Discrete Choice Models for Location and Travel in the Context of Developing Countries

K.V. Krishna Rao
Associate Professor
Civil Engineering Department
Indian Institute of Technology Bombay

M. Kumar
Research Scholar
Civil Engineering Department
Indian Institute of Technology Bombay

Decisions relating to location, travel and other related choices, such as residential location, shopping destination, recreational destination, trip frequency, mode of travel, route, car ownership, etc., have increasingly been modelled by using the discrete choice theory developed based on the concept of utility maximisation. Attempts have been made by many researchers to develop integrated urban activity and transport models. In the above context, the authors have taken up a study to explore the applicability of discrete choice models in arriving at a realistic decision framework for the various alternative choices involved in location and travel aspects in the cities of developing countries. As a first step towards this goal, in this paper, the experiences of the authors in designing the stated preference (SP) experiments, their execution, calibration of the choice models, validation and prediction will be discussed. These SP experiments have been carried out in Mumbai Metropolitan Region of Maharashtra State of India. Also a framework for integrated land use transport model is suggested.

1. Introduction

Travel demand in general depends on land use activity levels and their spread. Land use means the utilisation of land for different activities like residential, commercial, industrial, educational etc., and travel is the link between these activities. More commonly the land use models are used to project demographic and socioeconomic inputs to travel demand models. The transportation infrastructure, including public transportation system is an important determinant of the pattern and intensity of land use. In turn, the land use pattern determines the direction, purpose, magnitude and spatial distribution of travel to be accommodated by the overall transportation system. Therefore, the urban system is composed of different types of spatially separated human activities connected by movement of people and goods etc. The potential for movement between different land uses is subject to usual obstruction of traveling through the transport network. Therefore, the land use transportation interaction is significant and must be understood, analysed, and accounted for in order to ensure that land use and transportation plans and policies are effective and can succeed. Most important, there is a growing appreciation of the idea that transportation and land use policies cannot be formulated independently of one another. Over the last four decades, numerous attempts have been made to model land use and transport changes and their interactions. The main theme of this research describes household's choice of location together with a bundle of other choices for construction of integrated land use transportation in a disaggregate way.

A brief review of the literature on land use transport modeling and a suggested frame work for an integrated land use transport model with disaggregate choice models are provided in the next two sections. The importance of Stated Preference (SP) and Revealed Preference (RP) techniques in the development of constituent choice models of the integrated model, design of SP experiment, its execution, use of computers in SP data collection, calibration of choice models, details of choice models developed by the authors are presented in the subsequent sections.

2. The Land Use Transport Modelling Tradition

In the field of urban economics, Von Thunen (1826), Losch (1954), Wingo (1961) and Alonso (1964), among others, tried to understand the functioning of a city from an analytic point of view with as few and general assumptions as possible. The purpose was to describe the aggregate behavior of a city in terms of land prices, lot sizes, commuting patterns, location of citizens of different categories, etc. This helped planners and scientists to achieve a better understanding of what planning could accomplish and what it could not, and which kinds of measures that could improve the functioning of a city. However, this resulted in virtually no operational models. This should be no surprise, since the basic assumptions were clearly unrealistic and too far-reaching. The main contribution to the initial operational models was the bid-rent approach, introducing (among other things) the possibility to model the land rent market in a consistent way. This idea has been successfully implemented in urban models such as MUSSA (e.g. Martinez, 1992) and RURBAN (e.g. Miyamoto, 1996).

The origin of the wellknown aggregate spatial interaction land use models of gravity type can be said to be Hansen (1959). Some of the earliest models were developed by Lowry (1964) and Echenique (1968). It must be considered as a remarkable accomplishment that this type of models, having been operational for almost a decade, could later be "filled with theory" by the works of Wilson (1967, 1970), Senior and

Wilson (1974), Erlander (1977), Snickars and Weibull (1977) and Smith (1978) among others. With both theory and a firm practical tradition, spatial interaction models were constructed during the 1970s that were both possible to estimate, useful for predictions and stood on a solid theoretical ground. These included the works of Wilson (1974), Putman (1973, 1975a, b), Coelho and Williams (1977), Lundqvist (1975) to mention a few. Many of these models are still very much operational and under continuing development.

