POLICY ANALYSIS USING A MICROSIMULATION MODEL OF LAND USE AND TRANSPORT SYSTEMS

Roger L.MACKETT
Reader in Transport Studies
Transport Studies Group
University College London
London - United Kingdom

INTRODUCTION

This paper describes the use of a micro-level model for the analysis of land-use and transport systems. The model is MASTER model (Micro-Analytical Simulation of Transport, Employment and Residence) which represents the relationships between transport and the location of activities taking into account the processes of residential choice, employment choice, the associated search processes, car purchase and the travel decisions of individuals and households. The model has been designed in such a way that it can be tested against data representing the real world.

In the next section the methodology underlying the model is discussed, particularly how the processes are represented. In Section 2 the structure of the model is described, followed by a discussion on how the model can be used for policy analysis.

1. METHODOLOGY

MASTER model uses micro-analytical simulation (or microsimulation) to represent the actions and decisions of a population through time. In MASTER model the population is set of households and their members, who go through a series of processes associated with the long-term response to changes in transport and related professes. The model focusses on the relationship between the residential location, employment location, car ownership, mode choice and route choice processes. However, because the analysis is carried out for a period of twenty years, it is not sensible to ignore the dynamics of demographic change, so it is necessary to represent this and the resulting household formation.

Microsimulation has been used in a number of fields, for example socio-economic systems (Orcutt et al, 1961), the housing market (Wegener, 1983) and residential mobility (Hayashi and Tomita, 1990).

One obvious requirement of this approach is an initial population to simulate. The demographic and household characteristics could be obtained from a survey. However, such an approach is expensive, and not possible in this work. Consequently, the initial population is simulated, by means of microsimulation techniques (Mackett, 1990a).

The processes in the model are represented by four types of mechanism: deterministic, rule-based, probabilitic and search.

Deterministic processes are those which have only one possible outcome. For example, in the aging process, at each time point each person becomes older by the length of the preceding time period, thus

$$a_p[t] = a_p[t-1] + \Delta a_p[t-1,t] \quad (1)$$

where $a_p[t]$ is the state of person p at time t
and $\Delta a_p[t-1,t]$ is the change in the value of a over the period t-1 to t .

Rule-based processes take the form:

$$\text{if } X \text{ then } b_p \Rightarrow b'_p$$

where X is an event.
b_p is the state that person p is in before the event,
and b'_p the state after.

For example, if a person's spouse dies, the marital status of the person changes from 'married' to 'widowed'.

A probabilistic process is one in which there are several possible outcomes. The probability of each is established for the characteristics of the person or household. Monte Carlo simulation is used to ascertain which outcome actually occurs. This involves the cumulation of the probabilities, so that the final one has a value of 1.0. A random number between zero and one is generated and compared with the cumulated probability distribution. The outcome is then taken to be the one in the range of which the random number lies. With a large sample the distribution of the outcomes will tend to the probability distribution. Monte Carlo simulation may be shown as

$$R(p \mid z_p^{(1)}, z_p^{(2)}) \leq P^X(z^{(1)}, z^{(2)}) \quad (2)$$

where R is a random number between 0 and 1
p is the person being considered
$z^{(1)}, z^{(2)}$ are the relevant characteristics from a look-up table
$z_p^{(1)}, z_p^{(2)}$ are the values of $z^{(1)}$ and $z^{(2)}$ for person p
P^X is the probability of event X occurring. If the value is unique to a household or person this is indicated by a subscript h or p respectively.

The probabilities are either given by values based on empirical data, for example for the various demographic processes, of by estimating a utility function which is used in a multinomial logit model. The latter method is used for the explicit choice processes such as mode choice and the decision to consider moving home or changing job. The utility function is assumed to be linear and additive, in the absence of other information, for example:

$$U_{pk}^X = \beta_1^X + \beta_2^X z_p^{(1)} \; \beta_3^X z_k^{(2)} \quad (3)$$

where U_{pk}^X is the value of utility of alternative k for person p for process X
β_1 to β_3 are coefficients the values of which are obtained by calibration
$z_p^{(1)}$ is a variable that is specific to person p
$z_k^{(2)}$ is a variable not specific to person p , for example a zonal value

A multinomial logit model is given by

$$P^X_{pk} = \exp(U^X_{pk}) / \Sigma_l \exp(U^X_{pl}) \qquad (4)$$

where P^X_{pk} is the probability of person p choosing alternative k in process X .

