.4)	(0.2)	(−10.7)			
ENT	10.0	5.0	12.9	0.001	14.3	20.0	79.1	0.43	0.15
	(−0.7)			(0.2)	(0.0)	(0.7)			
Pediatrics	5.6	43.7	19.4	8.9	5.8	1.0	9.1	0.41	1.28
	(0.7)			(0.4)	(0.7)	(−4.8)			
Ophthalmology	20.0	3.1	11.9	10.0	8.3	10.0	9.4	0.60	0.24
	(0.0)			(−2.8)	(0.3)	(−2.3)			

NOTE: Confidence coefficients as defined in Table 3 appear in parentheses.

[a] Includes urology.

[b] Includes cardiology.

TABLE 7 Validation results for Example 2.

	Relevant catchment population (thousands)	Admissions per 1,000 people		Average hospital stay (days)[a]	Half-day consultant sessions per admission[b]	
		Inpatient	Outpatient		Inpatient[c]	Outpatient
Actual resource allocation in SWRHA in 1975[h]						
General surgery[d]	3,003.7	19.6	16.7	8.54	0.253	0.166
General medicine[e]	3,003.7	14.3	8.5	10.91	0.252	0.372
Obstetrics and gynecology	1,555.8[f]	35.8	31.1	6.08	0.115	0.147
T & O surgery	3,003.7	8.3	18.9	14.25	0.246	0.139
ENT	3,003.7	4.2	9.3	4.46	0.405	0.152
Pediatrics	654.7[g]	31.0	14.2	7.20	0.279	0.398
Ophthalmology	3,003.7	2.7	10.3	7.44	0.533	0.194
Predicted resource allocation in SWRHA in 1975						
General surgery[d]		20.6	18.9	8.22	0.244	0.138
General medicine[e]		15.1	8.4	10.41	0.220	0.335
Obstetrics and gynecology		38.5	36.1	5.96	0.082	0.118
T & O surgery		9.0	21.5	14.05	0.314	0.120
ENT		4.5	10.4	4.58	0.365	0.136
Pediatrics		29.4	14.9	6.83	0.285	0.381
Ophthalmology		2.7	10.7	7.24	0.463	0.194

[a] 922 bed-days available per 1,000 people in 1975.
[b] Assuming each full-time consultant works the equivalent of 450 half-day sessions per year; 48 half-day consultant sessions available per 1,000 people in 1975.
[c] Derived by subtracting actual outpatient sessions from total number of sessions.
[d] Includes urology.
[e] Includes cardiology.
[f] Excludes males.
[g] Excludes people more than 15 years old.
[h] From Department of Health and Social Security (1977b).

outpatient sessions worked during that year in that specialty. The ratio of the cost of a doctor to the cost of a bed is assumed to be 1.57 : 1 (Hughes 1978a). In deriving this figure, the cost of each bed includes all associated costs *except* the cost of the doctor.

Table 6 shows the model parameters that were estimated by the methods of Section 4 from historical allocation data from 1976 and 1977, and disaggregated for the five hospital areas of the SWRHA. With only ten data points we would not expect to estimate a complete parameter set with great confidence, and some of the figures in Table 6 are very uncertain. Nevertheless, the variations between parameters are as expected. In obstetrics and gynecology most of the demand is met (high α_j) but the need for outpatient treatment is very elastic (low β_{j22}). In general medicine, the reverse is true. Many patients do not receive hospital care, but the supply of resources to those who do is rather inelastic.

Table 7 compares the predictions made by the model using these parameters with the actual allocations in 1975. The agreement between model and reality is better than that found in Example 1, but this is partly because of relatively small changes in the SWRHA during the 3 years. Further calibration tests would be desirable.

Meanwhile, however, we consider how to use this model to answer the question at the beginning of this section. We want to increase the numbers of doctors, but this can be afforded only by decreasing the number of beds. We imagine that, from the 1975 resource levels, doctors are increased by 10 percent and beds decreased by 10 percent. (With only tentative parameter estimates, predictions for larger changes may be suspect.) What will happen? The response of the HCS could be to

- Treat different numbers of patients
- Use more or fewer resources per patient
- Change the specialty mix of patients treated
- Change the mix of resources used to treat different patients
- Change the mode of treatment between inpatient and outpatient care for different patients

The simple proportional changes do not indicate which effect will dominate; the model can.

Table 8 shows the predicted results of decreasing beds and increasing doctors, each by 10 percent. As might be expected, these changes result in fewer inpatients and more outpatients. Because of the several population divisors, the total percentage shifts are difficult to quantify, but inpatients decline by about 8 percent, and outpatients increase by about 6 percent. The remaining changes take place in the average lengths of stay and in the distribution of doctor's time among patients.

It is interesting to examine whether, when inpatients and outpatients are

TABLE 8 Predicted results for a decrease in beds and an increase in doctors (Example 2).

	Admissions per 1,000 people		Average hospital stay (days)[a]	Half-day consultant sessions per admission[b]	
	Inpatient	Outpatient		Inpatient[c]	Outpatient
General surgery[d]	20.1	19.2	8.02	0.255	0.161
General medicine[e]	11.7	11.2	10.15	0.240	0.344
Obstetrics and gynecology[f]	38.0	36.3	5.81	0.093	0.161
T & O surgery	8.8	22.1	12.15	0.321	0.123
ENT	4.4	10.7	3.43	0.373	0.138
Pediatrics[g]	28.3	15.4	6.63	0.298	0.445
Ophthalmology	2.7	10.9	7.05	0.479	0.200

[a] 830 bed-days available per 1,000 people (10 percent less than in 1975).
[b] Assuming each full-time consultant works the equivalent of 450 half-day sessions per year; 52 half-day consultant sessions available per 1,000 people (10 percent more than in 1975).
[c] Derived by subtracting actual outpatient sessions from total number of sessions.
[d] Includes urology.
[e] Includes cardiology.
[f] Excludes males.
[g] Excludes people more than 15 years old.

added together, more or fewer patients are treated in each specialty. The model suggests increases in T & O surgery, ENT, and pediatrics, and decreases in the other specialties. The specialty with the largest change from inpatient to outpatient care is general medicine. Naturally, all the lengths of stay decrease, most notably in T & O surgery (by 2 days) and ENT (by 1 day). Naturally, all the levels of doctor care rise, but some of them hardly at all (e.g., T & O surgery and ENT). The largest increases occur in obstetrics and gynecology, with the implication that doctors are under most pressure in these specialties.

