
Monograph:
Mackett, R.L. (1980) The Data Requirements of an Integrated Land Use Transport Model. Working Paper. Institute of Transport Studies, University of Leeds , Leeds, UK.

Working Paper 135

Working Paper 135

July 1980

THE DATA REQUIREMENTS OF AN INTEGRATED LAND USE - TRANSPORT MODEL

by

R. L. Mackett

Sponsored by the Transport and Road Research Laboratory

ABSTRACT

MACKETT, Roger L. (1980) The data requirements of an integrated land use - transport model. Leeds: Univ.Leeds, Inst. Transp.Stud., Work.Pap. 135

This paper contains the description of the data used in the development and testing of a model that represents the interaction between transport costs and the location of housing, population, jobs, employment, shopping and land. The data is based almost entirely on secondary sources, and emphasis is placed in the paper on the practical issues involved in the assembly of the required data base for the city of Leeds. Each topic is considered in turn, with comments on both the sources and the manipulation necessary to convert the data to the required format.

CONTENTS

THE DATA REQUIREMENTS OF AN INTEGRATED LAND USE - TRANSPORT MODEL

1. INTRODUCTION

In this paper is the description of the data used in the development of a model that represents the interactions between transport costs and the location of housing, population, jobs, employment, shopping and land (Mackett, 1979b, 1980a). The model has been applied to the City of Leeds and used to examine the impact of policies on the city, for example, in the inner city (Mackett, 1980b) and in terms of social equity (Mackett, 1980c). The model has been calibrated and tested against the comprehensive data base described here (Mackett, 1980d), and found to give a good representation (Mackett, 1979a).

In this paper, emphasis is placed on the practical issues in assembling a data base using secondary data almost exclusively. It will be seen that a variety of devices have had to be adopted. However, the construction of the data base by one individual ensures a degree of consistency, and means that the model could be set up and run elsewhere fairly easily, using similar data sources.

A further general comment is that the model has been run at an urban scale, that is, a city disaggregated into zones. This introduces some fairly severe data problems which could be avoided by working a regional scale where each zone is a whole urban unit, but this would mean that a large proportion of the spatial interaction would not be represented explicitly.

It is accepted that some of the data used is less accurate than is ideal. On some occasions, the data has been estimated by algebraic methods, and then is used as an input to the model. This is not completely satisfactory, but the very limited resources available made this necessary. It may well be possible to substitute better data later. At several stages in the development of the model this happened. It is only by establishing an explicit need for better data that an effective demand can be generated.

In some places below, a reference is made to a balancing factor technique. This means that it is required to obtain some cross-tabulation which is consistent with the numbers in each category. The method is similar to the trip distribution model devised by Wilson (1967), and may be shown as:

$$N_{ij} = A_i B_j E_i F_j W_{ij} \tag{1}$$

where N_{ij} is the required cross-tabulation being the number of items in both category i and category j, W_{ij} is a weighting representing some form of relationship between i and j, E_i is the number of items in category i and F_j is the number of items in category j, and A_i and B_j are balancing factors defined as

$$A_i = \{\sum_j B_j F_j W_{ij}\}^{-1} \tag{2}$$

and

$$B_j = \{\sum_i A_i O_i W_{ij}\}^{-1} \tag{3}$$

to ensure that

$$\sum_j N_{ij} = E_i \tag{4}$$

and

$$\sum_i N_{ij} = F_j \tag{5}$$

respectively. The term W_{ij} could be the values of N_{ij} for a different point of time. If the term W_{ij} is not present (that is, is implicitly unity) the balancing factors are not required since the values of N_{ij} could be obtained by scaling. Often this would mean assuming that the disaggregation of some item is identical in all zones, which is not very interesting.

In the next section, some of the disaggregations used in the model will be defined, then the sources used for each item in turn will be discussed. In many cases the description is straightforward, in others a fairly complex explanation is required.

2. DEFINITION OF CATEGORIES

2.1 Zones

There are 28 internal zones and 12 external zones. The internal zones are the wards that existed at the time of the 1961 and 1966 Censuses of Population. The 1971 Census data were aggregated to these zones using programs supplied by Leeds M.D. Planning Department. The 12 external zones are the adjacent local authority areas surrounding Leeds in 1977 plus some others such as Harrogate that have strong ties with the city. The external zone system ensures about 95% closure of the system in terms of the journey to work. A substantial increase in their percentage could only have been obtained by a very large increase in the number of external zones.