Discrete choice models have played an important role in transportation modeling for the last 30 years. Discrete choice theory which was derived from random utility theory was first introduced in the field of travel demand by McFadden (1974), and Domencich and McFadden (1975). The first applications dealt with transport problems at a "lower" level, such as mode choice and destination choice, but soon the classic "four-step model" was reinterpreted as an individual discrete choice model (Senior and Williams, 1977), along with the generalization of the multinomial logit to the Generalised Extreme Value (GEV) model, with the nested multinomial logit model as a special case (Williams, 1977a,b; McFadden, 1978). This type of models divides the decisions involved with a trip into four steps: whether to travel (choice of trip frequency), destination choice, mode choice and route choice. The first three steps are modeled with a random utility approach almost always as s logit model. The last step is most often handled by a deterministic network equilibrium model based on the Wardrop user equilibrium condition (Wardrop, 1952). Occasionally a first step where people choose location is added (Boyce, 1980). In retrospect, it is hard to understand why it took so long to realize that the logit and the entropy approach were identical for all practical purposes (Anas, 1983; Mattsson, 1984). Today, it does not make much sense to distinguish between "logit models" and "gravity models", the only difference being the formulation of the assumptions.

The combination of the spatial interaction model with the logit transport modeling tradition has created a variety of powerful models. Some are very comprehensive, modeling as many aspects as possible of a region's development. Examples are MEPLAN (Echenique et al., 1988), TRANUS (de la Barra, 1989) and the Dortmund model by Wegener (1985, 1986) to mention just a few. While comprehensiveness may be desirable or necessary for many analyses, these models are seldom easy or even possible to overview or evaluate for someone not deeply involved in its construction. Some modelers therefore relinquish comprehensiveness for the virtue of transparency. Some models that can be said to be examples of this are NYSIM (Anas, 1995), which also contains elements of micro-economic theory, IMREL (Anderstig and Mattsson, 1991), and the work started in Chicago by Boyce (1980). A review of the state-of-the-art of urban modeling can be found in Wegener (1994, 1998).

The purpose of this paper is to propose an integrated framework for location and travel choices considering the fact that the travel demand is a result of interaction of several decisions the individuals make. Such a framework considers the interactions among the various sub models of urban activity and travel, and depicts the decision hierarchies in a realistic way. These decisions usually relate to residential location choice, choice of work place, service destination choice, choice of mode, choice of time of travel, car ownership choice, etc. TRESIS is one such model developed by Hensher et al (2001) focusing on the interdependencies between land use, transport and the environment.

3. Ideal Integrated Land Use Transportation Models

At present there are more than twenty land use transportation models being used around the world (TMIP, 1997; US EPA, 2000; Wagener et al., 1999; Wilson, 1998). The operational models may be classified into three types according to their formulation structure as shown in Fig.1 (Miyamoto et al., 2000).

Firstly, Type A modeling system is a connected system of separate land use model and transportation model. In this case land use model represents land use only and transportation model represents transportation only. The land use framework ranges widely from Lowry model to disaggregated micro-simulation model; the transportation framework is mostly based on the conventional four-step model. The interactions between land-use and transportation is represented by externally interfacing the two models. In one direction the socioeconomic data are transferred from land-use model to transportation model; in the opposite direction the travel time/cost or accessibility measures are transferred back from transportation model to land-use model with time lagged. Example of Type A system includes the combination of DRAM/EMPAL, UrbanSim and EMM/2 in the US; DELTA and START in the European countries.