Of course, a decision about changing resource levels may be more complicated than represented above. In England, for example, approval for new consultant posts is granted in specific specialties. But a model run in which total consultant posts are increased is still useful in suggesting the specialties for which approval should be sought. The response of the system is also likely to be more complicated than represented above. For example, utilization measures such as bed occupancy may change, thereby upsetting DRAM's predictions. If this happens, a model of the more critical resources may be more appropriate, and DRAM is sufficiently flexible to allow this. Whenever data from past years are available which show how resources were distributed between categories and modes, such data can be used to test DRAM's hypothesis and, if possible, to calibrate a relevant model.

6 SUMMARY

Health care systems are unlike the more common engineering systems that are investigated by mathematical modelers. They are social systems, inaccessible for experiment, where many different agents act according to personal preferences, and without any operational definition of the principal output – health. The chances of using mathematical analysis to study resource allocation would seem to be slight. How then have we done so much algebra?

In fact, nearly all the algebra derives from just two equations – Eq. (5), which says that all resources are used, and Eq. (1), which says that the system tries to give the most care to the most people. Section 2 showed how these two equations are sufficient to derive Eqs. (7) and (8), which say which individuals get what sort of care. These equations constitute DRAM, and the rest of the report looks at the results that they predict.

The predictions will be good ones only if the two underlying equations are realistic. Because justification by common sense can be wrong, we have investigated in Section 3 the sorts of resource allocation patterns that DRAM can imply. This analysis found that the model cannot reproduce increasing levels of service and decreasing levels of supply simultaneously, but that it will always make use of all the available modes of care. Such results make DRAM applicable in many different sectors of health care, and perhaps elsewhere.

For DRAM to be useful, it must be possible to put numbers into the equations on the basis of observed data. Section 4 presented methods that use routine statistics, but that take into account that all sources of data may reflect inherent parameter variations. It is also possible to put numbers into the model on intuitive or professional advice, and some of our procedures indicate which of the parameters might be improved by intelligent guesswork.

Practical application of the model requires cheap and speedy solutions. The computing times reported in Appendix B indicate a very efficient solution algorithm. Even a program with full error handling and diagnostics is still quite small and easy to install.

For what purposes can we use DRAM? Section 5 discussed large and small applications, and two problems amenable to DRAM were investigated in two examples. The first was concerned with allocation of beds among patients with different diseases. The second dealt with the question: Will more or fewer individuals be treated in South Western England if hospital beds are decreased by 10 percent and hospital consultants are increased by 10 percent? The answer (more in some specialties, fewer in others) could be the beginning of a more detailed analysis.

Questions like these are not easy to answer from tables of statistics alone, and DRAM can be seen as a way of organizing information to help in problems of resource allocation. Section 4 therefore examined ways to make DRAM easier to set up when a lot of data are available. These methods are attractive because they derive from an ideal approach to estimating parameters, yet seem feasible and even efficient. Testing them within case study applications is a task for the future.

REFERENCES

Burton, R. M., D. C. Dellinger, W. M. Damon, and E. A. Pfeiffer (1978) A Role for Operational Research in Health Care Planning and Management Teams. Journal of the Operational Research Society 29 (7): 633–641.

Cardus, D. and R. M. Thrall (1977) The Concept of Positive Health and the Planning of Health Care Systems. In D. D. Venedictov, ed., Health System Modeling and the Information System for the Coordination of Research in Oncology. Proceedings of the IIASA Biomedical Conference. CP-77-4. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Department of Health and Social Security, U.K. (1972) Report on Hospital In-Patient Enquiry for the Year 1968. London: Her Majesty's Stationery Office.

Department of Health and Social Security, U.K. (1977a) Report on Hospital In-Patient Enquiry for the Year 1973. London: Her Majesty's Stationery Office.

Department of Health and Social Security, U.K. (1977b) Hospital Medical Staff Tables, SH3 Hospital Return Summaries, 1975, 1976 and 1977. Internal reports.

Feldstein, M. S. (1967) Economic Analysis for Health Service Efficiency. Amsterdam: North-Holland.

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

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

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

Gibbs, R. J. (1978b) The IIASA Health Care Resource Allocation Sub-Model: Mark 1. RR-78-8. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Hughes, D. J. (1978a) The IIASA Health Care Resource Allocation Sub-Model: Mark 2 – The Allocation of Many Different Resources. RM-78-50. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Hughes, D. J. (1978b) The IIASA Health Care Resource Allocation Sub-Model: Formulation of DRAM Mark 3. WP-78-46. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Hughes, D. J., E. A. Nurminski, and G. Royston (1979) Nondifferentiable Optimization Promotes Health Care. WP-79-90. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Kaihara, S., I. Fujimasa, K. Atsumi, and A. A. Klementiev (1977) An Approach to Building a Universal Health Care Model: Morbidity Model of Degenerative Diseases. RM-77-6. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Lemarechal, C. and R. Mifflin, eds. (1978) Nonsmooth Optimization. Proceedings of a IIASA Workshop. Oxford: Pergamon.

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

Nackel, J. G., J. Goldman, and W. L. Fairman (1978) A Group Decision Process for Resource Allocation in the Health Setting. Management Services 24 (12): 1259–1267.

Newhouse, J. P. and C. E. Phelps (1974) Price and Income Elasticities for Medical Care Services. R-1197-NC/OEO. Santa Monica, California: The Rand Corporation.

