WORKING PAPER

RAMOS: A MODEL OF HEALTH CARE RESOURCE ALLOCATION IN SPACE

L. Mayhew
A. Taket

August 1980
WP-80-125

IIASA
International Institute
for Applied Systems Analysis

RAMOS: A MODEL OF HEALTH CARE RESOURCE ALLOCATION IN SPACE

L. Mayhew
A. Taket

August 1980
WP-80-125

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 eventual aim of health care research at IIASA is to develop a family of submodels replicating components of the health care system in a meaningful way. These models - in the contexts which they are applied - are for use by health planners to assist them in taking rational decisions in what is an extremely complex operating environment. The models developed thus far deal with population, disease prevalence, resource need, resource allocation,and resource supply.

The model presented in this paper comes into the resource allocation category. Known as RAMOS (Resource Allocation Model Over Space), it provides a simple method for choosing between different resource configurations on congested regions (very large urban areas, industrial agglomerations, etc.) when the population size and structure, and the resource availability are changing simultaneously in space and time.

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

Andrei Rogers
Chairman
Human Settlements
and Services Area

ACKNOWLEDGMENTS

The authors are extremely grateful to the four Thames Regional Health Authorities in England for their assistance in this work and for their permission to reproduce the data, and to the Greater London Council (GLC) for supplying data on travel times in London. Thanks go also to Lucy Tomsits, secretary to the Health Care Task, for typing and collating the material.

ABSTRACT

This paper sets out the background and initial results of a resource allocation model called RAMOS. It was developed to explore the consequences on hospitalization rates resulting from one or more of the following: hospital building programs, treatment trends in in-patient care, population changes, or transport developments affecting the accessibility of the population to health care supply. For decision makers the control variables in the model are principally the resource levels in each geographical area of in-patient treatment. A typical question as asked of the model might be: what rearrangement of health care facilities would redress the regional imbalance in health care provision? RAMOS takes as inputs the current or projected morbidity in each area of the region (based on the sex and age structure of the population), a 'test' configuration of health care facilities, and data on patient accessibility. It then outputs the anticipated hospitalization rates by area of residence (admissions per 1000), and other information, so enabling the evaluation of many different allocation plans by the decision maker.

RAMOS is a behavioral model based on extensive data relating to southeast England, an area containing 13.5 million people. It represents a continuation of the work begun in the Department of Health and Social Security in 1979. RAMOS is especially suited to applications in rapidly changing regions, crowded urban settlements, and wherever the locations of health care facilities or of other types of service provision is an important issue.

CONTENTS

RAMOS: A MODEL OF HEALTH CARE RESOURCE ALLOCATION IN SPACE

1. INTRODUCTION AND BACKGROUND

The Health Care Task at IIASA is developing a range of models, each dealing with substantially independent portions of the Health Care System (HCS). These models are designed for use by decision makers and health planners in different countries and at different levels in the decision making process.

One theme developed at IIASA, in conjunction with the Operational Research Service of the Department of Health and Social Security in England, concerns the health care resource allocation process and the interactions which occur between different patient categories and modes of care. This research gave rise to the model DRAM (Gibbs 1978; Hughes and Wierzbicki 1978). The objective of this study is to present the initial findings of another model, which like DRAM,considers the interactions between resource supply and demand, but at a geographical level. More specifically, this model, called RAMOS (Resource Allocation Model Over Space), has been designed to explore the effects between hospitalization rates and patient flow patterns in a region or a country resulting from changes in:

- The number and location of hospital beds in various specialities which result from hospital closures, new developments, or other forms of reorganization
- The population size and structure
- The relative morbidity of the population
- The 'throughput' per bed (i.e., the rate at which hospitals are able to treat patients)
- The availability and efficiency of transport services and car travel over time

Although the model—developed at the Department of Health and Social Security (Mayhew and Taket 1979)—is applied in a United Kingdom context, it is believed that the results will be of much wider interest, as it is known that similar work is being conducted in other IIASA countries both in health and in other fields.

The impetus for this study came as a result of our earlier work on behalf of the London Health Planning Consortium (LHPC 1979). The aim of this work was to identify and quantify in broad terms the level of acute hospital services likely to be needed in various parts of the four Thames Regional Health Authorities (RHAs) which serve London and much of southeast England. The results showed that relative to the remainder of England and Wales, London is over-provided with acute hospital beds. In the report of the Resource Allocation Working Party (RAWP 1976) it is also shown that these four regions are relatively over-provided with financial resources. Furthermore, the population of the inner and outer parts of London has been declining and is expected to decline further, while the population of the counties in the Thames Regions surrounding London is expected to increase. Thus, there is considerable pressure on the London Health Authorities to reduce the level of acute services, and to develop instead services in the counties and services for other groups such as the elderly and mentally handicapped.

In meeting the challenge of providing an efficient hospital system in the 1980s, the four Thames RHAs are obviously concerned that patients do not suffer in the interim and that the costs of implementing plans are kept within resource constraints. The problem facing the RHAs, however, is that it is extremely difficult to know beforehand precisely what effects implementing such measures as hopsital closure or capital developments will have in an area comprising some 13.5 million people. This paper examines whether a model can be developed for the RHAs to deal with these and related problems and if such methodologies can be applied in other countries. The type of model (RAMOS) which has been considered, emerges from a family of gravity models developed elsewhere over many years and is of the singly-constrained kind.

The emphasis is on the specification, calibration, and validation of RAMOS rather than its theoretical basis, since the latter is already well documented (Wilson 1967, 1970, 1971). Two distinct variants are developed and tested using a purpose-written computer program, the details of which will be set out in another paper to be produced at IIASA. The first variant (Model 1) covers southeast England in an area served by the four Thames RHAs (NE, SE, NW, SW); the second (Model 2) covers only the greater London portion of the southeast Thames RHA. This comprises the administrative boroughs of Lambeth, Lewisham, Southwark, Bexley, Greenwich, and Bromley, which form part of the Greater London Council (GLC) region. Following some introductory background in section 2 to the factors affecting hospitalization rates and flow patterns, the model is presented in section 3 and the zoning systems are discussed. In section 4 the variables are defined in detail and certain refinements are made. The calibration procedures, results, and validation of the models are the subjects of sections 5 and 6, while in section 7 some conclusions are drawn.

2. DISCUSSION OF THE FACTORS AFFECTING HOSPITALIZATION RATES AND FLOW PATTERNS IN THE REGION OF INTEREST

The London region is probably unique in having several hundred hospitals dealing to greater or lesser extent with acute medical services as well as playing a very important role in the field of medical education. Figure 1 shows the location of London's acute hospitals within 50 kms of the city center in the period 1901 to 1971. Partly because of their proximity to each

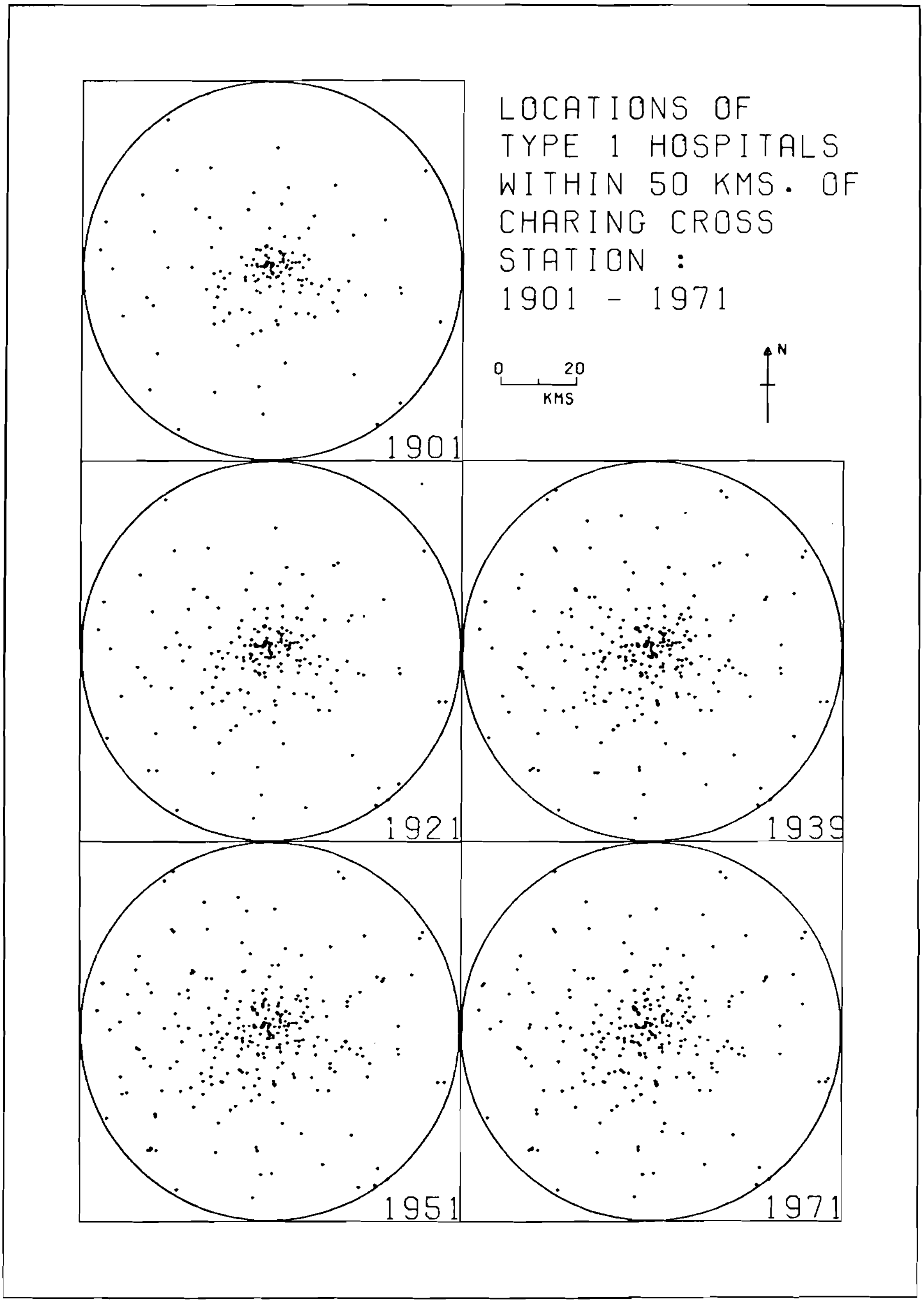


Figure 1. The locations of acute hospitals in London: 1901 - 1971 (Source: Mayhew, 1979).

other and partly because of the sophistication of London's transport system, the hospitals are highly interdependent in terms of the services they provide and the areas they serve. For example, a change in the population of one locality tends not only to affect patient flows to the neighborhood hospital, but also it affects flows to other hospitals nearby and these in turn affect others farther afield, so creating an interaction effect through the system.

While it is probably impossible to know all the reasons why individuals choose or are referred by their general practitioners to particular hospitals, our analysis showed that in the London region (and probably for the UK system as a whole) the bulk of observed patient flows from one area to another could be explained on the basis of three factors: the capacity of hospitals in an area to treat patients, the relative morbidity of the population, and the accessibility of the population to supply. The first factor reflects a generally held view—particularly in countries with free health care services (Feldstein 1965)—that supply fuels demand: whatever is provided gets used. The second factor is determined mostly by the age and sex structure of the population, although certain socioeconomic and environmental considerations are known also to be important (LHPC 1979). The third factor, accessibility, is the tendency for usage to reflect geographical availability. Substantial variations in hospitalization rates exist both regionally and nationally which cannot be accounted for in any other way.

Two empirical illustrations of geographical dependency are shown in Figures 2 and 3. Contained in Figure 2 is a plot of discharge rates against hospital bed availability for the London region in 1977. The correlation between hospital usage and local bed availability is clearly very strong ($r = 0.85$). In Figure 3 a histogram is shown of the percentage variation with distance of patient journey origins to a sample of 14 London hospitals for general medical and surgical specialities. Although it is seen that some patients do travel long distances, it is plain from this diagram that the great majority reside within only a few kilometers of the hospital they use.

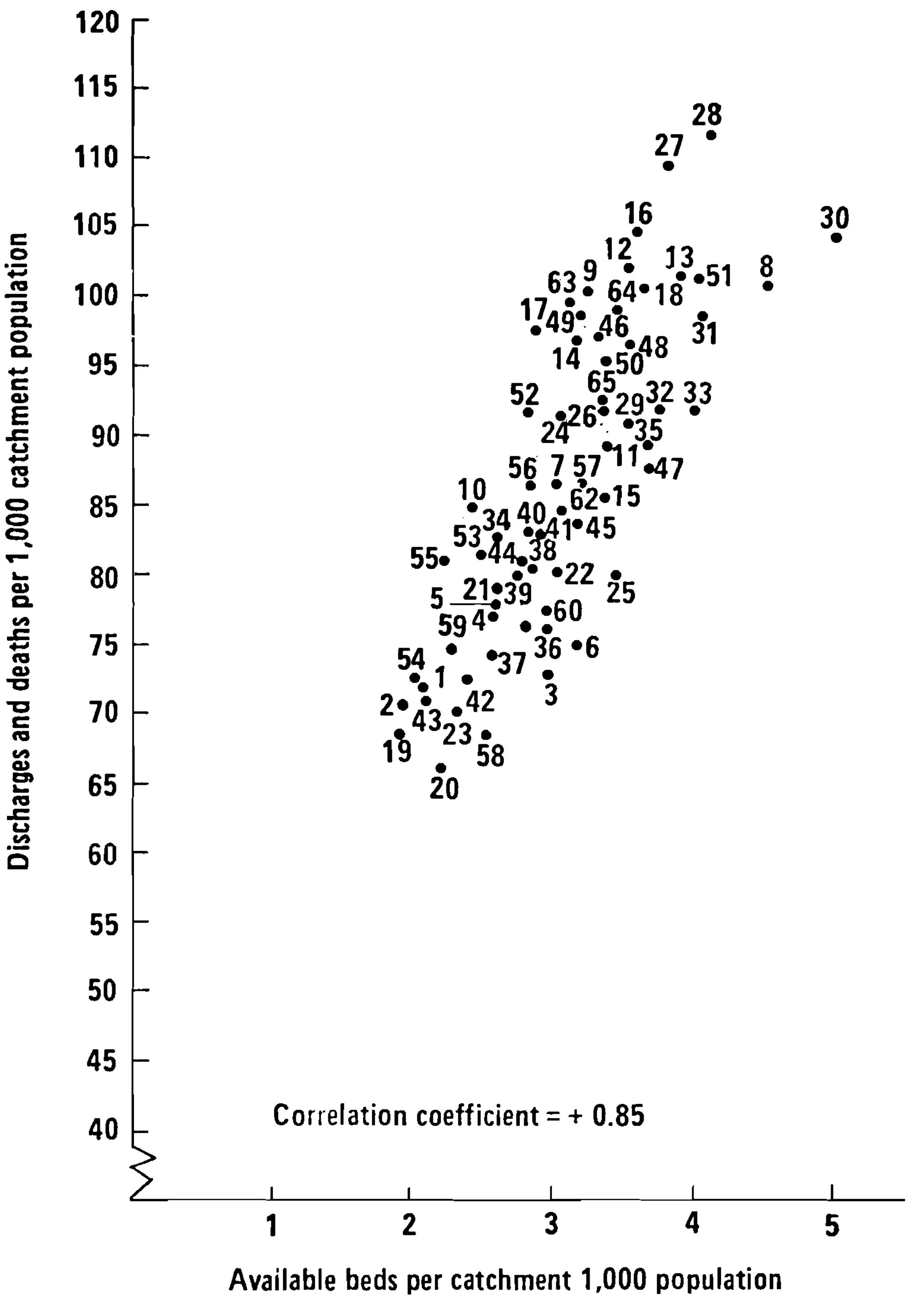


Figure 2. Relationship between hospitalization rate and level of provision in health districts within the four Thames Regions: 1977 number of nonregional acute cases and available beds per 1000 catchment population (a catchment population is the number of persons dependent on a health district. See page 22, Section 4.5.2). The key to numbered health districts is on page 13. (Source: LHPC,1979:26).

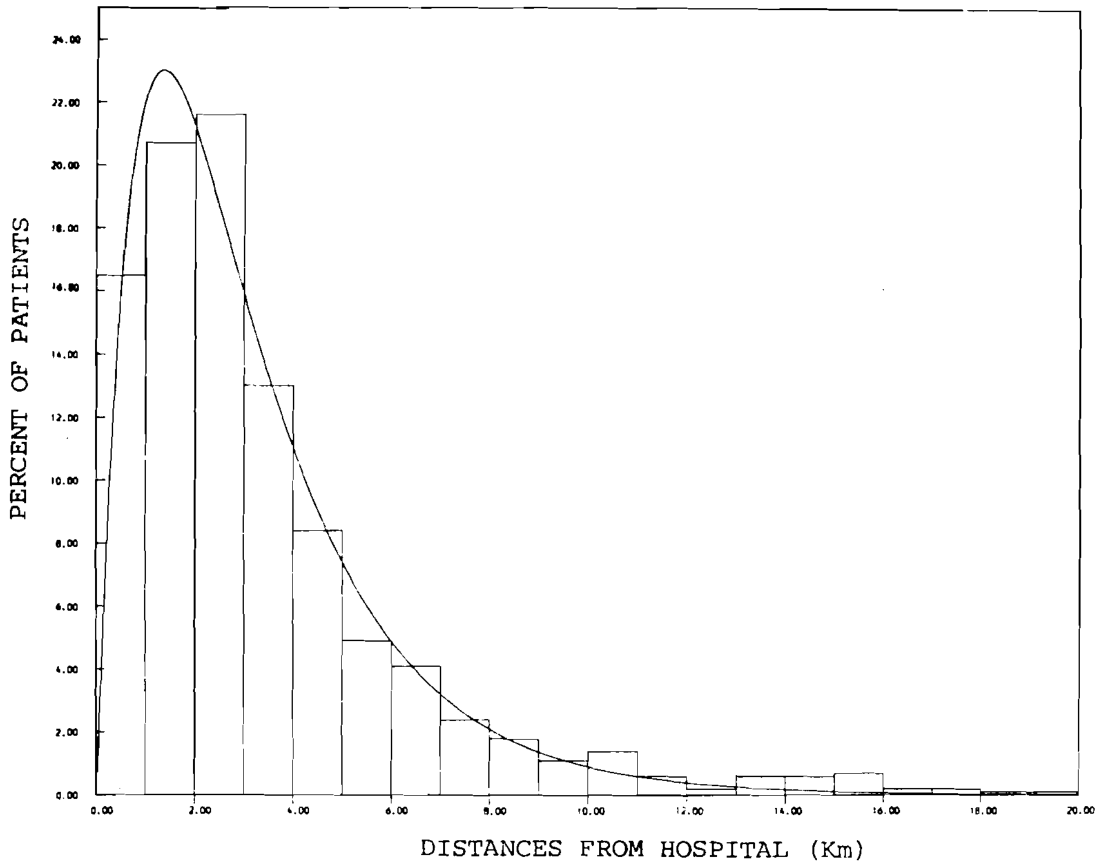


Figure 3. The relationship between patient journey origins and distance from hospital in London for general medical and surgical specialities. The equation of the fitted curve is $y = 100\ x^{1.325} \exp(-1.508\ x^{0.711})$, $R^2 = 0.98$. (Source: Mayhew, 1979)

When a new hospital opens therefore we would expect local hospitalization rates to rise. This presumption is borne out by the experience at Northwick Park Hospital, Harrow, a large acute hospital built on a 'greenfield' site near the periphery of the city which opened in 1969. The main effect of this hospital, whose construction created a seven-fold increase in the caseload capacity of Harrow between 1967 and 1977, was to increase the hospitalization rate in this borough and Brent by almost 50 percent as compared with increases averaging only 20 percent for other boroughs in northwest London over the same period of time. One suitable test for the model presented here therefore is to try to 'back-forecast', using the model, the impact of Northwick Park Hospital on hospitalization rates and to compare that 'backcast' with what actually happened, taking into account the substantial changes in population and other hospital caseloads in the study region which have occurred in the last ten years. This exercise is carried out in section 6. If the impact of these changes on Brent, Harrow and neighboring areas can be predicted over this period with reasonable accuracy, then the model may be applied with more confidence to events expected to take place in the future.

The importance of geographical availability in determining hospitalization rates is thus apparent from these examples. In under-provided areas it must be accepted that patients who would otherwise be admitted to hospital must seek treatment in some other form or not at all. In this study only the in-patient and day-patient sectors of the health care services are considered. Parallel facilities for obtaining treatment are to some extent available in the community (mostly in general practice, health centers, or clinics), in the out-patient departments of hospitals, or in the private medical sector. These alternatives and the interactions between them are not considered in this study, but they are clearly important in determining the overall balance of care (McDonald, Cuddeford, and Beale 1974). Nevertheless, it may be possible to incorporate into the scheme these parts of the health care system at a later date.

3. THE BASIC MODEL

The model used is a behavioral one and is of the singly-constrained gravity kind. It argues that patient flows from an area are in proportion to the morbidity in that area, and to hospital bed availability in all areas, but are in inverse proportion to the difficulty of geographical access in terms of travel time or distance. In order that the ability of hospitals to treat patients is not exceeded, a single constraint is included so that hospitals can treat up to their caseload capacities and no more. Realistically, some fluctuation, say ±5 percent, is likely in the caseloads either through higher throughput or because of slack in the system. This can be built into forecasts as desired, but basically the model assumes that resources are always used to capacity.

