RAMOS: A MODEL VALIDATION AND SENSITIVITY ANALYSIS

L.D. Mayhew
A. Taket

July 1981
WP-81-100

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

THE AUTHORS

Leslie Mayhew is an IIASA research scholar working within the Health Care Task of the Human Settlements and Services Area. He is on secondement from the Operational Research Unit of the Department of Health and Social Security, UK.

Ann Taket is a scientist in the Operational Research Unit of the Department of Health and Social Security, UK.

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.

The work presented in this paper uses validation techniques and sensitivity analysis to examine critically the predictive performance of the model RAMOS (Resource Allocation Model Over Space). This model is designed to predict the impact on hospitalization rates of changes in population and resource availability over time and space.

Related publications in the Health Care Systems Task are listed at the end of the paper.

Andrei Rogers
Chairman
Human Settlements
and Services Area

ABSTRACT

This paper focuses on the predictive capabilities of the model RAMOS (Resource Allocation Model Over Space). This model, which is being jointly developed by IIASA and the Operational Research Services of the UK Department of Health and Social Security (DHSS), is designed to predict the impact on hospitalization rates when population and resource availability are changing simultaneously in time and space. The performance of the model is critically examined using validation techniques and sensitivity analysis. The validation part is based on an experiment that tries to simulate the outputs of a regional health care system at a point earlier time. This "back-prediction" is then compared for accuracy with what actually occurred. It is shown that the model functions very well in achieving the purposes for which it was designed. Different model specifications are then tested in order to seek further improvements that remove some small but consistent biases in the outputs. Following this, a detailed sensitivity analysis is carried out on the main input variables and parameter, in order to check the internal consistency of the model when it is exposed to unrealistic extremes of change. The paper concludes by noting the mostly satisfactory performance of the model in both the validation tests and the sensitivity analysis but with some caveats and recommendations for further research.

CONTENTS

RAMOS: A MODEL VALIDATION AND SENSITIVITY ANALYSIS

1. INTRODUCTION

The size and spacing of health care facilities is a fundamental consideration in ensuring those in need of medical attention to have reasonable access to the supply of available services. The problem is that needs vary in time and space, mostly according to the relative size, structure, morbidity and spatial distribution of the population, whereas the facilities at supply points (hospitals, clinics, etc.) remain fixed in position for the duration of their functioning. In certain planning environments, the providers of health care services often experience much difficulty in equating the supply of resources in different locations with the relative needs of the local populations over time (RAWP, 1976). Particularly in densely populated regions or large urban agglomerations where changes in demographic structure can be rapid and substantial, these problems become sufficiently complex and potentially costly so as to warrant the development of better, more effective decision-making tools for determining the spatial consequences of different patterns of allocation and reallocation. In this way, the system can be made to respond more effectively to the relative needs of the population at medical risk (LHPC, 1979).

The reallocations in a regional health care system take several forms. Only occasionally, do they entail the opening of an entire new facility or the closure of an old one. In the short term, at least, it is more common for facilities to be simply updated, enlarged, or reduced in size according to the availability of hospital beds or manpower, for example. These reallocations can nevertheless be substantial (-30% to +16% in different treatment districts in south-east England between 1975 and 1977) indicating the necessity for planning tools with both long- and short-range perspectives.

At IIASA, a group of models is being developed that enables users to simulate the consequences of different resource configurations when there are simultaneous changes in demand and resource availability of the type described. This work is being carried out in conjunction with the Operational Research Services of the UK Department of Health and Social Security (DHSS). Currently, information is available on how to specify, construct and calibrate the basic model (Mayhew and Taket, 1980) and on how to apply it, or one of its close variants, in particular decision-making contexts (Mayhew, 1980, 1981).

The purpose of this paper is to investigate the accuracy of the model in its ability to predict change, and from this to obtain an accurate indication of confidence with which the model can be used for decision-making purposes. The two main components of this investigation are a set of detailed validation tests and a sensitivity analysis of the model parameters. Together, the tests show that the model is indeed suited to the purposes for which it was designed but that some further empirical work in refining the input variables is needed to remove some small though consistent biases in the output.

The empirical effort that has gone into validation of the outputs seems from published sources to be rare for this class of model. Thus the results are also of general interest, having implications for a broad range of applications in the spatial interaction field.

2. THE MODEL

The basic model is known as RAMOS (Resource Allocation Model Over Space). In its simplest form it hypothesizes that the number of hospital patients generated in an origin zone i (place of residence) and treated in treatment zone j (a hospital district) is in proportion to the morbidity or "patient generating potential" of i and the resources available in j but is in inverse proportion to the accessibility costs of getting from i to j.

Mathematically, the model is stated as follows:

$$T_{ij} = B_j D_j W_i f(\beta, c_{ij}) \quad (1)$$

where

$i = \overline{1,I}$, the number of origin zones
$j = \overline{1,J}$, the number of treatment zones

T_{ij} = the predicted patient flow from origin zone i to treatment zone j

D_j = the available resources as measured by the caseload capacity in j for treating patients in a medical specialty or group of specialties

W_i = the patient generating factor (pgf), which is an index of the propensity of the population in i to generate patients in the same group of specialties

$f(\beta,c_{ij})$ = a deterrence function, monotonic and declining, representing the fall in demand for health care services with decreasing accessibility (e.g., $\exp(-\beta c_{ij})$, $c_{ij}^{-\beta}$). In later sections, $f(\beta,c_{ij})$ is abbreviated for convenience to f_{ij}.

c_{ij} = a measure of accessibility expressing the difficulty of a person in i to be admitted as a patient in j. It is normally represented by distance, travel time or a related surrogate

β = a parameter to be determined empirically from the existing pattern of patient flows

$$B_j = \left[\sum_i W_i \ f(\beta, c_{ij}) \right]^{-1} \tag{2}$$

Equation (2) is a constraint, known as a balancing factor. It ensures that the resources in each location are used to capacity. With some reformulation this assumption can be relaxed to take account of slack or other factors in particular systems, but this is not considered in the current application.

The model functions in two modes: calibration and prediction. The first consists of finding a value for β such that the model most accurately recreates an observed matrix of patient flows $\{\overline{T}_{ij}\}$; the second, is concerned with the prediction of patient flows, hospitalization rates, and other outputs using forecasted values for D_j and W_i, the resources and patient generating potential. In this investigation we are concerned mostly with the second mode.

3. VALIDATION

The method of validation is based on an experiment that back-predicts the output variables of the model using input data consistent with the time of back-prediction and then compares these outputs with what actually occurred. In the experiment the model parameter β is assumed unchanged. This is because it is an empirically derived constant, specific to the system under investigation, that is usually assumed to be unchanged over a typical planning period. Clearly, if the model outputs accurately portray the realized outputs of the system, then the model can be used with more confidence to predict a wide range of possible planning scenarios.

The validation exercise is divided into three parts. Part I takes a model, calibrated on 1977 data, and then compares the model predictions with the actual performance of the system two years earlier; Part II examines different specifications of certain of the input variables and compares the accuracy of the resulting predictions with those obtained with the original model specifications; and Part III gives a detailed error analysis and suggestions for further improvement.

The use of 1975 as a test year was determined by data availability, and although it is near to 1977 in terms of time, the changes in data values were found sufficient in this two-year period for validation purposes. Some additional data of much less quality were also available for 1967, but only for a smaller part of the region of interest (Figures 1 and 2). Accordingly, less emphasis must be placed on the results obtained. Nevertheless, these results are presented where useful for comparisons.

3.1. 1975 Back-prediction

The version of the model discussed in this section has the following specifications:

$D_j(t)$ = The resources in each treatment zone are defined as the number of acute patient admssions to hospitals in time-period t (for list of included specialties, see Mayhew and Taket, 1980, p.16).

$f(\beta, c_{ij})$ = The deterrence function, defined as $\exp(-\beta c_{ij})$, where $\beta = 0.367$ and $\{c_{ij}\}$ is "Matrix 3" in the above reference.

$W_i(t)$ = The patient generating factor defined as $\sum_l \sum_m P_{il}(t) U_{lm}(t)$, where $P_{il}(t)$ is the population in i in age-sex category l at time t and U_{lm} is the national discharge rate in l for clinical specialty m.

The zoning system over which the model is applied covers 44 origin zones and 69 destination zones concentrated in London and southeast England. This system is shown in the two accompanying maps (Figures 1 and 2); a key to the numbered zones is shown in Table 1.