Rousseau, J. (1977) 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. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Shigan, E. N., D. J. Hughes, and P. I. Kitsul (1979) Health Care Systems Modeling at IIASA: A Status Report. SR-79-4. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Van der Gaag, J., F. Rutten, and B. van Praag (1975) Determinants of Hospital Utilization in the Netherlands. Health Services Research 10 (3): 264–278.

Venedictov, D. D., A. A. Blyusin, T. G. Berezneva, K. M. Kelmanson, A. S. Kiselev, A. A. Klementiev, Yu. Komarov, V. N. Novoseltsev, A. M. Petrovsky, M. A. Schnepps-Schneppe, E. N. Shigan, P. P. Volkov, and A. I. Yashin (1977) Health Care: A Systems Approach. In D. D. Venedictov, ed., Health System Modeling and the Information System for the Coordination of Research in Oncology. CP-77-4. Laxenburg, Austria: International Institute for Applied Systems Analysis.

APPENDIXES

A AN ALTERNATIVE FORMULATION OF DRAM

The formulation of the DRAM model depends upon the definition in Eq. (1) of the function $U(x, y)$. An alternative definition

$$U(x, y, z) = \sum_j g_j(z_j) + \sum_j \sum_k \sum_l x_{jk} h_{jkl}(y_{jkl}) \tag{A1}$$

was investigated by Hughes (1978b), in which

$$g_j(z) = \frac{c_j Z_j}{\alpha_j} \left[1 - \left(\frac{z}{Z_j} \right)^{-\alpha_j} \right] \tag{A2}$$

$$h_{jkl}(y) = \frac{C_l Y_{jkl}}{\beta_{jkl}} \left[1 - \left(\frac{y}{Y_{jkl}} \right)^{-\beta_{jkl}} \right] \tag{A3}$$

and where the HCS is assumed to want to increase the total number of individuals in category j who receive care (per head of population per year)

$$z_j = \sum_k x_{jk} \qquad \forall j \tag{A4}$$

irrespective of the numbers x_{jk} in each mode of care. The parameters Z_j represent the total number of individuals in category j who need care. The parameters c_j are the marginal costs of treating one individual in category j, when all demands are met

$$c_j = \sum_k \sum_l C_l Y_{jkl}$$

The other parameters are as defined for the original DRAM.

An important property of $U(x, y)$ in the original DRAM is that $x_{jk} \rightarrow 0$ for any j, k, causes $U \rightarrow -\infty$. Because the solution to DRAM maximizes U, this condition automatically excludes solutions in which any $\hat{x}_{jk}$ is zero. However, this condition is not true of $U(x, y, z)$ in Eq. (A1), which can be maximized when some x_{jk} are zero. For this reason, the constraint $x_{jk} \geqslant 0$, $\forall j, k$, must be explicitly applied when solving this alternative formulation, and this leads to expressions in the solution that have "corners" or which are "nondifferentiable".

We do not wish here to solve this alternative formulation of DRAM, but only to investigate the number m of category-mode pairs $(j, k) \in \{1, 2, \ldots, J; 1, 2, \ldots, K\}$ such that $\hat{x}_{jk} > 0$. From Eq. (A4), this number satisfies

$$J \leqslant m \leqslant JK \tag{A5}$$

but stronger conditions on m can be found.

Using Lagrange multipliers λ_l, $l = 1, 2, \ldots, L$, to adjoin an equality resource constraint

$$R_l - \sum_j \sum_k x_{jk} y_{jkl} = 0 \qquad \forall l \tag{A6}$$

to function (A1), which is to be maximized, gives

$$H(x, y, z, \lambda) = \sum_j g_j(z_j) + \sum_j \sum_k \sum_l x_{jk} h_{jkl}(y_{jkl}) + \sum_l \lambda_l \Big(R_l - \sum_j \sum_k x_{jk} y_{jkl}\Big)$$

Solutions for $\hat{y}(\lambda)$ satisfy

$$\frac{\partial H}{\partial y_{jkl}} = x_{jk} h'_{jkl}(y_{jkl}) - \lambda_l x_{jk} = 0$$

which gives

$$\hat{y}_{jkl} = h'^{-1}_{jkl}(\lambda_l) \qquad \forall j, k \text{ such that } x_{jk} > 0 \tag{A7}$$

Solutions for $\hat{x}(\lambda)$ are zero or satisfy

$$\frac{\partial H}{\partial x_{jk}} = g'_j(z_j) + \sum_l h_{jkl}(y_{jkl}) - \sum_l \lambda_l y_{jkl} = 0$$

which gives

$$\hat{z}_j = (g'_j)^{-1} \left\{ \sum_l [\lambda_l \hat{y}_{jkl} - h_{jkl}(\hat{y}_{jkl})] \right\}$$

Using Eqs. (A2), (A3), and (A7), this becomes

$$\hat{z}_j = Z_j(\mu_{jk})^{-1/(\alpha_j+1)} \qquad \forall j, k \text{ such that } x_{jk} > 0 \tag{A8}$$

where μ_{jk} is a function of λ similar to that defined by Eqs. (9) and (10). Because the left-hand side of Eq. (A8) is independent of k, it implies $(m - J)$ identities of the form $\mu_{jk_1} = \mu_{jk_2}$, in which there are only L unknowns λ_l, $l = 1, 2, \ldots, L$. In general therefore, solutions exist only if $(m - J) \leqslant L$. Combining this result with Eq. (A5) gives the following condition on the number of category-mode pairs that can be active:

$$J \leqslant m \leqslant \min(JK, J + L) \tag{A9}$$

For the data in Example 2 in Section 5, $J = 7$, $K = L = 2$, and inequality

(A9) is $7 \leqslant m \leqslant 9$. This implies that of the seven patient categories, not more than two can use more than one mode of care. For some definitions of categories and modes this result may be realistic, and we have made progress in solving models like this using nonsmooth optimization methods (Lemarechal and Mifflin 1978, Hughes *et al.* 1979). For Example 2, however, this result is unrealistic, because all categories of patients use both modes oi care. Therefore, we have not pursued this formulation here.