2.2 Social groups

The model is run for three social groups, a small top social group containing professional and managerial workers, a small bottom group containing unskilled manual workers and a fairly large (about 70%) middle group. The definitions are aggregations of the Census of Population socio-economic groups as at 1971 and are shown in table 1. These groups are intended to be well-defined in terms of residential choice, particularly as no income information is available. It would be very difficult to establish this without a large household survey.

2.3 Industries

Twelve industries are used in the model. These are divided into three categories - primary (industries 1-3), secondary (industries 4-7) and tertiary (industries 8-12). The Census of Population uses the Standard Industrial Classification which is changed about every ten years, including 1968. The SICs are divided into Minimum List Headings (MLH). The allocation of SICs and MLHs to the industries used in the model are shown in table 2. Where an SIC is shown all the constituent MLHs are included. MLHs are only shown where an SIC is being split between two or more model industries.

Table 1 Allocation of socio-economic groups to social groups

Social group	Socio - economic groups
1 top	1, 2, 3, 4 and 13
2 middle	5, 6, 7, 8, 9, 10, 12, 14, 15, 16 and 17
3 bottom	11

Table 2 Allocation of SICs and Minimum List Headings (MLH) to industries used in the model

Industry		Standard industrial classification	
		1966	1971
	Primary		
1	Agriculture, forestry and fishing	SIC I	SIC I
2	Mining and quarrying	SIC II	SIC II
3	Gas, water and electricity	SIC XVIII	SIC XXI
	Secondary		
4	Manufacturing	SIC III - XVI	SIC III - XIX
5	Construction	SIC XVII	SIC XX
6	Transport and communications	SIC XIX	SIC XXII
7	Public administration and defence	SIC XXIV	SIC XXVII
	Tertiary		
8	Convenience retail	MLH 820	MLH 820, 821
9	Durable retail		
10	Business services	MLH 810, 831,832, SIC XXI, MLH 871, 873 879	MLH 810, 811, 812, 831, 832, SIC XXIV, MLH 871, 873,876 879
11	Educational services	MLH 872	MLH 872
12	Personal services	MLH 874, 875, SIC XXIII	MLH 874, 875 SIC XXVI

Table 3. Allocation of modes of transport from the Census of Population to headings used in the model

Mode used in model	Mode used in Census of Population	
	1966	1971
1 Private transport	Car Motor cycle	Car Motor cycle
2 Public transport	Train Bus Goods vehicle Pedal cycle On foot Other None Not stated	Train Bus Pedal cycle On foot and none Other Public transport (as described) Not stated

2.4 Mode of transport

Two modes of transport are used in the model; the first is motorised private transport and includes car and motor cycle. The second includes all other modes, including walking, cycling, none and not stated. The allocation is shown in table 3. The method of incorporating the non-motorised public transport modes is discussed in Mackett (1980a, Section 3). As explained there, all workers have to be taken into account so the non-motorized categories are included in the intrazonal trips.

2.5 Car ownership

There are two categories of car ownership - group 1 own a car, group 2 who do not. For the journey to work the Census of Population category of car owned is taken to mean car available. For flows to shop the car availability is computed using the methodology in Mackett (1980a, Section 11).

3. TRANSPORT

3.1 Distances

Distances along the road links were measured on Ordnance Survey maps at a scale of 1 to 25000. The same distances were used for the two modes. This means that changes in the distance travelled are independent of modal choice.

3.2 Travel times

Travel times for cars were computed using speeds from two sources. For interzonal trips, average speeds were based upon research at the Transport and Road Research Laboratory (Marlow, 1971) for a comparable area and are a function of the type of road. The following speeds were used:

a)	the Ring Roads	48 kph
b)	other A roads	38.4 kph
c)	B roads	32 kph
d)	unclassified	24 kph

Bus travel times were taken from the published timetables. For intrazonal trips, average speeds were taken from WYTCONSULT (1976). These were:

a)	car	30 kph
b)	bus, inner urban areas	16 kph
c)	bus, outer urban areas	20 kph
d)	walk	5 kph

3.3 Number of buses used

The number of buses used to travel between pairs of zones was found from the published timetable and accompanying map using the assumption that people change buses as little as possible. In Leeds most routes are to the centre, and nearly all changes occur there. In no case were more than two buses required.