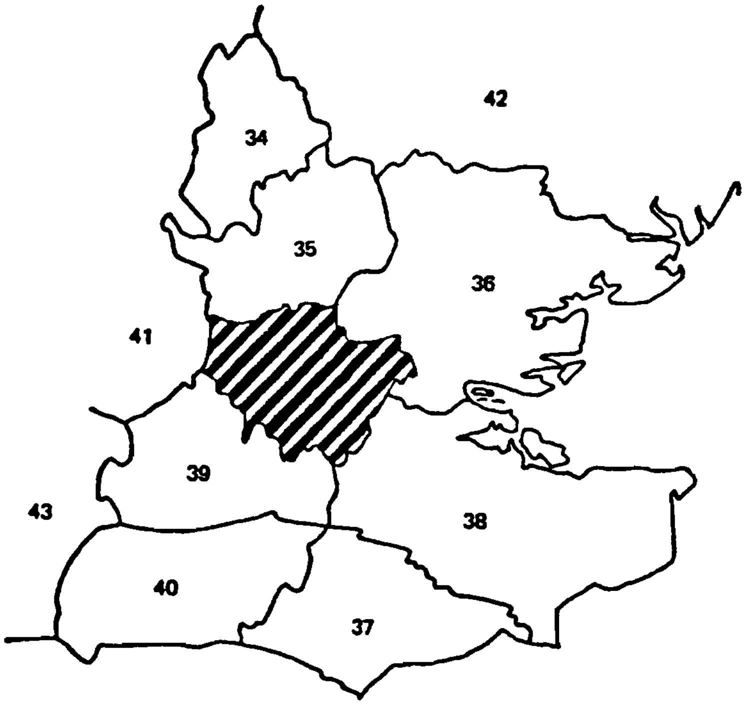
3.1.1. The data

The data available for the 1975 validation consisted of

(a) a 44 x 18 origin destination matrix of actual patient flows in 1975, the destinations covering the portion of the region served by the North West Thames Regional Health Authority (i.e., the northwest quadrant in Figure 2A)

(b) the total number of hospital admissions generated in the 40 origin zones covered by the four Thames Regional Health Authorities (i.e., excluding origin zones 41 to 44 in Table 1)

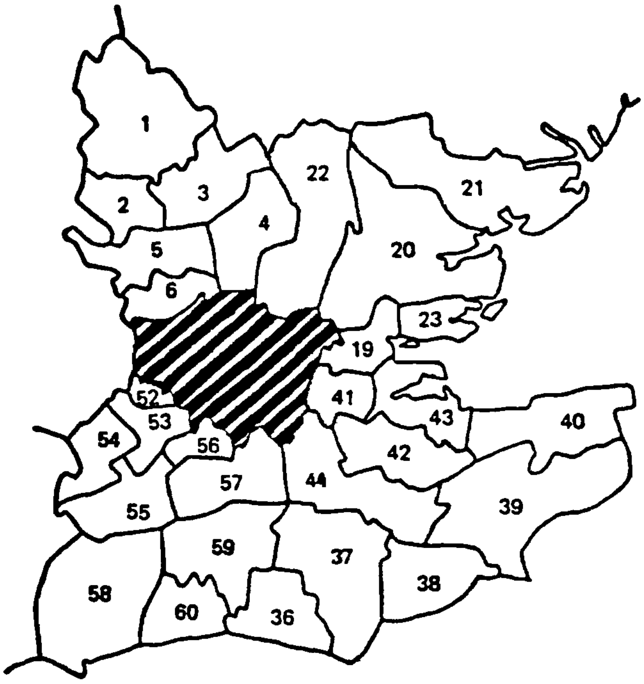
A) Southeast England



B) Greater London Council (GLC)

Figure 1. Model 1 origin zones. Key on page 9; zone 44 (rest of England) is not shown.

A) Southeast England



B) GLC

Figure 2. Model 1 destination zones. Key on page 9; zone 69 (other RHAs) is not shown.

Table 1. Key to Figure 1.

Origin		Destination			
1	Barnet	1	N Bedfordshire	45	Bexley
2	Brent	2	S Bedfordshire	46	Greenwich
3	Harrow	3	N Hertfordshire	47	Bromley
4	Ealing	4	E Hertfordshire	48	St Thomas'†
5	Hammersmith	5	NW Hertfordshire	49	Kings'
6	Hounslow	6	SW Hertfordshire	50	Guys'
7	Hillingdon	7	Barnet*	51	Lewisham
8	Kens + Chelsea	8	Edgware*	52	N Surrey
9	Westminster	9	Brent	53	NW Surrey
10	Barking	10	Harrow	54	W Surrey
11	Havering	11	Hounslow	55	SW Surrey
12	Camden	12	S Hammersmith	56	Mid Surrey
13	Islington	13	N Hammersmith	57	E Surrey
14	City	14	Ealing	58	Chichester
15	Hackney	15	Hillingdon	59	Crawley
16	Newham	16	K/C/W NW*	60	Worthing
17	Tower Hamlets	17	K/C/W NE	61	Croydon
18	Enfield	18	K/C/W S	62	Kingston
19	Haringey	19	Basildon	63	Roehampton
20	Redbridge	20	Chelmsford	64	Wandsworth
21	Waltham Forest	21	Colchester	65	Sutton
22	Bexley	22	Harlow	66	Oxfore
23	Greenwich	23	Southend	67	E Anglia
24	Bromley	24	Barking	68	Wessex
25	Lambeth	25	Havering	69	Other RHAs
26	Lewisham	26	N Camden		
27	Southwark	27	S Camden		
28	Croydon	28	Islington		
29	Kingston	29	City		
30	Richmond	30	Newham		
31	Merton	31	Tower Hamlets		
32	Sutton	32	Enfield		
33	Wandsworth	33	Haringey		
34	Bedfordshire	34	E Roding		
35	Hertfordshire	35	W Roding		
36	Essex	36	Brighton		
37	E Sussex	37	Eastbourne		
38	Kent	38	Hastings		
39	Surrey	39	SE Kent		
40	W Sussex	40	Thanet		
41	Oxford	41	Dartford		
42	E Anglia	42	Maidstone		
43	Wessex	43	Medway		
44	Other	44	Tunbridge		

*K/C/W = Kensington, Chelsea, and Westminster

†Destinations 48,49,50 are named after teaching hospitals within the districts.

3.1.2. Changes in model inputs 1975 and 1977

Table 2 provides an indication of the change in the main input variables, W_i and D_j, that occurred between 1975 and 1977. It shows that the patient generating factors, W_i, were smaller in 1975, which was partly a reflection of the lower national hospital utilization rates at that time. It also shows some interesting geographical variations in W_i with the largest increases (5% to 10%) occurring in the peripheral parts of the region. These are mostly an indication of the growth in the elderly population over this period in these areas, although the long-term trend for a deconcentration of people from the central area contributes to this difference.

For the resource variable, D_j, the proportionate changes in values are much larger (-30% to +16%), with the biggest increases concentrated outside the London area. Caseload capacities, the resource measures, are a function of capital developments, trends in treatment (length of hospital stay), differential utilization rates in each clinical specialty, manpower availability, and other factors.

Finally, Table 3 shows, for a sample of origins, the percentage changes in hospitalization rates (the number of hospital admissions per thousand resident population) -- one of the principal variables that we would like the model to predict accurately. The values indicate a considerable distributional change (-21% to +7%), which suggests that they should provide a good test for the model. A closer examination of this table also reveals the important observation that some of the origin zones in the inner London area had higher hospitalization rates in 1977 than in 1975 despite lower patient generating factors, indicating some important dependency of demand on supply that we would also like the model to predict. Thus, it may be concluded from these tables that the 1975 data will provide a very suitable basis for the main investigation.

Table 2. Changes in input variables: 1975 validation compared with 1977 (calibration year).