B COMPUTER PROGRAMS AND SOLUTION EFFICIENCY

The procedures for model solution and parameter estimation described in Sections 3 and 4 have been implemented as computer programs. They are written in simple FORTRAN with many in-line comments, error handling, and full but suppressible diagnostic printout. They use no special software beyond simple matrix manipulation routines. Input and output files are read and written sequentially, and all files are formatted for easily understandable display. The programs are best used interactively, and a small utility program can quickly modify the input file when many model runs with different resource levels are required. Batch operation is equally possible.

Table B1 gives some statistics for the three principal programs, which solve the model with given parameters, estimate the level parameters X, Y, and estimate the power parameters α, β. They show that the average length of each routine is low (less than 60 statements) and that the fraction of comment code is high (more than 0.5). The total core load of each program is reasonable (less than 55K decimal bytes).

All three principal programs use an iterative solution and the running times therefore depend upon the starting values, the accuracy required of the solution, and the conditioning of the problem. For the model solution program, the running time additionally depends upon whether the dual constraint, Eq. (17) or (21), is applied and binding. Section 2.4 described how this constraint is handled computationally.

Table B2 gives typical running times for the three principal programs, used on problems of different sizes, when no diagnostic printout was requested and with arbitrary starting values (typically a first guess of $\lambda = 5$). Convergence is measured by the fractional change of

- The dual function $\hat{H}(\lambda)$, in the model solution program
- $\partial\hat{H}(\lambda)/\partial\lambda$, in the X, Y estimation program
- λ, in the α, β estimation program

and is usually fast. It is especially so for the model solution program (less than 15 CPU for a medium-sized problem), even when the solution lies on a dual constraint. No attempt has been made to speed up the parameter estimation programs, the second of which may converge slowly or not at all. But the fast

TABLE B1 Computer program statistics.

	Separate routines[a]	Number of FORTRAN statements[b]	Number of		Maximum numbers of				Total core load on PDP 11/70 (decimal bytes)
			Input files	Output files[c]	Patient categories	Treatment modes	Resource types	Data points	
Model solution	38	1,730 (1,061)	1	3	20	3	5		47,818
Estimation of X, Y	14	1,012 (557)	1	3	12	3	5	14	53,874
Estimation of α, β	16	1,241 (702)	1	3	12	3	5	14	44,264

[a] Including main line program, FUNCTION procedures, and matrix manipulation routines.
[b] The number of statements excluding COMMENTS is given in parentheses.
[c] One of these is suitable for display on a terminal during program execution.

TABLE B2 Typical running times of computer programs.

	Table showing run results	Dimensions of problem J	K	L	N	Precision of solution	Number of iterations to solution	CPU time to solution (sec)[a]
Model solution	3	6	1	1		10^{-5}	6	2.9
	4	6	1	1			4	2.7
	4	6	1	1			8	2.9
	7	7	2	2			8	12.8
	8	7	2	2			4	12.8
		5	1	2			4	5.9[b]
		7	2	2			6	14.1[b]
Estimating X, Y	3	6	1	1	14	10^{-3}	2[c]	14.7
	6	7	2	2	5		2[c]	18.7
Estimating α, β	3	6	1	1	14	$5 \cdot 10^{-2}$	528[d]	143.2
	6	7	2	2	10	10^{-3}	8	17.2

[a]With no diagnostic printout.
[b]In these runs the dual constraint was binding. In others it was not.
[c]Average number of iterations per data point.
[d]Very badly conditioned problem. Convergence is usually faster.

model solution program means that improved parameter estimation methods such as the one described in Section 4.5 are highly practical.

C FITTING FOUR PARAMETERS TO FOUR DATA POINTS

In this appendix we consider how to estimate the four model parameters X, Y, α, β from four data points (x_i, y_i), $i = 1, 2, 3, 4$, for the simplest possible DRAM when $J = K = L = 1$. This analysis extends and completes the discussion of Sections 3.2 and 3.3. We assume without loss of generality that $x_1 > x_2 > x_3 > x_4$ and $y_1 > y_2 > y_3 > y_4$. (If such an ordering of the data is not possible, it means that they do not satisfy the fundamental condition (26) on admissible solutions of DRAM.)

Equations (27) and (28) in Section 3.2 determine X, Y from x_1, x_2, y_1, y_2 when α, β are known. Substituting these results and the other two data points into Eq. (25), we get two nonlinear equations

$$\gamma_3(\alpha, \beta) = \gamma_4(\alpha, \beta) = 0 \tag{C1}$$

which determine α, β implicitly, where

$$\gamma_3(\alpha, \beta) = \left(\frac{y_1}{y_2}\right)^\beta \left[\left(\frac{x_1}{x_3}\right)^{\alpha+1} - 1\right] - \left(\frac{y_1}{y_3}\right)^\beta \left[\left(\frac{x_1}{x_2}\right)^{\alpha+1} - 1\right] - \left(\frac{x_1}{x_3}\right)^{\alpha+1} + \left(\frac{x_1}{x_2}\right)^{\alpha+1} \tag{C2}$$

$$\gamma_4(\alpha,\beta) = \left(\frac{y_1}{y_2}\right)^{\beta}\left[\left(\frac{x_1}{x_4}\right)^{\alpha+1} - 1\right] - \left(\frac{y_1}{y_4}\right)^{\beta}\left[\left(\frac{x_1}{x_2}\right)^{\alpha+1} - 1\right] - \left(\frac{x_1}{x_4}\right)^{\alpha+1} + \left(\frac{x_1}{x_2}\right)^{\alpha+1} \tag{C3}$$

Equations (C1) also define implicit functions $\beta_3(\alpha)$, $\beta_4(\alpha)$ which in turn define solutions $\tilde{\alpha}$, $\tilde{\beta}$ for α, β

$$\tilde{\beta} = \beta_3(\tilde{\alpha}) = \beta_4(\tilde{\alpha}) \tag{C4}$$

For successful solution of Eq. (C4), two sets of existence conditions need to be established. First, we must find conditions for Eqs. (C1) to have a solution $\beta > 0$, assuming the existence of a solution $\alpha > 0$. Second, we must find conditions for Eq. (C4) to have a solution $\alpha > 0$. When the second condition is satisfied, the first condition will ensure $\beta > 0$.