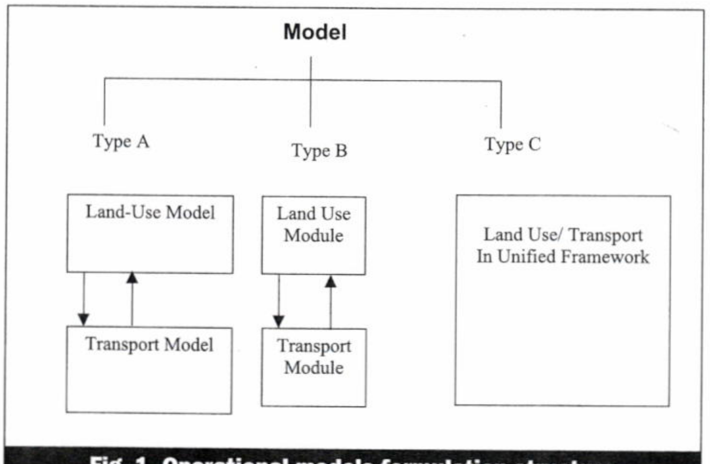

Fig. 1 Operational models formulation structure

Secondly, Type B system is an interaction or composite system of land-use module and transportation module connected in the same physical software package. Land-use module determines land-use and also derives transportation (travel demand), which is then input to transportation module. Transportation module consequently works for mode choice and route assignment only. The interface between two modules is done by converting the flow of activity into the flow of passenger and freight in the appropriate time unit in one direction; and to convert travel time/cost into the appropriate time unit in the opposite direction. Examples of Type B system are obviously the MEPLAN and TRANUS models.

Thirdly, Type C system is a unified or integrated system of land-use and transportation. The ideal integrated model system consists of four interrelated components; Land development, location choice, activity/travel and automobile ownership. The land development models the evolution of the build environment and includes the initial development of previously vacant land and the redevelopment over time of existing land uses. This component could also be labeled "Building supply," because building stock supply functions are included. The location choice includes the location choices of households (for residential dwellings), firms (for commercial locations), and work places (for jobs). The activity/travel models involves predicting the trip making behavior of the population, ultimately expressed in terms of Origin and Destination flows by mode by time of day. Finally the automobile ownership component models household auto ownership levels, an important determinant of household travel behavior. Fig.2 presents a highly idealized representation of land use transportation modeling system. The behavioral core of this system consists of four interrelated components.

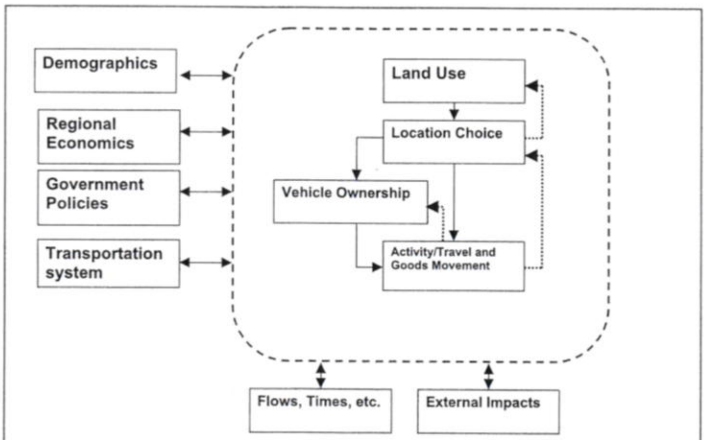

Fig. 2 Idealized integrated urban modeling system

As explained, an integrated land use transport model consists of several sub-models that explain the decisions of households/individuals relating to location, travel, vehicle ownership, etc. Each sub-model is a discrete choice model calibrated with either RP or SP data or both. The calibrated sub-models can be integrated using microsimulation. Stated Preference techniques may be very handy in building the sub-models of an integrated disaggregate land use transport model.

4. Stated Preference Techniques

The most frequent type of choice data corresponds to revealed preference (RP) information which is the data about actual or observed choices made by individuals. In the case of stated preference (SP) individuals are asked about what they would do in a hypothetical situation. As the opportunities for undertaking real life controlled experiments within transport systems are very limited, SP surveys, a quasi-experiment based on hypothetical situations set up by the analyst, provide an approximation to this. The degree of artificiality of these situations may vary according to the rigour and needs of the exercise. A very basic problem with this type of information is how much faith we can put on individuals actually doing what they stated they would do when the case arises. In fact, experience in the 1970's was not very good in this sense, with up to 100% differences between predicted and actual choice found in many studies (Ortuzar, 1980).

The situation improved considerably in the 1980's and recently good agreement with reality has been reported from models estimated using SP data (Louvier, 1988). However, this has occurred because SP data collection methods have improved enormously and are now very demanding, not only in terms of survey design expertise, but also in terms of their requirements for operational resources. Table 1 gives the clear difference between stated preference and revealed preference data.