The use of the gravitational analogy is well-known and dates back many years (Carrothers 1956). A significant advance in the theoretical basis of gravity models was published by Wilson in 1967. Published applications in health care systems are, however, extremely rare in the UK. Similar models have been developed in the United States (e.g., Morrill and Kelley 1970), but they are generally unsuited for use in a UK context because of the totally different approach in the former to the provision of health services. Nevertheless, there is scope for unifying these separate perspectives to produce versions of the same model that can be applied in both market and planned health care systems. Some additional comments on this are made in section VI.

Mathematically, the basic model used here is as follows:

$$T_{ij} = B_j D_j W_i \exp(-\beta c_{ij}) \tag{1}$$

where

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

D_j = the caseload capacity in j for treating patients in a specialty or group of specialties

W_i = a patient generating factor (PGF) which is an index of the propensity of an area to generate patients in the same group of specialties

c_{ij} = the time-cost or distance of travel between i and j

and

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

This is a constraint which ensures that,

$$\sum_i T_{ij} = D_j$$

that is, flows from all i to j exactly equal the case capacity in j

β = a parameter to be determined

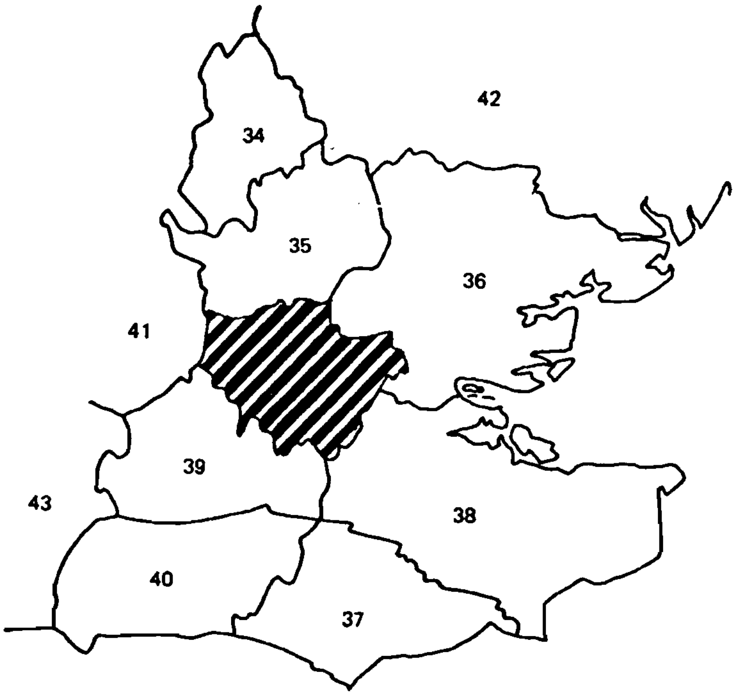
The model operates in two distinct modes: the first is the calibration mode, which consists of finding a value of β such that predicted patient flows $\{T_{ij}\}$ match the observed flows $\{N_{ij}\}$ as well as possible; the second is the forecasting mode which examines the flow consequences of changes in the input variables assuming β is unchanged. This assumption is the behavioral basis of the model.

The first stage of study in both versions of the model involves the definition of zones. There are basically three types of patient flows which must be represented: those from zones in the external world to zones in the internal study region and vice versa; those between zones in the study region, and between zones in the external world; and those within individual zones. Data availability is a major constraint on the suitable geographical delimitation of zones. In the first model (model 1) origin zones are in fact different from destination zones for this reason. Accordingly, there are 44 origin zones and 69 destination zones (see Figures 4 and 5). Four of these zones (Oxford RHA, East Anglia RHA, Wessex RHA, and the rest of England) are outside the four Thames RHAs and these are regarded as the external zones in this model. The internal study region thus has 40 origins (London administrative boroughs and counties outside the GLC in the Thames Regions) and 65 destinations (the Health Districts in the four Thames Regions).*

In the second model (model 2) there are 46 internal zones based on traffic districts used for planning purposes by the GLC (Crawford et al, 1975), and 13 external zones based partly on Area Health Authorities (AHAs) covering the remainder of Thames regions (see Figure 6). Traffic districts are divisions of boroughs, but with suitable adjustment they can be readily aggregated to the Health District (HD)

*For administrative use England is divided into 14 Regional Health Authorities (RHAs), 40 Area Health Authorities (AHAs), many of which in turn are divided into Health Districts (HDs).

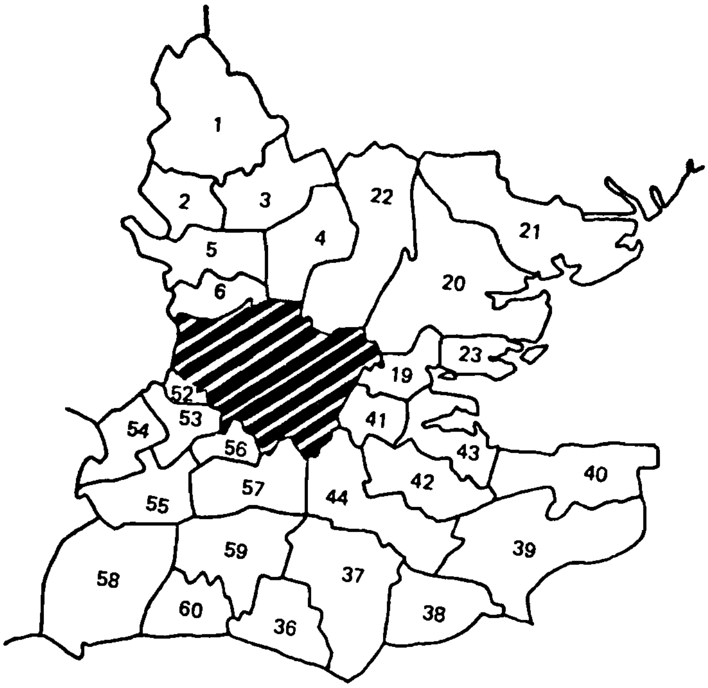
A) Southeast England



B) Greater London Council (GLC)

Figure 4. Model 1 origin zones.
Key on page 13; zone 44 (rest of England) is not shown.

A) Southeast England



B) GLC

Figure 5. Model 1 destination zones. Key on page 13; zone 69 (other RHAs) is not shown.

Key to Figures 4 and 5

Origin

1 Barnet
2 Brent
3 Harrow
4 Ealing
5 Hammersmith
6 Hounslow
7 Hillingdon
8 Kens + Chelsea
9 Westminster
10 Barking
11 Havering
12 Camden
13 Islington
14 City
15 Hackney
16 Newham
17 Tower Hamlets
18 Enfield
19 Haringey
20 Redbridge
21 Waltham Forest
22 Bexley
23 Greenwich
24 Bromley
25 Lambeth
26 Lewisham
27 Southwark
28 Croydon
29 Kingston
30 Richmond
31 Merton
32 Sutton
33 Wandsworth
34 Bedfordshire
35 Hertfordshire
36 Essex
37 E Sussex
28 Kent
39 Surrey
40 W Sussex
41 Oxford
42 E Anglia
43 Wessex
44 Other

Destination

1 N Bedfordshire
2 S Bedfordshire
3 N Hertfordshire
4 E Hertfordshire
5 NW Hertfordshire
6 SW Hertfordshire
7 Barnet *
8 Edgware *
9 Brent
10 Harrow
11 Hounslow
12 S Hammersmith
13 N Hammersmith
14 Ealing
15 Hillingdon
16 K/C/W NW *
17 K/C/W NE
18 K/C/W S
19 Basildon
20 Chelmsford
21 Colchester
22 Harlow
23 Southend
24 Barking
25 Havering
26 N Camden
27 S Camden
28 Islington
29 City
30 Newham
31 Tower Hamlets
32 Enfield
33 Haringey
34 E Roding
35 W Roding
36 Brighton
37 Eastbourne
38 Hastings
39 SE Kent
40 Thanet
41 Dartford
42 Maidstone
43 Medway
44 Tunbridge

Destination

45 Bexley
46 Greenwich
47 Bromley
48 St Thomas'+
49 Kings'
50 Guys'
51 Lewisham
52 N Surrey
53 NW Surrey
54 W Surrey
55 SW Surrey
56 Mid Surrey
57 E Surrey
58 Chichester
59 Crawley
60 Worthing
61 Croydon
62 Kingston
63 Roehampton
64 Wandsworth
65 Sutton
66 Oxford
67 E Anglia
68 Wessex
69 Other RHAs

---

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

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

A) Southeast England

B) Southeast GLC

Figure 6. The zoning system for model 2. There are 59 origin and destination zones, 46 within the southeast GLC and 13 in the rest of the four Thames RHAs.

administrative level. Health Districts serving the second study region are Kings, Guys, St. Thomas, Lewisham, Bromley, Bexley, and Greenwich, the first three of which are also teaching districts. Traffic districts are also aggregations of census wards, and for the southeast only patient flows at ward level are known.

All zones in both studies were each allocated a centroid from which distance or travel time could be measured. The centroids in the first model, which uses distance, were defined initially either as weighted centers of population or, if available, by other suitable nodal points. In the second model, which uses time, the centroids were already defined by the GLC for each traffic district, but for external zones weighted mean centers of population were used.

4. VARIABLE SPECIFICATION

4.1. Caseloads

Caseloads (D_j) are defined as the combined case capacities of hospitals in each zone to treat patients in particular groups of specialties. For calibration purposes the data were obtained from patient flow information in the Hospital Activity Analysis (a comprehensive statistical annual review of in-patients by RHAs). For example, if N_{ij} is the observed flow from i to j, then the caseload of j is defined as

$$D_j = \sum_i N_{ij} \qquad (2)$$

Caseloads for both models were based on 1977 data. The list of specialties considered in each is shown in Table 1. This list combines both regional and sub-regional specialties: that is no distinction is drawn between them in terms of their differential geographical availability. For some applications of the model, it makes sense to disaggregate on these lines. This was done in the case of model 1, to produce a regional and sub-regional model. (Regional specialties service much larger populations and are indicated by an *.) The results of the all-specialty and disaggregated models are compared in section 6.

Table 1. Specialties list in models 1 and 2.

Four Thames Regions Model	Southeast Model
Specialties included:	
General Medicine	As for model one plus:
Paediatrics	Geriatrics
Infectious Diseases	Special care babies
Chest Diseases	Staff wards
Dermatology	Convalescent
Neurology*	Acute mental illness
Cardiology*	
Rehabilitation/Physical Medicine	
STD	
Rheumatology	
General Surgery	
ENT	
Traumatic and Orthopaedic Surgery	
Ophthalmology	
Radiotherapy*	
Urology	
Plastic Surgery*	
Thoracic Surgery*	
Dental Surgery (including Orthodontics)	
Neurosurgery*	
Gynaecology	
GP Medicine	
OSU*	

*Regional Specialties

In the forecasting mode case capacities can be determined from trends in treatment patterns over longer periods, and from proposed developments such as hospital construction. For example, many specialties—because of improving treatment and a better organization of resources—are experiencing falling lengths of stay, enabling more cases to be treated in one year with the same given number of beds. Similarly the average length of time between successive bed occupants can be reduced, thus enabling more cases to be treated. These effects can be built into forecasts using the following formula as an example:

$$Cm(t+n) = \frac{Bm(t+n)}{Bm(t)} \times \frac{[lm(t) + tm(t)]}{[lm(t+n) + tm(t+n)]} \times Cm(t)$$

where

$Cm(t)$ = cases treated in specialty m in year t

$Bm(t)$ = available beds in specialty m in year t

$lm(t)$ = length of stay in specialty m in year t

$tm(t)$ = turnover interval in specialty m in year t

4.2. Patient Generating Factors (PGFs)

PGFs are an index of a zone's ability to generate patients in the specialties of interest. Ideally, we would need an assessment of the morbidity in a population; however, accurate and undisputed measures of this are hard, if not impossible to come by. IIASA is developing some morbidity models that offer potential (Kitsul, 1980), and they may be used in future work but for the present the method used in this study relies on the relative national pattern of hospital usage by specialty by persons of different age and sex. Thus, if U_{lm} is the national discharge rate in age/sex category l, and specialty m and P_{il} is population in i, also in age/sex category l, then

$$W_i = \sum_m \sum_l P_{il} U_{lm} \tag{3}$$

defines the PGF for zone i. This index takes no account of socio-economic and environmental factors likely to influence patient generating potential. These could be incorporated by an appropriate weighting of the W_is. Standardized mortality ratios (SMRs),* for example, may be used as measures of relative need. The use of SMRs in obtaining patient generating factors has been investigated and the results are discussed in section 6. These attempts at devising suitable PGFs do not exhaust the possibilities however. Improvements to include more factors can be made as experience with the model grows.

For forecasting purposes, PGFs are dependent on population change and trends in relative hospital utilization rates. For the former population projection can be used; for the latter trends national patterns over time are the best indication (e.g., LHPC 1979).

4.3. Travel Costs

Two measures of travel cost were used: simple distance in models 1 and 2, and travel time in model 2. Simple distance is defined as:

$$c_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \qquad (4)$$

where x_i, y_i and x_j, y_j are the centroids of zones i and j, and c_{ij} is the cost-distance between them in kilometers. For intra-zonal distances (i.e., when i = j), a formula based on the proximity of the next nearest centroid was used. A drawback with distance is that it is not always a reliable measure of accessibility, particularly in urban areas where travel is affected by a variety of factors. One prominent hindrance to travel, for instance, is the River Thames, and it was found necessary to weight inter-zonal distances which crossed it.

---

*$SMR_i = \sum_l M_{il} / \sum_l r_l P_{il}$ where M_{il} is the actual number of deaths in i in age/sex category l, P_{il} is the population in i in category l, and r_l is the national age/sex specific death rate.

Model 2 uses travel times in addition to distance in an attempt to overcome these features. Inter-district travel times were supplied by the Greater London Council from the 1972 Greater London Travel Survey (GLTS) for public and private transport. In using two measures of cost for each origin-destination pair it becomes necessary to decide what proportion of the patient population will travel by each form of transport. People live in households, and the number of households with one or more cars is generally known. This information is used to split the PGFs into two streams: (a) those patients with potential car access, and (b) those without. Not everyone in a household will be qualified to drive, or have access to a car for a given hospital journey, however. This further reduces the first stream by a factor assumed to lie between 50 and 75 percent. Those with access and contemplating car travel will then weigh the advantage of traveling by private or public transport. The actual number of persons involved is then determined within the model as follows.

4.4. Modal Split

Restating the basic model, we have

$$T_{ij}^{kn} = B_j D_j W_i^n \exp(-\beta c_{ij}^k) \tag{5}$$

where

$$B_j = \left[\sum_i \sum_n \sum_{k \in \gamma(n)} W_i^n \exp(-\beta c_{ij}^k) \right]^{-1} \tag{6}$$

ensures

$$\sum_i \sum_n \sum_{k \in \gamma(n)} T_{ij}^{kn} = D_j \tag{7}$$

Here n is the class of traveler who has available a set of modes given by $\gamma(n)$, while k is the mode of travel. In our case there are two modes and two classes. If we consider car-owners ($n = 1$) the proportion who use public transport ($k = 2$) between i and j is determined by calculating the patient flows generated by each mode individually and dividing out, i.e.,

$$\frac{T^{21}_{ij}}{T^{11}_{ij} + T^{21}_{ij}} = \frac{1}{1 + \exp\{-\beta(c^1_{ij} - c^2_{ij})\}} \qquad (8)$$

The key factor from equation (8) in determining the proportion is hence the difference in journey times by each mode. This proportion determines the modal split, and the results were checked for validity against a sample hospital travel survey carried out in London. The details are given later in section V but a graph of the relationship in equation (8) is shown in Figure 7.

In the forecasting mode, distance does not change but time might because of changes in the transport system. These will normally be slight over a typical forecasting period. The car ownership factor will be more important, however, and any expected changes can be incorporated once forecasts for W_i have been established. A greater access to cars, for example, implies more mobility, and one consequence of this in model 2 is that patients will travel longer distances. This may eventually permit the provision of fewer, though larger, hospitals.

4.5. Other Considerations

4.5.1. Hospitalization Rates and Elasticities

The criterion health administrators will be most interested in is the effect a particular plan of action will have on hospitalization rates. These are defined for each origin zone as,

$$R_i = \sum_j \frac{T_{ij}}{P_i} \qquad (9)$$

That is, the row total of predicted flows divided by the total population of i. This compares the actual hospitalization rates which are defined in terms of N_{ij}.

Elasticity, by contrast, is an index of hospitalization rates' sensitivity to a change in caseload. It is useful as a measure of a zone's reliance on a group of hospitals. The hospitalization rate in equation (9) can be re-written using equation (1) thus,

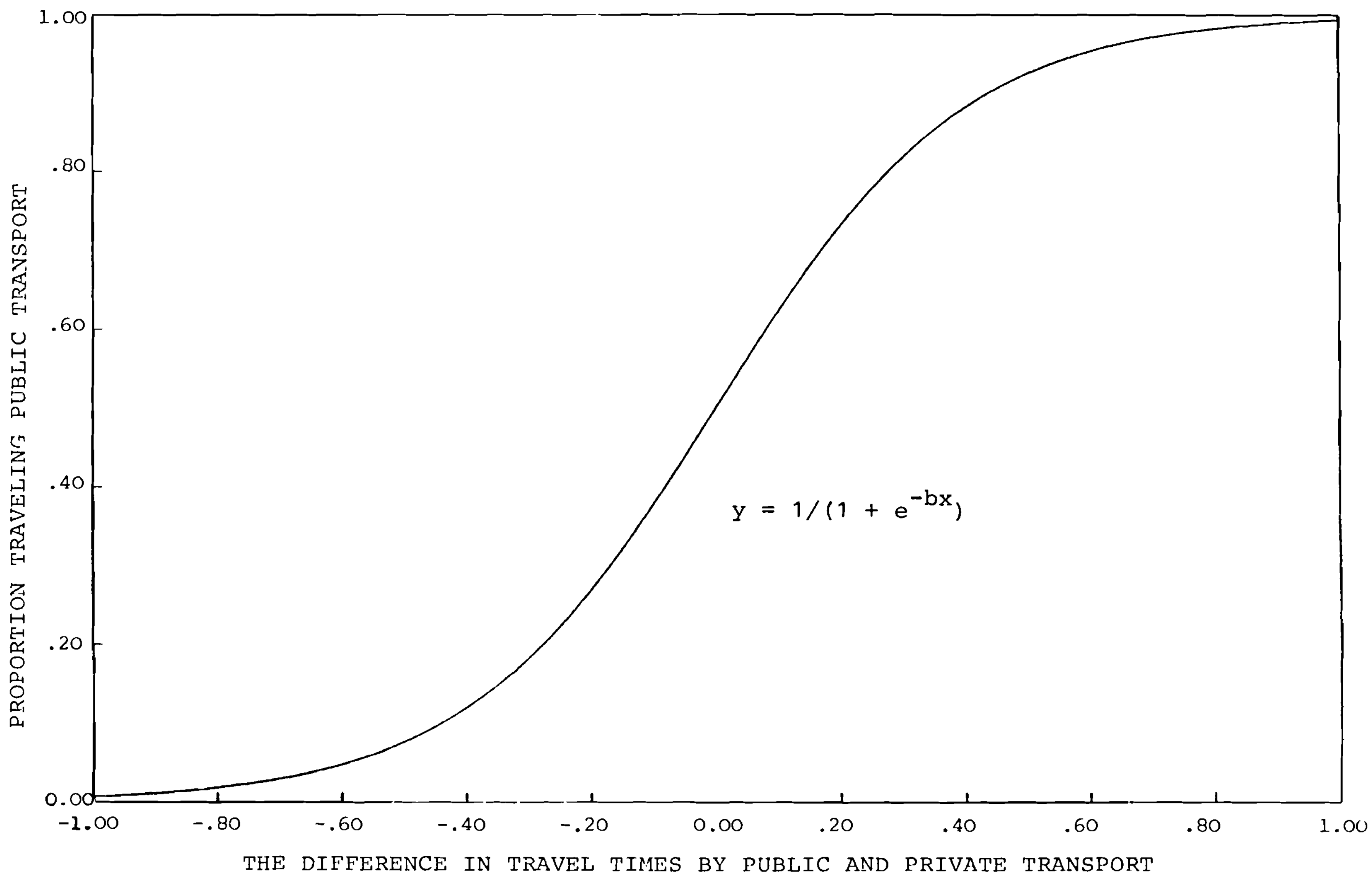


Figure 7. Modal split function.

$$R_i = \sum_j \frac{T_{ij}}{P_i} = \frac{W_i}{P_i} \sum_j B_j \, D_j \exp\,(-\beta c_{ij}) \tag{10}$$

therefore, defining

$$E_{ij} = \frac{D_j}{R_i} \frac{\partial R_i}{\partial D_j} = T_{ij} / \sum_j T_{ij} \tag{11}$$

where E_{ij} is the required elasticity. E_{ij} varies from 0 to 1 and is the ratio of the predicted flow to the row total. It expresses the proportionate change in the hospitalization rate expected in i following a small change in the caseload of j. It is best interpreted as the dependency of a population on a specific destination zone. Typically it is highest when i = j, or there is considerable overlap between the zones i and j, indicating that zonal populations are generally more reliant on their local hospitals than on hospitals in any other zones.