For a binary choice the model may be shown as

$$P^X_{pk} = 1 / [1 + \exp(-U^X_{pk})] \qquad (5)$$

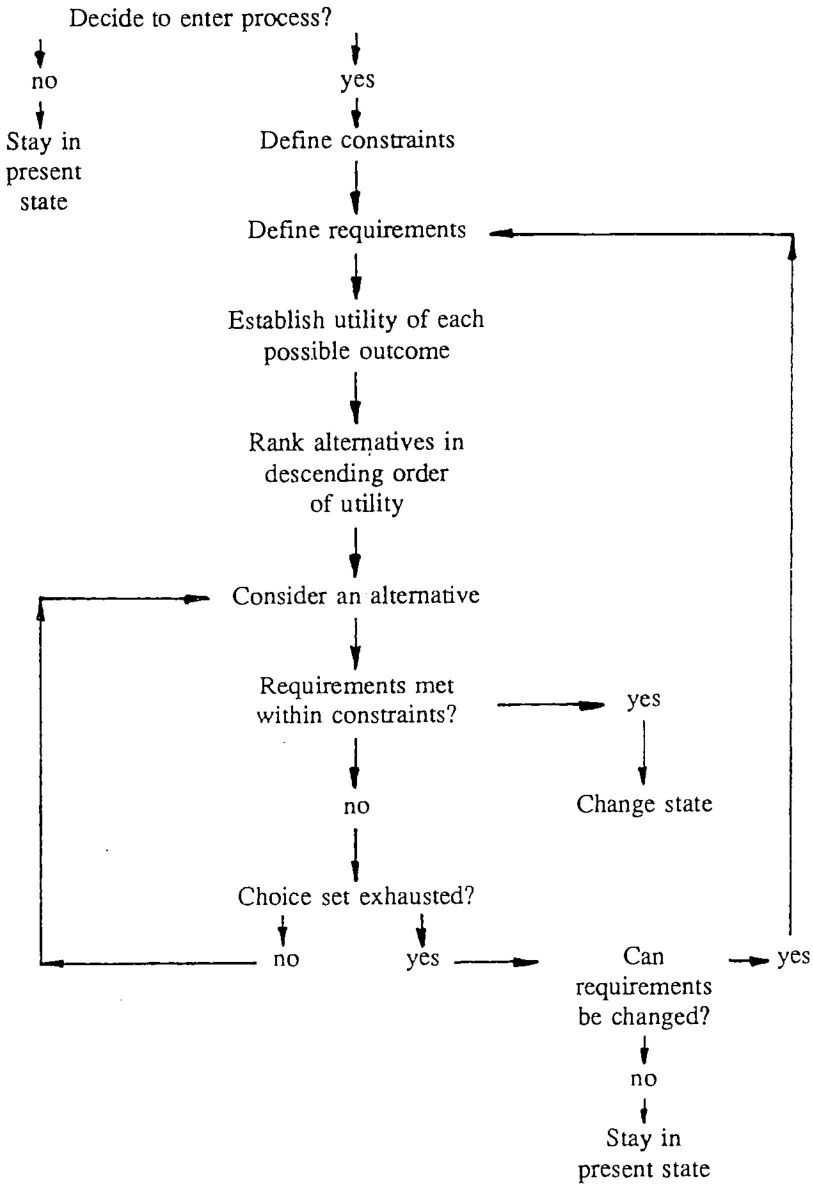
With a binary choice the event occurs if the value of R is less than the probability.