Revealed Preference data	Stated Preference data
Based on actual market behavior	Based on hypothetical scenarios
Attribution measurement error	Attribute framing error
Limited attribute range	Extended attribute range
Attributes correlated	Attributes uncorrelated by design
Hard to measure intangibles	Intangibles can be incorporated
Cannot directly predict response to new alternatives	Can elicit preferences for new alternatives
Preference indicator is choice	Preference indicators can be rank, rate or choice intention
Cognitively congruent with market demand behavior	May be cognitively non-congruent

Table 1 Difference between Stated Preference and Revealed Preference Data

Stated preference data collection methods have improved enormously and are now very demanding, not only in terms of survey design expertise but also in their requirements for trained survey staff and quality assurance procedures (Pearmain et al, 1991). The main features of an SP survey may be summarized as follows.

• It is based on the elicitation of respondents' statements of how they would respond to different hypothetical (travel) alternatives.

• Each option is represented as a 'package' of different

attributes like travel time, travel cost, headway, reliability and so on.
• The individual effect of each attribute can be estimated using experimental design techniques that ensure the variations in the attributes in each package are statistically independent from one other.
• The respondents state their preferences towards each option by ranking them in order of attractiveness, rating them on a scale indicating strength of preference, or simply choosing the most preferred option from a pair or group of them.

4.1 Attributes and Alternatives

A stated preference experiment has as one of its main elements, the construction of a set of hypothetical options which are referred to as technologically feasible alternatives; these are defined on the basis of the factors assumed to influence most strongly the choice problem under consideration. The design of these technologically feasible alternatives requires four distinct tasks (Ortuzar, 2000):

a) the identification of the range of choices (broad options will be included),
b) the selection of the attributes to be included in each broad option,
c) the selection of the measurement unit for each attribute, and
d) the specification of number and magnitudes of the attribute levels

4.2 Stages in SP Data Collection

Practical experience recommends contemplating at least the following stages in SP data collection (Kocur et al, 1982).

• Identify the range of choices, the attributes to be considered and their likely levels of variation.
• Design an initial version of the experiment and survey instrument. The design should include the way in which the options will be presented to the respondents and how they will be allowed to express their preferences.
• Develop a sampling strategy to be followed to ensure a rich and representative data set. Pre-test the survey instrument using a small stratified sample in order to consider the opinion of the largest possible number of interesting sectors of the population.
• Evaluate the pre-test results both in terms of the quality of the survey instrument and of the intuitive quality of the responses obtained by population strata; correct the data before its final distribution.

4.3 Experimental Design

The number of attributes (a) and the number of levels each one can take (n), determine a factorial design (n^a). Tables exist (Kocur et al, 1982) which give the number of hypothetical options needed to test most designs of interest and guarantee independence among options (orthogonality). A problem with the latter is that they require the construction of many more options; for this reason fractional factorial designs, which assume that some or all variable products are negligible, are often used. Consider a situation with five attributes, two at two levels and the rest at three levels (i.e. a $2^2 3^3$ design). In this case, depending on the number of interactions to test, the number of options required would vary as follows (Kocur et al, 1982);

• 108 to consider all effects (a full factorial design)
• 81 to consider principal effects and all interactions between pair of attributes, ignoring the effects of higher order
• 27 to consider principal effects and interactions between one attribute and all the rest
• 16 only if no interactions are considered.

Once the design has been decided, the technologically feasible options are constructed and eventually the experiment conducted and the data collected. Fowkes and Wardmen (1988) give practical recommendations for the desirable variation of the attribute levels; their experience indicates that it is often beneficial to sacrifice some purity in the experimental design (loose complete orthogonality) if one gains in realism.

4.4 Sampling Strategy

As in any other data collection exercise, issues such as sample composition and size are very important in the design of an appropriate SP experiment. Also, in common with RP studies, a basic requirement is to obtain a sufficiently large and representative sample. On other hand, SP experiments are statistically efficient, in the sense that each interviewee produces not just one observation but several on the same choice context; therefore, samples are typically smaller than for comparable RP studies (Bradley 1988). In fact, an early rule-of-thumb seems to have stated that around 30 interviews per market segment might be sufficient. However, more recent work suggests that 75-100 interviews per segment would be more appropriate (Pearmain and Swanson 1990; Bradley and Kroes 1990; and Swanson et al. 1992).