4.5.2. Catchment Populations

Health administrators will also be interested in the catchment population of each destination zone using a measure which takes into account the effects of cross-boundary flows. Catchment populations are related to the total population of an origin zone and to the elasticity of the hospitalization rates defined in section 4.5.1. Thus

$$C_j = \sum_i E_{ij} \, P_i \tag{12}$$

where C_j is the required catchment population of j and,

$$E_{ij} = T_{ij} / \sum_j T_{ij} \tag{13}$$

By using the predicted elasticities, therefore, the catchment implications of changes either in caseload or population can be determined. This measure would be particularly useful for instance in assessing the likely impact in terms of population served of a new hospital.

4.5.3. Deterrence Function

The basic model distributes hospital flows in accordance with a negative exponential function [exp $(-\beta c_{ij})$ in equation 1] sometimes called the deterrence function: other functions which are likewise monotonic-declining and asymptotic to the horizontal axis have been used in gravity modeling. Although the present program expects the negative exponential form, input cost matrices $\{c_{ij}\}$ can be simply transformed to obtain other functions which may give a better fit to the observed flows. Table 2 lists some examples that were tried in the course of this study.

Table 2. Input transformations for changing deterrence function.

Function	Transformation	Restrictions
$\exp(-\beta c_{ij})$	None	None
$c_{ij}^{-\beta}$	$\{c_{ij}\} \dashrightarrow \{\log c_{ij}\}$	$c_{ij} \neq 0$
$c_{ij}^{-\beta}\exp(-\beta c_{ij})$	$\{c_{ij}\} \dashrightarrow \{c_{ij}+\log c_{ij}\}$	$c_{ij} \neq 0$
$\exp(-\beta c_{ij}^{k})$	$\{c_{ij}\} \dashrightarrow \{c_{ij}^{k}\}$(k=constant)	None

The purpose of using different deterrence functions can be appreciated from the curves in Figure 8. For instance, for the same set of data the power function $(c_{ij}^{-\beta})$ will give more emphasis to patients generated at low rather than high travel costs. The exponential function [exp $(-\beta c_{ij})$] in contrast emphasizes intermediate travel costs, but has a negligible effect when these costs are very high. The mixed function [$c_{ij}^{-\beta}$ exp$(-^{\beta}c_{ij})$] offers more flexibility, but raises the question of developing a two-parameter instead of a one-parameter model of deterrence [i.e. $c_{ij}^{-\alpha}$ exp$(-\beta c_{ij})$].

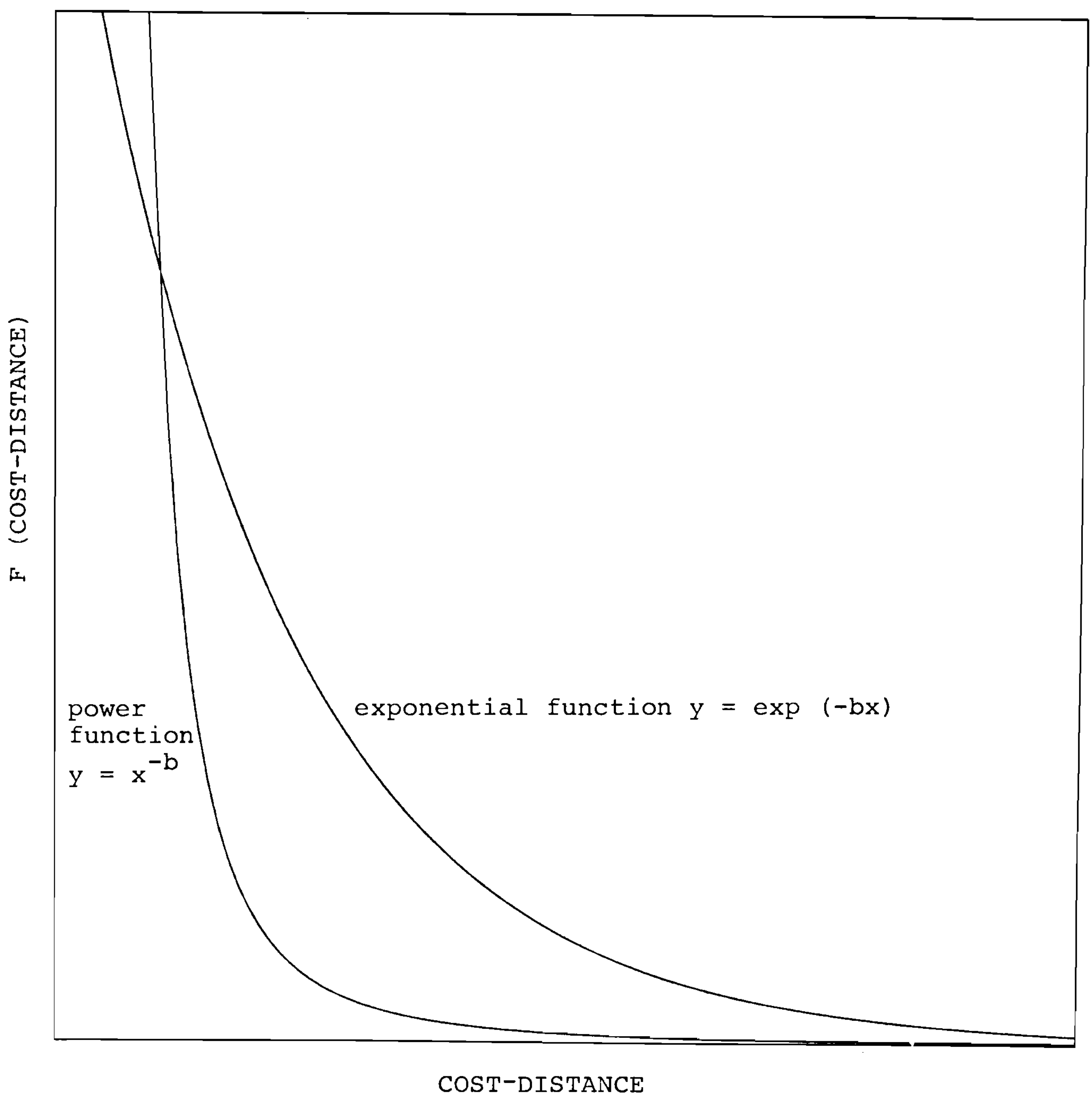


Figure 8. Examples of deterrence function S, where b is an illustrative parameter.

5. CALIBRATION

Calibration is finding a value for β in equation (1) such that predicted flows $\{T_{ij}\}$ most accurately portray observed flows $\{N_{ij}\}$. Several methods of calibration exist and are documented in the literature. They nearly all involve some form of search procedure which stops either when a calibration statistic assumes a particular value or when it reaches some maximum or minimum value. The calibration statistic is calculated over some subset of the trip matrix, which is referred to as the region of calibration. Questions concerning the choice of region of calibration are covered in a later section. It is disconcerting that different calibration statistics produce different values of β. However, there is no way of telling which method or statistic is best except by exhaustive testing. The basic calibration procedure is more or less the same irrespective of the calibrating statistic, and the way it is handled by the program is shown in Figure 9. Experience reduced the number of calibration methods to three of which the third was generally found to be most suitable.

Calibration method [1] was based on the principle of maximum likelihood (Batty and Mackie 1972). If the deterrence function is a negative exponential, this method states that predicted flows are most likely to be correct when the mean predicted travel cost equals the actual mean cost, that is

$$\bar{c}_p = \bar{c}_{obs} \qquad (14)$$

where

$$\bar{c}_p = \frac{\sum_i \sum_j T_{ij} c_{ij}}{\sum_i \sum_j T_{ij}} \qquad (15)$$

and

$$\bar{c}_{obs} = \frac{\sum_i \sum_j N_{ij} c_{ij}}{\sum_i \sum_j N_{ij}} \qquad (16)$$

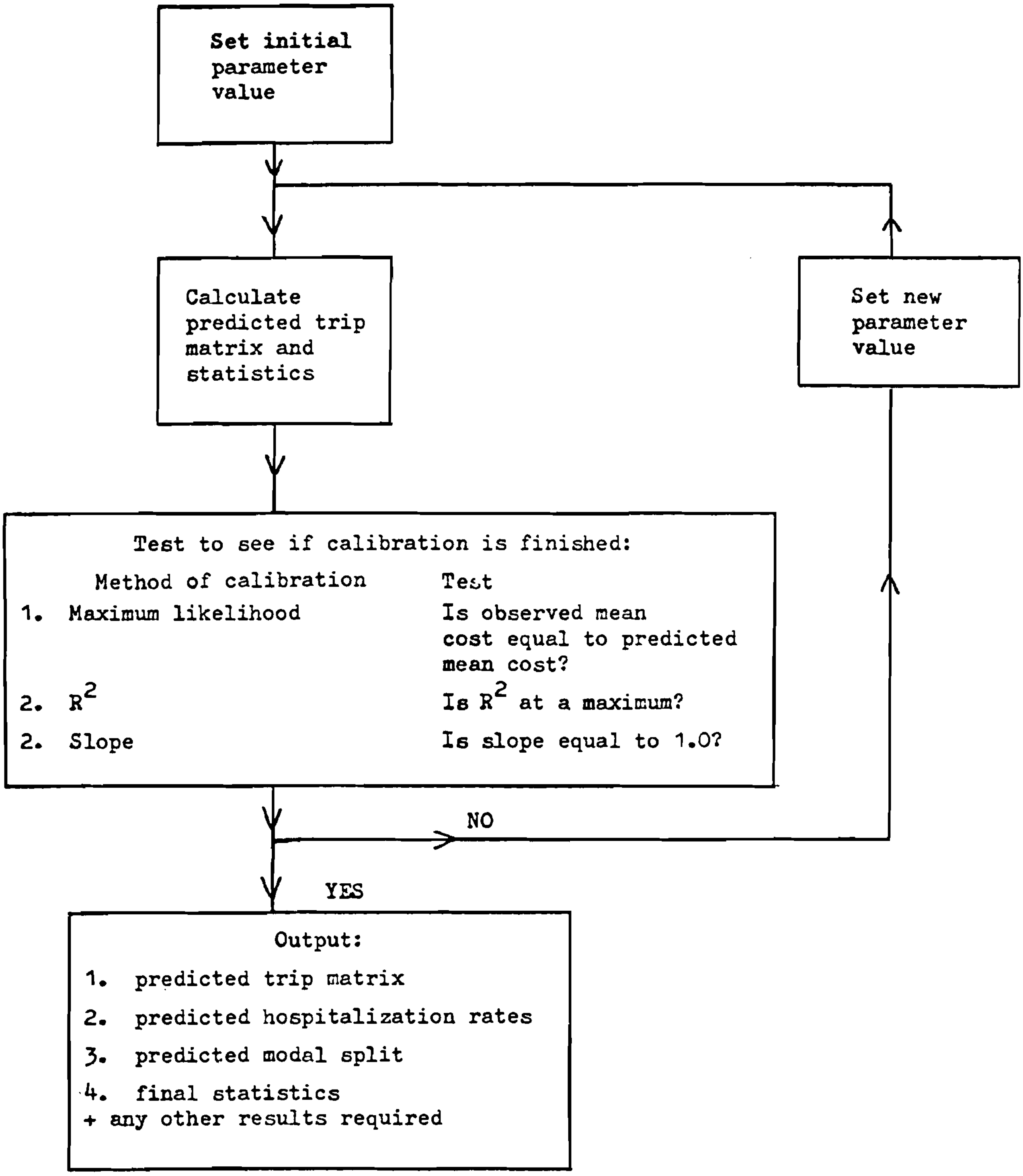


Figure 9. RAMOS calibration procedure: flow diagram.

The program uses Hyman's linear interpolation convergence formula (Hyman, 1969) for iterating towards a solution for β. If n is the number of iterations β^{n+1} is found by

$$\beta^{n+1} = \beta^{n} + \frac{(\bar{c}_{obs} - \bar{c}_p^{\,n})\ (\beta^{n} - \beta^{n-1})}{(\bar{c}_p^{\,n} - \bar{c}_p^{\,n-1})} \qquad (17)$$

For the first iteration only,

$$\beta^{2} = \beta^{1}\bar{c}_p / \bar{c}_{obs}$$

where β^1 is the initial estimate supplied by the user. Setting β^1 to $(\bar{c}_{obs})^{-1}$ normally resulted in successful convergence after only a few iterations. Accuracy is determined from a tolerance value which can be set to any value by the user.

Several problems came to light in using this method in both models 1 and 2. It was found that the value of β obtained was sensitive to the number of zones over which calibration took place. Ideally one would have expected little or no change in β whether calibration was based on flows over all the RHAs for example or just parts of them. It was further found that $\bar{c}_{obs}$ was very senstive to the definition of centroids, particularly those in external zones which are heavily weighted by large patient flows. In model 2 an additional difficulty was in obtaining a value for $\bar{c}_{obs}$. Clearly equation (16) is inappropriate in the two-mode case because it requires prior knowledge of the modal split by public and private transport of patients traveling to hospital. A survey value based on travel to London hospitals was therefore used instead (Mayhew, 1979), but this too had its drawbacks.

The use of this method is also conditioned by the functional form of the deterrence function. Equations 14-16 apply only to the ordinary negative exponential deterrence function. For the power function, for example, it is necessary to substitute in equations (15) and (16) log c_{ij} for c_{ij} before the method will work.

The second method of calibration was based on maximizing the statistic R^2, which is the proportion of variance explained by the regression of predicted on observed patient flows. It is written

$$R^2 = \frac{\sum_i \sum_j (\hat{T}_{ij} - \bar{T})^2}{\sum_i \sum_j (T_{ij} - \bar{T})^2} \tag{19}$$

where $\hat{T}_{ij}$ is the expected flow (predicted by the regression), T_{ij} is the predicted flow by the model, and $\bar{T}$ is the mean predicted flow.

Two problems detracted from the use of this statistic: firstly, it is insensitive to the value of β; and secondly, it is often very close to one, the upper limit of the R^2 range. While a value of one would indicate a perfect fit, it was found that, for model 2, calibration runs with R^2 values only slightly less than this could still have many undesirable properties.

The third method of calibration proved the most suitable. This method was based on the slope of the regression of predicted on observed flows rather than on the proportion of variance explained. When the value of the slope is equal to one, it means that on average predicted and observed flows are the same. The regression slope b is defined in terms of T_{ij} and N_{ij} as

$$b = \frac{\sum_{i,j} N_{ij} T_{ij} - \frac{\sum_{i,j} N_{ij} \sum_{i,j} T_{ij}}{N}}{\sum_{i,j} N_{ij}^2 - \frac{(\sum_{i,j} N_{ij})^2}{N}} \tag{20}$$

where N is the number of cells in the matrix for which $D_j \neq 0$.

A search technique is used to find the required value of β, but unlike method 1 the procedure used is not of the convergence type. Experience showed however, that a good starting value for β could be obtained using maximum likelihood, and this greatly shortened the search by the slope method.

Accompanying the mean cost, R^2 and slope statistics were other statistics, which though lacking in calibrating potential acted as good measures of fit. These statistics, which were output at each iteration, are summarized in Table 3.

To illustrate the points made in this section we conclude by showing in Table 4 an example of a typical sequence of iterations towards a solution based on method 3. Attention is drawn to the fact mentioned above that the statistics concerned have very diverse behaviors, and that extreme caution should be exercised in selecting the appropriate one for calibration purposes.

6. MODEL 1 RESULTS

6.1. Introduction

In this section the results obtained with model 1 are discussed and the calibrations using three different cost matrices are compared. The first calibration uses a cost matrix consisting of the unmodified straight line distances between the centroids of the origin and destination zones (Matrix 1). Model 1 uses a different zoning system for origins than for destinations, and for this reason, the crude distance matrix obtained was found inadequate in its estimation of distances between origin and destination zones where there was considerable overlap between the zones. The distance between each such origin-destination zone pair was altered to give a more realistic assessment of the actual mean distance for the trip concerned and a second cost matrix (Matrix 2) was produced incorporating these modifications. This matrix also contained one other refinement; that is, increases were made in the distances for trips between zones separated by the River Thames, where some detour from a straight line path would be necessary to reach a crossing point. This was effected by the use of a single factor increasing all such distances by a constant proportion.

Table 3. Other statistics used in measuring goodness-of-fit.

Symbol	Statistic	Formula for calculation
a	intercept of regression line of predicted flows against observed flows	$a = \dfrac{\sum_{i,j} T_{ij} - b \sum_{i,j} N_{ij}}{N}$
X^2	chi-squared statistic	$X^2 = \sum_{i,j} \dfrac{(N_{ij}-T_{ij})^2}{T_{ij}}$ such that $T_{ij} \neq 0$
$\lvert e \rvert$	mean absolute error	$\lvert e \rvert = \sum_{i,j} \dfrac{\lvert N_{ij} - T_{ij} \rvert}{N}$
$\lvert pe \rvert$	mean absolute percentage error	$\lvert pe \rvert = \sum_{i,j} \dfrac{\lvert N_{ij} - T_{ij} \rvert}{N_{ij}} \times \dfrac{100}{M}$ such that $N_{ij} \neq 0$, where $M = \sum_{\substack{i,j \\ N_{ij} \neq 0}} 1$
RMSQ	root mean square error	$RMSQ = \left[\sum_{ij} \dfrac{(N_{ij} - T_{ij})^2}{N} \right]^{\frac{1}{2}}$

Table 4. An example of an iteration sequence using the power function (model 2).

iteration	calibration parameter	mean cost	chisquare	rmsq error	r square	regression coeffs		mean abs er	mean abs pc er
1	0.30000	7.40751	0.107140e+20	402.8	0.9668	52.86	0.8208	122.3	155.1
2	0.31000	7.27141	0.494569e+20	359.3	0.9711	46.52	0.8501	113.1	147.4
3	0.32000	7.14439	0.228744e+21	319.8	0.9746	40.27	0.8786	104.2	140.7
4	0.33000	7.02560	0.105981e+22	285.2	0.9774	34.11	0.9063	95.99	134.9
5	0.34000	6.91433	0.491788e+22	257.3	0.9797	28.06	0.9331	89.15	129.8
6	0.35000	6.80995	0.228523e+23	237.5	0.9813	22.13	0.9590	84.02	125.2
7	0.36000	6.71186	0.106321e+24	227.5	0.9824	16.32	0.9841	80.53	121.1
8	0.37000	6.61955	0.495211e+24	227.9	0.9830	10.64	1.008	79.47	117.4

Results obtained with this second cost matrix indicated a substantial improvement over those obtained with Matrix 1. However, an examination of the calibration results continued to show the inadequacy of using a cost matrix which was still heavily based on straight line distance. This led to the production of a third cost matrix (Matrix 3), incorporating a second set of modifications designed to reflect factors such as congestion in the GLC area, particularly in central London, the relative ease of access from the counties to central London health districts (compared to similar straight line distances to other health districts outside London), and so on. The modifications used are empirically derived, and consist of a set of multiplying factors used to:

A. Decrease "distance" from origins outside London to destinations inside the GLC (mainly central London)
B. Increase "distance" between zones in the GLC area

Results obtained using Matrix 3 showed a substantial improvement over those obtained with Matrix 2, both in terms of the goodness-of-fit of the calibration to the actual 1977 data and in terms of the accuracy obtained when the predictive ability of the model was tested using data for 1967 in the North West Thames RHA area. In the sections that follow the results obtained with the three cost matrices are compared, and the results obtained with Matrix 3 are examined in particular detail.

6.2. Overall Statistics

Table 5 presents a comparison of the results obtained using the three different cost matrices in terms of some overall statistics. The results shown were all obtained using the slope method of calibration, and the statistics referring to the trip matrix are calculated over the region of calibration only -- in this case all flows from origins in the GLC area to destinations in the Thames Regions. This choice of region of calibration is discussed later in this section. As Table 5 shows Matrix 3 produces a better value for each of the statistics considered. In

Table 5. Model 1 - Comparison of overall statistics.

	COST MATRIX USED		
	Matrix 1	Matrix 2	Matrix 3
Parameter	0.428	0.434	0.367
i. Trip matrix statistics			
R^2	0.724	0.850	0.983
slope of regression line, b	1.0001	1.0013	1.0010
intercept of regression line, a	111.88	62.26	12.30
root mean square error	1063.2	724.3	226.4
mean absolute error	283.3	186.1	79.3
mean absolute % error	424.6%	159.5%	118.5%
ii. Hospitalization rate statistics			
mean absolute error	38.8	31.5	5.7
mean absolute % error	33.8%	26.8%	5.0%
number of areas of residence with <10% error	14	14	36

particular the value of R^2, the percentage of variation in the observed flows explained by the model, is 0.98 using Matrix 3, a great improvement over the values of 0.85 obtained with Matrix 2 and 0.72 with Matrix 1.

The other statistics relating to the predicted trip Matrix, (root mean square error, mean absolute error, and mean absolute percentage error) all suffer from defects when used as an overall measure of goodness-of-fit, due to the fact that there is an enormous variation in the range of cell values in the trip matrix (from 0 to 30,000). The values of root mean square error and mean absolute error are dominated by cells with large flows. Although some of these have large absolute errors, the percentage error is often small. On the other hand, the mean absolute percentage error is dominated by cells with small flows for which a small absolute error is obtained, but in percentage terms this can be very large. In the case of model 1 this second defect is likely to be the most serious. Thus, too much significance should not be attached to the actual values of these statistics shown in Table 5, but it is important to notice that they all show a considerable improvement using cost Matrix 3 over Matrix 1 and Matrix 2.