PATIENT GENERATING FACTORS (W_i)

Zone number	Area of Residence, i	1975 patient generating factor as percentage of 1977 patient generating factor		
5	Hammersmith	99		
13	Islington	98		
33	Wandsworth	97		
22	Bexley	95		
37	East Sussex	95		
12	Camden	94		
35	Hertfordshire	93		
		Average:	Inner London	97
			Outer London	95
			Other*	94

AVAILABLE RESOURCES (D_j)

Zone number	Health District, j	1975 capacity as percentage of 1977 capacity		
27	South Camden	116		
61	Croydon	104		
5	North West Herdfordshire	97		
64	Wandsworth	95		
25	Havering	90		
28	Islington	86		
45	Bexley	72		
37	Eastbourne	70		
		Average:	Inner London	94
			Outer London	91
			Other**	88

*Zones 34, 35, 36, 37, 39, 40.
**Zones 1-6, 19-23, 36-44, 52-60.

Table 3. Change in output variables: 1975 validation compared with 1977 (calibration year).

HOSPITALIZATION RATES

Zone number	Area of Residence	1975 hospitalization rate as percentage of 1977 hospitalization rate
5	Hammersmith	107
17	Tower Hamlets	102
28	Brent	96
2	Surrey	91
31	Islington	89
11	Havering	89
22	Bexley	86
37	East Sussex	85
31	Merton	82
3	Harrow	79
		Average: Inner London 96
		Outer London 89
		Other 89

3.1.3. Reproduction of the patient flow matrix

Table 4 gives several statistics showing the goodness-of-fit of the predicted 1975 flow matrix $\{T_{ij}\}$ to that observed $\{\bar{T}_{ij}\}$ as described in section 3.2. The test carried out to produce these results was based on a regression analysis of flows between origin-destination pairs predicted by the model and those that were actually observed. The most important statistics shown are R^2, the coefficient of explanation, and the slope and intercept of the regression. When R^2 and the slope equal one and the intercept is zero, a perfect correspondence is indicated between the model predictions and reality (Mayhew and Taket, 1980). As is seen, the realized values match these criteria very well.

The 1977 calibration statistics are also included for comparative purposes. The results for both dates are thus in close correspondence, suggesting that the model performs very well with respect to these measures and is successful in back-predicting the flow matrix.

3.1.4. Reproduction of hospitalization rates

The second level of validation concerns the model's ability to recreate the 1975 hospitalization rates. Contained in Table 5 is a list of the actual rates by origin zone and those predicted by the model. In Figure 3 the results are plotted with the 10% error margins also added. As is seen, errors in 30 out of the 39 zones shown are less than 10%, while overall the absolute percentage error is only 6%. This compares very favorably with the calibration year model in which 32 out of the same 39 zones had less than 10% error and where the average absolute error was 5.7%. The results of these two tests--based on back-predicting the flows and rates --are thus highly satisfactory, the model performing almost identically in 1975 as it did in the calibration year, 1977. We shall now seek further improvements by testing alternative specifications of the model inputs.

Table 4. Reproduction of section of 1975 trip matrix.

	1975 goodness-of-fit over destinations in northwest quadrant	1977 calibration statistics
Coefficient of explanation, R^2	0.9626	0.983
Slope of regression line	0.9766	1.001
Intercept of regression line	10.05	12.30
Root mean square error*	325.2	226.4
Mean absolute error**	114.0	79.3
Mean absolute % error***	137.7%	118.5%

* $$\mathrm{RMSQ} = \left[\sum_i \sum_j \frac{(\bar{T}_{ij} - T_{ij})^2}{N}\right]^{\frac{1}{2}}$$

** $$|e| = \sum_i \sum_j \frac{|\bar{T}_{ij} - T_{ij}|}{N}$$

*** $$|pe| = \sum_i \sum_j \frac{|\bar{T}_{ij} - T_{ij}|}{\bar{T}_{ij}} \times \frac{100}{M}$$

such that $N_{ij} \neq 0$

where $M = \sum_i \sum_j (N_{ij} \neq 0)$

Table 5. 1975 validation.

Zone number	Area of Residence	Hospitalization Rates Actual	Model	%Error	Overall average absolute % error = 6.0%
5	Hammersmith	145.6	111.0	-23.8	Inner London: average absolute % error 7.5
17	Tower Hamlets	141.5	135.0	- 4.6	
9	Westminster	133.7	135.2	1.1	
33	Wandsworth	124.5	124.7	0.2	
8	Kensington and Chelsea	122.9	135.5	10.3	
14,15	City and Hackey	122.3	137.7	12.6	
13	Islington	121.5	140.2	15.4	
12	Camden	118.3	110.6	- 6.5	
27	Southwark	117.0	116.8	- 0.2	
19	Haringey	113.2	125.9	11.2	
16	Newham	109.3	112.2	2.7	
25	Lambeth	106.7	104.1	- 2.4	
26	Lewisham	106.0	112.5	6.1	
2	Brent	124.6	133.5	7.1	Outer London: average absolute % error 5.5
23	Greenwich	113.3	110.9	- 2.1	
1	Barnet	107.1	112.6	5.1	
6	Hounslow	102.4	101.0	- 1.4	
21	Waltham Forest	102.4	113.2	10.5	
10	Barking	102.1	94.9	- 7.1	
18	Enfield	99.1	99.6	0.5	
4	Ealing	96.6	108.7	12.5	
3	Harrow	95.9	95.9	0.0	
24	Bromley	94.6	98.1	3.7	
31	Merton	94.5	107.3	13.5	
22	Bexley	91.4	94.5	3.5	
32	Sutton	90.4	93.6	3.5	
30	Richmond	89.9	94.7	5.3	
28	Croydon	89.4	85.9	- 3.9	
29	Kingston	85.7	85.3	- 0.5	
11	Havering	85.2	74.4	-12.7	
20	Redbridge	83.7	79.9	- 4.5	
7	Hillingdon	81.4	75.3	- 7.5	
39	Surrey	93.5	100.0	7.0	Other: average absolute % error 4.4
28	Kent	84.5	85.6	1.3	
36	Essex	83.1	78.4	- 5.7	
35	Hertfordshire	81.1	88.4	9.0	
37	E Sussex	80.0	77.9	- 2.6	
40	West Sussex	77.2	80.1	3.8	
34	Bedfordshire	71.5	72.3	1.1	