The form of the search processes is shown in Figure 1. An important feature of the method is the initial stage, in which the household or person decide whether or not they wish to enter the process, for example to start to look for a new home or job. This is based on a calibrated utility function, which in essence expresses whether they are satisfied with present state, that is, home or job. The more dissatisfied they are, the more likely they want to change. The dynamics of demographic change are significant. For example, a couple might buy a house which suits them. However, once they start having children the house will become less suitable in terms of size as the family grows, hence the household is more likely to want to move. Those entering the process determine their constraints and requirements. In the case of residential search the former is the amount of money available, and the latter defined in terms of number of bedrooms and the desired tenure category. The utility of each area for the relevant purpose is established using the utility function. The zones are then ranked in descending order of utility. The household then checks to see if there is a suitable dwelling that can be afforded, or the worker checks to see if there is a job of suitable salary level. If so, the home or job is accepted and then the appropriate characteristics updated. If this zone does not provide a suitable opportunity the next is considered. This process is continued until a suitable opportunity is found or the choice set is exhausted.

Based on empirical evidence (Mackett and Johnson, 1985) about six seems to be appropriate number of areas to be considered. There are those who have entered the search process through force of circumstance rather than choice. In the case of residential search, these are people whose homes have been demolished or one partner in a divorce. In-migrants are treated in a similar manner since, by definition they do not have a home within the study area and so must be given one. The unemployed enter the job search process involuntarily. These people include all areas in their choice sets. They are not, of course, guaranteed to find a suitable opportunity. Once all the items in the choice set have been considered, the searcher may change his or her requirements. In the case of residential search this means a change of tenure, and in the case of employment it means a change of type of job sought.

Rules determine the change sequence. The process continues until all the suitable opportunities have been considered. Anybody who has failed to find a suitable opportunity stays in the previous state. These people are regarded as 'frustrated'. This is a useful concept as it means it is possible to identify which people have unmet needs. It is then possible to see whether any of the policies being examined help those needs to be met. Unemployed people who fail to find a job stay unemployed. 'Forced' residential searchers who fail to find a home are assumed to out-migrate.

Figure 1: The structure of the search processes



3. THE STRUCTURE OF THE MODEL

3.1. The overall form of the model

The overall model structure as it is run on the computer, is shown in Figure 2. This starts with the input of data, the updating of some values, such as some of the supply-side variables and probabilities. This is followed by a series of processes that a household and its members can follow. In practice it is unlikely that any one household will go through all the processes in any one time period.

3.2. Demographic process

The demographic processes are aging, death, marriage, divorce and giving birth. The aging process is represented deterministically, as in Equation (1). The processes of dying, wishing to marry, divorcing and giving birth are all determined by means of Monte Carlo simulation, as shown by Equation (3), using probabilities based on aggregate rates. Some rules are then applied where appropriate, for example to ensure that when a person dies, his or her spouse becomes a widow or widower. It would be possible to estimate utility functions for the demographic processes, but there is little point in doing so, because they are not policy sensitive. The objective of the demographic model is to provide the population to go through the rest of the processes.

3.3. Changes in economic activity state

The age at which a young person enters the labour market is determined by Monte Carlo simulation as a function of their education level, sex and social status. Those who enter the labour market and those made redundant are labelled 'unemployed'. They become employed by finding a suitable job, as discussed in Section 3.6. People leave the labour market by retiring, or in the case of women, to have a child. In the latter case return to the labour market is determined by Monte Carlo simulation. The stage in the personal life-cycle is a function of economic activity state, and age, both absolute and relative to the head of the household.