No values for the chi-squared statistic are shown in Table 5. This is because experience during the calibration of the model showed that this statistic was very misleading as a measure of goodness-of-fit. In cases where the actual trip matrix is fairly sparse (in the sense that the number of trips in many of the cells is very small), and the predicted trip matrix underestimates these values, very large values of chi-squared can easily be obtained, even though the fit of the predicted matrix in cells with a significant number of trips may be excellent. As the trip matrix for model 1 is one in which about 75% of the cells have values less than 100, and the model typically underestimates these flows, the values of this statistic were considered unlikely to be helpful in any way.

Perhaps the most important statistics shown in the table are those relating to the model's prediction of the actual hospitalization rates (see section III) in the various areas of residence, as one of the main uses of the model is likely to be in predicting change in hospitalization rates consequent upon change in any of the input variables of the model. With both Matrix 1 and 2 the values of the mean absolute error and mean absolute percentage error obtained are unsatisfactorily high. Matrix 2 does, however, show a clear improvement over Matrix 1. Matrix 3 shows a much better performance in reproducing the actual hospitalization rates, with a mean absolute error of 5.7 (on rates in the range 80 to 140 roughly) and a mean absolute percentage error of 5%. This aspect of the model's performance is also discussed in greater detail later.

6.3. Reproduction of Actual Trip Matrix

This section examines the performance of the model in reproducing the actual flow matrix observed in 1977. Figures 10 to 12 show graphically for each of the three cost matrices, a plot of the predicted flow for each cell in the Thames Regions against the actual flow in that cell. The large number of cells with only a small number of trips associated are not distinguishable on the scale of the graph. The graphs clearly demonstrate the better performance of Matrix 3 to Matrix 2 and Matrix 2 to Matrix 1. The final graph shows a much closer clustering of points around one line at 45^0 to each axis, demonstrating a much better replication of the actual trip matrix than that obtained with the two other matrices.

The cells that are badly predicted using Matrix 2 (those lying far away from the diagonal line in Figure 11) consist of elements from three distinct types of flows. Firstly flows from the counties into London health districts (all underestimated using Matrix 2), secondly flows from the inner London boroughs to health districts in the GLC (all over-estimated using the model), and lastly flows from outer London boroughs to neighboring counties (overestimated using the model). In Figure 10 these features are also present, as well as additional elements which are estimated badly; these consist of flows between zones with considerable overlap. These aspects of the performance of the model using matrices 1 and 2 are not present to any great extent in Figure 12, where there are no longer any particular types of flows which are being consistently over- or underestimated.

6.4. Patterns of Patient Flow to Health Districts

Figures 13 to 15 show in more detail the model's performance in reproducing actual patterns of patient flow. Three different health districts have been chosen and the actual percentage distribution of area of residence for patients treated in each district is compared to that produced by the model using (a) Matrix 1 - the crude distance matrix and (b) Matrix 3 - the final modified distance matrix. In each of the figures on the column representing

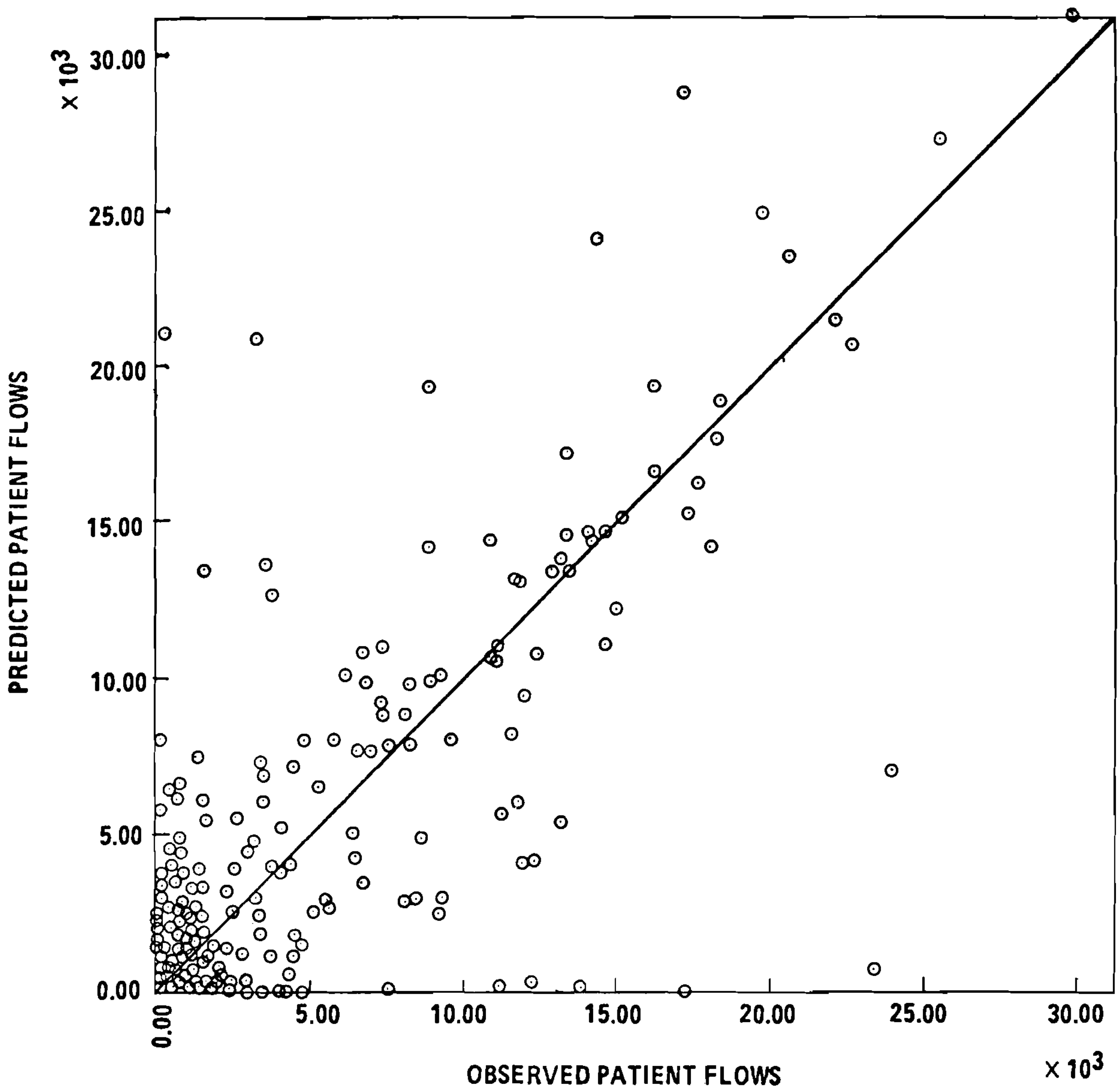


Figure 10. All specialties, 1977, Thames Regions; model 1, Matrix 1.

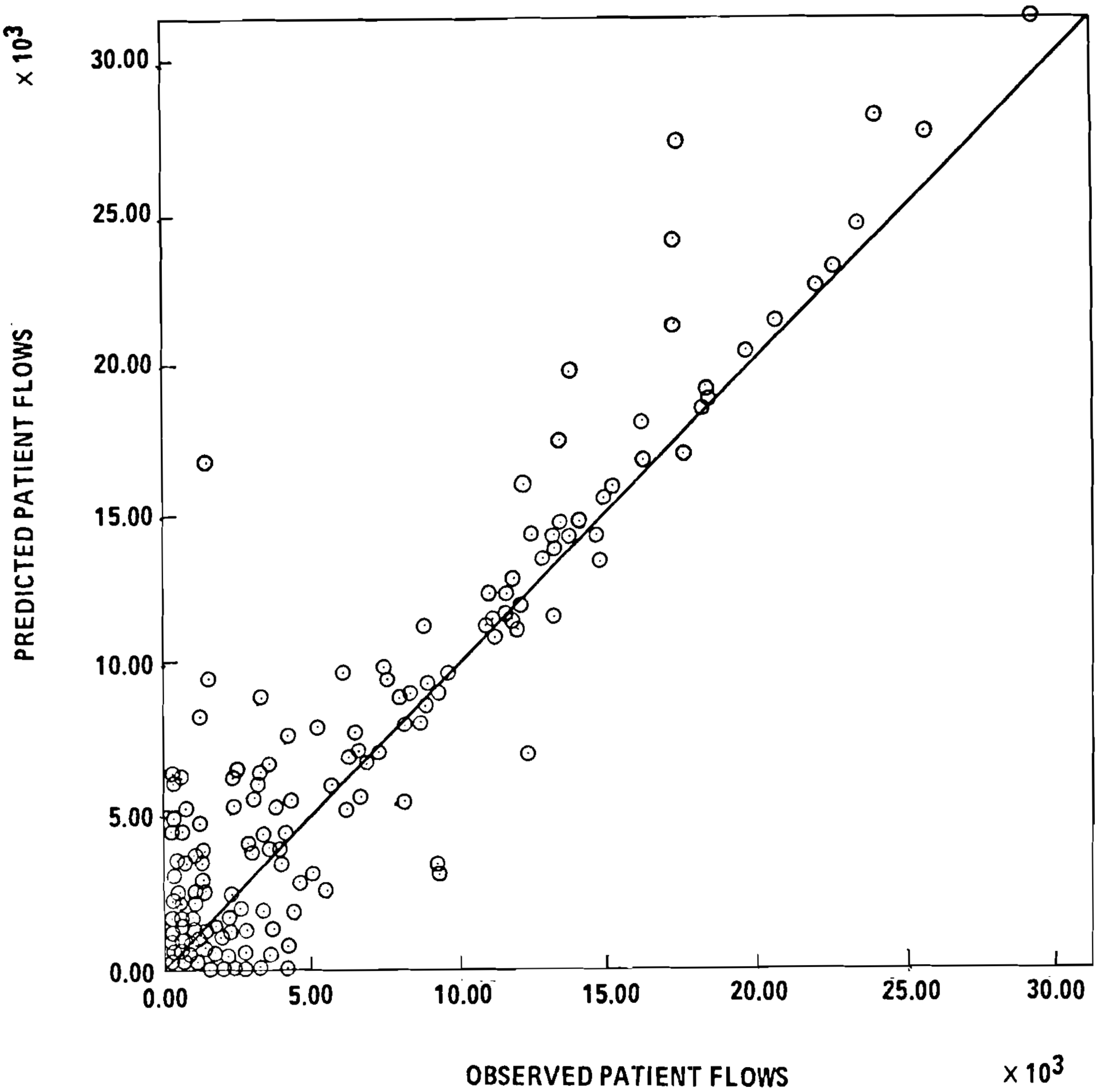


Figure 11. All specialties, 1977, Thames Regions; model 1, Matrix 2.

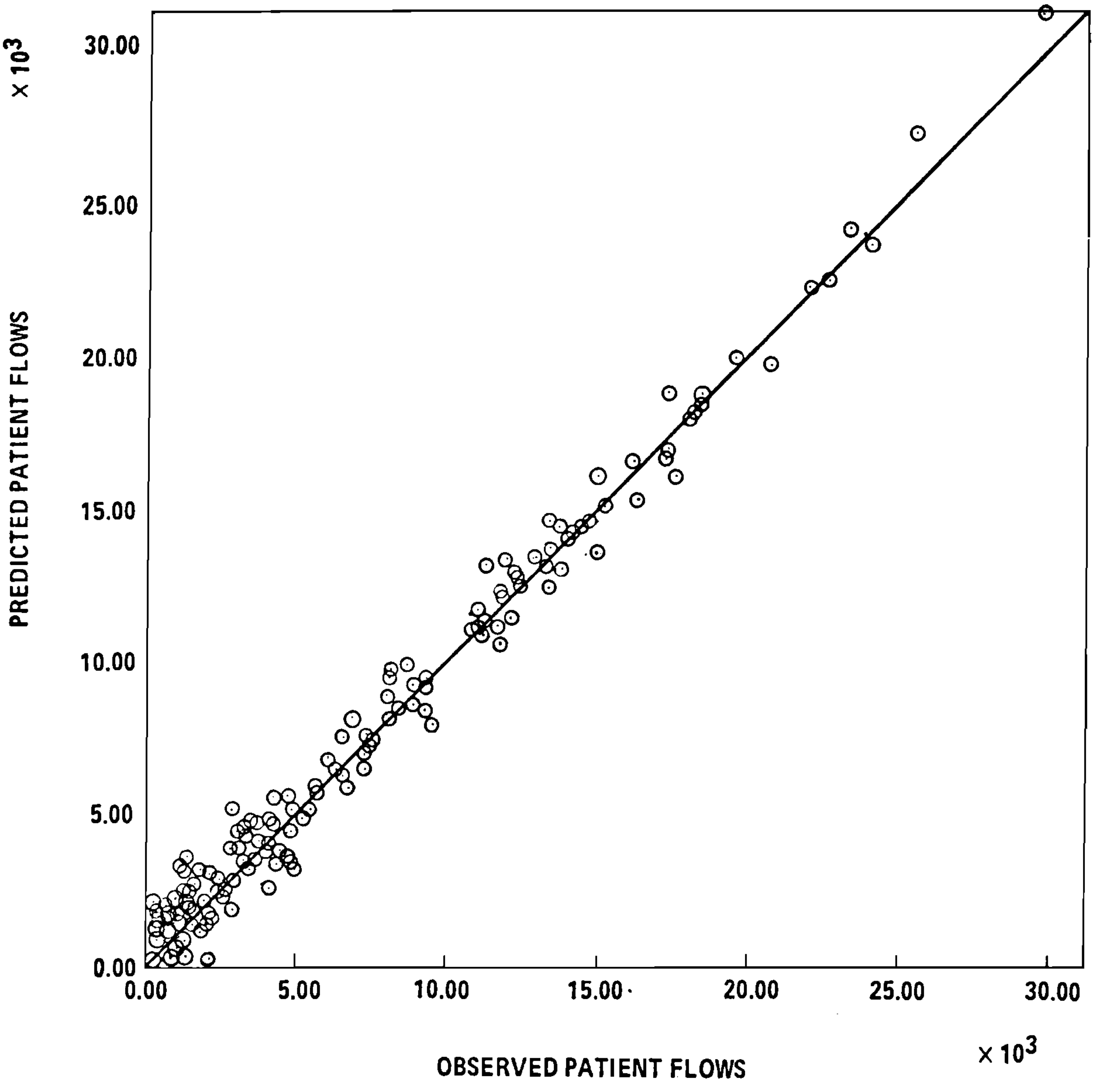


Figure 12. All specialties, 1977, Thames Regions; model 1, Matrix 3.

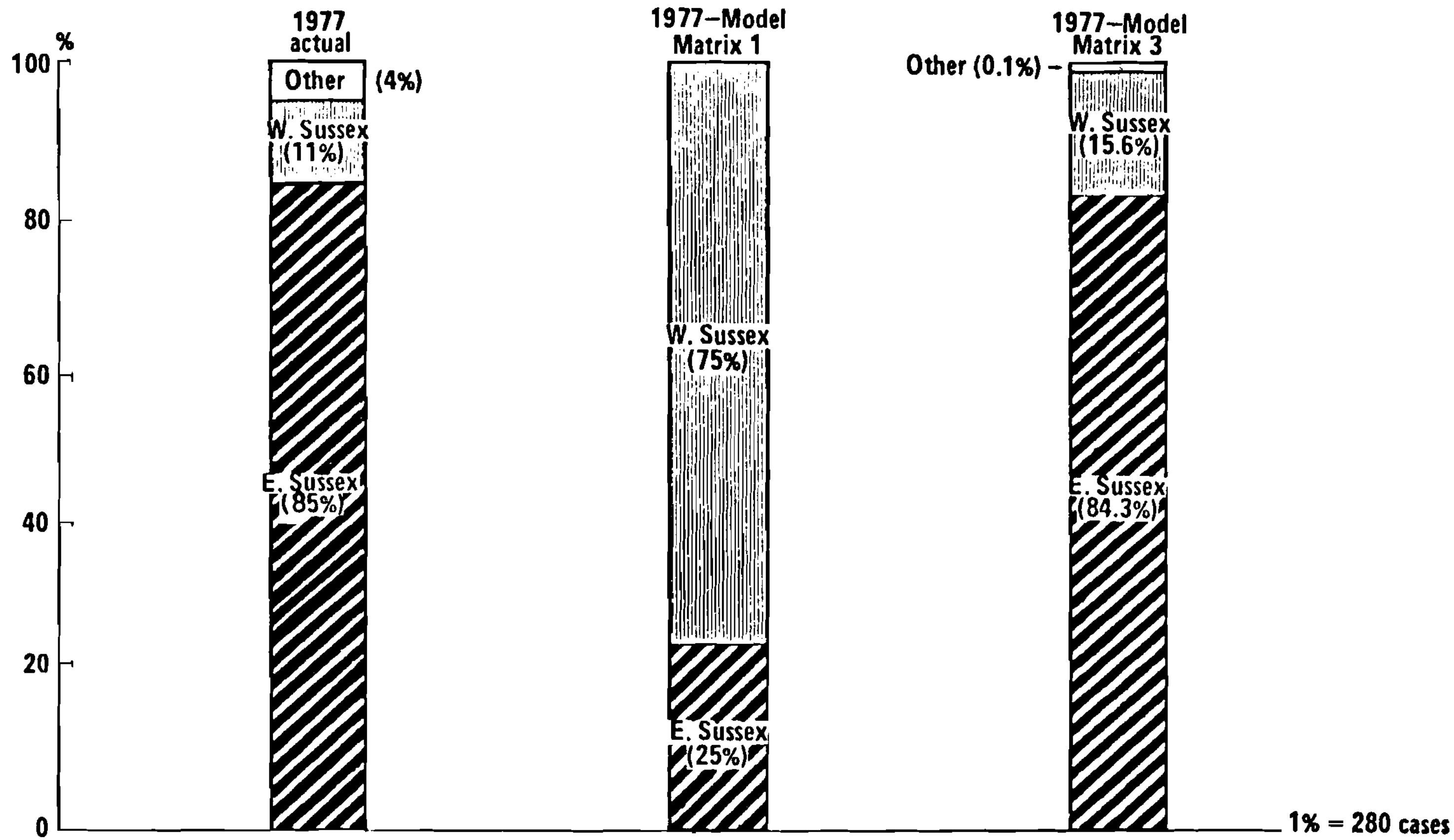


Figure 13. Model 1 - Brighton Health District: % distribution of area of residence for patients. COmparison of actual figures with those produced by the model.

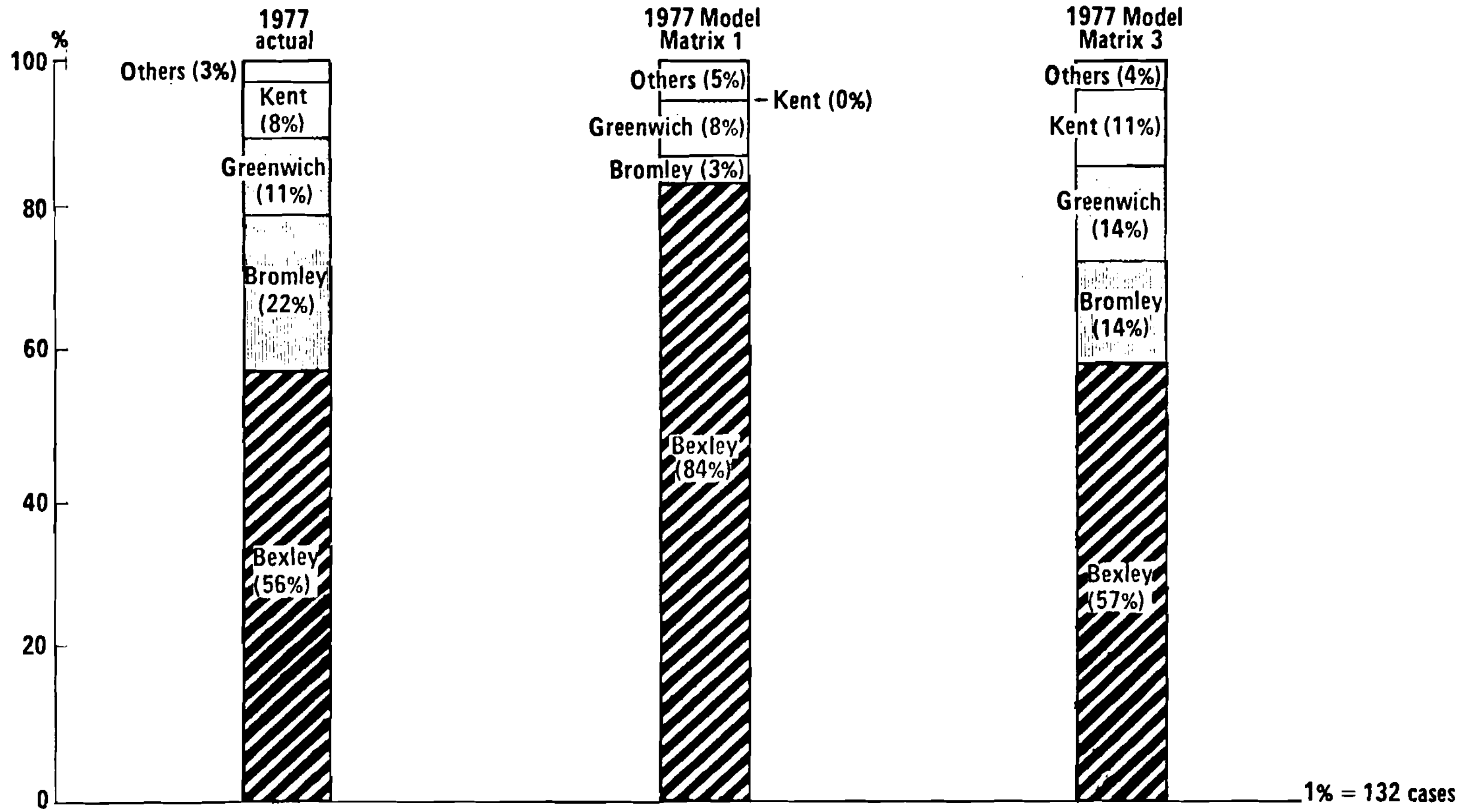


Figure 14. Model 1 - Bexley Health District: % distribution of area of residence for patients. Comparison of actual figures with those produced by the model.