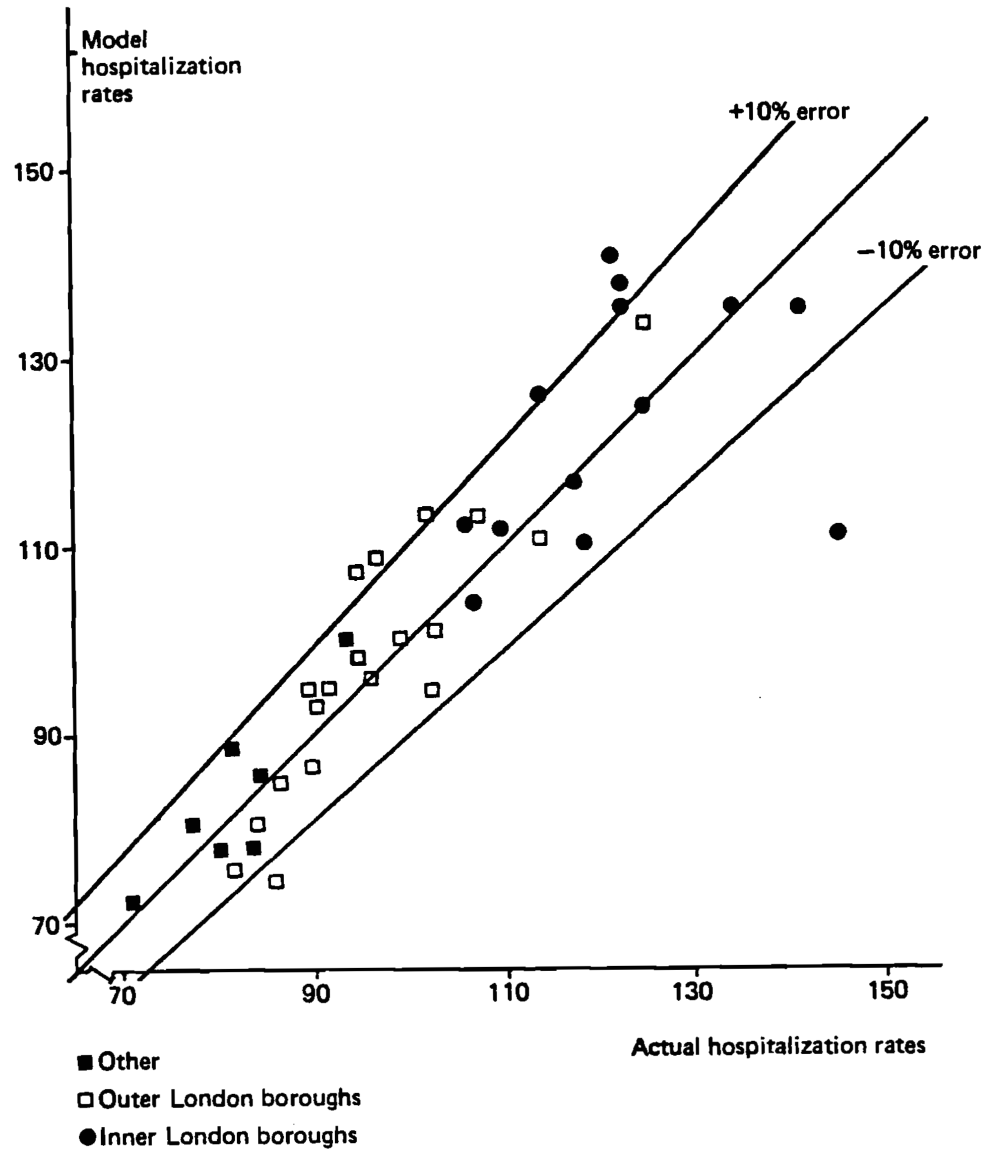


Figure 3. Plot of 1975 hospitalization rates predicted by the model on actual hospitalization rates.

3.2. Alternative Model Specifications

Thus far the model has been validated against 1975 data. In this section slightly different model specifications are attempted to check whether the prediction errors can be further reduced. Consideration is focused on W_i, the patient generating factor, and on f_{ij}, the deterrence function.

3.2.1. The patient generating factor

The propensity to use health care services is mostly a function of age and sex, but it is also believed to be influenced by social, economic, environmental, and other factors. In investigations (LHPC, 1979) it has been shown that death ratios are highly correlated with key social and economic indicators of deprivation. The proposal, therefore, is to modify the existing W_i by a zone-specific death ratio and then to re-run the model to see whether better predictions result. A death ratio in zone i is calculated from routinely published statistics at time t as follows

$$R_i(t) = \frac{ACF_i(t) R_i^*(t)}{R(t)} \tag{3}$$

where

$R_i^*(t)$ = the crude death rate in origin zone i. Deaths in psychiatric or other long stay institutions are apportioned over all areas of the country according to the sizes of the non-institutional populations before the calculation of this rate

$R(t)$ = the national death rate

$$ACF_i = \frac{\sum_l f_l(t)P_l(t)/P(t)}{\sum_l f_l(t)P_{il}(t)/P_i(t)} \qquad (4)$$

where ACF_i = the area comparability factor for place of residence i

f_l = the national death rate in age-sex category l

P_l = the national population in category l

P_{il} = the population in origin zone i category l

and where

$$P(t) = \sum_l P_l(t)$$

$$P_i(t) = \sum_l P_{il}(t)$$

The death rate $R_i(t)$ is hence a type of standardized mortality ratio whose use is hypothesized to reflect those regional variations in patient generating potential unaccounted for by the local age and sex structure. It is applied by multiplying it with the existing value of $W_i(t)$ as follows

$$W_i^*(t) = R_i(t)\ W_i(t) \qquad (5)$$

New calibration and prediction runs of the model were carried out using the modified vector of generating factors. Table 6 shows the calibration statistics, whereas Table 7 gives the broad results of the back-predictions with additional comparisons for 1967.

Table 6. Comparison of calibration statistics.

	Method of calculation of pgfs	
	with death ratios	no death ratios
Parameter	0.363	0.367
Flow matrix statistics		
R^2	0.981	0.983
Slope of regression b	1.0000	1.0010
Intercept a	11.65	12.30
Root mean square error	239.8	226.4
Mean absolute error	83.4	79.3
Mean absolute % error	119.1%	118.5%
Hospitalization rate statistics		
Mean absolute error	6.5	5.7
Mean absolute % error	5.6%	5.0%
Number of origins with <10% error	34	38

†The calibration procedure is fully described in Mayhew and Taket (1980). Briefly, the predicted flows are regressed on the observed. The parameter value β is systematically adjusted until the slope of the regression b equals one.

Table 7. Comparison of prediction runs.

Run/Statistic	Method of calculation of pgfs: with death ratios	no death ratios
*1967 Hospitalization rates**		
Mean absolute error*	13.9	13.5
Mean absolute % error*	14.8	14.5
1975 Hospitalization rates		
Mean absolute error	6.8	6.4
Mean absolute % error	6.6%	6.0%
Trip Matrix (destination in north-west quadrant only)		
R^2	0.962	0.963
Root mean square error	325.9	325.2
Mean absolute error	115.3	114.0
Mean absolute % error	139.5%	137.7%

*Calculated over origins in northwest quadrant. These were the only origin zones for which actual 1967 data were available.

As is seen, death ratios make almost no difference to the goodness-of-fit statistics in the calibration run. As for the exercise in back-prediction, the errors are marginally worse at both times. The conclusion, therefore, is that death ratios do not add to the explanatory power of the model, and that if social, economic, and other factors do alter the propensity to use hospital services, then death ratios are not a good way of representing them.

3.2.2. A derived deterrence function

If an actual flow matrix $\{\bar{T}_{ij}\}$ is available then it becomes possible to derive the deterrence function $\{f_{ij}\}$ directly. In conventional calibrations of the model, it is more normal to work with a cost matrix $\{c_{ij}\}$ and hence with particular functional forms for f_{ij}, for example $\exp(-\beta c_{ij})$ or $c_{ij}^{-\beta}$. An advantage of the first approach, however, is that it enables a user always to obtain a "perfect fit" to the calibration year data. If it is assumed that the empirically derived $\{f_{ij}\}$ remains constant over time, then the model can be used for prediction in the usual way. More importantly, it is a reasonable assumption that, if changes do occur in f_{ij}, they will almost certainly be smaller than those occurring either in D_j or in W_i. Because of the "perfect fit" property, therefore, validation tests assume a great importance, enabling the user to test rigorously different model specifications.

To obtain such a deterrence function, $\{f_{ij}\}$ must be defined, a deterrence matrix, where f_{1j} is abitrary and has no unit of measurement. Then using simple substitution, we get

$$f_{ij} = \frac{\bar{T}_{ij} W^1 f_{1j}}{W_i \bar{T}_{1j}} \qquad i \neq 1 \tag{6}$$

where

$\overline{T}_{ij}$ = the observed flow from i to j

$\overline{T}_{1j}$ = the observed flow from origin zone 1 to j

[If $\overline{T}_{ij} = 0$, set $\overline{T}_{ij}$ to some small number, here $0.4/\sum_i 1$. This is to ensure $f_{ij} \neq 0$, and to avoid problems with a zero denominator in (6).]

W_j, W_1 = the pgf in zone i and zone 1 respectively

and where f_{1j}, the first element in each row of the deterrence matrix is fixed arbitrarily to a suitable value > 0.

Using the above method, two sets of deterrence functions (A and B) were obtained for calibration year data

(i) Function A based on the usual pgfs, i.e.,

$$W_i = \sum_l \sum_m P_{il} U_{lm} \tag{7}$$

(ii) Function B based on the use of death ratios, i.e.,

$$W_i^* = R_i W_i \tag{8}$$

The model was then re-run using 1975 and 1967 data and the outputs were compared with what actually occurred. The results are shown in Table 8. These indicate a significant improvement in accuracy at both times over the results obtained with the conventional calibration procedure (Table 4). They also show that the inclusion of death ratios (Function B) tends to detract from the explanatory power of the model, confirming the results of section *3.2.1.* above. The conclusions of these respecification procedures are hence threefold: first, age and sex are confirmed as the dominant criteria influencing the potential demand for health care services; second, the

Table 8. Prediction runs using empirically derived deterrence functions.