3.4 Parking charge

The parking charge is imposed only in the central area (City Ward) and Bradford. The values of 10p in Leeds and 5p in Bradford for peak and 5p and 2.5p respectively off-peak are based upon WYTCONSULT (1976) data and are at 1975 prices.

3.5 Relative weighting on walking time

The value of 1.7 was taken directly from WYTCONSULT (1976).

3.6 Relative weighting on waiting time

The value of 2.3 was taken directly from WYTCONSULT (1976).

3.7 Waiting times for buses

It was found to be impractical to make the waiting time for buses a function of bus headway, because so many bus routes use the same roads. Extremely detailed modelling would be required to use that method so values of 5.0 and 7.5 minutes for peak and off-peak respectively were adopted.

3.8 Value of time

The value of time of 0.515 pence/minute for all purposes that was used in WYTCONSULT (1976) was adopted for this study. 1975 prices have been used in this study, as in WYTCONSULT. The value of time has been kept constant over the forecast period. The same value is used for each social group to facilitate equity considerations in the analysis.

3.9 Perceived operating cost for cars

The value of 1.77 p/vehicle km used in WYTCONSULT (1976) has been adopted in this study.

3.10 Walking time to buses

The walking time to buses was taken as half the intrazonal walk time, subject to a maximum of 10 minutes. The maximum value tended to be used in the external zones, and since the zones are much larger than those in WYTCONSULT, the value there of 5 minutes was doubled.

3.11 Car occupancy rate

The values of the car occupancy rate used were 1.35 and 1.37 for the peak and off-peak respectively. The values were based upon the West Yorkshire Transportation Study (Traffic Research Corporation Ltd., 1969) since WYTCONSULT make no distinction between the peak and off-peak in this case.

3.12 Bus fare function

Bus fares in Leeds have changed quite dramatically from being very cheap to be very expensive in real terms over the period 1966 to 1971. Consequently it was not sensible to assume that the WYTCONSULT figures could be used. Fare tables are published with the timetables and from these the fares were computed. The fare stage in Leeds is 1.25 km. The fares are shown in the table below.

Table 4. Bus fares for 1966, 1971 and 1976.

	monetary prices in p			real (1975) prices in p		
	1966	1971	1976	1966	1971	1976
boarding element	0.4167(1d)	1	3	0.9245	1.6851	2.5738
distance element/ stage	0.4167(1d)	1	2	0.9245	1.6851	1.7159
distance element/ km	—			0.7396	1.3481	1.3727

The 1976 values are used in the forecasts to 1991, that is, in the absence of other information, fares are assumed to rise in line with inflation. The bus fares in 1961, for which no fare tables were available, were found by deflating the 1966 fares using the Retail Price Index and national changes in public transport expenditure.

3.13 Changes to the transport networks

The past changes to the road network were obtained from the Leeds Development Plan Review (City and County Borough of Leeds, undated) and Ordnance Survey maps. The proposed changes were based upon proposals in the 1978/79 submission of the Transport Policies and Programme of West Yorkshire Metropolitan County Council (1977). Past changes in the bus network were taken from the bus timetables and maps.

3.14 Transport cost levels over time

National figures of the operating costs of vehicles and bus fare levels were used to find the appropriate values for the car operating cost for 1961, 1966, 1971 and 1976 and bus fares for 1961. These were taken from Transport Statistics (formerly Highway Statistics) published by the Department of Transport.

3.15 Inflation indices

As already shown, inflation has an important effect on the monetary cost of travel, so all costs have to be adjusted to a common year. The Retail Price Index (RPI) is the most commonly used measure to adjust costs. Its value was taken from the Annual Abstract of Statistics (Central Statistical Office, 1977).