The first conditions follow from inspecting the derivatives $\partial\gamma_3/\partial\alpha$, etc. We find that sufficient (but more than necessary) conditions for $\beta > 0$ given $\alpha > 0$ are

$$\left(\frac{(x_1/x_3) - 1}{(x_1/x_2) - 1}\right) > \left(\frac{\ln(y_1/y_3)}{\ln(y_1/y_2)}\right) \tag{C5}$$

$$\left(\frac{(x_1/x_4) - 1}{(x_1/x_2) - 1}\right) > \left(\frac{\ln(y_1/y_4)}{\ln(y_1/y_2)}\right) \tag{C6}$$

The second conditions follow from lower, upper, and asymptotic estimates of the functions $\beta_3(\alpha)$, $\beta_4(\alpha)$.

$$\beta_j^{\min}(\alpha) = \left\{\ln\left[\frac{(x_1/x_j)^{\alpha+1} - 1}{(x_1/x_2)^{\alpha+1} - 1}\right] + \ln\left[\frac{\ln(y_1/y_j)}{\ln(y_1/y_2)}\right]\right\} \Big/ \left(\ln\frac{y_2}{y_j}\right) \tag{C7}$$

$$\beta_j^{\max}(\alpha) = \left\{\ln\left[\frac{(x_1/x_j)^{\alpha+1} - 1}{(x_1/x_2)^{\alpha+1} - 1}\right]\right\} \Big/ \left(\ln\frac{y_2}{y_j}\right) \tag{C8}$$

$$\beta_j^{\alpha\to\infty}(\alpha) = (\alpha + 1)\frac{\ln(x_2/x_j)}{\ln(y_2/y_j)} \qquad j = 3, 4 \tag{C9}$$

If $\beta_3^{\infty}(\alpha) > \beta_4^{\infty}(\alpha)$, then $\beta_4(0) > \beta_3(0)$ will guarantee a solution $\tilde{\alpha} > 0$ to Eq. (C4). This condition is depicted in Figure C1. Conversely, if $\beta_4^{\infty}(\alpha) > \beta_3^{\infty}(\alpha)$, then $\beta_3(0) > \beta_4(0)$ will guarantee a solution. Both of these sufficient and necessary conditions (which must be computed numerically) can be approximated by sufficient but more restrictive conditions (which need not be computed numerically)

$$\beta_4^{\min}(0) > \beta_3^{\max}(0) \qquad \text{or} \qquad \beta_3^{\min}(0) > \beta_4^{\max}(0)$$

In order to illustrate the approach, we consider the data shown in Table C1 which satisfy conditions (26), (C5), and (C6). In addition, $\beta_3^{\infty}(\alpha) > \beta_4^{\infty}(\alpha)$ and $\beta_4(0) > \beta_3(0)$, thereby guaranteeing solutions $\tilde{\alpha}, \tilde{\beta} > 0$. On Figure C1 are plotted values of $\beta_3(\alpha)$, $\beta_4(\alpha)$ obtained by solving Eqs. (C1) by the following iteration

$$\beta_j(i+1) = 2\bar{\beta}_j(i+1) - \beta_j(i) \qquad j = 3, 4, \quad i = 1, 2, \ldots$$

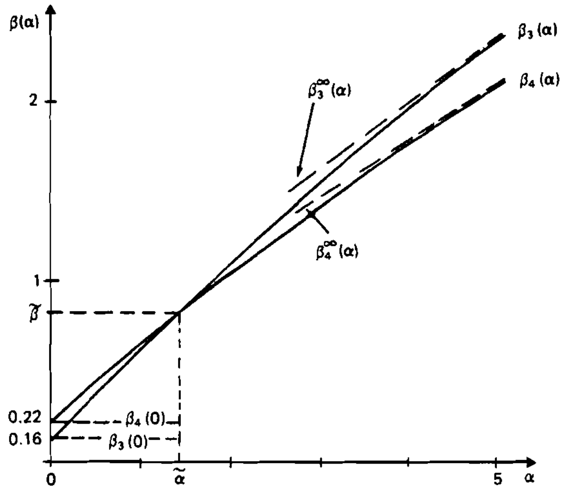

FIGURE C1 Solution of Eq. (C4) for data from Table C1.

where

$$\tilde{\beta}_j(i+1) = \frac{\ln\left[1 + \frac{(x_1/x_j)^{\alpha+1} - (x_1/x_j)}{(x_1/x_2)^{\alpha+1} - 1} - \frac{(x_1/x_j)^{\alpha+1} - (x_1/x_2)^{\alpha+1}}{(x_1/x_2)^{\alpha+1} - 1}(y_2/y_1)^{\beta_j(i)}\right]}{\ln(y_2/y_j)}$$

The solution to Eq. (C4) is found at $\tilde{\alpha} = 1.60$, $\tilde{\beta} = 0.83$, $X = 1.05$, and $Y = 1.08$, although it is not very accurately determined because the problem is rather ill-conditioned. This is seen in the approximate equality of $\beta_3^\infty(\alpha) = 0.386(\alpha + 1)$ and $\beta_4^\infty(\alpha) = 0.355(\alpha + 1)$, and in the very flat intersection in Figure C1. Nevertheless, the estimated values are close to the true parameter values (shown in Table C1) used to derive the four data points.

TABLE C1 Data for test of parameter fitting. Solutions of the simplest possible DRAM with $x = 1, y = 1, \alpha = 2, \beta = 1$.