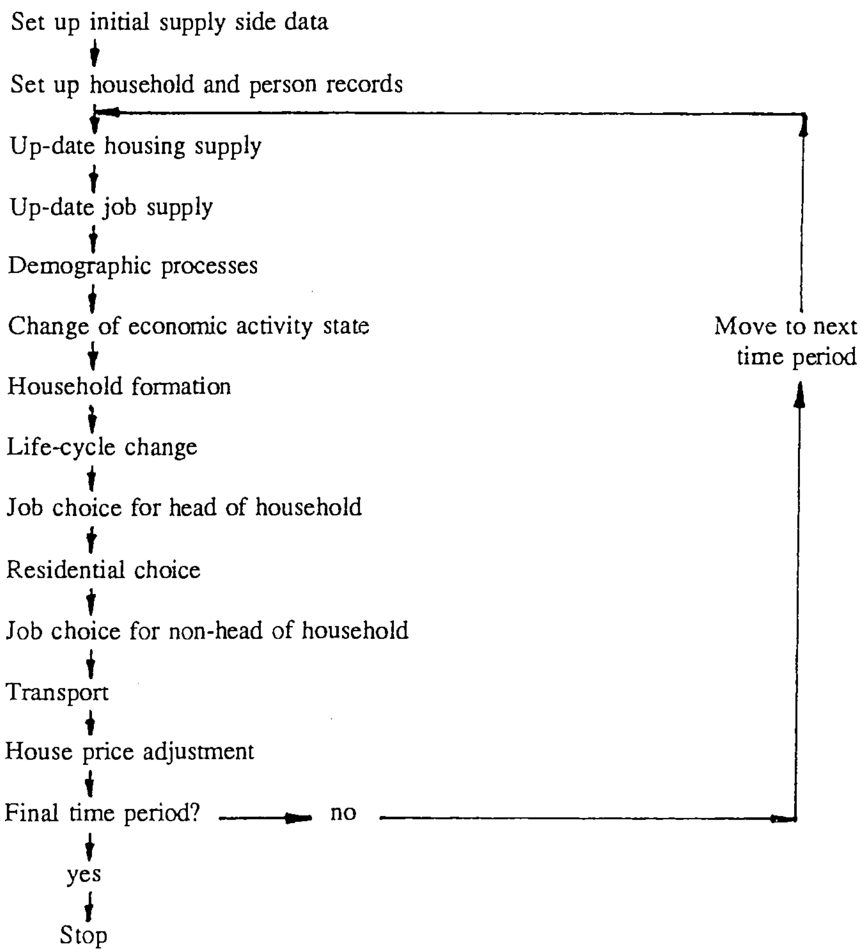
3.4. Household formation

Households are formed when one of the following occurs:

(a) a couple marries and finds a dwelling of their own, or ceases to live with the parents of one of them,
(b) a young person leaves home to live alone,
(c) several people come together to form a 'group of friends', or
(d) a couple divorces and forms two new households.

Households progress through a series of life-cycle stages. The stage is determined by a set of rules based on the number of people in the household, the marital and economic activity status of the head and the age of the youngest person.

Figure 2: MASTER Model



3.5. Job supply

Ideally, the model would represent the actions of industrialists and politicians in determining the location of jobs. This is not feasible, so heuristic methods are used. Redundancies are determined by Monte Carlo simulation at the individual level. New job location is determined as a function of the existing pattern and the overall level of change.

3.6. Employment choice

Job choice for an individual is a two-stage process. The first is a binary choice as to whether the consider changing job. The second is a search process to find a suitable job. The job choice process is carried out before the residential choice process for heads of households but after it for non-heads. This is to reflect the likely sequence of events.

The decision to consider looking for a job is a binary choice using a model of the form of Equation (5), and is a function of the age, sex, marital status, income, social status, and in the case on non-head of households, whether or not there has been a recent change of residence. Monte Carlo simulation is used to determine whether a person wishes to consider changing job. These people, plus all the unemployed, enter the job search process. This is similar to that shown in Figure 1. The utility of each zone is a function of the cost of travel and the travel time from their residential location, the distance from the present job, and the number of jobs available. The first two factors reflect the fact that people would prefer a job that is easy to reach from home. The third factor represents the fact that jobs tend to cluster and that the person is more likely to be aware of a job nearby than one far away, and the final factor reflects the fact that the more jobs there are in an area, the more likely somebody is to find a job there.

The vacant jobs are considered in descending order of utility. Each zone has a set of vacant jobs, each with a socio-economic classification and salary. Rules are used to determine whether the person finds a suitable one. If no suitable job is found the person considers the next zone. Unemployed people search all zones, whereas those with a job only consider a limited number. When a suitable job is found the person takes it, and the relevant characteristics are updated, and the previous job, if any, vacated, and made available for somebody else.