4. THE JOURNEY TO WORK

4.1 Journey to work trips

The journey to work pattern, by mode, is required for the calibration procedure. Fortunately, such data have been obtained for the West Yorkshire area by a consortium of the local authorities for 1966 and 1971. In the Census of Population, householders were asked the workplace and main mode of travel to work for all members of the household in employment. The workplace was then coded to the appropriate enumeration district (ED). The 1966 data were purchased for flows between pairs of ED's by sex, mode and industry. The 1971 data were for flows between pairs of EDs by mode,

with accompanying 'back-up tables' which give the number of employed residents and workers in various categories. From these two data sets it was possible to obtain suitable trip matrices of the journey to work for the zoning-system being used, by the two modes defined in section 2.4 above. As discussed in Mackett (1980d) there are many problems with these data, in particular, many people are missing because the workplace address that they gave could not be coded to an enumeration district, and 'Barnardization', the process of adding zero, plus or minus one at random to preserve confidentiality, appears to have a cumulative effect on aggregation.

4.2 Public transport used for intrazonal trips

The intrazonal costs for public transport are designed to take into account walking trips and those working at home, so these trips are required. These are obtained from the workplace analysis described above.

4.3 Commuters in and out of Leeds

The number of people crossing the boundary in and out of the city of Leeds, plus the number of people both living and working in the city are used as part of the external zone mechanism (Mackett, 1974). The data were obtained from the Census of Population Workplace and Transport Tables for 1966 and 1971, and disaggregated by social group using the methods described in Mackett (1980d). These are also used to calculate the rate of change in the number of people entering and leaving this city for use in the model (Mackett, 1980a Section 10).

4.4 Relationship between social group and car ownership (W_{qn}^{QN})

This value is used in the calibration process and was obtained from the Workplace and Transport Tables for 1966 and 1971 for the West Yorkshire conurbation. The figure is not available for Leeds, but since this is only a weighting taking the figure for the whole conurbation is not likely to lead to serious error.

4.5 Relationship between car ownership and mode used (W_{kn}^{KN})

This value, also used in the calibration process was obtained from the Workplace and Transport Tables and, again, only a figure for the conurbation was available and so was used.

4.6 Relationship between social group and mode (W_{qk}^{QK})

The value was obtained, for 1971, from the Ward and Parish Library of the Census of Population. The value for 1966 was obtained using a balancing factor routine, as defined in section 1, with the totals for each social group and mode taken from the County Report, and Workplace and Transport Tables respectively.

5. POPULATION

5.1 Population in each zone

The population in each zone, was obtained from the Ward and Parish Library for internal zones and the County Report of the Census of Population for 1966 and 1971 for external zones.

5.2 Proportion of population in each social group

The Census of Population for 1966 only gives the socio-economic group for males so this figure was used and applied to the whole population. For 1971 the figure used is that for members in a household whose head is in each social group. This was chosen since it was felt this was the most appropriate category for the residential location decision. For both 1966 and 1971 the figures come from the Ward and Parish Library for internal zones and the County Report for external zones.

5.3 Population surviving five years in each zone

The 1971 Census of Population gives the number of people who had an address outside the zone five years previously. Those people who were aged five or over were subtracted from the population who were living in the zone five years earlier to give the number of people who have survived, that is, have not out-migrated or died during the period. These are used to compute the survival rates used in Mackett (1980b section 9). Forced migrants, that is those whose houses have been demolished, are subtracted to prevent double counting so these people must be added on to the survivors before the calculation of the survival rate. In other words, it is assumed that the people who moved only because

their houses were demolished would othewise have survived. These data are only available for 1971, so the same rates are assumed to apply in 1966. The number of in-migrants in each social group whose head has in-migrated are scaled to the total who have in-migrated. These are then subtracted from the population in each social group in 1966 to give the number who actually survived. The number of forced migrants by social group is found by applying the same rate as in the model: the lowest social group are assumed to be the first to have their houses demolished, then the middle social group and finally, the top social group. This is subject to not more than half the dwellings occupied by a particular social group being forced to more unless over half the total housing stock is being removed, in which case the same proportion is applied to all social groups. These numbers of people are added on to give the values which are input to the model to give the survival rates.

5.4 Population aged 0-4

People are born during the five years of a model forecast and these are found by applying a zonal-specific birthrate to the population in the model. This rate is found by dividing the zonal population into the number of people aged 0-4 at the end of the period. To be consistent with the survival rates the values for 1971 are used, even in forecasts from 1966. The values are taken from the Ward and Parish Library.