Model run/statistic	Function A no death ratios in pgfs	Function B death ratios in pgfs
*1967 Hospitalization rates**		
Mean absolute error*	8.7	10.5
Mean absolute % error*	9.1%	10.8%
1975 Hospitalization rates		
Mean absolute error	3.5	4.0
Mean absolute % error	3.2%	3.7%
Flow matrix (destinations in north-west quadrant only)		
R^2	0.989	0.988
Root mean square error	178.2	182.1
Mean absolute error	46.1	48.6
Mean absolute % error	96.6	98.9

*Calculated over origins in northwest quadrant only. These were the only origin zones for which actual 1967 data were available.

effects of socio-economic factors on additional unexplained variations in the use of health services cannot be described using death ratios; and three, the enhanced accuracy of the model using derived deterrence functions indicate that there is scope for improving the specification of $\{c_{ij}\}$, the cost matrix, as used in conventional calibration methods (Mayhew and Taket, 1980).

3.3. Further Error Analysis

One of the findings of the calibration analysis described in Mayhew and Taket (1980) was a tendency for the model to over-predict hospitalization rates in the inner urban zones. When the errors resulting from the 1975 back-prediction were closely examined, this bias seemed to recur in the same form, thus raising two questions for research.

1) Can the input variables, both pgfs and accessibility costs, be improved to remove the source of this bias?

2) Given the apparently consistent nature of the biases, is it possible to derive empirically based correction factors that can remove them?

The first question was partially dealt with earlier in section 3, and currently more research is in progress to identify improved measures of both potential demand (W_i) and accessibility costs (c_{ij}). We now examine the second possibility in more detail.

Bias in the prediction of hospitalization rates

Figure 4 gives a comparison of the actual change in hospitalization rates by origin zone with that predicted by the model. It is apparent in this diagram that the model correctly predicts the direction of change (and usually the percentage too) in most cases (two serious exceptions are zones 12 and 17). The absolute values, however, are often wrong, though not by very much. As noted in section 2, the magnitude of the prediction errors are very similar to those in the calibration stage, implying therefore, that errors in calibration will be repeated during prediction runs. Figure 5, a plot of 1975 errors on those in 1977, shows a marked correlation ($r = 0.80$), substantiating this hypothesis. A similar exercise using 1967 data gave a comparable result ($r = 0.81$). The conclusion is, therefore, that until more research is available that improves the specification of input variables, there seems to be an empirical basis for making small adjustments to the model outputs in order to improve further the accuracy of the predictions.

4. SENSITIVITY ANALYSIS

Sensitivity analysis consists of examining the changes in the model outputs when perturbations are made to the input variables and parameters. Unlike the validation tests, sensitivity analysis is concerned with the theoretical behavior of the outputs when the model is exposed to extremes of change rather than with the accuracy of the predictions. For current purposes the sensitivities analyzed are with respect to the patient generating factors, resources and parameter value, β. The first two are of direct concern to health care planners reflecting the dimensions of demand and supply, whereas the third, the sensitivity of β to change, is important from the standpoint of the model's assumptions (section 3).

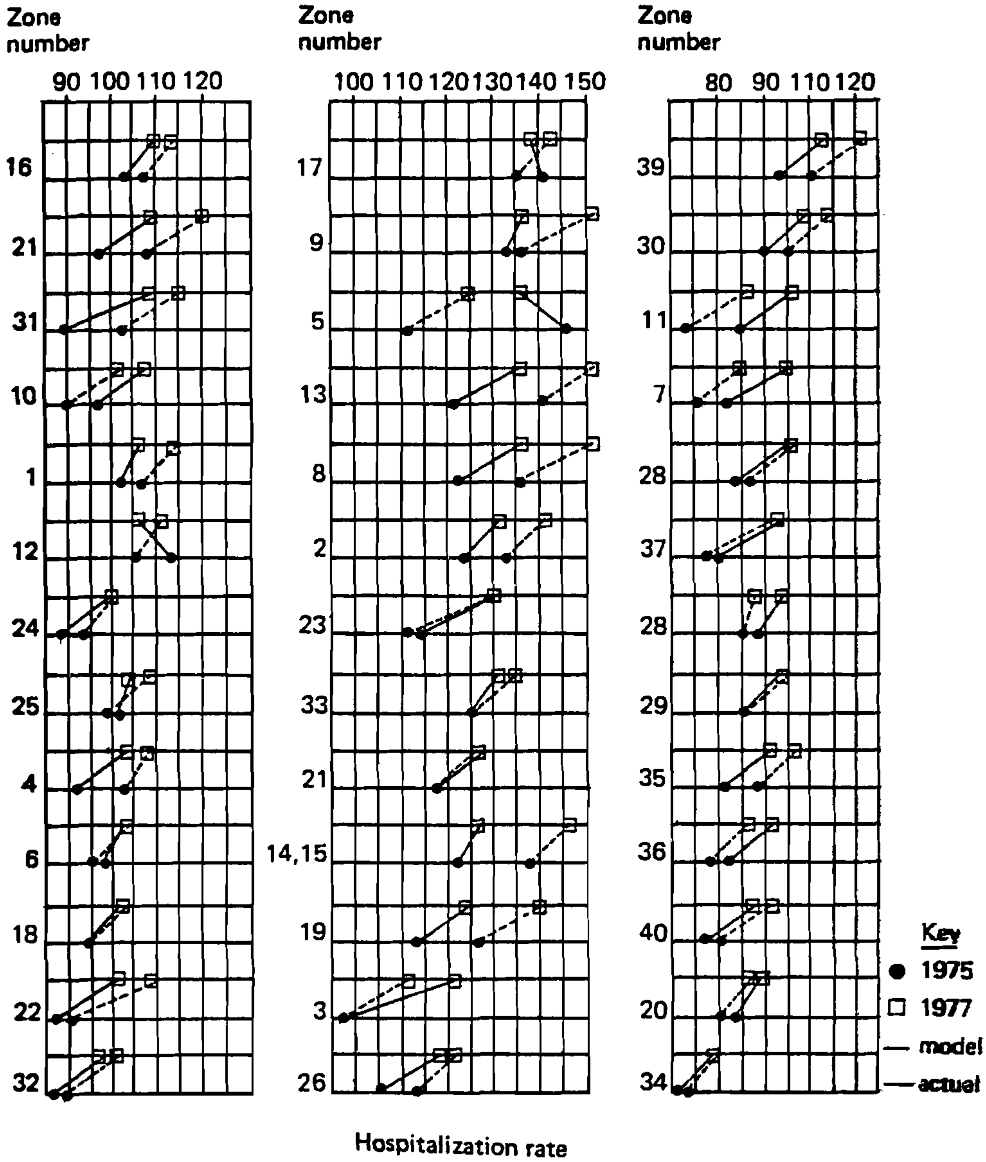


Figure 4. Changes in hospitalization rates 1975 to 1977: "model" and "actual". Rates, on the horizontal axis, are in cases per thousand.