3.7. Housing supply

The housing supply varies over time because of demolition, conversion and construction. Demolitions are determined by Monte Carlo simulation, using exogenously supplied data as the basis of the probabilities.

New housing is located on the basis of a multinomial logit model of the form:

$$H^N_{if} = H^N_{*f} A^A_i \exp(U^N_{if}) / \Sigma_j A^A_j \exp(U^N_{jf}) \quad (6)$$

where H_i^N is the number of new dwellings in tenure group f built in zone i
H_{*f}^N is the total number of new dwellings in tenure category f
A_i^A is the area of land available for housing in zone i
U_{Nf}^i is the utility of zone i for the building of dwellings in tenure category f.

The value of the utility of each zone is a function of the distance and cost of travel from the city centre, the density of housing, and the number of households in the tenure category. For the owner-occupied sector, the price of dwellings per unit area and the proportion of households in the sector are also included.

There is conversion of housing stock for some large dwellings in inner areas to small dwellings for rent and sale, and the sale of publicly-rented dwellings to their occupiers.

3.8. Residential choice

Three processes are represented in the residential choice process: the decision to consider moving, the choice of tenure category, and the search process.

The decision to consider moving home is similar to the decision to consider changing job, except that there are separate versions for whole households and for those leaving a household. In each case the utility is a function of age, the existing tenure category, and the number of bedrooms in the existing dwelling compared with the number desired. The latter factor is based on the number of people in the household, and the ages and sexes of the children, if any.

The desired tenure is a nested logit model using two binary choices. The first decision is between buying and renting a dwelling. The second decision is for those who are renting to choose between renting in the public or private sector. The utilities are functions of the household income and social status, the age of the household head, and the existing tenure.

The residential search process is similar to that for job choice as shown in Figure 1. The utility of each zone is a function of the cost and travel time to work for all workers in the household, the distance from the city centre, the number of households in the zone and various zonal characteristics of the population living in the zone. For potential buyers, the cost of dwellings relative to the household's budget is also included.

3.9. Property prices

The price of each property is calculated as a function of its type, size and location, using a regression equation. The values are adjusted in each time period to reflect market conditions, using the residential utility functions, which incorporate both locational and travel variables.

3.10. Transport supply

Transport supply is represented by a road network upon which cars and buses operate, plus a simple rail network. Walking is also included.

The shortest path through the road network is found using Dijkstra's (1959) algorithm. Bus routes are included, with a factor included to represent the effects of overcrowding on route choice.

3.11. Travel demand

The transport demand processes that are modelled are concerned with estimating the numbers using each mode, and the routes they take, allowing for the effects of congestion. Since these are strongly influenced by the use of the car, it is necessary to model car acquisition and car availability within the household. This implies the necessity to establish which members of the household hold a driving licence.

Driving licence holding is determined by Monte Carlo simulation, using age and sex specific rates based on data about those passing the driving test.

Car ownership is treated in terms of the transition between different levels of car owning. Utility functions have been calibrated, and are functions of the income and number of economically active people in the household, the age of the head, the proportion of the household who hold a driving licence, and whether or not there is a new car licence holder.

Car availability within the household is a rule-based process which interacts with the modal choice process. Cars are allocated to members of the household who have a driving licence and drive to work, in descending order of age. If a person has a car made available, but chooses not to use it, it passes to the next person. This process continues until either all the eligible people have been considered, or all the cars allocated.

Modal choice is a multinomial logit model, in which the modal shares are a function of travel cost, travel time, and the income, age and sex of the person concerned.

Route choice for car is also determined by multinomial logit, based on the cost and travel time along several possible routes. The use of a probabilistic approach means that the trip makers are spread between the routes in a manner similar to that of Dial (1971). The model is being run for a sample population (typically one per cent). Congestion, by definition, requires the modelling of all trips so it is necessary to scale up to the 100 per cent level. Instead of using Monte Carlo simulation, a set of people equal in number to the inverse of the sampling rate (100 for a one per cent sample) is allocated to the routes on the basis of the probabilities given by the multinomial logit model. If any links are above capacity, the speed-flow relationships are used to adjust the speeds, and hence the travel times.