4.5 Identification of Preferences

An important issue is how the respondents will be asked to express their preferences for each option offered to them. As is detailed below, there are three main ways of collecting information on preferences about alternatives: asking respondents to rank them in order of preference, rate them on an arbitrary scale or to choose between them in choice experiments (Ortuzar and Garrido, 1994).

• Ranking responses: This approach presents all the options at once to the respondents and they are then asked to order them by preference. The main attraction of this approach is that all options are presented together but this also limits the number of alternatives that may be considered without fatiguing the respondent. Furthermore, note that the data collected represent judgments that not necessarily correspond to the type of choices faced by respondents in real life.
• Rating: These techniques have been used for many years by market research practitioners. In this case respondents are asked to express their degree of preference for an option using an arbitrary scale (such as, between 1 and 5 or between

1 and 10). In this case the respondents are allowed to express his or her degree of preference between two options on a semantic scale, typically with five points.

• Choice: Choice experiments require the respondent to select an option either from a pair (binary choice) or group of alternatives. In its pure form the respondent only chooses his or her preferred option, thus expressing his or her preference in a form analogous to a RP survey. In its extended form the respondent is allowed to declare preferences using a semantic rating scale as outlined above. To increase realism it may be possible to allow the choice of " none of the above" to avoid forcing the selection of a least bad but still unacceptable alternative (Olsen and Swait, 1998)

4.6 Use of Computers in SP Surveys

In principle computers offer significant advantages over "paper and pen' methods in terms of aiding the conduct of SP surveys. The most attractive element of computer based interviewing is the possibility of tailoring the experiment to the subject. Most SP interviews will include an initial questionnaire in which data about the respondent and a recent journey is collected and used to build a subsequent experiment. This questionnaire can be reproduced in software with the added advantage of automatic entry validation and automatic routing. With a computerized system the responses to this questionnaire can be used to generate SP experiments and options automatically for each individual, following a specific design. Automatic routing can be used to select the appropriate experiment depending on the circumstances of each individual. Furthermore, range and logic checks on the responses and pop-up help screens or look up information windows can be incorporated to improve the quality of the surveys.

The use of computers for SP surveys makes it possible to design more complex interviews than might be attempted manually, although this complexity may never be apparent to the respondent or even to the interviewer. It has also become possible to design experiments including graphic material in built in the SP game. Moreover, good software permits randomization of the order in which the options are offered to each person thus removing a further potential source of response bias. Finally, as all responses are stored directly on disk there are no entry costs nor are coding errors and data available immediately for processing. On more subjective side, it has been found repetitively in practice that using computers has two further advantages. Firstly, respondents tend to consider the interaction with the machine as more "serious" matter, thus giving more attention to the task and hopes participating more responsibly in the experiment. Second, and even more subjective, some respondents seem to have less trouble in punching sensitive information (income) than to declare it to a human interviewer (Ortuzar, 1996).

A number of software packages offer excellent facilities for designing and coding complex interviews with a minimum of understanding of computing itself. Among the best known in the transport field are ALASTAIR (Steer Davis Gleave), MINT (Hague Consulting group) and ACA (Sawtooth Software) and there are many more packages available in the market research field. The most recent development is EXPLICIT (Steer Davies Gleave) which has the feature of allowing the inclusion of powerful graphic aspects in the design.

4.7 Modelling with Mixed RP and SP Data

In econometric theory estimation of models with different data sources is called mixed estimation. These data are classified into two types; primary data and secondary data, the primary data provide additional information about the main modeling parameters, the revealed Preference data constitute the primary set, since these data capture the actual behavior of the individuals, and Stated Preference data constitute the secondary set. The difference between the error in RP and SP to be represented as a function of variances of each type of error ε and τ. It can be written as

$$\sigma_\varepsilon^2 = \mu^2 \sigma_\tau^2 \tag{1}$$

Where, μ is an unknown *scale coefficient* which leads to the following utility functions for a certain alternative A_i:

$$U_i^{RP} = \theta X_i^{RP} + \alpha Y_i^{RP} + \varepsilon_i \tag{2}$$

$$\mu U_i^{SP} = \mu\,(\theta X_i^{SP} + \phi Z_i^{SP} + \tau_i) \tag{3}$$

Where, α, ϕ and θ are parameter (vector) to be estimated; X_i^{RP} and X_i^{SP} are attribute vectors (of both alternatives and individuals) at the RP and SP levels, respectively.