6. HOUSING

6.1 Housing stock

The housing stock in each zone was obtained from the Census of Population. For 1961 the values were taken from the County Report, for 1966 and 1971 from the Ward and Parish Library. For 1971 only, both vacant and occupied housing are specified.

6.2 Housing demolished and built

Housing demolitions are very difficult to obtain information about, since, by definition, the houses are not there and so are not included in the Census of Population. The total number of demolitions in Leeds each year is given in Local Housing Statistics, but because of local government reorganization in 1974, the records of Leeds M.D. planning department of where demolitions actually occurred only go back to 1971. Consequently, an estimation technique had to be applied to allocate

demolitions to zones. It has been assumed that demolitions occurred only in those zones in which the housing stock showed a net decline over the five year period. There is a further difficulty: the data for 1961 and 1966 only applies to occupied housing, consequently, only occupied housing has been used in the housing calibration process. This would only lead to serious error if there were some zones with vast proportions of empty housing. This is not the case in 1971 at least. The demolitions in each zone were allocated in proportion to the net decline in the level of housing stock. The number of houses built were found by adding the number of demolitions onto the net charge in housing stock.

Data for the period of 1971 to 1991 on demolitions and to 1981 for completions at a very detailed spatial scale have been supplied by Leeds M.D. Planning Department. These have been coded to the zones within Leeds. More recent figures on the total number of demolitions have been obtained from the Structure Plan (West Yorkshire M.C.C. 1978). The house building rate of 2,500 per annum for the future has been taken from the same documents.

6.3 Housing occupied by each social group

It is very difficult to obtain information relating housing occupancy to houshold characteristics. Consequently, some assumptions have had to be made about the number of houses occupied by each social group. The top social group are taken to have one dwelling per household, so that sharing households all belong to the two lower social groups. The sharing households are allocated between the two lower social groups such that the proportion of households in the lowest social group that share is twice that in the middle social group.

7. CAR OWNERSHIP

7.1 Number of households owning a car

Information on the numbers of households owning a car was obtained from the Census of Population. For the internal zones this was obtained from the Ward and Parish Library and for the external zones from the

County Report. For 1966, the values for Wharfedale, which has a population below the 15,000 threshold, had to be obtained from the 'C' tables, which are not included in the published County Report.

7.2 Number of cars owned

The numbers of cars owned in each zone were obtained from the same sources as the number of households owning a car.

7.3 Change in car ownership rate over time

The proportional change in car ownership, strictly the proportion of non-car owners at one time point who change to being car owners over a five year period, was found by comparing the proportion of households in each social group not owning a car in 1966 and 1971 in Leeds obtained from the Census of Population.

8. ECONOMIC ACTIVITY

8.1 Employment by industrial sector in each zone

The employment in each zone in each of the twelve industrial sectors is an important input to the model, since employment location is a significant factor in the determination of the location of other urban activities. Unfortunately, such data are far less reaily available than population and housing data and so various devices have been used to obtain suitable information. The Census of Population is the main source, but as the questions are asked at the residence, employment information is only available indirectly, in effect by aggregation of the journey to work matrix over residential areas. It is not practical to utilize other data sources such as the records of the Factory Inspectorate because they only cover limited sectors, ignore firms employing less than five workers, are not coded to zones (only by address), are not readily accessible and are not available historically.

The Workplace Analyses from the Census of Population for 1966 and 1971 provide the basis of the employment data. The 'back-up tables' from the 1971 Census include employment in every ED by SIC category and Minimum List Heading. Consequently these data have been aggregated to the model zoning system and the employment categories. The 1966 data

only uses eight industrial grouping, seven of which correspond to each of the primary and secondary sections used in the model and the eighth to the total tertiary employment. As discussed above, there are problems of underenumeration and 'Barnardization' with the data. To overcome the problems the total in each cell was scaled to the total for the sector for Leeds from the Economic Activity Tables, which tend to be higher because some responses to the workplace question in the Census could be coded to Leeds, but not to an E.D. This partly overcomes the effect of Barnardization, because the adjustments are made at the level of publication and it has been found that such adjustments can accummulate. There is a further problem: retail employment is not divided into convenience and durable in the Census. In 1971 the categories 'Retail distribution of food and drink' and 'Other retail distribution' were used. In 1966 and 1961 no division of retail employment was made. However, the Census of Distribution does provide such information for the whole of Leeds for 1961, 1966 and 1971, and so has been used to divide up the retail sectors to give values for Leeds. There are discrepancies between the totals from the two sources, no doubt due to the very different methods of data collection.