	$i = 1$	$i = 2$	$i = 3$	$i = 4$
x_i	0.90	0.70	0.40	0.20
y_i	0.84	0.51	0.12	0.015
$R_i = x_i y_i$	0.756	0.357	0.048	0.003

D UNBIASED REGRESSION ESTIMATORS

In the estimation of power parameters α, β in Section 4.3, we assumed that α, β are constant across the areas of a region, and then we performed regression analysis on the cross-sectional data. However, even if this assumption is incorrect and α, β are different in different areas, we can show that this procedure still yields useful regional estimates.

We define the indices $j = 1, 2, \ldots, J$ areas or subregions, and $i = 1, 2, \ldots, N$ observations in each area, and suppose that data $x_j(i), y_j(i)$ satisfy the linear model

$$y_j(i) = b_j x_j(i) + \epsilon_j(i) \tag{D1}$$

in which $\epsilon_j(i)$ are uncorrelated random disturbances with zero mean and variance σ^2. The unknown parameter b_j is different for different areas. Nevertheless, we assume that it is constant and we form the usual least-squares estimate

$$\hat{b} = \left(\sum_j X_j' X_j\right)^{-1} \sum_j X_j' Y_j \tag{D2}$$

in which $X_j = \{x_j(1), \ldots, x_j(N)\}'$ and $Y_j = \{y_j(1), \ldots, y_j(N)\}'$. We now investigate the properties of $\hat{b}$ when the unknown parameters b_j are actually random samples from a normal or Gaussian probability density function with mean m and variance v^2:

$$b_j \sim N(m, v^2)$$

Combining Eqs. (D1) and (D2) gives

$$(\hat{b} - m) = \left(\sum_j X_j' X_j\right)^{-1} \left\{\sum_j X_j' X_j (b_j - m) + \sum_j X_j' E_j \sigma\right\}$$

whence the result: EXPECTATION $(\hat{b} - m) = 0$; the estimator b is an unbiased estimator of the mean regional parameter m.

Equations (55) and (56) in Section 4.3 are like Eq. (D1). The functions corresponding to b_j are $(\alpha_j + 1)^{-1}$ and $(\beta_{jkm} + 1)^{-1}$ which are estimated without bias, subject to the above assumptions. Additionally, we may show that

$$\text{EXPECTATION } (\hat{b} - m)^2 \sim \frac{\sigma^2}{JN} + \frac{v^2}{J}$$

The first term on the right-hand side is the usual residual variance term, and the second arises from the uncertainty about b_j.

E SENSITIVITY OF THE SOLUTION TO PARAMETER CHANGES

The parameter estimation procedure described in Section 4.5 needs expressions for the sensitivities of the solutions $\hat{x}, \hat{y}$ to a change in a parameter $p \in P = \{X, Y, \alpha, \beta\}$. These expressions are derived below.

The total sensitivity derivatives can be written as the sum of two sets of partial derivatives

$$\frac{d\hat{x}_{jk}}{dp} = \frac{\partial \hat{x}_{jk}}{\partial p} + \sum_m \frac{\partial \hat{x}_{jk}}{\partial \hat{\lambda}_m} \frac{\partial \hat{\lambda}_m}{\partial p}$$

$$\frac{d\hat{y}_{jkl}}{dp} = \frac{\partial \hat{y}_{jkl}}{\partial p} + \sum_m \frac{\partial \hat{y}_{jkl}}{\partial \hat{\lambda}_m} \frac{\partial \hat{\lambda}_m}{\partial p}$$

The first term in each equation is the partial derivative when the Lagrange

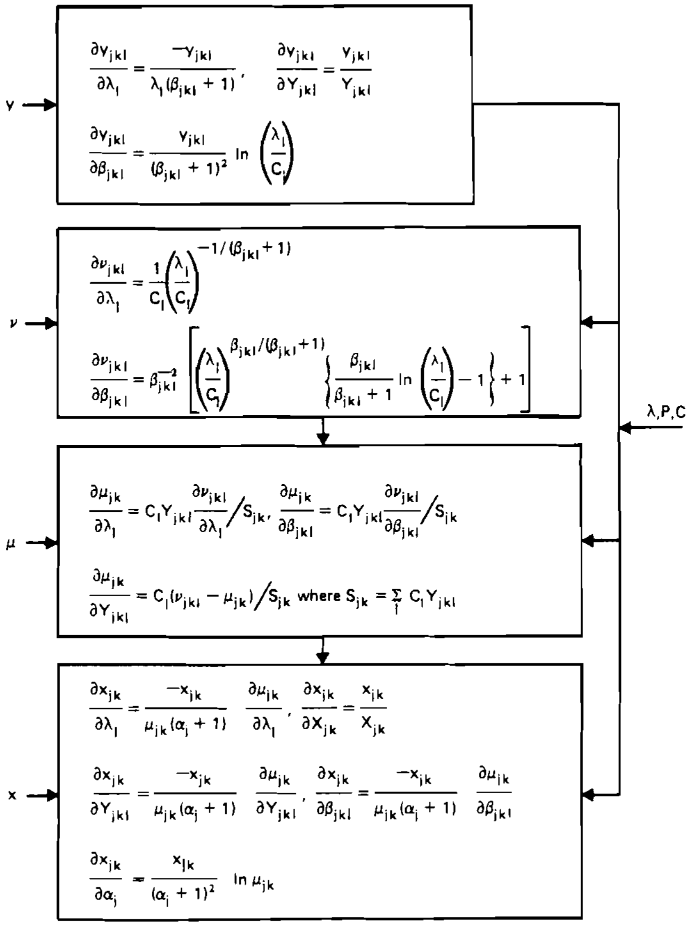

FIGURE E1 Calculation of sensitivity derivatives (superscript carets (ˆ) are omitted for clarity).

multipliers λ are held constant. The second term in each equation reflects the sensitivity of the solution to changes in the Lagrange multipliers.