Y_i^{RP} and Z_i^{SP} are attributes, which only belong to RP or the the RP or SP sets, respectively.

The above equation, in which the SP parameters were multiplied by μ, makes the associated stochastic error to have the same variance as that of RP.

In general the SP data have more noise than the RP data, therefore it is sensible to expect that μ should lie in between one and zero. If not, it may be assumed that the RP data is noisier than SP and it can be tested by re-estimating the model with the converse structure. The choice probabilities would be given by the following expressions (Morikawa et al. 1992):

$$P_i^{RP} = \frac{\exp(\theta X_i^{RP} + \alpha Y_i^{RP})}{\sum_j \exp(\theta X_j^{RP} + \alpha Y_j^{RP})} \tag{4}$$

$$P^{SP} = \frac{\exp \mu\,(\theta X_i^{SP} + \phi Z_i^{SP})}{\sum_j \exp \mu\,(\theta X_j^{SP} + \phi Z_j^{SP})} \tag{5}$$

from the above expressions it is possible to obtain the following joint likelihood function:

$$L(\theta,\mu,\alpha,\phi) = \left[\prod_{n=1}^{N^{RP}} \prod_{Ai\varepsilon A(q)} P_{iq}^{RG}\right] * \left[\prod_{n=1}^{N^{SP}} \prod_{Ai\varepsilon A(q)} P_{iq}^{SP}\right] \qquad (6)$$

which should be maximized to yield the parameter estimates. The above equation is a non-linear function because μ is multiplying not only the attributes but also the SP parameters. One approach to solve this problem is to apply software especially designed to deal with non-linear likelihood functions directly. Two methods are available in the literature for jointly estimating a model with Sp and RP data. These are outlined here.

Sequential estimation method

This method has the advantage of allowing the use of ordinary logit or probit estimation software. The algorithm is as follows (Ben-Akiva and Morikawa, 1990):

Step 1 Estimate the SP model according to utility functions given in equation (3) in order to obtain the estimators of $\mu\theta$ and $\mu\phi$. Then, define a new variable:

$$\hat{V}_i^{RP} = \mu\theta X_i^{RP} \qquad (7)$$

Step 2 Estimate the following RP model with the new variable included, in order to estimate the parameters λ and α:

$$U_i^{RP} = \lambda\hat{V}_i^{RP} + \alpha Y_i^{RP} + \varepsilon_i \qquad (8)$$

Where, $\lambda = 1/\mu$.

Step 3 Multiply X^{SP} and Z^{SP} of SP data by μ to obtain a modified SP data set. Pool the RP data and the modified SP data and then estimate the two models jointly.

Simultaneous estimation method

This method, developed by Bradley and Daly (1991), consists of constructing an artificial tree which has twice as many alternatives as there are in reality, as shown in Fig.3. Half of these are labeled RP alternatives and the other half as SP alternatives. For the RP alternatives, SP choices are unavailable, and therefore, these are estimated as standard logit model. The problem of non-linearity can be solved by using existing tree logit estimation software like ALOGIT (Hague Consulting Group). The tree has as many elementary alternatives as there are in RP and SP sets combined. The RP alternatives emerge directly from the root, whereas, the SP alternatives are placed in nest, emerging from the root. The RP alternatives are modeled using the nest structure and SP alternatives are modeled using the tree structure. In case of SP alternatives, each nest comprises of only one alternative. The main utility of the dummy-alternative can be computed as suggested by Daly (1987) and is given by