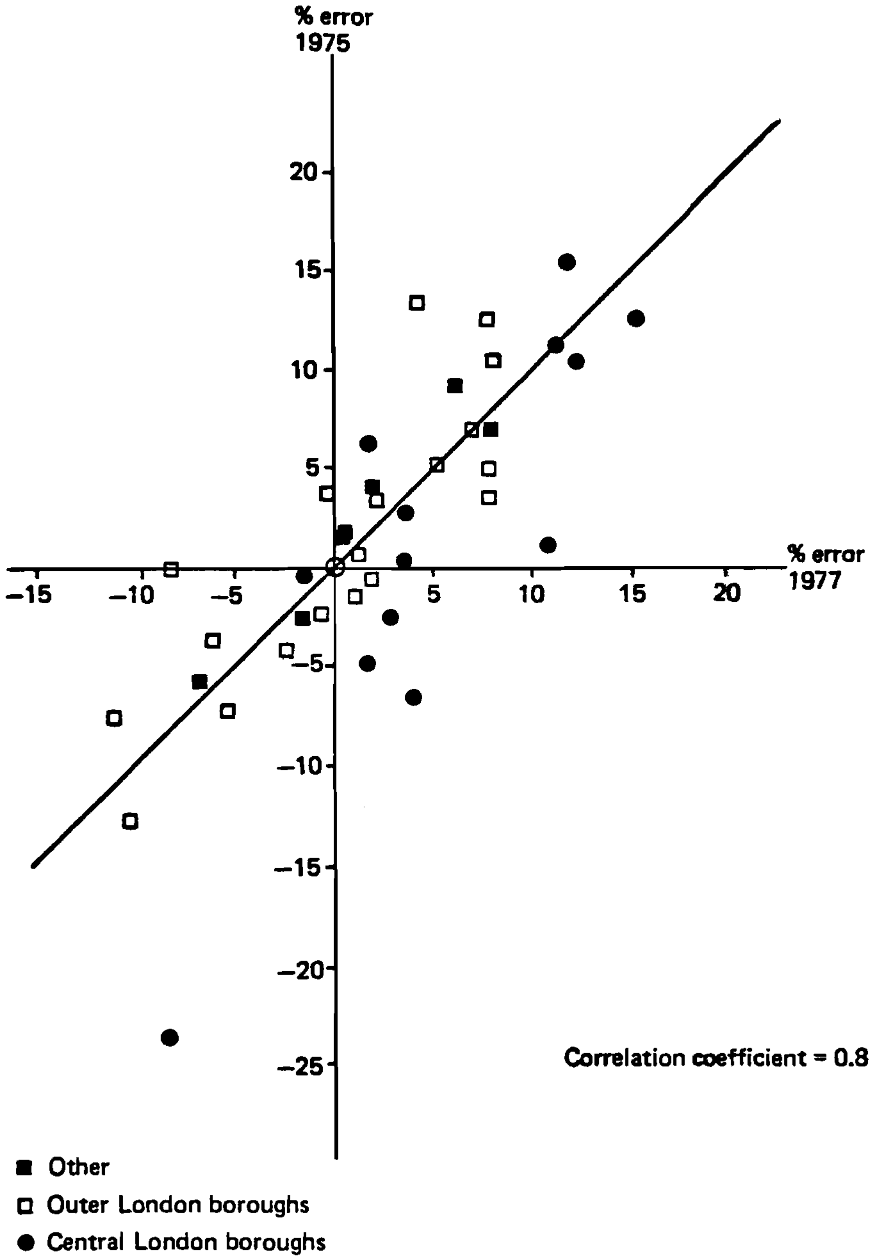


Figure 5. A plot of prediction errors in hospitalization rates in 1975 on those obtained for the calibration year, 1977.

4.1. Patient Generating Factors

The user of the model is concerned to know how changes in values of the pgfs affect the number of patients generated in a zone ($\sum_j T_{ij}$). There are three cases to consider.

1) The effect on $y_i (= \sum_j T_{ij})$ caused by a change in W_i

2) The effect on y_i caused by a change in W_k, $k \neq i$

3) The effect on y_i caused by simultaneous changes in all W_i

Case 1: From (1), summing over j, y_i may be written in the form

$$y_i = \sum_j \frac{W_i D_j f_{ij}}{W_i f_{ij} + \sum_{\substack{l \\ l \neq i}} W_l f_{lj}} \qquad (9)$$

The first derivative of (9), $y_i'(W_i)$ (i.e., dy_i/dW_i) is positive; the second, $y_i''(W_i)$, (i.e., $d^2 y_i/dW_i^2$) is negative ($0 < W_i < \infty$). At infinity there is an upper bound given by $\sum_j D_j$; at this point, then, i theoretically commands all the resources in the system. These facts describe a concave function of the saturation type.

Some examples for different values of i are shown in Figure 6. An examination shows that zones peripheral to the metropolitan center increase rapidly for small W_i, but with further increases the rate of growth drops sharply (e.g., 36, 34, 37). The centrally positioned zones in contrast experience

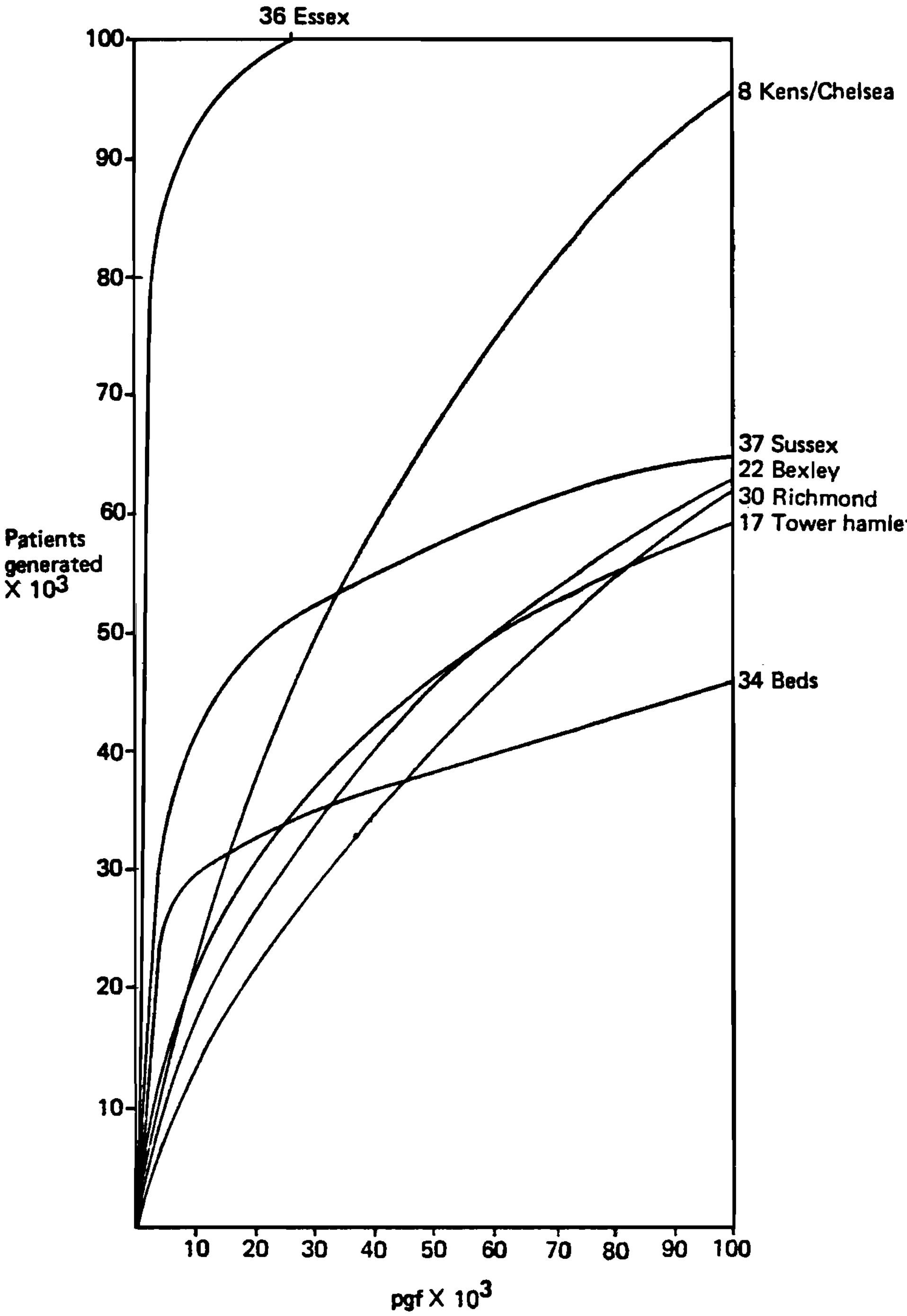


Figure 6. Case 1: Sensitivity of numbers of cases generated to changes in W_i, the pgf, for different origin zones.

a slower rate of growth initially but a slackening off in these cases is not apparent in the range of W_i considered (zones 8, 17, 22, 30). These results are consistent with the differential patterns of facility access one expects in urban and non-urban areas. In the urban case, access is better but the spatial competition for resources is more intense; in the non-urban case, there is less external competition from other zones, but the populations are more highly dependent on their local facilities.

Case 2: From (1), summing over j, y_i may be written as a function of W_k, $k \neq i$

$$y_i(W_k) = \sum_j \frac{W_i D_j f_{ij}}{W_k f_{kj} + \sum\limits_{\substack{l \\ l \neq k}} W_l f_{lj}} \qquad (10)$$

Here, the value of y_i goes to zero as W_k increases to infinity. Thus, the number of patients generated by a zone declines when there is an increase in the pgfs of another zone and where all zones compete for the same resources. Conversely, there is an increase in y_i when the pgf of another zone declines. The size of the change is governed also by the values of f_{kj}, and f_{ij}. If k is remote from i, y_i will -- other values being constant -- change less than if k is near. This is seen from an inspection of the expression for $y_i'(W_k)$, which is

$$y_i'(W_k) = - W_i \sum_j \frac{D_j f_{ij} f_{kj}}{(W_k f_{kj} + \sum\limits_{\substack{l \\ l \neq k}} W_l f_{lj})^2} \qquad (11)$$

In equation (11), f_{ij} and f_{kj} are largest when i and k are close to j and thus close to each other. Hence, changes in y_i in these situations will be relatively greater.