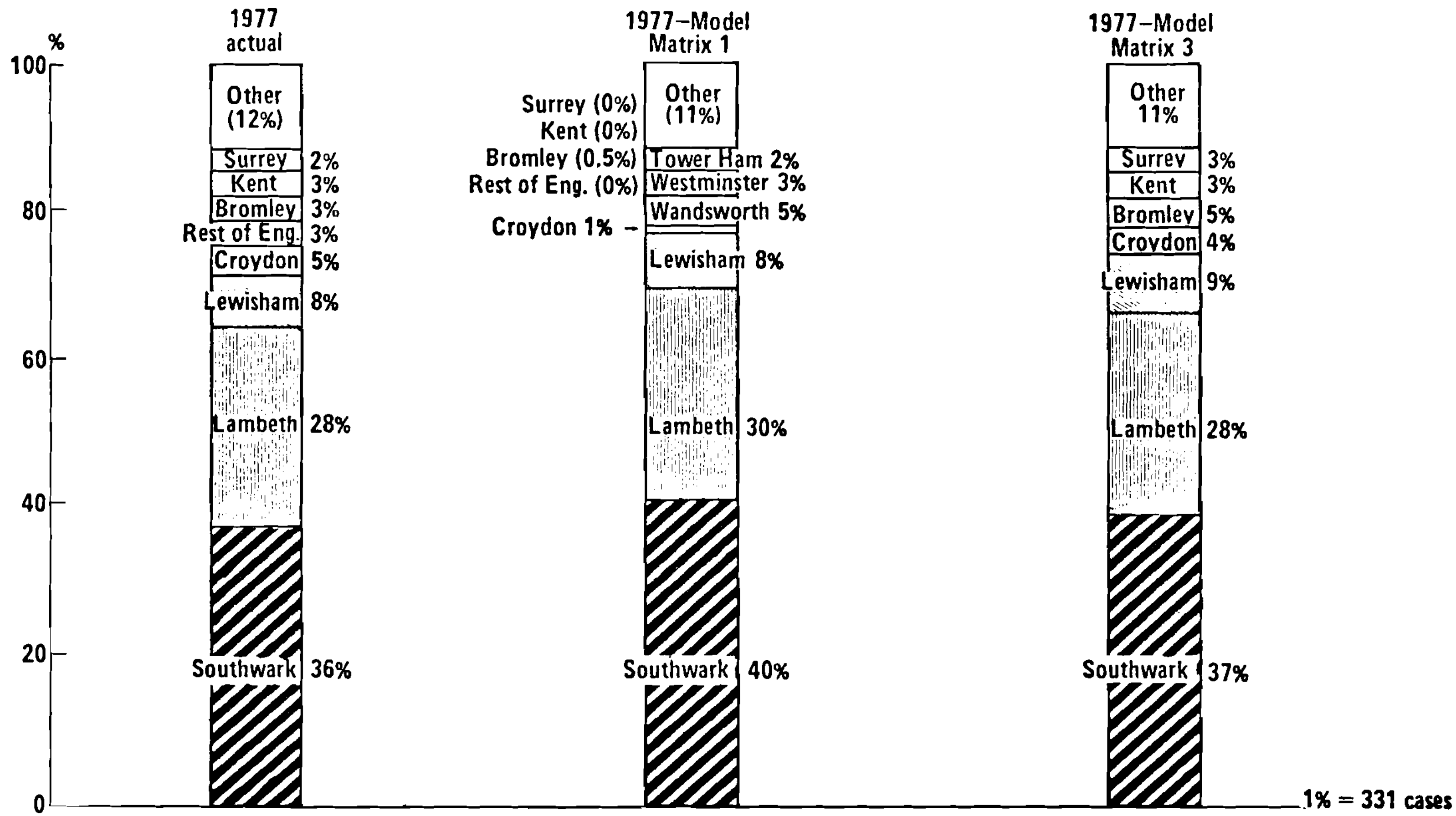


Figure 15. Model 1 - Kings' Health District: % distribution of area of residence for patients. Comparison of actual figures with those produced by the model.

the actual pattern of patient flow, only areas of residence contributing at least 2% of the health district's patients are shown. The other columns then show each origin shown in the first column plus any others with predicted contribution of over 2%. The three health districts have been chosen to demonstrate the variability which exists in the pattern of patient flow.

At one extreme is Brighton Health District (Figure 13) treating a total of 28,081 cases in 1977, where there are only two areas contributing more than 2% of the patients. Furthermore, a clear majority (85%) of patients come from one of these zones, East Sussex. The model using Matrix 1 greatly misrepresents this pattern, predicting that East Sussex contributes 25% and West Sussex 75% (compared with an actual figure of 11%). This result is clearly due to the use of crude distances between centroids and does not appear with either Matrix 2 (not shown in the figure) or Matrix 3, where (as the last column in the figure shows) the split between West and East Sussex is accurately reproduced.

The other two health districts shown, Bexley and Kings, exhibit a more complex pattern of patient flow. (This is partly a consequence of the smaller scale of the zoning system in this part of the study region as compared with the Brighton area). In Bexley (Figure 14) where total cases were 13,162 in 1977, the majority of patients (56%) came from the immediate surrounding origin zone, the London borough of Bexley, with remaining contributions from the three neighboring zones on the south side of the River Thames - Bromley, Greenwich, and Kent. This pattern is considerably distorted in the predictions using Matrix 1. Firstly, the contribution of Bexley is grossly overestimated. This is due to problems associated with the use of different, overlapping zoning systems for areas of residence and places of treatment (see Figures 4 and 5). The centroids of Bexley health district and Bexley borough were separated by a distance of only 0.5 kilometers, a substantial underestimate of the mean distance of Bexley residents from hospitals in the health district. When this distance was modified in Matrices 2 and 3 the figure for Bexley residents was much closer to the actual value. Matrix 1 also seriously underestimates the flow from Kent to Bexley; Matrix 2 is no better in this respect; while as can be seen, the use of Matrix 3 gives improved results.

The third health district, Kings' (Figure 15), with a total of 33,096 cases in 1977, shows the most complex pattern of actual flows, with 8 zones each contributing over 2% of the patients treated in Kings. Of these, three are outside the RHA in which the Kings' health district is situated. Kings shows most dependence for its patients on the local areas of residence of Southwark (36%) and Lambeth (28%), but treats a considerable number of patients from distant localities - zones such as Surrey (2%), Kent (3%) and rest of England (3%). Although the contributions from Southwark, Lambeth and Lewisham are reproduced very well by Matrix 1, the remaining areas shown in the "actual" column are all underpredicted, with zero value for Surrey, Kent, and rest of England. Instead, this version of the model produces flows from Wandsworth, Westminister, and Tower Hamlets, places which in terms of distance are close to Kings' but which in fact each contribute less than 1% of Kings' patients. The use of Matrix 3 avoids these problems: predictions for Lambeth and Southwark are improved and contributions from other zones are better represented, particularly those from the more distant locations. Matrix 3 is still unable, however, to reproduce the flow from the rest of England. This problem was found generally in all those health districts for which large numbers of patients came from the "rest of England". Generally such health districts were teaching districts, or had postgraduate hospitals in them.

6.5. Hospitalization Rates

This final section on the performance of model 1 in reproducing the calibration year data examines the prediction of hospitalization rates predicted by the model for each origin zone which were obtained using equation 4.9. Hospitalization rates were also calculated for larger areas of aggregation. Nine in all, these consisted of two areas for each of the four Thames

Regions (namely, that part of each region lying in the GLC and then the rest of the region itself), and finally all the external zones together (Oxford, East Anglia, Wessex, and rest of England).

Table 6 compares the performance of the model using each of the three matrices in reproducing the actual hospitalization rates in each of the nine aggregate zones. With each of the matrices the hospitalization rate of the external zones is slightly underpredicted. The results for Matrix 1 and Matrix 2 both show a clear pattern of overestimation in each of the four GLC zones and underestimation in each of the four non-GLC zones. This is a consequence of the flows from the county zones being generally underestimated as described earlier. The results produced by Matrix 2 are a definite improvement over Matrix 1, while as expected the predicted hospitalization rates associated with Matrix 3 are much closer to the actual values observed. In the latter case, there is no longer any consistent underestimation of hospitalization rates in the zones outside the GLC although each of the 4 quarters of the GLC is still slightly overestimated.

Examination of the results for individual origin zones shows a more complicated pattern. With Matrices 1 and 2, although each of the county zones is underestimated, not all of the London boroughs are overestimated. The boroughs in the center of London are all overestimated (by as much as double in some cases), while of the remaining boroughs, some are predicted fairly accurately, but others are underestimated by as much as the county zones. As the statistics in Table 5 show, the level of accuracy obtained with the first two matrices is generally bad. The results obtained with Matrix 3 meanwhile are shown in Figure 16 and Table 7. The figure shows a graph of the predicted against the actual hospitalization rates for each of the origin zones. Also plotted on the graph are the bands corresponding to $\pm$ 10% error in the predictions. As is seen, most of the predictions fall within 10% of the actual value, but as Table 7 indicates, there is still a tendency to overestimate the hospitalization rates in the center of London.

Table 6. Model 1 - Performance in reproducing actual hospitalization rates.

Area of residence	Actual 1977	Model 1 predictions		
		Matrix 1	Matrix 2	Matrix 3
GLC Area				
North West Thames RHA	119.1	153.9	139.9	122.7
North East Thames RHA	114.4	154.3	135.1	119.0
South East Thames RHA	116.8	148.5	135.3	118.6
South West Thames RHA	106.8	138.5	117.9	108.3
Outside GLC				
North West Thames RHA	87.4	53.8	58.2	91.4
North East Thames RHA	91.6	48.0	79.1	85.8
South East Thames RHA	94.8	67.4	90.3	95.0
South West Thames RHA	97.7	67.5	81.0	103.4
Rest of England (external zones)	84.7	83.8	83.7	83.8

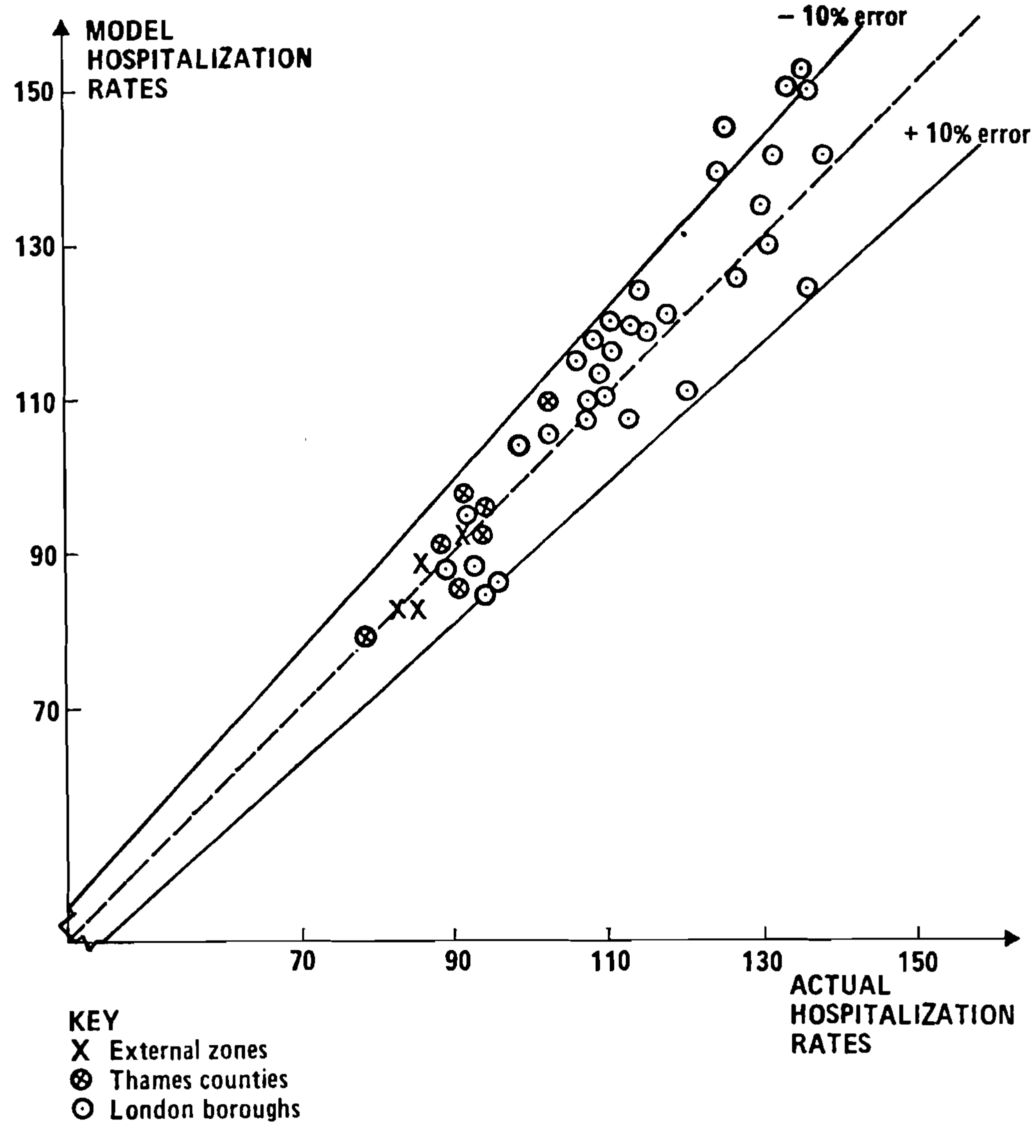


Figure 16. Model 1 - Graph of predicted hospitalization rates (Matrix 3) against actual hospitalization rates.

Table 7. Model 1, Matrix 3, Replication of hospitalization rates.

	Hospitalisation Rates*		
Central London Boroughs	Actual 1977	Model 1977	% error
Tower Hamlets	139.4	141.8	1.7
Westminster	136.6	151.1	10.6
Hammersmith	136.2	124.9	-8.3
Islington	136.1	152.2	11.8
Kensington and Chelsea	134.1	150.5	12.2
Wandsworth	130.7	135.3	3.5
Southwark	127.7	126.3	-1.1
City and Hackney	126.6	145.9	15.2
Haringey	125.2	139.2	11.2
Lewisham	118.9	121.2	1.9
Newham	115.1	119.1	3.5
Camden	111.5	116.1	4.1
Lambeth	109.9	113.1	2.9
Thames Region Counties			
Surrey	103.2	111.4	7.9
Kent	95.1	95.6	0.5
East Sussex	94.2	92.9	-1.4
Hertfordshire	91.9	97.7	6.3
Essex	91.6	85.8	-6.3
West Sussex	88.9	90.7	2.0
Bedfordshire	78.9	79.1	0.3

	Hospitalisation Rates*		
Other London Boroughs	Actual 1977	Model 1977	% error
Brent	132.2	141.8	7.3
Greenwich	131.1	130.3	-0.6
Harrow	120.7	111.0	-8.0
Waltham Forest	115.1	124.3	8.0
Merton	114.7	119.6	4.3
Barking	113.7	107.7	-5.3
Barnet	111.5	119.9	7.5
Bromley	110.7	110.3	-0.4
Ealing	109.6	118.3	7.9
Hounslow	108.2	109.4	1.1
Enfield	107.8	108.4	0.6
Bexley	106.8	115.3	8.0
Sutton	103.2	105.6	2.3
Richmond	99.2	104.0	5.1
Havering	95.9	86.2	-10.1
Hillingdon	95.6	85.3	-10.8
Croydon	93.5	88.1	-5.8
Kingston	93.3	95.1	1.9
Redbridge	88.8	86.9	-2.1
External zones			
Wessex	92.9	93.3	0.4
East Anglia	87.1	87.4	0.3
Oxford	85.8	83.9	-2.2
Rest of England	83.6	82.6	-1.2

*Cases per 1000 resident population.

Five of the seven zones where the percentage error is greater than 10% lie in the center of London, and in each of these five the model prediction is not wholly satisfactory.

6.6. Other Aspects of Calibration Using Model 1

6.6.1. The Region of Calibration

The region of calibration employed in the main series of results consisted of all flows with origins in the GLC area and destinations in the four Thames Regions. This region includes 2145 origin-destination zone pairs, just over two-thirds of the total number (3036) of cells in the trip matrix. This particular region was chosen because it is the largest area over which the straight line distance measures could be judged reasonably accurate, and because it omits all the very large origin and destination zones where the choice of centroid was imprecise. It is necessary, however, to ensure that the fit of the calibrated model in areas outside the region of calibration is adequate, if the model is to be used in these areas. The trip matrix statistics were therefore calculated for (a) trips lying within the Thames Regions (2600 cells) and (b) the whole trip matrix, using the calibrated version of the model with cost Matrix 3. These statistics are compared in Table 8 and as can be seen there is no evidence that the model is performing significantly worse in the areas outside the region of calibration.

6.6.2. The Use of Different Deterrence Functions

The performances of three alternate forms of deterrence functions were investigated. These are noted in Table 2 and consist of a power function, a mixed (exponential and power) function, and a modified exponential function. (In the last case the value k=2 was used.) The use of these functions was explored using cost Matrix 2, as the modifications incorporated in Matrix 3 were derived especially for the exponential deterrence function and hence were considered inappropriate for other functions. Both the power function and the modified exponential function gave substantially worse fits to the calibration data than that obtained with the exponential function. It is interesting

Table 8. Model 1 - Comparison of trip matrix statistics over various sections of the trip matrix using Matrix 3.

Statistic	Section of the trip matrix: Calibration region	Thames regions	Whole matrix
R^2	0.983	0.989	0.999
slope of regression line b	1.0001	1.0082	1.0024
intercept of regression line a	12.30	13.97	- 3.34
root mean square error	226.4	246.8	335.1
mean absolute error	79.3	91.1	108.4
mean absolute % error	118.5%	134.6%	129.2%

to note that this result for the power function is in exact contrast to that found with model 2 discussed below.

The fit obtained with the mixed exponential was very similar to that of the straight-forward exponential, giving a slightly better performance on some statistics, but slightly worse on others. As there was no indication that the mixed function could significantly improve the overall performance of the model, its use was not investigated any further. In the case of both the mixed and the modified exponential function these might better be used in the model as two parameter functions, i.e.:

$$c_{ij}^{\alpha} \exp\ (-\beta c_{ij}) \qquad (21)$$

and

$$\exp\ (-\beta c_{ij}^{\alpha})$$

where now both parameters α and β must be determined during the calibration process. The disadvantage of using such two parameter versions of the model is that the calibration process becomes much more complex (see Batty and Mackie, 1972) and without further investigation it remains uncertain whether any significant improvements would be obtained.

6.6.3. The Use of Standardized Mortality Ratios in the Calculation of Patient Generating Factors

The patient generating factors (PGFs) used in the earlier results did not make allowance for factors (other than the size and age/sex structure of resident populations) which may influence an area's propensity to generate patients. Other factors arguably ought to be included to reflect, for instance, the impact of environmental and socio-economic conditions on health care needs. Standardized mortality ratios (SMRs) of various kinds have often been used as measures of such relative need between residents of different areas (LHPC, 1979), and it was decided that their inclusion in the calculation of the PGFs could lead to an improvement in the performance of the model.

This approach was investigated using cost Matrix 2. Two different types of SMR were tried - overall SMRs and aggregated SMRs (which consist of an average of condition-specific SMRs, each weighted by the national bed usage in the clinical specialty).

The results obtained showed a slight improvement in the majority of the statistics with the use of overall SMRs, but a slight worsening in all of the statistics with the use of aggregated SMRs. While the changes in goodness-of-fit were not large enough for any definite conclusions to be drawn, the use of overall SMRs seemed to offer the greatest possibility of improvement, but this conclusion would need to be confirmed by further investigation. The set of modifications incorporated into cost Matrix 3 is dependent not only on the deterrence function used but also on the set of patient generating factors and thus a new set of modifications should be derived for use when any SMRs are included in the PGFs and this recalibrated version then compared with the Matrix 3 calibration with no SMRs.