In order to obtain the terms $\partial\hat{\lambda}/\partial p$ we note that, at the solution point $\lambda = \hat{\lambda}$,

$$\frac{\partial \hat{H}(\lambda)}{\partial \lambda_l} = F_l[\hat{x}(\hat{\lambda}), \hat{y}(\hat{\lambda})]$$

$$= R_l - \sum_j \sum_k \hat{x}_{jk}\hat{y}_{jkl} = 0 \qquad \forall l$$

Differentiating this result with respect to $p \in P$ gives

$$\frac{\mathrm{d}}{\mathrm{d}p}\left[\frac{\partial \hat{H}(\lambda)}{\partial \hat{\lambda}_l}\right] = \frac{\partial}{\partial p}\left[\frac{\partial \hat{H}(\lambda)}{\partial \hat{\lambda}_l}\right] + \sum_m \frac{\partial^2 \hat{H}(\lambda)}{\partial \hat{\lambda}_l \partial \hat{\lambda}_m}\frac{\partial \hat{\lambda}_m}{\partial p} = 0 \qquad \forall l$$

whence $(\partial\hat{\lambda}/\partial p) = -(\hat{H}_{\lambda\lambda})^{-1}(\partial\hat{H}_\lambda/\partial p)$ or

$$\frac{\partial\hat{\lambda}_m}{\partial p} = \sum_l (\hat{H}_{\lambda\lambda})^{-1} \sum_j \sum_k \hat{x}_{jk} \frac{\partial\hat{y}_{jkl}}{\partial p} + \hat{y}_{jkl} \frac{\partial\hat{x}_{jk}}{\partial p}$$

$(\hat{H}_{\lambda\lambda})^{-1}$ is the inverse Hessian matrix which is calculated during the solving of the model. The other terms are simply the other group of partial derivatives that follow straightforwardly from Eqs. (7)-(10). The only difficulty is in organizing the computation in the most convenient way. Figure E1 depicts a possible scheme.

The calculations are considerably simplified by the presence of many zero terms. Most obviously,

$$\frac{\partial y_{jkl}}{\partial X_{jk}} = \frac{\partial y_{jkl}}{\partial \alpha_j} = 0$$

$$\frac{\partial \nu_{jkl}}{\partial X_{jk}} = \frac{\partial \nu_{jkl}}{\partial Y_{jkl}} = \frac{\partial \nu_{jkl}}{\partial \alpha_j} = 0$$

$$\frac{\partial \mu_{jk}}{\partial X_{jk}} = \frac{\partial \mu_{jk}}{\partial \alpha_j} = 0 \qquad \forall j, k, l$$

Less obviously

$$\frac{\partial y_{jkl}}{\partial Y_{\bar{j}\bar{k}\bar{l}}} = \frac{\partial y_{jkl}}{\partial \beta_{\bar{j}\bar{k}\bar{l}}} = 0$$

$$\frac{\partial \nu_{jkl}}{\partial \lambda_{\bar{l}}} = \frac{\partial \nu_{jkl}}{\partial \beta_{\bar{j}\bar{k}\bar{l}}} = 0$$

$$\frac{\partial \mu_{jk}}{\partial \beta_{\bar{j}\bar{k}\bar{l}}} = \frac{\partial \mu_{jk}}{\partial Y_{\bar{j}\bar{k}\bar{l}}} = 0$$

$$\frac{\partial x_{jk}}{\partial X_{\bar{j}\bar{k}}} = \frac{\partial x_{jk}}{\partial Y_{\bar{j}\bar{k}\bar{l}}} = \frac{\partial x_{jk}}{\partial \alpha_{\bar{j}}} = \frac{\partial x_{jk}}{\partial \beta_{\bar{j}\bar{k}\bar{l}}} = 0$$

for $(j, k, l) \neq (\bar{j}, \bar{k}, \bar{l})$. Unfortunately, the matrices of total derivatives

$$\left\{\frac{\mathrm{d}\hat{x}_{jk}}{\mathrm{d}p}\right\}, \left\{\frac{\mathrm{d}\hat{y}_{jkl}}{\mathrm{d}p}\right\}$$

have in general no zero terms because of the dependence of each Lagrange multiplier upon every parameter. Together, they have $(JK + JKL)(2JKL + JK + J)$ terms, but Section 4.5 shows that not all these elements need to be stored.

F LIST OF PRINCIPAL SYMBOLS

Symbols used only in the appendixes are not included here.

Symbol	Definition	Page of first appearance
i =	1, 2, . . . , N, iterations, times, regions, data points	9, 21, 21, 24
j =	1, 2, . . . , J, patient categories	3
k =	1, 2, . . . , K, modes of care	3
l =	1, 2, . . . , L, resource types	3
A, B	decomposition of $\hat{H}_{\lambda\lambda}$	6
a_{lm}, b_{lm}	elements of A, B	6
A_{jkl}, B_{ml}	expressions relating α to γ and β to η	24
a^x_{jk}, a^y_{jkm}	constant terms in regression estimation of γ, η	24
C_l	marginal unit cost of type l resource, when all needs are met	4
d^i	Newton direction at iteration i	9
$F_l(x, y)$	function in constraint equation	4
$f_l(\lambda)$	function specifying θ_l	21
$g_{jk}(x)$	functions measuring the benefits of increasing service levels	4
$h_{jkl}(y)$	functions measuring the benefits of increasing supply levels	4
$H(x, y, \lambda)$	Lagrangian function	5
$\hat{H}(\lambda)$	Lagrangian when $x = \hat{x}(\lambda)$, $y = \hat{y}(\lambda)$	5
$\hat{H}_\lambda, \hat{H}_{\lambda\lambda}$	gradient and Hessian of $H(\lambda)$ when $x = \hat{x}(\lambda)$, $y = \hat{y}(\lambda)$	5
$\overline{H}_{ml}$	ml-th element of inverted Hessian matrix	23
H_{zz}	matrix of second derivatives of Lagrangian with respect to primal variables	6
$J(P)$	function of squared prediction errors	27
$\overline{L}$	set of active resource constraints	9
M	finance for purchasing resources	7
$P(i)$	population in region i	21
P	parameter set $\{X, Y, \alpha, \beta\}$	26
p, q =	$(x/X), (y/Y)$	12
R_l	available resource of type l	4
r_l	excess resource of type l	7
S	statistics of ξ_1, ξ_2 processes	27
s	convergence coefficient	10
t	step-size coefficient	9
$U(x, y)$, $U(x, y, P, R, C)$	function which is maximized by DRAM	4, 26
v_m	sensitivity derivative vectors	28
$X_{jk}, x_{jk}, \hat{x}_{jk}$	ideal, actual, optimal service levels	4, 5
$Y_{jkl}, y_{jkl}, \hat{y}_{jkl}$	ideal, actual, optimal supply levels	4, 5
$\overline{X}, \overline{Y}$	regional ideal levels	21