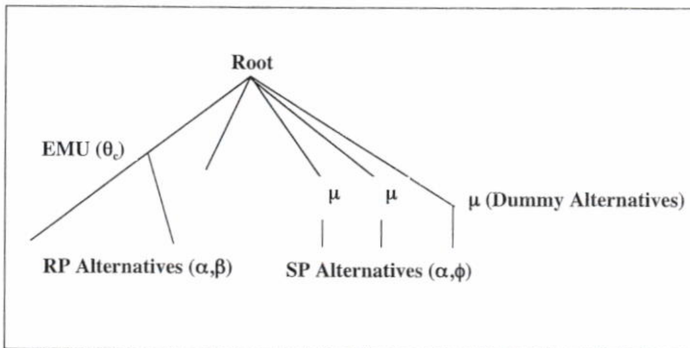

Figure 3 Artificial nested-tree structure for mixed RP/SP estimation

$$V^{COMP} = \mu\log\sum\exp(V^{SP}) \qquad (9)$$

As there is only one alternative in the nest the expected maximum utility (EMU) of the nest becomes equal to the utility of the alternative itself and can be given as

$$V^{SP} = \theta X^{SP} + \phi Z^{SP} \qquad (10)$$

Therefore, the utility of the nest will become

$$V^{SP} = \mu X^{SP} + \phi Z^{SP} \qquad (11)$$

which is exactly the same required and presented in the equation (3). The scale factor should take the same value for all the SP alternatives. Also as the individuals are not modeled as choosing from among the RP and SP alternatives simultaneously, the assumption of scale factor not exceeding unity, does not apply. If the scale factor is higher than unity, then it implies that SP data has less noise than the RP data and opposite is true if scale factor is less than unity (Ortuzar and Willumsen 1994).

Two stated preference experiments that are designed and executed for finding the preference of the travelers towards the proposed new modes viz., mass rapid transit system and water transport system in Mumbai Metropolitan Region, India are presented in the next section.

5. Discrete Choice Models Based on SP Data

5.1 SP Model for a Proposed Passenger Water Transport

An SP experiment was floated to know the preferences of people towards the attributes of the proposed Passenger Water Transport (PWT) system along the Western coast of Mumbai city of India. The attributes of PWT system (Catamaran or Hovercraft) like travel cost, travel time, waiting time, comfort, convenience and reliability of service would play major role in attracting the passengers from the existing modes like taxi, car, autorickshaw, two wheeler, etc. The SP experiment was designed as a rating experiment by constructing several options by Hovercraft/Catamaran with different attribute levels. The basic format of stated preference experiment used in the present study along with the rating scale is shown in Table 2.

Existing Mode		PWT Mode	
Waiting Time	Stated	Waiting Time	3 Levels
Travel time	Stated	Travel Time	Computed
Travel Cost	Stated	Travel Cost	3 Levels
Discomfort	Stated	Discomfort	1 Level

		CHOICE SCALE		
Definitely Existing	Probably Existing	Can't Say	Probably PWT	Definitely PWT

Table 2 Typical Stated Preference Experiment Design

The attribute *waiting time* is considered at three levels, i.e., 5 minutes, 10 minutes and 15 minutes for both catamaran and hovercraft. These are arrived at considering the worst service headway of 30 minutes. The attribute *travel cost* is considered at three levels, i.e., Rs. 2.0, 3.0, 4.0 per km for hovercraft and Rs.1.75, 2.5, 3.0 per km for catamaran. These fare levels are worked out based on operation and maintenance costs of the vessels and the fares charged by the operators on existing routes. In fact, these fare levels are administered in the experiment as overall fares for the entire trip and not as fare per unit distance. The attribute *travel time* is computed for each respondent at the time of the interview by knowing the origin and destination of his trip. The terminal-to-terminal travel times by hovercraft/catamaran are added to the access and egress travel times to get the overall travel time by PWT mode. Average speeds of 60 km/hr (33 knots/hr) and 45 km/hr (25 knots/hr) respectively for hovercraft and catamaran are used for computing these travel times. The attribute *discomfort* is used on a scale of 0 to 4. Discomfort level of 0 indicates comfortable sitting in air-conditioned environment. Discomfort level 4 indicates standing in a overcrowded bus/train. The discomfort level of catamaran and hovercraft are taken as 0 (Comfortable seat in air-conditioned environment) while administering the SP experiment.
Each respondent was asked to rate 9 options (3 waiting time levels × 3 travel cost levels × 1 travel time level 1 × 1 comfort level) for hovercraft and 9 options for catamaran on a rating scale. As the preference of individuals towards a comfortable and convenient service varies with respect to the purpose of trip, the respondents were asked to give their rating separately for work trip that they have made and recreational trip with similar attributes which they would consider making in the near future. In addition two monthly pass options were also included in the experiment. SP questionnaires were prepared for four representative routes of the proposed PWT system. These routes cover more or less the entire range of travel times and fares of the proposed PWT system. In addition a leaflet describing the general features of the proposed PWT system and showing pictures of interior of the vessels, terminal building and other amenities was prepared to adequately convince the respondent that this would be a real system in the near future.