6.6.4. Disaggregation of the Model

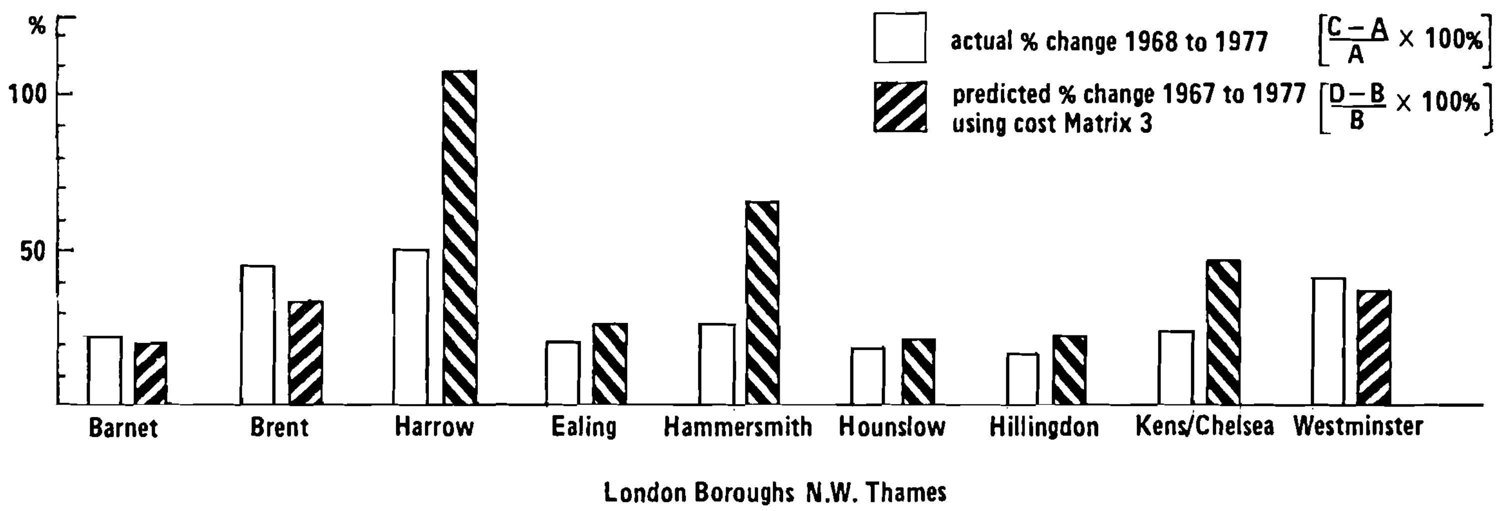
Model 1 has given results for all the acute specialties together. For many possible uses of the model results may be required for smaller groups of specialties or even for individual specialties. In particular, the regional specialties, such as radiotherapy and neurosurgery, which are provided in only a small subset of the health districts, generate a very different pattern of patient flow compared with other specialties such as general medicine which is provided in every health district. The acute specialties included in model 1 were split into two groups, regional specialties and subregional specialties (the precise division is indicated in Table 1) and calibrations for each group obtained using cost Matrices 1 and 2.

The new parameter values for each group were, as expected, higher than the all-specialty parameter for the subregional groups, and lower for the regional group, indicating that in general, patients are drawn from a much wider area for the latter group of specialties. The fit obtained for the subregional specialties was good, with higher R^2 values than those in the corresponding all-specialties model. On the other hand, the fit obtained with the regional specialties was much worse, with an R^2 value of 0.57 obtained with Matrix 2 (compared with 0.85 for all specialties and 0.87 for subregional specialties with the same cost matrix). It is likely that the reason for this bad result is that each regional specialty is provided in a different subset of health districts as opposed to the situation for the subregional specialties where the majority of districts provide services in all such specialties. This point may be clarified by investigating the use of the model for each regional specialty on its own.

6.7. Model 1 Validation - Back-Predictions for 1967

The three calibrated versions of model 1 were each used to back-predict hospitalization rates and patterns of patient flow for the year 1967 in the GLC part of the North West Thames Regional Health Authority. Considerable changes in population structure and the availability of beds took place in this area between 1967 and 1977, the greatest change to beds being the opening of a large new hospital, Northwick Park, in Harrow Health District. Caseload capacities and patient generating factors were recalculated for 1967, and it was assumed that the 1977 values of the other input variables of the model (namely,model parameter and elements of the cost matrix) would be appropriate for 1967. The predictions obtained were then compared to actual data available for the year 1968.

Figure 17 compares the model predictions using cost Matrix 3 with the actual changes occurring. Of the nine boroughs for which 1968 data were available, the model predicts both the 1967 hospitalization rate and the percentage change to 1977 satisfactorily in six boroughs. Of the remaining three boroughs, Hammersmith and Harrow both have predicted 1967 hospitalization rates which are very low in comparison with the actual figures for 1968 and this causes the predicted percentage changes to be over twice as large as those which actually occurred. In the final borough (Kensington and Chelsea), the percentage



HOSPITALIZATION RATES

	Barnet	Brent	Harrow	Ealing	Hammersmith	Hounslow	Hillingdon	Kens/Chelsea	Westminster	
ACTUAL 1968	90.9	91.3	81.0	90.2	109.8	92.6	82.0	110.2	97.7	A
MODEL 1967	99.8	106.5	53.6	94.0	75.4	93.4	70.5	104.8	112.1	B
ACTUAL 1977	111.5	132.2	120.7	109.6	136.2	108.2	95.6	134.1	136.6	C
MODEL 1977	119.9	141.8	111.0	118.3	124.9	109.4	85.3	150.5	151.1	D

Figure 17. Changes in hospitalization rates from 1967 to 1977 (model 1).

change is again over double the actual change; this, however, is due mainly to the poor replication of the 1977 hospitalization rate in the borough at the calibration stage, as the rate for 1967 is predicted accurately. This set of predictions is obviously not entirely satisfactory. In particular, the results for the two boroughs associated with the largest changes in local caseload capacity, Harrow and Hammersmith, are very poor and generally the predictions get worse as the amount of change in local capacity increases. Results obtained with cost Matrices 1 and 2 showed only a slightly lower level of accuracy with regard to the percentage change from 1967 to 1977, but were considerably worse in reproducing actual rates for 1967.

Data for 1968 were also available showing the percentage distribution of place of treatment for residents of Brent and Harrow. This, and the actual 1977 distributions, are compared with those predicted by the model using Matrix 3 in Figure 18. As the figure shows, the predicted distributions are close to those actually observed and the changes in the distributions from 1967/1968 to 1977 are in general very well represented by the model. In contrast, results obtained with Matrices 1 and 2 (not shown) considerably misrepresented the actual distributions in many places.

The results obtained, despite being very good in places, do not fully validate the model and several issues remain to be settled. Although the actual data for 1968 is based on a sample survey and is thus subject to some error, it seems unlikely that the results obtained could be entirely due to this, particularly since the mismatch is so large in two of the boroughs, Harrow and Hammersmith. One possibility is that parts of the model may not be correctly specified; for example, the use of alternate forms of deterrence function and the inclusion of SMRs in the calculation of PGFs have already been partially investigated. The use of other deterrence functions gave no improvement in the back predictions obtained, however (and as discussed earlier, the other functions also gave worse fits to the calibration year data). In contrast, the modification of the

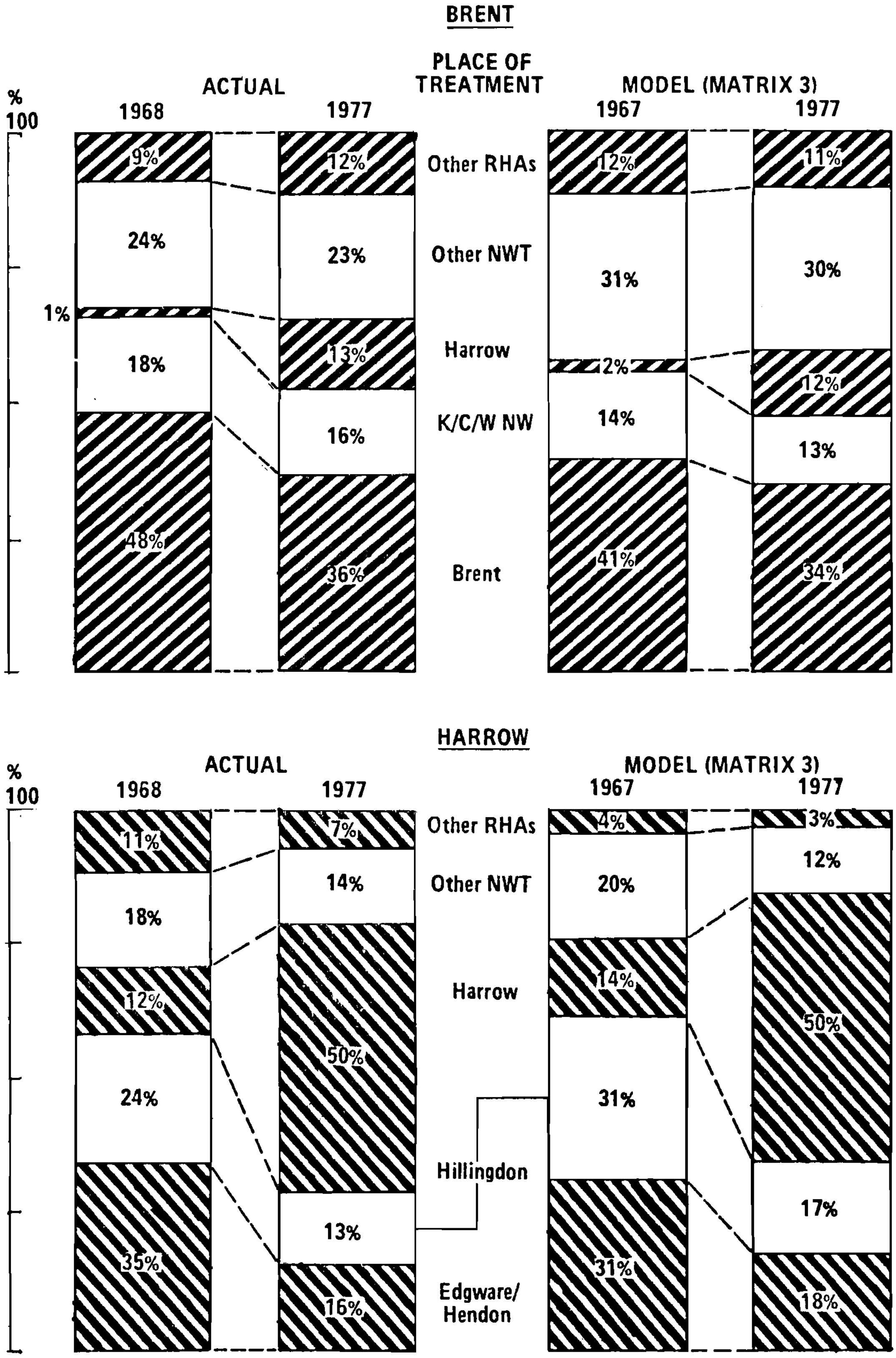


Figure 18. Percentage distribution of place of treatment for residents of Brent and Barrow.

patient generating factors using overall SMRs showed encouraging initial results. There are still refinements necessary, however, since any change in deterrence function or method of calculation of PGFs will affect the modifications to the distance matrix needed to obtain a final calibration. Further improvement may also be obtained by using 1967 SMR values for the 1967 prediction runs, rather than the 1977 values that have been used in the preliminary tests.

A basic assumption made for the purposes of prediction is that the cost matrix remains unchanged. This is certainly questionable, particularly in the case of Harrow Health District where the building of a completely new hospital, together with the changes in the public transport system and the ambulance services that accompanied it, seem unlikely to have left the 'cost' of receiving care in Harrow unaltered. However, as model 1 uses a cost matrix based on distance there exists no readily apparent method for systematic modification to reflect such changes. If a cost matrix based instead on travel time were used, then such modifications, based perhaps on surveyed changes in travel time, could be incorporated for use in prediction runs.

If none of the suggested alterations in the formulation of the model removes the large prediction errors obtained, there remains the possibility that the model is unable to reproduce changes of great magnitude or that there are time lags between the introduction of a new resource and its full utilization. The fact that the prediction errors are worse for areas where there have been large changes in the local input variables through investments, certainly seems to support this idea. The level of change occurring in Harrow Health District in particular was very high with a sevenfold increase in the case capacity of the district over the period 1967 to 1977. If this is so, then it might be possible to determine limits to the amount of change in the input variables, within which the model will reasonably predict the consequences of such change, but outside of which the predictions will be subject to uncertainly large errors.

Alternatively, it may be possible to define time lags, which have the objective of dampening down the effects of very large changes such as these. However, considerable further testing against actual data would be needed before any such procedures could be determined.

6.8. Results Model 2

Model 2 differs form model 1 in two ways: the zones in the area of interest are smaller and travel time is used as well as distance as measures of accessibility. Of particular interest therefore, is whether travel time predicts patient flows better than distance.

Before examining the results in detail the main findings can be summarized as follows. Firstly, the versions of model 2 which used travel time predicted patient flows better than the model 2 version based on distance. Secondly, unlike model 1 which used a negative exponential deterrence function, model 2 operated better with a simple power function. This finding was probably due to the finer zoning system used for this study area (see Figure 6). Thirdly, patient flows in the finer zoning system employed in model 2 are generally much harder to predict than in the zoning system in model 1. Finally, the proportions of patients predicted to travel to hospital by public and private transport were plausible, and fairly consistent with the limited data available for hospital trips in the London area.

In more detail, Figure 19 compares the observed flows from one representative health district (Kings' HD) with those predicted by four calibrated versions of model 2. Kings Health District generated 23,525 patients in all specialties (see Table 1) in 1977 of whom 20,216 were treated in hospitals in the southeast GLC. It is the GLC component only which is broken down in the diagram. As is seen the majority of patients sought treatment in Kings Health District with the remainder going to hospitals in the neighboring health districts of Guys' and St. Thomas'. From the calibration versions of model 2, it is plain that the only one able to portray this pattern with any

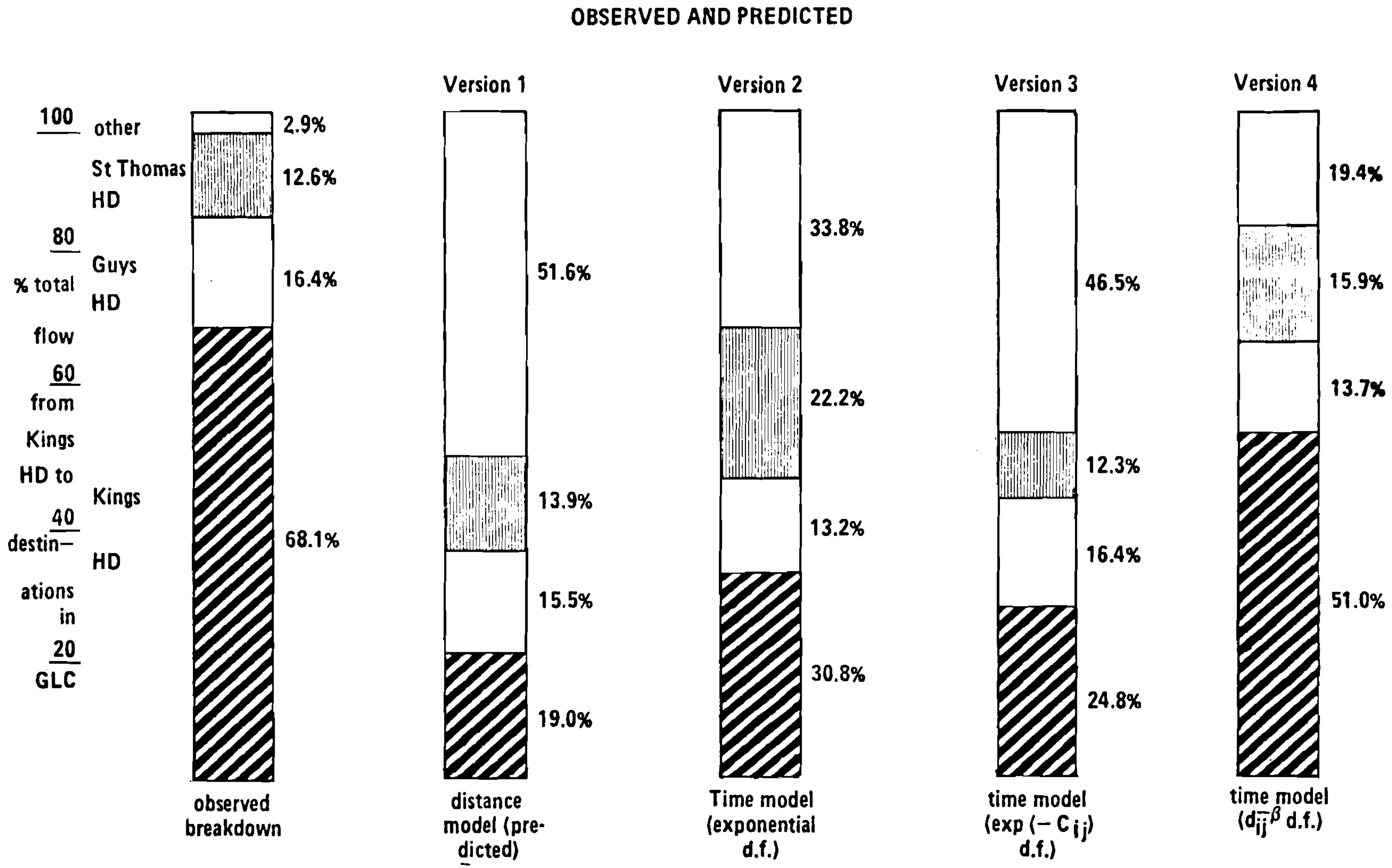


Figure 19. Model 2 - Breakdown of patient flows from Kings' Health District within the internal study area.

accuracy is the fourth, which uses the power function (versions 4). This version is especially successful in predicting the proportion as there is a strong tendency in all health districts in the study area for the largest proportion of patients to use locally available facilities. The reason why the power function predicts this better than the other function presented is that it gives more weight to flows over short distances. This can be seen from the graphical comparison of deterrence functions presented in Figure 8.

The distance version of model 2 uses unmodified distances and, from the experience of calibrating model 1, the results presented in Figure 19 could undoubtedly be improved upon by making similar assumptions about the effects on travel of traffic congestion and so forth. One modification that was made to this version, however, involved the redefinition of destination centroids, based not on the weighted center of population, but on centers defined by the weighted average caseloads of the hospitals. Although more testing is required, initial results were mixed, showing improved results in some zones and worse results in others. However, a combination of modifications similar to those used in model 1 would greatly increase the predictive accuracy of this version.

The interesting feature of versions 2 and 3 was their ability to portray accurately the total number of patients generated by a health district. Unfortunately, both failed in correctly apportioning the resultant flows among the various destinations. The impression gained, however, was that "travel time" is a significant improvement on "crude distance" used in version 1 in several aspects of the model 2's performance. Journey time to and from external zones outside the GLC could not be determined from available data in the Greater London Traffic Survey (GLTS), and so had to be estimated from public transport time tables and other sources. Thus, small modifications to these times ($\pm$ 10 - 15 minutes, say) were allowed in order to improve predictions between external and internal zones. In the longer run, however, a detailed travel survey would be the best answer to this problem. The effect of using the deterrence function in version 3 was to generate patients from farther afield, mostly at the expense of patients living in or near a destination zone. Thus, unlike version 4, this version regularly underestimated the flow of patients to neighborhood hospitals.

A comparison of the calibration statistics for all four versions is shown in Table 9.

Table 9. Model 2 - A comparison of calibration statistics for for different versions of model 2 using southeast GLC data.

Model	Distance(kms)	Travel(hours)	Travel time	Travel time
Deterrence function	$\exp(-\beta c_{ij})$	$\exp(-\beta c_{ij})$	$\exp(-\beta_{ij}^{2})$	$c_{ij}^{-\beta}$
Calibration method	max.like	max.like	slope	slope
	0.11	5.72	1.665	2.25
$\bar{c}$	12.11	0.75	0.87	-1.03^{a}
$\chi^2 (\times 10^6)$	6.24	12.79	2.84	1.10
RMSQ error $(\times 10^3)$	4.32	1.99	1.35	0.92
R^2	0.94	0.97	0.96	0.99
Intercept	-52.85	-87.29	-1.00	-7.89
Slope	0.99	1.13	1.00	1.01
Mean abs.err. $(\times 10^2)$	1.31	2.86	2.91	2.25
Mean abs.% err. $(\times 10^2)$	6.47	6.17	10.90	6.83

$$^{a}\ \frac{\sum_i \sum_j T_{ij} \log c_{ij}}{\sum_i \sum_j T_{ij}}$$

A close examination of the results further underlines the warnings expressed in the sections on calibration (see section 4). For example, R^2 is consistently high, and yet the above discussion has shown that substantial variations exist in the predictive abilities of each version. Furthermore, the mean predicted time of travel is clearly related in versions 3 and 4 to the form of the deterrence function. In version 1 and 2, calibrated using method 1 (section IV), the observed mean distance, $\bar{c}$ (calculated), and mean travel time (estimated) were supplied as is necessary for this form of calibration. As for the parameter β, this varies with the units of c_{ij} (i.e. kilometers or hours), with the form of the deterrence function, and with the value of $\bar{c}$ which increases as β decreases.

The advantages of the slope calibration over maximum likelihood are also brought out by this table. The expected value of the intercept term should ideally be zero, and in this respect versions 3 and 4 come closer than either versions 1 or 2, which use maximum likelihood. Slope estimates using maximum likelihood likewise depart from one, but are still close enough to raise problems of deciding which calibration method is best. A diagram showing the predicted versus observed flows is contained in Figure 20 for version 4. The more condensed scatter of observations compared with model 1 (Figures 10 to 12) is due to the substantial transition in zonal area between the internal and external study region (Figure 6).

A further characteristic of versions calibrated using travel time is that output with all the other statistics are estimates of the proportion of patients traveling by public and private transport. To some extent these estimates are a side-issue, arising only because of the necessity of using two sets of journey times. However, a broad check for accuracy can be obtained by comparing predictions with results obtained by Mayhew (1979) in a travel survey conducted at 14 hospitals in the London area - not just southeast GLC considered here.