In the ways described above the total in each sector in Leeds was obtained. The employment in each zone in 1971 was obtained by summation over the sectors and scaled to be equal to the total obtained by summing over each sector for the city as a whole from the Economic Activity Tables. A balancing factor method was then used to obtain a consistent matrix, using the corrected zonal and sectional totals as the row and column totals and the original Census matrix of employment by zone and sector as the weightings. The values for the two retail sectors used in the Census were used to correspond to 'convenience' and 'durable' respectively, but with the sectoral totals amended to correspond with the values from the Census of Distribution.

The 1966 figures for the employment in each primary and secondary sector from the aggregation of the trip matrix were scaled to agree with the totals from the Economic Activity Tables. The employment in each zone in the five tertiary sectors was obtained by using a balancing

factor method, using the sectoral totals from the Economic Activity Tables and the zonal totals from the aggregation of the journey to work matrix scaled to the sum over the sectoral totals. Once again the Census of Distribution was used to allocate the total retail employment between the sectors convenience and durable. The final 1971 matrix of employment by sector was used as the weightings, but with some manual adjustments to allow for changes in the size of new shopping centres which would otherwise be 'blurred over' by the estimation technique. These are shown for the periods 1966-1971 and 1961-1966 in table 5 and represent the factors applied to the employment figure for the end of the period to obrain a first estimate for the beginning of the period to be used as the weighting. Such a factor was also used for the central area (zone 7), using the figures from the Census of Distribution for the central area to scale the appropriate weightings.

The distribution of employment by zone and sector for 1961, which is only used in the calibration process, was obtained by similar means, except that no zonal totals were available, so the method described in Mackett (1973a) was used. This involved estimating the rate of change of the mean trip length over the period 1961 to 1966 and then applying a singly constrained employment location model. The zonal values were used in the same way as in 1966 to calculate the distribution of employment by zone and sector. Clearly there is some tautology in using a spatial interaction model to estimate the inputs to a model that itself contains spatial interaction, but the 1961 data are only used for calibration purposes, and not for forecasting directly. Clearly, care is required in interpreting the results based on such methods.

The methods described above were used for the internal zones. Employment in each industry in the external zones was obtained from the Economic Activity Tables of the Census of Population for 1961, 1966 and 1971.

Table 5. Scaling factors applied to the weightings to represent large changes in the size of shopping centres

1961 - 1966		
Shopping centre	Zone	Factor
Seacroft	8 - Crossgates	0.500
Crossgates	8 - Crossgates	0.700
City centre	7 - City	1.008
1966 - 1971		
Shopping centre	Zone	Factor
Crossgates	8 - Crossgates	0.700
Armley	2 - Armley	0.750
City centre	7 - City	1.008

8.2 Employment by social group in each zone

The employment in each social group in each zone is used in the estimation of the trip matrix for calibration purposes. The values for 1971 were obtained from the 'Back-up Tables' of the Workplace Analysis of the Census of Population, in the same was as the industrial figures described above. The 1971 values are used for 1966 since a weighting is required, rather than an exact value.

8.3 Sales in shops in Leeds

The total volume of sales in Leeds is used in the calculation of the in-flow of cash from residents outside Leeds to shopping centres within. The value was obtained for convenience and durable type shops from the Census of Distribution for 1966 and 1971 by dividing the annual turnover by 52 to obtain a value per week. The values were adjusted to allow for non-response using a factor given in the Census of Distribution. The sales were converted to 1975 prices by using the retail price index.

8.4 Population served by non-retail tertiary activites in Leeds

The equivalent to sales for non-retail tertiary activities is the population served. For a city like Leeds, which serves as a regional centre, one would expect this figure to be larger than the population.

However, this is rather a difficult concept to obtain a realistic value for, so values have been assumed. The values used are 550000, 500000 and 550000 for business, educational and personal services, since the population of Leeds is slightly under 500000. A lower figure is used for education because relatively few school-children cross the Leeds city boundary because of the way education is organized under the Education Authorities. Even if these figures are not very accurate they will not make a great deal of difference to the spatial distribution of employment in these sectors since both the total for each time point and the base year spatial distributions are specified.