Case 3: This is the most complex case, and it is difficult to make general statements about it except when all the changes take place in one particular direction. This complexity is due to the extensive interaction effects that occur in the system that the model is attempting to simulate. An illustration of this difficulty is given if we try to evaluate small change in y_i by considering the total differential of y_i. This change, dy_i, is

$$dy_i = \sum_j B_j D_j f_{ij} dW_i - W_i \sum_k \sum_j D_j B_j^2 f_{ij} f_{kj} dW_k \tag{12}$$

where dW_k is the change of W_k. Clearly, dy_i is dependent in many other interactions taking place elsewhere in the system, interactions that are reflected by the second term in (12).

4.2. Resources

The effect on the predicted number of patients generated in a zone due to changes in resource levels is more straightforward. From (1), differentiating with respect to D_j,

$$y_i'(D_j) = \frac{W_i f_{ij}}{\sum_i W_i f_{ij}} \tag{13}$$

Equation (13) is a constant, and it means that growth in i is proportional to the share of the total potential demand on j discounted by accessibility costs. For example, if facilities are expanded in a location near i, the largest proportion of new demand will be generated in the locality of i rather than elsewhere (zone k, say) since almost certainly $W_i f_{ij} > W_k f_{kj}$ providing W_k is of the same order as W_i. The sensitivity of the model to changes in D_j are thus simple and intuitively reasonable.

A useful measure to derive from this property of the model is the elasticity of the hospitalization rate in i with respect to the resource level in j. This is

$$E_{ij} = \frac{D_j}{R_i} \frac{\partial H_i}{\partial D_j} = T_{ij} / \sum_j T_{ij} \tag{14}$$

where H_i is the hospitalization rate for a population P_i

$$R_i = \sum_j \frac{T_{ij}}{P_i} \tag{15}$$

Equation (14) expresses the proportionate change expected in i following a change in the resources in j. It is of particular value in determining a catchment population -- the resident population in a region dependent on a treatment zone -- which is defined as

$$C_j = \sum_i E_{ij} P_i \tag{16}$$

Equation (16) is one of several possible ways of representing catchment populations. This particular one has the advantage of being easily related to the model outputs.

4.3. Discount Parameter, β

The model parameter β is assumed constant in the predictive mode of the model. Thus, it is necessary to test the effects on the model outputs in the event that this assumption breaks down. These effects are not easy to predict as the first derivative suggests

$$y_i'(\beta) = \sum_j T_{ij}(B_j \sum_i c_{ij} W_i e^{-\beta c_{ij}} - c_{ij}) \qquad (17)$$

This result also depends on the form of the deterrence function [here $f_{ij} = \exp(-\beta c_{ij})$]. Some experiments were therefore carried out on hospitalization rates for different zones in the range β = 0 to β = 2.0. This range has been deliberately exaggerated to see how the model performs when it is stretched. (In fact the maximum change that could be expected if the model were recalibrated would only be around ± 0.1.) In interpreting the results, an increasing β is associated with diminishing accessibility as would occur if the real costs of transportation increased. A decreasing value of β would imply the converse. When β is zero, f_{ij} goes to 1.0 (since $e^{-\beta c_{ij}} = 1 \; \forall ij$) and so, as is seen from equation (1), patients will be allocated to treatment zones by their share of the total patient generating potential, $W_i/\sum_i W_i$. Figure 7 shows the results for several urban and non-urban zones. For large β, centrally positioned urban zones (8 and 17) experience sharp increase in rates; less central and non-centrally located zones usually experience decreases. For the range β = 0.2 to β = 0.4, the portion in which some change could be realistically expected, a second diagram is shown (Figure 8). Most sensitive to this variation here are zones 8 and 17, the two most central zones in this sample. This is perhaps not surprising since some difficulty is usually experienced in

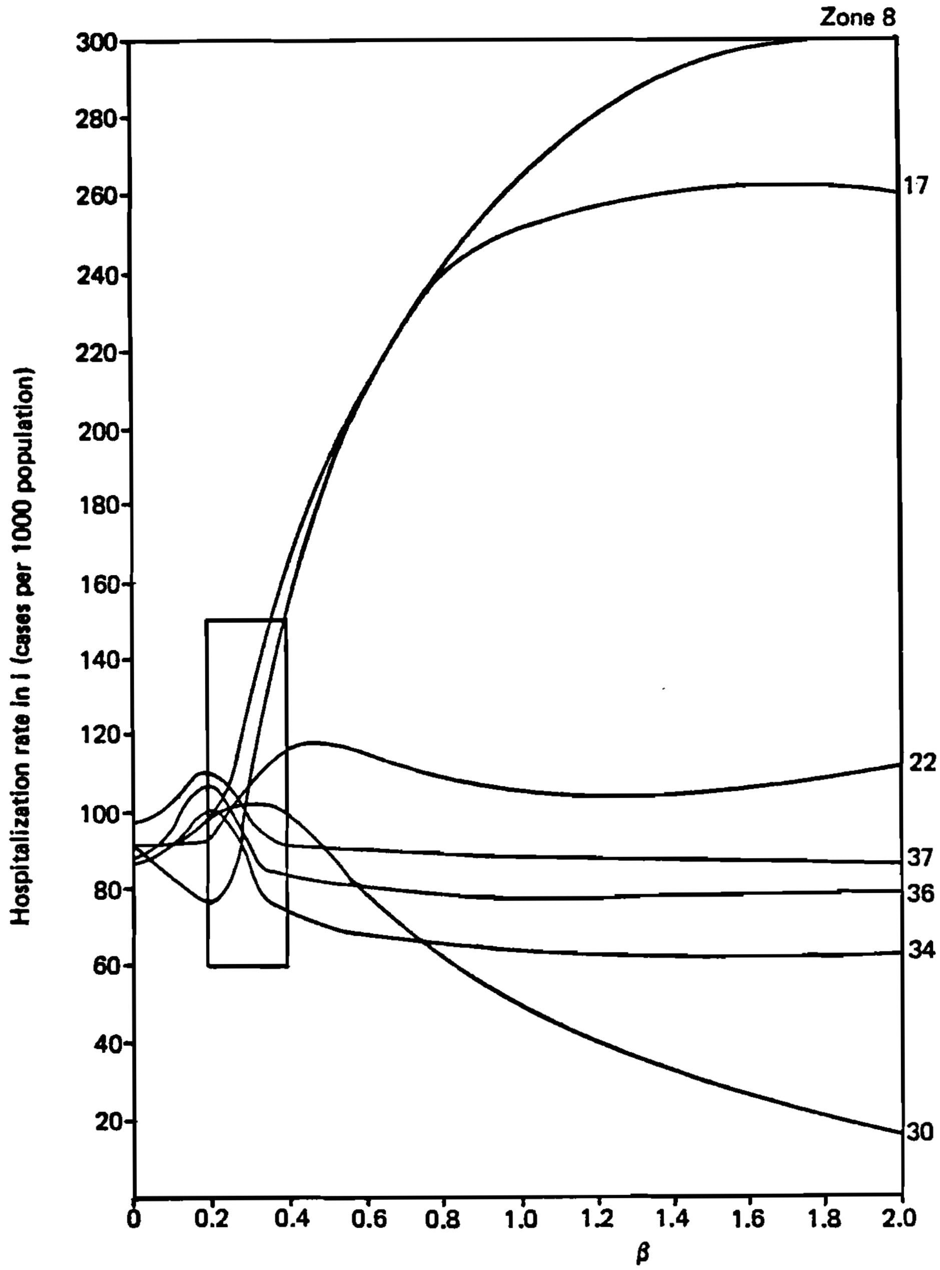


Figure 7. Variation in hospitalization rates in different origin zones as a function of β, the model parameter (see also inset in Figure 8).