Table 10 compares the mean proportions obtained in the survey with those predicted in the versions 2, 3, and 4 of model 2 for car travelers. In the table London is divided into three rings: the inner, the suburban,and the outer suburban. The conclusions are that all three versions give broadly

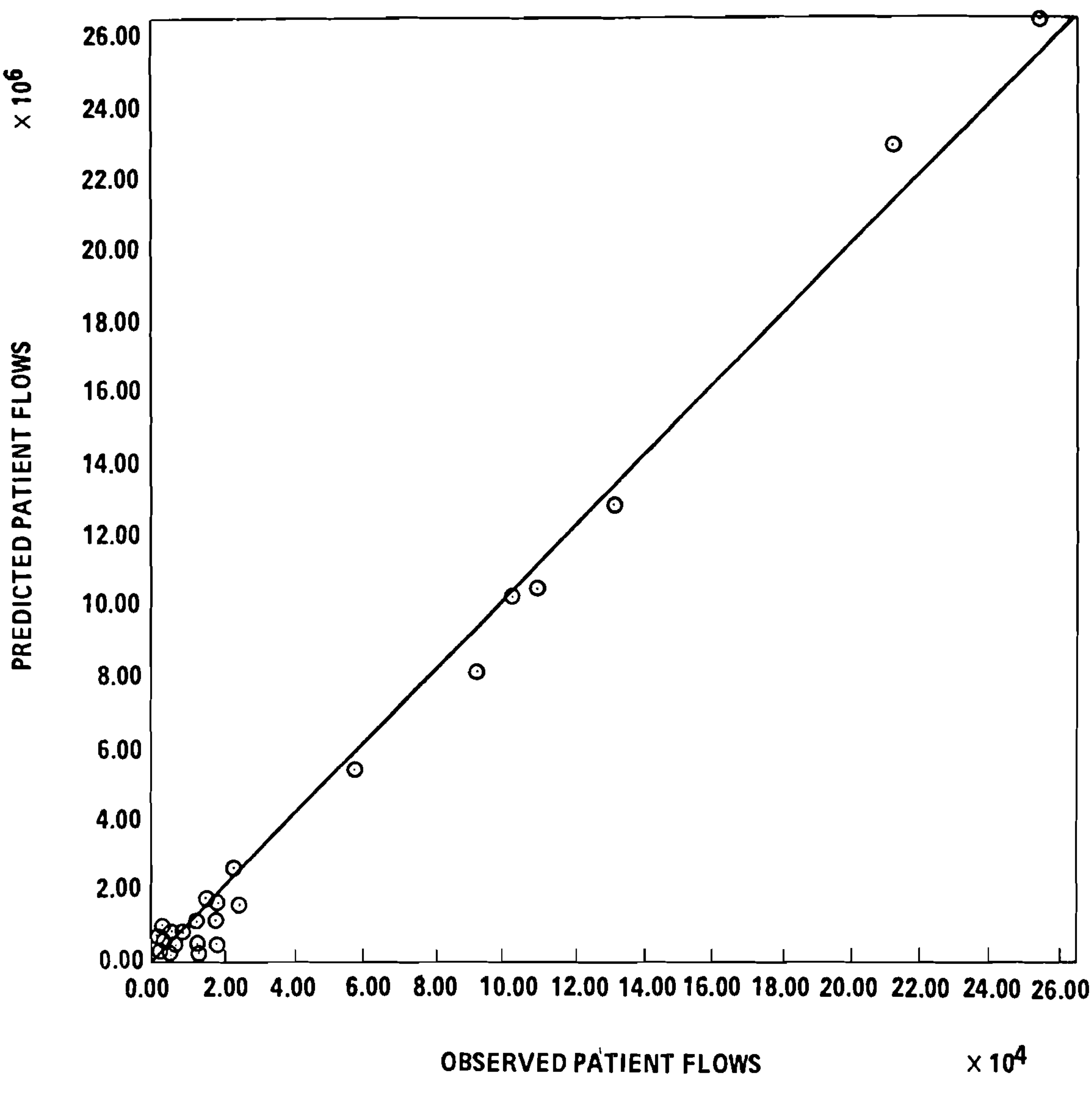


Figure 20. All specialties, 1977; southeast model (model 2).

Table 10. Percentage of patients traveling to hospital by private transport: Observed and predicted (model 2).

Ring	Survey	Version 2	Version 3	Version 4
Inner	27.5	51.69	29.0	40.0
Suburban	38.9	65.31	41.5	51.3
Outer Suburban	58.1	81.52	56.2	60.1

similar results which at this level are in the correct relation with the survey.

A smaller proportion use cars in central London because

(a) The general availability of public transport is higher
(b) The proportion of car-owning households is smaller
(c) The congestion and parking charges are an added inconvenience

The opposite is true of the outer suburbs, and so it is encouraging that the model correctly predicts this trend.

None of the versions shown assumes parking charges (in the form of a time penalty). This could be used to correct the tendency of the model to overpredict inner London car flows. A second method of reducing mobility, if too many are predicted to travel by car, is to lower the percentage car availability in households. In the above table, this is assumed to be 25% among car-owning households (see section III). A smaller value (say 20%) would probably suffice to overcome the overprediction apparent in version 4. Better data are necessary, however, to make a fuller evaluation of the 'modal split submodel' (section 3).

Table 11 compares the observed and predicted hospitalization rates for health districts in the internal study area. The results at this level of aggregation seem reasonable; but split down to the level of the Traffic district the predictions are far less satisfactory.

Table 11. Comparison of observed and predicted hospitalization rates (model 2).

	Observed	Predicted
Bromley	116.2	116.5
Greenwich	142.4	135.2
Bexley	112.4	103.2
St. Thomas	117.2	106.8
Kings	114.9	130.7
Guys	131.7	142.9
Lewisham	121.7	118.1
Externals	111.2	111.4

This is due to the greater difficulty the model has in correctly predicting small flows. One of the least satisfactory results in the table is for Kings' Health District where hospitalization rates are over-predicted. The reason for this is apparent from Figure 19 which shows under version 4 that far too many patients are allocated to areas not in the Kings', St. Thomas' or Guys' health districts.

7. GENERAL CONCLUSIONS AND OTHER CONSIDERATIONS

In this paper we have presented an analysis of the initial results obtained using the model, RAMOS, to predict trips to hospital for acute in-patient treatment. These results were in general very good, particularly in terms of the ability of the model to replicate the observed flows for the calibration year. Nevertheless, there are several issues which remain unresolved or need clarification, and so warrant further research. These relate to the specification of the input variables and whether they can be improved in anyway. For instance, it may be possible to find more appropriate measures for the patient generating factors, such as composite health indices or aggregate morbidity estimates, which are able to perform better than the approach used here based on relative utilization rates and SMRs. Also it was shown that simple distance acts as a poor distributor of trips unless it is substantially modified. The use of travel time indicated an improvement but raised the level of model complexity. The eventual solution to this problem is uncertain but it may involve a form of generalized cost which takes into account not only time and distance, but also the opportunity costs involved in entering hospital. The latter will depend on patient income and other factors.

In the validation experiment, the model was able to forecast correctly the direction of change in hospitalization rates in all the zones considered. In addition, the patient flows were apportioned to each destination with reasonable accuracy. In the case of the two zones where caseloads changed most, the model over-predicted the resultant charges in hospitalization rates to a significant degree. Several reasons were suggested for this earlier, including the incorrect specification of some of the variables. The lesson from this validation exercise is that it pays to give very careful consideration to the results which are forecast from a given set of measures, and to the extent to which all these measures change in time (for example, no consideration was given to possible changes in travel cost during the back-predicting exercise).

How will the model be used in a planning context? RAMOS acts in the input variables; namely, the patient generating factors, the caseload capacities, and the cost matrix, to produce the basic output of the model - a matrix of patient flows between origin-destination pairs. The output matrix can be manipulated in a large variety of ways to produce information of considerable value to decision makers. Thus by varying the assumptions concerning population structure, resource availability, and transport services, it is possible to gauge the likely impact of change on such diverse indices as the hospitalization rates in individual zones of residence or the catchment populations of particular destinations, and to measure the effects on patient flows due to the opening or closure of facilities in the region of interest. It is envisaged that the assumptions concerning change which are put into the model will be provided by other submodels concerned with either resource supply, demographic change, or morbidity.

In the Health Care Task at IIASA there have been developed a number of models which are admirably suited to these purposes (Shigan et al., 1979). On the output side, it should be a simple matter to transform, if necessary, the results into financial terms. Currently the model is designed for application in health care systems in which service availability is free. It can be argued that rationing in these systems takes place not through any market mechanism, but principally through factors such as accessibility to supply. Even so, the model as presented considers only one layer in the multitude of interactions that take place. It ignores, for example, the trade-offs which occur by treating patients in different ways and with different resources, or the interactions which arise between patient categories, resources, and modes of care (Gibbs, 1978; Hughes, 1978) due to hospital admissions policies. It may, however, be possible for these shortcomings to be remedied at a later date.

A question which arises is whether this approach can be used in different types of health care systems. The signs are that it can, but that changes will be needed depending on the system. While the gravity formulation would remain essentially intact, it is considered probable that variables will need to be respecified in order to reflect the different motivational apsects associated with, for example, market-based health care systems as compared with planned systems. In the former, income is known to be an important determinant in the consumption of certain types of health care services, and it would be appropriate to incorporate this fact, for instance, in the definition of patient generating factors. Also it is possible that the constraint on resources would have to be taken off supply, and put instead on demand. The resultant model would then be similar to that applied, for example, by Morrill and Kelley in the United States (Morrill and Kelley, 1970). In sum, therefore, the gravity model approach is thought to have considerable potential both in decision making and forecasting the resulting demands on health care services when resource supply and population structure are changing simultaneously over space.

APPENDIX

The following sections give an overview of the RAMOS computer program and an example of the output obtained in a typical run. The data in this print-out refer to model 1; while calibration is by the slope method (see page 28). Not all the print-out is included as the matrices (when all the options are employed) are extremely large. The program was written for use on a CDC 7600 machine and is capable of handling systems with up to 80 zones. The PDP 11/70 at IIASA is a smaller machine and so the program had to be adjusted accordingly*. Currently the maximum size of system accepted at IIASA is 45 x 70 zones, data space being the main limitation. Two types of singly constrained gravity models may be run using the program.

A. Attraction Constrained Model

This is the model investigated in this paper. Trips to destination zones are constrained so that the capacity of each zone is not exceeded. A supply driven model, it is formulated as follows:

---

*The authors are extremely grateful to Peer Just, IIASA, for making the required conversion.

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

where

T_{ij} = predicted trips between zones i and j
D_j = caseload capacity of zone j
W_i = patient generating factor for zone i (in an index of propensity of residents of i to generate patients)
$f(\beta,\ c_{ij})$ = deterrence function. This is a function of some measure of the cost of travel, c_{ij}, from zone i to j. Usually, it is the negative exponential [exp $(-\beta c_{ij})$].

$$B_j = \frac{1}{\sum_i W_i\ f(\beta,\ c_{ij})}$$

which ensures,

$$\sum_i T_{ij} = D_j$$

B. Production Constrained Model

Here, trips are constrained so that the demand arising from each zone is met exactly. This is a demand driven model of the 'shopping' type (see for instance, 'Urban and Regional Models in Geography and Planning', by A.G. Wilson, 1974). It is written,

$$T_{ij} = A_i\ W_i\ D_j\ f(\beta,\ c_{ij})$$

where T_{ij}, $f(\beta,\ c_{ij})$ are as in (i) but now,

D_j = attractiveness factor for hospitals in zone i
W_i = demand from zone i in terms of the number of cases requiring treatment

$$A_i = \frac{1}{\sum_j D_j \; f(\beta, \; c_{ij})}$$

which ensures that

$$\sum_j T_{ij} = W_i$$

With each model, using different assumptions about the cost-distances in the system, a variety of versions can be developed. For the CDC program, the full range is shown in Table A1.

Table A1. Model versions available using RAMOS program.

Type	Version			
	1	2	3	4
A. Production constrained B. Attraction constrained	single mode cost=distance (centroids supplied)	single mode cost=private transport time (matrix supplied)	single mode cost=public transport time (matrix supplied)	two modes public and private transport times (matrices supplied)

For the IIASA program, only versions 1 to 3 are available.

In addition to this program which can be used for both calibration and forecasting runs, another program has been developed and tested at IIASA that is used only for forecasting purposes. The main difference is that it simply takes a parameterized model, and then tests the flow consequences of changes on the input variables. The print-out is more detailed, however, with estimates of catchment populations, catchment areas, and the average costs of travel between zones.

Data Requirements

The following table summarizes data requirements for the all versions set out in Table A1.

Table A2. RAMOS: the data requirements.

Variable	Description	Versions for which required
$\{N_{ij}\}$	Actual patient flows between i and j in specialties of interest	All
W_i	Patient generating factors	All
W_i^1	PGF for car-owners (i.e. W_i disaggregated)	A4, B4
W_i^2	PGF for non-car-owners	A4, B4
D_j	Case capacities in j (i.e. resource levels)	All
$\{c_{ij}^1\}$	Cost Matrix 1 (distance, time by mode 1 or some equivalent measure)	A2,A4,B2, B4
$\{c_{ij}^2\}$	Cost Matrix 2 (time by mode 2 or some equivalent measure)	A3, A4, B3, B4
$\{x,y\}$	Centroids	A1, B1
$\bar{c}$	Average cost of travel in same units as for c_{ij}	All (maximum likelihood calibration but not slope calibration)
P_i	Origin populations	All
P_j	Destination populations	All

Description of Sample Output

The sample output overleaf is structured in the following way. Page A1 is a summary of the run parameters and options desired in the run. A one-zero variable is a switch controlling the level of detail required in the output. Page A2 is a typical iteration sequence using the slope calibration procedure. It stops when the slope of the regression of prediction on observed patient flows is within the desired degree of accuracy of one (column 8). Page A3 shows the results for origin zones; and page A4 the results for destinations. As this output on page A4 is for model 1 no breakdown of flows by public and private transport is produced. On page A6 the results are aggregated into larger areas for ease of reference. This aggregation can be in any desired form. The next three pages provide sample outputs from three matrices: the actual flow matrix, the predicted flow matrix, and the cost matrix. Only 30 x 15 of the 44 x 69 zones are shown. The final page is a scattergram of observed and predicted flows here within the region of calibration (33 x 65 observations). The numbers refer to superimposed data points ($X \geq 10$).

REFERENCES

Batty, M., and S. Mackie (1972) *The Calibration of Gravity, Entropy and Related Models of Spatial Interaction.* Environment and Planning 4:205-233.

Carrothers, G.A.P. (1956) *An Historical Review of the Gravity and Potential Concepts of Human Interaction.* Journal of the American Institute of Town Planners 2:94-102.

Crawford, K.A.J., A.A. Stroud, and W. Tyrell (1975) *Greater London Transportation Survey, Internal Zone Coding.* GLC Research Memorandum.

Feldstein, M.S. (1965) *Hospital Bed Scarcity, An Analysis of Inter-regional Differences.* Economica 32:393-409.

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

Hughes, D.J., and A. Wierzbicki (1980) *DRAM: A Model of Health Care Resource Allocation.* RR-80-23. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Hyman, G.M. (1969) *The Calibration of Trip Distribution Models.* Environment and Planning 1:105-112.

Kitsul, P. (1980) *A Dynamic Approach to the Estimation of Morbidity.* WP-80-71. Laxenburg, Austria: International Institute for Applied Systems Analysis.

LHPC (1979) *Acute Hospital Services in London.* A profile by the London Health Planning Consortium HMSO, London.

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

Mayhew, L.D., and A. Taket (1980) *Modeling Patient Flows: A Gravity Model Approach.* An internal report for the Operational Research Service of the Department of Health and Social Security.

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

Morrill, R.L., and M. Kelley (1970) *The Simulation of Hospital Use and the Estimation of Location Efficiency.* Geographical Analysis 2:283-299.

RAWP (1976) *Sharing Resources for Health in England.* Report of the Resource Allocation Working Party, HMSO, London.

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

Wilson, A.G. (1967) *A Statistical Theory of Spatial Distribution Models.* Transportation Research 1:253-269.

Wilson, A.G. (1970) *Advances and Problems in Distribution Modelling.* Transportation Research 4:1-18.

Wilson, A.G. (1971) *A Family of Spatial Interaction Models and Associated Developments.* Environment and Planning 3:1-32.

Wilson, A.G. (1974) *Urban and Regional Models in Geography and Planning.* Wiley, London.

RECENT PUBLICATIONS IN THE HEALTH CARE SYSTEMS TASK

Shigan, E.N., ed. (1978) *Systems Modeling in Health Care.* Proceedings of an IIASA Conference, November 22-24, 1977 (CP-78-12).

Gibbs, R.J. (1978) *The IIASA Health Care Resources Allocation Sub-Models: Mark 1* (RR-78-08).

Gibbs, R.J. (1978) *A Disaggregated Health Care Resource Allocation Model* (RM-78-01).

Kaihara, S., N. Kawamura, K. Atsumi, and I. Fujimasa (1978) *Analysis and Future Estimation of Medical Demands Using A Health Care Simulation Model: A Case Study of Japan* (RM-78-03).

Fujimasa, I., S. Kaihara, and K. Atsumi (1978) *A Morbidity Submodel of Infectious Disease* (RM-78-10).

Propoi, A. (1978) *Models for Educational and Manpower Planning: A Dynamic Linear Programming Approach* (RM-78-20).

Klementiev, A.A., and E.N. Shigan (1978) *Aggregate Model for Estimating Health Care System Resource Requirements (AMER)* (RM-78-21).

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

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

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, 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., Gibbs, R., 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-80-23).

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

Sample Output

details of run

```
run title          Resource allocation model over space: a trial run
                   using the slope calibration. Data  are for the four
                   thames regions model (model 1).Cost matrix is "matrix 3".

 44  n number of origin zones

 69  m number of destination zones

 33  nz no of origins used in calibration

 65  mz no of destinations used in calibration

  9  nd no of districts after aggregation

  2  ic type of model  1 single mode cost=distance (centroids supplied)
                       2 single mode cost=distance (matrix supplied) or cost=private transport times
                       3 single mode cost=public transporttimes
                       4 two modes,public and private,cost=transport times

  1  type of model  1=attraction constrained,2=production constrained

  1  type of run  1=calibration,2=prediction

  1  kpn output of actual trip matrix

  1  kpt output of predicted trip matrix

  1  kpc output of cost matrix(s)

  1  jp output of results for origins and destinations

  1  is 1=statistics required for every step in calibration
        2=stats required only for final step in calibration

  0  js final statistics for prediotion run

  1  jg graphics

  0  je elasticities

  0  jq tij to perm file
```

Resource allocation model over space: a trial run
using the slope calibration. Data are for the four
thames regions model (model 1).Cost matrix is "matrix 3".

actual mean cost 5.83660

calibration

iteration	parameter	mean cost	chisquare	rmsq error	r square	regression coeffs		mean abs er	mean abs pc er
1	0.30000	7.40751	0.107140e+20	402.8	0.9668	52.86	0.8208	122.3	155.1
2	0.31000	7.27141	0.494569e+20	359.3	0.9711	46.52	0.8501	113.1	147.4
3	0.32000	7.14439	0.228744e+21	319.8	0.9746	40.27	0.8786	104.2	140.7
4	0.33000	7.02560	0.105981e+22	285.2	0.9774	34.11	0.9063	95.99	134.9
5	0.34000	6.91433	0.491788e+22	257.3	0.9797	28.06	0.9331	89.15	129.8
6	0.35000	6.80995	0.228523e+23	237.5	0.9813	22.13	0.9590	84.02	125.2
7	0.36000	6.71186	0.106321e+24	227.5	0.9824	16.32	0.9841	80.53	121.1
8	0.37000	6.61955	0.495211e+24	227.9	0.9830	10.64	1.008	79.47	117.4

Resource allocation model over space: a trial run

using the slope calibration. Data are for the four

thames regions model (model 1).Cost matrix is "matrix 3".