8.5 Expenditure per head convenience and durable goods

Data on expenditure per household are provided in the Family Expenditure Survey. These formed the basis of the required figures of expenditure per head on convenience and durable goods in Leeds and in the model. As in other cases described above several operations had to be performed on the data. Firstly it was decided to regard expenditure on food and tobacco as convenience and on alcohol, clothing, durable households goods and other goods as durable. This raises various issues as to the appropriate categories for some of these items, but, since some goods might be regarded as being in different categories for different people, depending upon their shopping habits, this is an issue that requires a decision to be made. A second decision concerned the social group disaggregation, since different categories are used in the Family Expenditure Survey. Workers in professional, technical, administrative, managerial and teaching occupations were taken to be equivalent to social group 1, workers in clerical occupations to social group 2 and workers in manual occupations to social group 3. The expenditures per household were converted to values per person by dividing by the mean household size for each social group in Leeds. Prices tend to be lower in Leeds thatn the national average, so this was taken into account by multiplying the expenditures by conversion factors for each sector, taken from the 1971 Census of Distribution. These values were then summed for each social group and converted to 1975 prices using the Retail Price Index.

8.6 Flows of money to shops

The flows of money from residences to shops are required for the calibration of the shopping model. The data are required to be disaggregated

into convenience and durable, and by car availability and mode of transport used. The only satisfactory way to obtain such data is to carry out a questionnaire survey, preferably at the home rather than shopping centre. A small household survey was carried out in January 1972 by a group of staff and students from the Department of Geography at the University of Leeds and is described in Mackett (1973b). Briefly, households were selected at random from the Electoral Register within areas defined by Enumeration Districts used in the 1966 Census of Population. Thirty addresses were selected in each of 15 E.D.s, giving a total of 450. Of these 106 were successful and were coded to origin and destination of the flows. The origin was simply the residential address of the respondent, the destinations were coded using records of Leeds C.B. Planning Department containing the name, address, grid reference and classification of every shop in Leeds. Questions were asked about five types of convenience goods and six types of durable goods, the mode used for the purchase and whether a car was owned. From this information matrices of flows of money to purchase goods, by mode and car ownership group were produced. Because of the variation in the response rate, and the fact that not every zone was represented in the sample, some scaling was required. The selected E.D.s had been chosen at random from thirteen pairs of adjacent wards, plus one from each of the remaining two. In the cases where one E.D. represented a pair of wards the values from the survey were divided between the zones in proportion to the total expenditure by residents of each zone. The purchases by residents of each zone were then scaled to sum to the total expenditure by people in that zone. In effect, this meant that the survey was being used to give the spatial distribution and mode used for the flow of cash from each zone of residence to shops.

8.7 Proportion of workers in each industry that are in each social group

The proportion of workers in each industry that are in each social group is used to convert the spatial distribution of jobs by industry to the equivalent jobs by social group. The Ward and Parish Library of the 1971 Census of Population includes a suitable cross-tabulation for residents so this was used as the weightings for a balancing factor routine with the numbers employed in Leeds in each industry and in each social group as the row and column totals. The cells of the resulting matrix were

divided by the appropriate total in the industrial sector to give the proportion in each social group. The same method was used for 1966, but with different row and column totals. Unfortunately, tertiary activity was not disaggregated into sectors so values for Great Britain had to be used. Once again the numbers in each industry and social group were divded by the total in each industry to give the required proportions.

8.8 Employment survival rates

In the model only a proportion of workers are assumed to be newly locating in a time period, so it is necessary to know the survival rate, that is the proportion of workers who remain in the same job over the time period. Such data appear to be very scarce. In fact the topic of length of duration in employment seems to have been little researched. The source used here is Harris and Clausen (1967) who provide a wide selection of tables of length of employment. There is not one by socio-economic group, but there is one by social class, so this has been used taking social classes 1 and 2 as equivalent social group 1, social classes 3 and 4 to social group 2, and social class 5 to social group 3. These are national figures, which is not very satisfactory, but no local figures were available.

8.9 Employed residents in external zones

The number of employed residents in each external zone is required for the journey to work calibration procedure and was obtained from the Workplace and Transport Tables of the Census of Population for 1966 and 1971.