While it would be possible to use iteration to reach equilibrium between the travel utilities and the flows, this conflicts with the behavioural approach being adopted in this model, since it would imply each individual making choices on successive iterations as the utilities were adjusted, which does not happen in reality. However, people do gain experience, which they use in subsequent decisions. This is replicated in the model by updating the travel times as people are loaded on the network, so that the travel times for routes for those further down the stack of households will be slightly different. The stack is treated as a continuous loop from the previous time period, thus ensuring stability. Thus the model is using a form of incremental assignment, with very small increments.

4. APPLICATION OF THE MODEL

The model has been applied to the city of Leeds in the north of England. It has been calibrated for 1971 and forecasts made in 20 one-year intervals to 1991. This facilitates comparison with the 1981 Census of Population.

The model has been applied to the same spatial and temporal system as the Leeds Integrated Land use Transport model (LILT) (Mackett, 1983) used in the work of the International Study Group on Land-Use Transport Interaction (ISGLUTI), so that the two models can be compared.

The model is currently being used for the analysis of a variety of the policies devised by ISGLUTI. These include changes to the cost of travel, population growth and changes to the employment pattern. Comparisons between an earlier version of MASTER and LILT yield some interesting results (Mackett, 1990b).

From MASTER model it is possible to calculate elasticity values for both car and public transport, disaggregated by a variety of sectors, such as sex, age, income and social status. A useful aspect of this model is that it permits analysis of those who wish to enter the processes of job and residential change, to see whether or not they are successful. If they are not they are labelled as 'frustrated'. This is a useful indicator of how well the housing and job markets are functioning. Policies can then be examined to see if they improve in the way the markets operate.

5. CONCLUSIONS AND FUTURE WORK

This paper has described a model that permits the examination of the impact of a variety of land use and transport policies. By working at the micro level the impact of the policies can be traced explicitly. However, the model uses considerable computing power. To overcome this problem, and to permit application of the method to a larger urban system, namely London, a new model, based on experience gained in the course of this work, is being developed using parallel processing methods.

ACKNOWLEDGEMENT

The development of this model has been carried out as part of the programme of work of the Transport Demand Division of the Transport and Road Research Laboratory. the Project Officer is Dr. A.Ash.

BIBLIOGRAPHY

Dial, R.B.. A probabilistic multipath traffic assignment model which obviates path enumeration. Transportation Research, 5, 1971, 83-111.

Dijkstra, E.W.. Note on two problems in connection with graphs (spanning tree, shortest path). Numerical Mathematics, 1, 1959, 269-271.

Hayashi, Y. and Tomita, Y.. A micro-analytical mobility model for assessing the effects of transport improvement. Transport Policy, Management and Technology - Towards 2001. Selected proceedings of the Fifth World Conference on Transport Research held in Yokohama, Japan. Ventura: Western Periodicals, 1989.

Mackett, R.L.. The Leeds Integrated Land-use Transport model (LILT). Supplementary Report SR 805. Crowthorne: Transport and Road Research Laboratory, 1983.

Mackett, R.L.. MASTER model (Micro-Analytical Simulation of Transport, Employment and Residence). Supplementary Report SR 237. Crowthorne: Transport and Road Research Laboratory, 1990a.

Mackett, R.L.. Comparative analysis of modelling land-use transport interaction at the micro and macro levels. Environment and Planning A, 21, 1990b, 323-338.

Mackett, R.L. and Johnson, I.. Residential search behaviour - the implications for survey and analytical design. Tijdschrift voor Economische en Sociale Geografie, 76, 1985, 173-179.

Orcutt, G.H., Greenberger, M., Korbel, J. and Rivlin, A.M.. Micro-Analysis of Socio-Economic Systems: a Simulation Study. New York: Harper and Brothers, 1961.

Wegener, M.. The Dortmund housing market model: a Monte Carlo simulation of a regional housing market. Arbeits Paper Number 7. Dortmund: Institut für Raumplanung, Universität Dortmund, 1983.