origin	patient generating factors		resident population	predicted cases	hospitalisation rate	error in hosp rate
barnet	0.24880	0.00000	292200.0	35062.9	119.9962	8.5
brent	0.21466	0.00000	255900.0	36314.1	141.9073	9.7
harrow	0.16828	0.00000	198900.0	22125.0	111.2367	-9.5
ealing	0.24523	0.00000	291900.0	34425.2	117.9348	8.3
hammersm	0.13960	0.00000	163900.0	20614.2	125.7731	-10.5
hounslow	0.16729	0.00000	200000.0	21870.7	109.3535	1.2
hillingn	0.18976	0.00000	229000.0	19643.7	85.7803	-9.8
kens+che	0.13405	0.00000	159500.0	24166.6	151.5145	17.4
westmter	0.18311	0.00000	209900.0	31906.2	152.0067	15.4
barking	0.13051	0.00000	153000.0	16476.4	107.6888	-6.0
havoring	0.19794	0.00000	240400.0	20786.3	86.4654	-9.4
camden	0.16066	0.00000	189300.0	22162.7	117.0772	5.6
islingtn	0.14189	0.00000	168400.0	25835.6	153.4178	17.3
city	0.00456	0.00000	5600.0	2449.1	437.3382	138.9
hackney	0.16412	0.00000	194000.0	26879.0	138.5517	17.0
newham	0.19151	0.00000	229700.0	27345.5	119.0487	4.0
tower h	0.12751	0.00000	150300.0	21496.6	143.0249	3.7
enfield	0.22149	0.00000	260500.0	28240.0	108.4067	0.6
haringey	0.18979	0.00000	227900.0	31816.5	139.6071	14.5
redbridg	0.19466	0.00000	229000.0	19831.4	86.6002	-2.2
waltham	0.18860	0.00000	221800.0	27596.2	124.4194	9.3
bexley	0.17810	0.00000	214700.0	24769.8	115.3693	8.6
grenwich	0.17261	0.00000	206200.0	26953.0	130.7128	-0.4
bromley	0.24612	0.00000	293100.0	32339.9	110.3375	-0.4
lambeth	0.23639	0.00000	282900.0	32108.3	113.4970	3.6
lewisham	0.20633	0.00000	243800.0	29585.5	121.3514	2.4
sthwark	0.19054	0.00000	224100.0	28513.2	127.2344	-0.5
croydon	0.26638	0.00000	321900.0	28313.9	87.9588	-5.5
kingston	0.11616	0.00000	136400.0	13007.8	95.3649	2.0
richmond	0.14196	0.00000	165100.0	17129.4	103.7520	4.6
merton	0.14252	0.00000	165800.0	19857.0	119.7647	5.0
sutton	0.14166	0.00000	166700.0	17699.0	106.1730	3.0
wandswth	0.23430	0.00000	278300.0	37850.0	136.0043	5.3
bedfords	0.39441	0.00000	489600.0	38629.1	78.8992	0.0
herts	0.76754	0.00000	941600.0	91349.2	97.0149	5.1
essex	1.17865	0.00000	1425900.0	122022.9	85.5761	-6.0
e sussex	0.58425	0.00000	652900.0	60593.4	92.8066	-1.4
kent	1.20948	0.00000	1445100.0	138222.4	95.6490	0.6
surrey	0.82827	0.00000	995800.0	110465.9	110.9318	7.7
w sussex	0.54404	0.00000	625100.0	56576.2	90.5075	1.7
oxford	1.81209	0.00000	2237100.0	187625.0	83.8697	-1.9
e anglia	1.51833	0.00000	1827400.0	159730.0	87.4083	0.3
wessex	2.20967	0.00000	2661000.0	248246.2	93.2906	0.3
other	21.61957	0.00000	26079700.0	2154148.0	82.5987	-1.0

Resource allocation model over space: a trial run

using the slope calibration. Data are for the four

thames regions model (model 1).Cost matrix is "matrix 3".

destination	caseload capacity	local population	balancing factor	cases per head of local population	per cent by private transport	per cent by public transport
n beds	14637.00000	226800.0	214.7293	64.5370	0.00	0.00
s beds	16690.00000	262800.0	195.4493	63.5084	0.00	0.00
n herts	14268.00000	176100.0	15.1169	81.0222	0.00	0.00
s herts	11496.00000	275000.0	16.7511	41.8036	0.00	0.00
nw herts	13826.00000	248300.0	16.5134	55.6827	0.00	0.00
sw herts	14015.00000	189400.0	3.7197	73.9968	0.00	0.00
barnet+	13257.00000	215600.0	7.4528	61.4889	0.00	0.00
edgware	20467.00000	176000.0	7.9231	116.2898	0.00	0.00
brent	19068.00000	198700.0	4.8577	95.9638	0.00	0.00
harrow	21315.00000	210800.0	5.7322	101.1148	0.00	0.00
hounslow	23024.00000	210500.0	5.1692	109.3777	0.00	0.00
s hamm	23453.00000	140400.0	8.6298	167.0441	0.00	0.00
n hamm	15439.00000	105400.0	11.6437	146.4801	0.00	0.00
ealing	7149.00000	212200.0	3.7436	33.6899	0.00	0.00
hillingn	34643.00000	229000.0	17.8375	151.2795	0.00	0.00
kcw nw	33012.00000	159900.0	12.9539	206.4541	0.00	0.00
kcw ne	33437.00000	74900.0	38.7606	446.4219	0.00	0.00
kcw s	45941.00000	159400.0	29.8386	288.2121	0.00	0.00
basildon	24633.00000	272200.0	33.6276	90.4960	0.00	0.00
chelmsfd	18905.00000	325000.0	34.0025	58.1692	0.00	0.00
colchest	27325.00000	271700.0	34.3136	100.5705	0.00	0.00
harlow	19698.00000	248600.0	36.2057	79.2357	0.00	0.00
southend	22393.00000	308400.0	34.1580	72.6103	0.00	0.00
barking	24566.00000	153000.0	10.3663	160.5620	0.00	0.00
havering	11343.00000	240400.0	39.7300	47.1839	0.00	0.00
n camden	21935.00000	107400.0	10.5436	204.2365	0.00	0.00
s camden	46414.00000	81900.0	26.3334	566.7154	0.00	0.00
islingtn	31479.00000	168400.0	8.2683	186.9299	0.00	0.00
city	40112.00000	199600.0	12.0082	200.9619	0.00	0.00
newham	18587.00000	229700.0	4.3171	80.9186	0.00	0.00
tower h	39285.00000	150300.0	10.1829	261.3772	0.00	0.00
enfield	16876.00000	260500.0	5.5534	64.7831	0.00	0.00
harigey	25087.00000	227900.0	6.5947	110.0790	0.00	0.00
e roding	13262.00000	170400.0	12.6994	77.8286	0.00	0.00
w roding	28948.00000	280400.0	5.4096	103.2382	0.00	0.00
brighton	28081.00000	301600.0	19.2730	93.1068	0.00	0.00
estbourn	13465.00000	201600.0	22.6688	66.7907	0.00	0.00
hastings	14243.00000	149700.0	21.8068	95.1436	0.00	0.00
e kent	16653.00000	248300.0	33.3051	67.0681	0.00	0.00
thanet	31266.00000	276600.0	33.4238	113.0369	0.00	0.00
dartford	21538.00000	222200.0	26.2207	96.9307	0.00	0.00
maidstne	14462.00000	183900.0	15.7233	78.6406	0.00	0.00
medway	23017.00000	319100.0	32.7977	72.1310	0.00	0.00
tunbridg	27244.00000	195000.0	33.5280	139.7128	0.00	0.00
bexley	13162.00000	214700.0	17.0709	61.3041	0.00	0.00
grenwich	38156.00000	206200.0	11.1315	185.0436	0.00	0.00
bromley	26362.00000	293100.0	13.4810	89.9420	0.00	0.00
st thom	26153.00000	189900.0	44.9305	137.7198	0.00	0.00
kings	33096.00000	224800.0	17.1732	147.2242	0.00	0.00
guys	28887.00000	147300.0	41.3412	196.1100	0.00	0.00

Resource allocation model over space: a trial run

using the slope calibration. Data are for the four

thames regions model (model 1).Cost matrix is "matrix 3".

destination	caseload capacity	local population	balancing factor	cases per head of local population	per cent by private transport	per cent by public transport
lewisham	20626.00000	188800.0	6.0171	109.2479	0.00	0.00
n surrey	11912.00000	134900.0	17.5087	88.3024	0.00	0.00
nwsurrey	18741.00000	202300.0	15.5340	92.6397	0.00	0.00
w surrey	19449.00000	260200.0	10.2008	74.7463	0.00	0.00
swsurrey	19403.00000	177600.0	12.9990	109.2511	0.00	0.00
midsurry	13791.00000	167900.0	13.1900	82.5808	0.00	0.00
e surrey	15971.00000	185300.0	18.9020	86.1900	0.00	0.00
chicstr	14637.00000	158300.0	31.4240	92.4637	0.00	0.00
crawley	17291.00000	240700.0	25.6302	71.8363	0.00	0.00
worthing	14669.00000	226100.0	33.9236	64.8784	0.00	0.00
croydon	19558.00000	321900.0	4.2353	60.7580	0.00	0.00
kingston	18176.00000	244100.0	7.9272	74.4613	0.00	0.00
roehampn	13758.00000	108500.0	15.0704	126.8018	0.00	0.00
wans+em	43573.00000	248500.0	10.5800	175.3441	0.00	0.00
sutton+w	34012.00000	288500.0	13.2328	117.8926	0.00	0.00
oxford	187625.00000	2237100.0	3.5097	83.8697	0.00	0.00
e anglia	159730.00000	1827400.0	4.1887	87.4083	0.00	0.00
wessex	233939.00000	2661000.0	2.8782	87.9139	0.00	0.00
others	2154148.00000	24079700.0	0.2942	89.4591	0.00	0.00

Resource allocation model over space: a trial run using the slope calibration. Data are for the four thames regions model (model 1).Cost matrix is "matrix 3".

aggregated results

district	resident population	predicted cases	predicted hosp rate	error in hosp rate
glc nwt	2001200.0	246128.5	122.9904	3.9
glc net	2269900.0	270915.2	119.3512	5.0
glc set	1464800.0	174269.7	118.9717	2.1
glc swt	1234200.0	133857.2	108.4566	1.7
rest nwt	1431200.0	129978.3	90.8177	3.4
rest net	1425900.0	122022.9	85.5761	-6.0
rest set	2098000.0	198815.9	94.7645	-0.0
rest swt	1620900.0	167042.1	103.0552	5.4
others	32805200.0	2749749.2	83.8205	-0.9

actual trip matrix (sample only)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	2.0	4.0	10.0	20.0	38.0	46.0	8168.0	8105.0	283.0	331.0	11.0	401.0	146.0	16.0	395.0
2	2.0	4.0	4.0	3.0	18.0	23.0	41.0	3993.0	11779.0	4409.0	33.0	448.0	786.0	9.0	448.0
3	2.0	4.0	2.0	5.0	7.0	183.0	17.0	3984.0	836.0	11622.0	8.0	242.0	137.0	13.0	3225.0
4	3.0	3.0	8.0	8.0	13.0	35.0	19.0	192.0	4119.0	2321.0	3136.0	1264.0	4532.0	6700.0	4721.0
5	2.0	1.0	0.0	2.0	3.0	7.0	5.0	41.0	231.0	30.0	63.0	10982.0	4161.0	21.0	43.0
6	1.0	2.0	0.0	1.0	1.0	15.0	4.0	69.0	215.0	62.0	12075.0	1986.0	487.0	142.0	346.0
7	2.0	5.0	1.0	3.0	13.0	20.0	13.0	126.0	418.0	695.0	208.0	263.0	265.0	77.0	17216.0
8	1.0	1.0	0.0	7.0	10.0	33.0	8.0	59.0	119.0	56.0	32.0	1354.0	812.0	3.0	65.0
9	4.0	3.0	0.0	8.0	8.0	13.0	16.0	74.0	216.0	133.0	21.0	585.0	161.0	11.0	94.0
10	0.0	0.0	1.0	1.0	1.0	3.0	3.0	18.0	2.0	0.0	3.0	15.0	19.0	0.0	5.0
11	0.0	3.0	2.0	0.0	2.0	3.0	3.0	2.0	2.0	4.0	1.0	23.0	26.0	0.0	11.0
12	1.0	0.0	1.0	8.0	5.0	2.0	24.0	82.0	76.0	60.0	12.0	90.0	70.0	0.0	61.0
13	0.0	0.0	3.0	3.0	3.0	32.0	13.0	33.0	13.0	13.0	7.0	52.0	72.0	0.0	17.0
14	0.0	0.0	0.0	0.0	1.0	0.0	0.0	2.0	0.0	0.0	0.0	1.0	2.0	0.0	2.0
15	0.0	1.0	0.0	1.0	1.0	9.0	11.0	32.0	11.0	17.0	4.0	32.0	54.0	0.0	14.0
16	0.0	3.0	0.0	1.0	0.0	0.0	4.0	39.0	9.0	17.0	5.0	34.0	30.0	0.0	16.0
17	0.0	0.0	3.0	1.0	1.0	0.0	1.0	15.0	13.0	7.0	3.0	19.0	38.0	0.0	9.0
18	1.0	8.0	3.0	28.0	11.0	7.0	280.0	61.0	20.0	23.0	6.0	37.0	70.0	1.0	79.0
19	0.0	3.0	7.0	2.0	3.0	5.0	87.0	95.0	13.0	44.0	9.0	55.0	102.0	0.0	39.0
20	0.0	3.0	0.0	7.0	2.0	3.0	10.0	10.0	2.0	27.0	0.0	26.0	40.0	0.0	11.0
21	2.0	1.0	2.0	2.0	3.0	7.0	12.0	26.0	5.0	0.0	3.0	41.0	26.0	0.0	15.0
22	0.0	0.0	1.0	0.0	2.0	1.0	2.0	6.0	0.0	4.0	0.0	13.0	19.0	0.0	4.0
23	1.0	0.0	2.0	1.0	0.0	3.0	2.0	16.0	3.0	7.0	2.0	46.0	15.0	0.0	17.0
24	3.0	5.0	1.0	2.0	3.0	5.0	1.0	16.0	1.0	15.0	2.0	31.0	27.0	0.0	14.0
25	1.0	1.0	3.0	3.0	5.0	1.0	4.0	27.0	14.0	14.0	17.0	110.0	58.0	0.0	13.0
26	0.0	1.0	0.0	3.0	3.0	12.0	2.0	13.0	4.0	7.0	4.0	23.0	27.0	0.0	6.0
27	1.0	0.0	4.0	0.0	2.0	2.0	5.0	18.0	22.0	21.0	11.0	44.0	25.0	0.0	14.0
28	2.0	3.0	1.0	0.0	4.0	5.0	6.0	32.0	15.0	16.0	18.0	233.0	65.0	1.0	13.0
29	2.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	5.0	1.0	63.0	86.0	76.0	4.0	12.0
30	1.0	2.0	0.0	1.0	0.0	1.0	4.0	9.0	19.0	26.0	6376.0	838.0	172.0	25.0	71.0

predioted trip matrix (sample only)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.0	0.9	0.3	40.9	19.2	58.5	8101.2	8785.4	549.1	1193.5	51.0	283.4	736.1	33.5	264.8
2	0.0	0.2	0.0	3.2	5.9	60.7	128.5	4071.0	12290.9	4200.9	200.6	715.0	2619.7	173.4	318.8
3	0.0	1.6	0.1	5.4	36.6	487.7	244.9	3984.6	357.9	11165.9	77.9	98.5	255.8	44.2	3876.5
4	0.0	0.1	0.0	0.5	2.2	35.6	34.7	654.4	4249.9	2606.4	3814.1	2460.2	3806.5	5808.7	5424.0
5	0.0	0.0	0.0	0.2	0.2	2.0	0.0	0.4	183.5	1.1	34.4	11080.1	2521.8	34.0	0.0
6	0.0	0.0	0.0	0.0	0.3	5.5	5.4	81.4	586.3	312.4	11430.0	1159.3	855.8	780.8	1091.3
7	0.0	0.4	0.0	0.2	6.0	108.4	13.3	179.7	166.1	773.6	345.2	95.8	142.8	80.5	17122.9
8	0.0	0.0	0.0	0.2	0.2	1.6	0.1	1.8	163.7	2.3	14.1	3126.3	2020.0	13.9	0.0
9	0.0	0.0	0.0	0.5	0.2	1.7	0.2	2.3	59.0	1.3	2.4	488.5	562.1	2.6	0.0
10	0.0	0.0	0.0	0.1	0.0	0.0	0.6	0.5	0.4	0.1	0.1	1.8	1.7	0.0	0.0
11	0.0	0.0	0.0	0.1	0.0	0.0	0.2	0.2	0.1	0.0	0.0	0.3	0.3	0.0	0.0
12	0.0	0.0	0.0	1.6	0.5	2.6	0.1	1.3	7.3	0.2	0.0	12.3	46.2	0.1	0.0
13	0.0	0.0	0.0	1.2	0.2	0.9	0.0	0.1	0.4	0.0	0.0	1.4	3.3	0.0	0.0
14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
15	0.0	0.0	0.0	1.4	0.1	0.4	0.2	0.3	0.4	0.0	0.0	1.3	2.3	0.0	0.0
16	0.0	0.0	0.0	0.3	0.0	0.0	1.2	1.4	2.0	0.4	0.3	9.7	9.0	0.2	0.0
17	0.0	0.0	0.0	0.3	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.3	0.3	0.0	0.0
18	0.0	0.2	0.5	61.1	2.4	2.3	373.7	234.4	29.4	33.4	2.7	30.6	61.4	1.8	7.0
19	0.0	0.1	0.1	9.5	1.1	2.4	91.0	76.7	20.9	9.4	0.8	23.5	54.6	0.7	0.5
20	0.0	0.0	0.0	0.8	0.0	0.0	2.7	1.7	0.9	0.3	0.1	2.7	3.2	0.1	0.0
21	0.0	0.0	0.0	4.7	0.2	0.3	12.3	6.6	2.3	0.9	0.1	4.6	7.4	0.1	0.1
22	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
23	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
24	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.1	0.0	0.0
25	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	12.5	1.7	0.0	0.0
26	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
27	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0
28	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	14.1	2.0	0.0	0.0
29	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	29.7	206.3	27.8	1.8	0.0
30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	15.6	1.1	6456.2	1456.1	362.2	67.3	1.1

private transport cost matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	59.6	37.2	32.5	19.1	21.6	14.6	3.0	4.1	10.1	8.8	17.2	14.0	11.1	14.3	17.2
2	64.3	40.2	38.6	25.6	24.4	14.1	13.8	5.8	1.3	4.9	13.1	11.1	7.3	9.5	16.3
3	59.3	34.5	35.4	23.5	18.8	7.8	11.4	5.2	10.2	1.6	15.0	15.8	12.9	12.5	8.9
4	68.0	43.0	43.5	30.8	27.4	15.9	17.7	11.1	4.5	6.6	5.5	8.1	6.6	0.3	9.0
5	72.4	48.5	46.0	32.6	32.7	22.1	40.3	29.9	11.5	26.2	16.7	2.5	6.2	12.7	41.6
6	72.6	47.1	48.8	36.4	31.8	19.9	21.7	15.7	8.9	11.3	1.5	9.1	9.6	4.7	12.3
7	64.6	38.5	43.2	32.2	24.0	12.2	19.6	13.9	12.6	9.2	11.3	16.2	14.8	11.2	5.2
8	72.0	48.6	45.1	31.5	32.7	22.6	33.2	25.4	11.7	24.0	19.0	5.8	6.7	15.0	39.0
9	71.3	48.5	43.8	30.1	32.7	23.4	31.4	25.6	15.3	26.4	24.6	11.7	11.0	20.4	42.8
10	76.7	59.0	46.7	34.4	44.8	39.6	26.9	28.6	27.6	31.6	32.6	26.0	25.8	30.6	40.5
11	77.2	61.3	47.2	35.9	47.9	43.7	30.6	33.1	33.0	36.5	38.3	31.6	31.3	36.1	45.5
12	68.2	45.9	40.4	26.7	30.2	21.8	31.3	26.8	20.6	31.0	36.0	21.3	17.4	29.8	52.8
13	69.3	47.7	40.9	27.2	32.2	24.4	35.0	33.0	28.0	38.3	42.5	26.8	24.2	36.5	60.3
14	72.4	50.7	44.0	30.3	35.1	26.8	51.3	47.1	36.0	51.3	50.4	30.3	30.0	44.1	76.8
15	69.6	48.9	40.7	27.0	33.6	26.7	30.4	31.0	28.6	36.2	40.2	27.4	25.6	35.4	54.0
16	74.7	55.2	45.2	32.0	40.2	33.6	26.2	27.0	24.6	29.9	30.4	22.4	22.3	28.0	40.4
17	73.6	52.8	44.6	31.0	37.5	30.0	42.0	41.4	35.2	45.5	45.5	31.0	30.8	41.1	64.7
18	60.4	41.1	31.1	17.7	26.9	23.0	11.0	13.6	17.7	18.1	24.8	19.7	17.5	21.9	26.7
19	64.7	43.9	36.0	22.3	28.7	22.5	14.4	16.2	18.2	21.1	27.7	20.0	17.4	23.9	33.7
20	71.6	53.7	41.7	29.1	39.6	34.8	24.0	26.5	26.7	30.3	33.0	25.9	25.1	30.5	40.2
21	66.9	47.8	37.4	24.2	33.3	28.1	19.8	22.8	24.2	27.4	32.5	24.4	22.8	29.1	39.1
22	127.0	98.1	82.5	63.0	75.6	64.8	47.9	48.0	42.6	50.6	47.3	37.7	39.1	45.3	63.2
23	121.0	90.8	77.0	57.0	67.9	56.5	56.4	55.7	47.6	58.9	54.6	41.0	42.8	51.7	76.4
24	133.0	101.0	89.4	69.1	77.3	63.8	50.3	47.7	38.9	48.2	40.3	32.2	35.2	39.6	59.1
25	118.0	83.5	76.0	55.4	59.7	45.1	58.1	50.5	33.7	49.6	37.4	22.3	27.4	35.0	67.8
26	122.0	90.0	79.0	58.6	66.5	53.4	67.2	63.1	49.6	64.8	55.3	39.6	43.4	52.5	84.5
27	116.0	82.8	73.1	52.6	59.3	45.9	65.2	59.4	43.8	61.2	52.6	33.2	36.6	48.4	85.0
28	131.0	96.3	89.6	69.1	72.5	56.9	57.2	51.1	37.4	49.0	34.9	22.3	27.2	35.6	59.4
29	123.0	85.4	84.9	65.0	62.0	44.6	50.8	41.4	26.2	35.6	16.6	12.8	17.9	20.2	40.9
30	115.0	76.9	77.8	58.4	53.7	36.1	43.1	32.9	18.2	26.2	2.6	8.1	11.5	10.9	30.5

vert axis=observed trips,horizontal axis = predicted trips