9. LAND

The allocation of land between uses is an important input to the model. The main sources of data are the Land Use Maps published for 1948 and 1971 the map for 1981 which formed part of the Development Plan Review for Leeds, all published by Leeds C.B. Council, and the Land Utilization Survey Map for 1966. The first three maps are at a scale of 1 to 10560, the last at 1 to 25000. Maps at these scales provide fairly detailed information, but

there were difficulties in obtaining very accurate information, partly because no distinction appears on the ground between some industrial sectors. Consequently the tertiary sectors are treated in aggregate for this purpose. Some other sectors required some scaling and disaggregation. The areas were measured using a planimeter and tracing graph paper. The main use of these data are to provide a check on the amount of development in a zone and as part of the general urbanization process. The definition of use is difficult even on the ground, but it is felt that the methods used to measure areas and allocate to uses are sufficiently accurate for the purposes of this model.

10. CONCLUSIONS

The objective in presenting this description of the data sources has been to demonstrate the difficulties encountered and the solutions developed, so that readers and potential users are aware of the weak points and can see where care is needed in interpretation. On the more positive side it can be seen that this model uses secondary data almost entirely, and so could be applied elsewhere very easily. The basic philosophy in this study has been to obtain a reasonable representation of the required data, rather than designing the model around the readily available data. If this latter course were followed there would be little innovation in the world of empirical model-building. By adopting a bold approach the demand for better data can be clearly established, which will, hopefully, lead to better representation and understanding of urban systems.

REFERENCES

CENTRAL STATISTICAL OFFICE (1977) Annual Abstract of Statistics (HMSO, London)

CITY AND COUNTY BOROUGH OF LEEDS (undated) Development Plan Review, Written Statement, Leeds.

HARRIS, R, CLAUSEN, A (1967) Labour mobility in Great Britain 1953-63, Government Social Survey (HMSO, London)

MACKETT, R.L. (1973a) The estimation of employment in small areas within the city, Working Paper 29, Department of Geography, University of Leeds, Leeds.

MACKETT, R.L. (1973b) Shopping in the city - the application of an intra-urban shopping model to Leeds, Working Paper 30, Department of Geography, University of Leeds, Leeds.

MACKETT, R.L. (1974) A residential location model incorporating spatially varying levels of information, Regional Studies, volume 8, pp.257-265.

MACKETT, R.L. (1979a) Modelling the impact of transport planning policy on land use, Proceedings of the 1979 PTRC Summer Annual Meeting, London, also Working Paper 115, Institute for Transport Studies, University of Leeds, Leeds.

MACKETT, R.L. (1979b) A model of the relationships between transport and land use, Working Paper 122, Institute for Transport Studies, University of Leeds, Leeds.

MACKETT, R.L. (1980a) The mathematical representation of a model of the relationships between transport and land use, Working Paper 123, Institute for Transport Studies, University of Leeds, Leeds.

MACKETT, R.L. (1980b) The relationship between transport and the viability of central and inner urban areas. Journal of Transport Economics and Policy, volume 14 (forthcoming).

MACKETT, R.L. (1980c) Modelling the impact of alternative transport strategies upon social groups, Working Paper 128, Institute for Transport Studies, University of Leeds, Leeds.

MACKETT, R.L. (1980d) The methodology for the calibration and testing of an integrated land use transport model, Working Paper 134, Institute for Transport Studies, Univerisyt of Leeds, Leeds.

MARLOW, M (1971) Repeat traffic studies in 1967 in eight towns previously surveyed in 1963/4, Laboratory Report 390, Road Research Laboratory, Crowthorne, Berks.

TRAFFIC RESEARCH CORPORATION LTD (1969) West Yorkshire Transportation Study, Leeds.

WEST YORKSHIRE METROPOLITAN COUNTY COUNCIL (1977) Transport Policies and Programme, 1978/79 Submission, Wakefield.

WEST YORKSHIRE METROPOLITAN COUNTY COUNCIL (1978) Structure Plan Written Statement, Wakefield.

WILSON, A.G. (1967) A statistical theory of spatial distribution models, Transportation Research, volume 1, pp. 253-269.

WYTCONSULT (1976) The structure and calibration of the county strategic model, Document 206, West Yorkshire Transportation Study, Leeds.