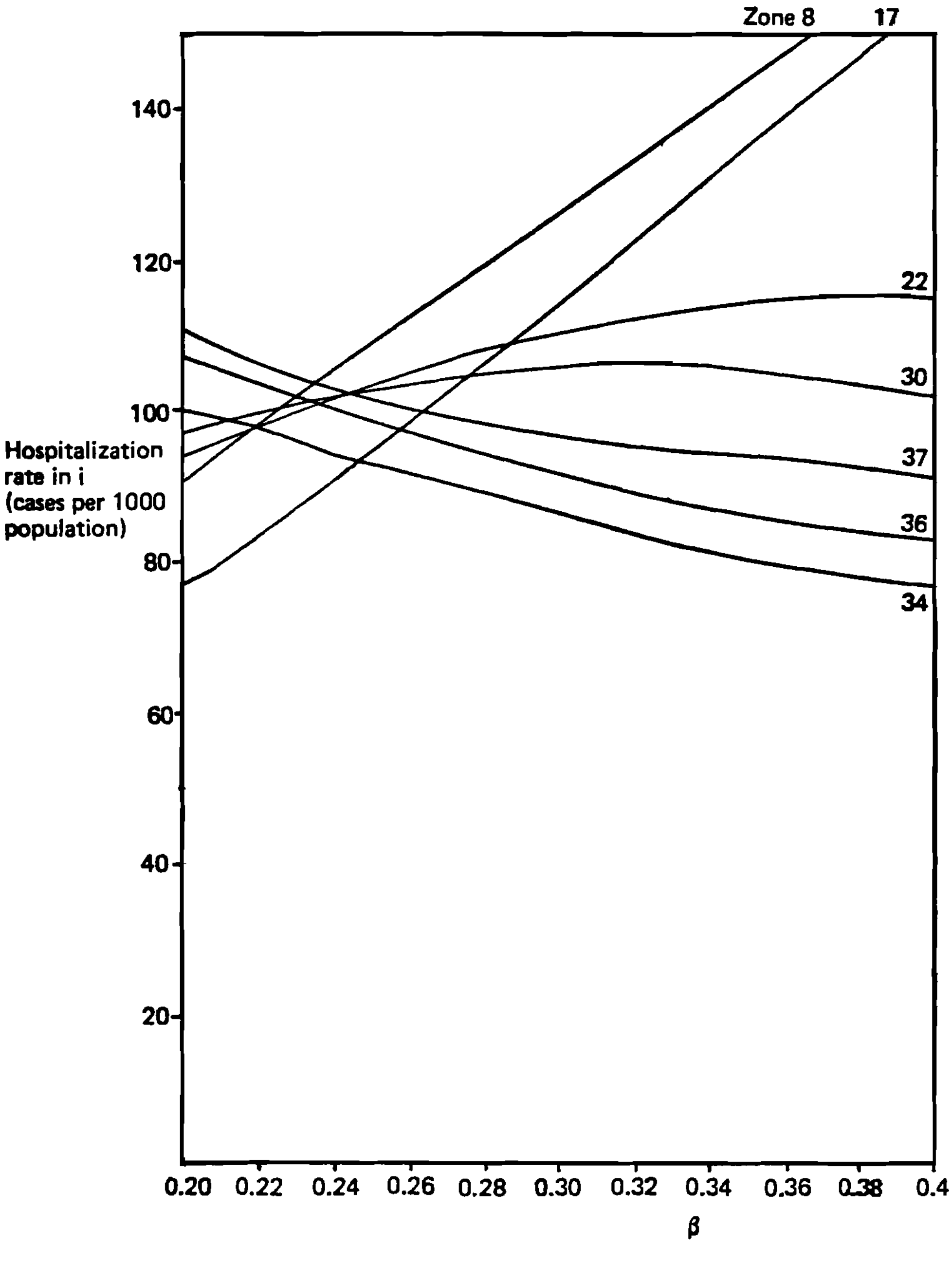


Figure 8. Hospitalization rates in different origin zones as a function of β in the range 0.20 to 0.4.

fitting the model to behavior in inner-urban zones at the calibration stage, and this sensitivity to β is one of the reasons for the difficulty.

5. CONCLUSIONS

This paper has described the results of validation experiments and a sensitivity analysis on the model RAMOS. This model is designed to assist decision makers in the planning of health care services at the regional level. Validation was accomplished in a back-prediction of the state of the system at a point earlier in time. It was found that the model was able to predict the outputs of the system with considerable accuracy but that further improvements were still possible. In the subsequent sensitivity analysis the logic of the model was exposed to small and large variations in the input variables and parameter values. The results were intuitively reasonable, although attention was drawn to the diverse sensitivities of different zones under parameter variation that need to be observed. The basic conclusion is, therefore, that the model achieves the purposes for which it was designed. The question arises whether the model can be used to tackle similar problems in other health care systems. The indications are that it can, although some small respecification may be necessary to take account of local conditions. It is, nevertheless, advisable that other applications should undertake routine validation experiments, since these can uncover aspects for improvement in the specification of the model while providing a check on its predictive power.

REFERENCES

LHPC, P. (1979) *Acute Hospital Services in London.* A profile by the London Health Planning Consortium. Her Majesty's Stationery Office, London, UK.

Mayhew, L.D. (1979) *The Theory and Practice of Urban Hospital Location.* Ph.D. Thesis Birkbeck College, University of London.

Mayhew, L.D., and A. Taket (1980) *RAMOS: A Model of Health Care Resource Allocation in Space.* WP-80-125. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Mayhew, L.D. (1980) *The Regional Planning of Health Care Services: RAMOS and RAMOS*$^{-1}$. WP-80-166. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Mayhew, L.D. (1981) *DRAMOS: A Multi-Category Spatial Resource Allocation Model for Health Service Management and Planning.* WP-81-39. Laxenburg, Austria: International Institute for Applied Systems Analysis.

RAWP (1976) *Sharing Resources for Health in England.* Report of the Resource Allocation Working Party. Her Majesty's Stationery Office, London, UK.

RECENT PUBLICATIONS IN THE HEALTH CARE SYSTEMS TASK

Hughes, D.J. (1979) *A Model of the Equilibrium between Different Levels of Treatment in the Health Care Systems: Pilot Version* (WP-79-15).

Fleissner, P. (1979) *Chronic Illnesses and Socio-Economic Conditions: The Finland Case 1964 and 1968* (WP-79-29).

Shigan, E.N., D.J. Hughes, and P. Kitsul (1979) *Health Care Systems Modeling at IIASA: A Status Report*(SR-79-4).

Rutten, F.F.H. (1979) *Physician Behaviour: The Key to Modeling Health Care Systems for Government Planning* (WP-79-60).

A Committee Report (1979) to IIASA by the participants in an Informal Meeting on *Health Delivery Systems in Developing Countries* (CP-79-10).

Shigan, E.N., P. Aspden, and P. Kitsul (1979) *Modeling Health Care Systems: June 1979 Workshop Proceedings* (CP-79-15).

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

Rousseau, J.M., R.J. Gibbs (1980) *A Model to Assist Planning the Provision of Hospital Services* (CP-80-3).

Fleissner, P., K. Fuchs-Kittowski, and D.J. Hughes (1980) *A Simple Sick-leave Model used for International Comparison* (WP-80-42).

Aspden, P., R. Gibbs, and T. Bowen (1980) *DRAM Balances Care* (WP-80-43).

Aspden, P., and M. Rusnak (1980)*The IIASA Health Care Resource Allocation Submodel: Model Calibration for Data from Czechoslovakia* (WP-80-53).

Kitsul, P. (1980) *A Dynamic Approach to the Estimation of Morbidity* (WP-80-71).

Shigan, E.N., and P. Kitsul (1980) *Alternative Approaches to Modeling Health Care Demand and Supply* (WP-80-80).

Hughes, D.J., and A. Wierzbicki (1980) *DRAM: A Model of Health Care Resource Allocation* (RR-80-115).

Aspden, P. (1980) *The IIASA Health Care Resources Allocation Submodel: DRAM Calibration for Data from the South West Health Region, UK* (WP-80-115).

Mayhew, L., and A. Taket (1980) *RAMOS: A Model of Health Care Resource Allocation in Space* (WP-80-125).

Mayhew, L.D. (1980) *The Regional Planning of Health Care Services: RAMOS and* $RAMOS^{-1}$ (WP-80-166).

Pauly, M.V. (1981) *Adding Demand, Incentives, Disequilibrium, and Disaggregation to Health Care Models* (WP-81-4).

Mayhew, L.D. (1981) *DRAMOS: A Multi-category Spatial Resource Allocation Model for Health Service Management and Planning.* (WP-81-39).