Z_j	ideal service levels summed across modes	21
z	arbitrary positive vectors	6
α_j, β_{jkl}	model power parameters	4
γ_{jkl}, η_{jkml}	empirical elasticities	23
δ_{lm}	Kronecker delta function	6
ϵ^x_{jk}, ϵ^y_{jkm}	error terms in regression estimation of γ, η	24
θ_l, ϕ_l	additional information for estimating X, Y	20, 22
$\tan \Theta$	$= (\beta + 1)/(\alpha + \beta + 2)$	12
$\tan \Phi$	$= (\alpha + 1)/(\alpha + \beta + 2)$	12
$\tan \Psi$	$= (\alpha + 1)/(\beta + 1)$	13
ω	$= \ln p_2/\ln p_1$	15
τ	$= (1 - \omega) \ln q_1$	15
μ_{jk}, ν_{jkl}	functions in expression for $\hat{y}$	5
λ_l, $\hat{\lambda}_l$	actual and optimal Lagrange multipliers	5
$\tilde{\lambda}_l$	minimum Lagrange multipliers	22
$\zeta(\beta, q)$	$= \ln \left[\left(1 + \frac{1}{\beta}\right)q^{-\beta} - \frac{1}{\beta}\right]$	15
ξ_1, ξ_2	random processes perturbing x, y	27
ρ^x_{ijk}, ρ^y_{ijkl}	weighting terms in $J(P)$	27

RELATED IIASA PUBLICATIONS IN HEALTH CARE SYSTEMS AND IN OPTIMIZATION

Mifflin, R. (1976) Semismooth and Semiconvex Functions in Constrained Optimization. RR-76-21.

Mifflin, R. (1977) An Algorithm for Constrained Optimization with Semismooth Functions. RR-77-3.

Propoi, A. I. (1977) On the Theory of Max–Min. RM-77-13.

Gibbs, R. J. (1977) Health Care Resource Allocation Models – A Critical Review. RM-77-53.

Lemarechal, C. (1978) Nonsmooth Optimization and Descent Methods. RR-78-4.

Gibbs, R. J. (1978) The IIASA Health Care Resource Allocation Sub-Model: Mark 1. RR-78-8.

Gibbs, R. J. (1978) A Disaggregated Health Care Resource Allocation Model. RM-78-1.

Mifflin, R., C. Lemarechal, and M. A. Keyzer (1978) Computing Economic Equilibria through Nonsmooth Optimization. RM-78-13.

Hughes, D. J. (1978) The IIASA Health Care Resource Allocation Sub-Model: Mark 2 – The Allocation of Many Different Resources. RM-78-50.

Mifflin, R. and C. Lemarechal (1978) Nonsmooth Optimization. CP-78-2.

Gibbs, R. J. (1978) Computer Programs for the IIASA Health Care Resource Allocation Sub-Model, Mark 1 – A User's Guide. WP-78-15.

Hughes, D. J. (1978) The IIASA Health Care Resource Allocation Sub-Model: Formulation of DRAM Mark 3. WP-78-46.

Nurminski, E. A. (1978) Nondifferentiable Optimization with ϵ-Subgradient Methods. WP-78-55.

Nurminski, E. A. (1978) On ϵ-Differentiable Mappings and their Application in Nondifferentiable Optimization. WP-78-58.

Ermoliev, Yu. M. (1978) Methods of Nondifferentiable and Stochastic Optimization and their Applications. WP-78-62.

Wierzbicki, A. (1978) A Quadratic Approximation Method Based on Augmented Lagrangian Functions for Nonconvex Nonlinear Programming Problems. WP-78-61.

Wierzbicki, A. (1978) Lagrangian Functions and Nondifferentiable Optimization. WP-78-63.

Hughes, D. J. (1978) The IIASA Health Care Resource Allocation Submodel: Estimation of Parameters. RM-78-67.

Hughes, D. J., E. A. Nurminski, and G. Royston (1979) Nondifferentiable Optimization Promotes Health Care. WP-79-90.

Nurminski, E. A. (1979) Conceptual Newton Method for Solving Multivalued Inclusions: Scalar Case. WP-79-50.

THE AUTHORS

David Hughes studied Engineering Science at Oxford University, graduating in 1971. He obtained his D.Phil. in Stochastic Control Theory from Oxford in 1974 and an M.Sc. in Statistics from London University in 1977. He has been with the Operational Research Service of the Department of Health and Social Security (UK) since 1974, working on resource allocation and regional planning. His main interests are models of the social services, planning and decision making under conditions of uncertainty, and coordination and control. David Hughes is a Fellow of the Royal Statistical Society, and a member of the Operational Research Society and the Institute of Statisticians.

Andrzej Wierzbicki received his Ph.D. in Automatic Control in 1964 and his D.Sc. in Mathematical Programming and Optimization in 1968, both from the Technical University of Warsaw. He has been with the Technical University of Warsaw since 1961. At present he is leading the System and Decision Sciences Area at IIASA, where his main interests are optimization theory and applications, including multiobjective optimization and decision making, augmented Lagrange functions, and nondifferentiable optimization. Professor Wierzbicki is a member of the High Council for Science and Education in Poland and of many other scientific councils, including that of the Institute of Systems Research, Polish Academy of Sciences. He is the author of several books and is associate editor of the journal *Automatica*.