Administering of SP experiment

A team of about 14 enumerators was thoroughly trained for administering the experiment on the respondents. The sample size requirement for stated preference experiments is very low because of the following two factors.

- The data collected on each individual is used in modeling in a disaggregate manner using discrete choice theory.
- In a stated preference experiment one gets many observations from a single individual.

It has been generally reported in the literature that 75-100 interviews per segment would be appropriate. In the present study travelers are segmented based on their present mode use, like, Bus, first rail, second rail, car, two-wheeler, etc. It was attempted to satisfy the above sample size requirement.
The survey was administered at work places, terminals, parking areas, etc. In majority of the cases the interviews were conducted by taking prior appointments from the necessary authorities. The enumerators would first explain the proposed PWT system with the help of the leaflet to the respondent and then collect his personal and trip information by filling the appropriate forms. The attributes waiting time, travel time, travel cost and discomfort by the existing mode obtained from trip information are then transferred to the appropriate place in he SP questionnaire. Travel time by hovercraft is then worked out by using the travel time tables and by assessing the access and egress times from the maps. Each individual is then asked to compare the attributes of the existing system with those of the hovercraft/catamaran option, and give his rating for work trip as well as for recreational trip. The responses obtained were checked for consistency and the errors were removed. The valid samples obtained after this exercise are 1123.
The number of options that were obtained from each respondent are as given below.

- 9 Hovercraft options
- 2 Monthly Pass options
- 9 Catamaran options
- Total options for work trip: 20
- Total options for recreational trip: 18

Thus one valid sample will give 20 observations for work trip model calibration and 18 observations for recreational trip model calibration.

Model calibration based on SP data

Mode wise binary logit models which give the probability of shift from existing mode to PWT mode were developed using the preferences indicated by the respondents. These models are of the following form.

$$\Pr(PWT/EM) = \frac{e^{V_{PWT}}}{e^{V_{PWT}} + e^{V_{PWT}}} \tag{12}$$

$$V_{PWT} = \alpha WT_{PWT} + \beta TT_{PWT} + \gamma TC_{PWT} + \phi DC_{PWT} + CONST \quad (13)$$

$$V_{EM} = \alpha\, WT_{EM} + \beta\, TT_{EM} + \gamma\, TC_{EM} + \phi\, DC_{EM} \quad (14)$$

Where,
$Pr(PWT/EM)$ = probability of shifting to PWT mode conditioned on existing mode (EM),

V_{PWT} = deterministic component of utility of PWT mode,
V_{EM} = deterministic component of utility of Existing Mode,
WT = waiting time,
TT = travel time,
TC = travel cost,
DC = discomfort,
a,b,g,f = parameters to be estimated using SP data, and
$CONST$ = constant that explains the unobserved effects.

The parameters of the mode wise logit models were calibrated by employing maximum likelihood method of estimation. The results of calibration of mode wise logit models for work trips are provided in Table 3. Similar models were developed for recreational trips also. The signs of all the parameters are found to be logical. All the variables entered the model are found to be statistically significant. The r2-statistic for all the models is found to be reasonably good. Any model with a r2 value of more than 0.4 is found to give a very good prediction success.

Mode	Parameter Estimate					ρ2 -statistic
	α	β	γ	ϕ	Const	
TW	- 0.047	- 0.055	- 0.081	- 0.401	0.000	0.69
CAR	- 0.055	- 0.068	- 0.048	- 0.140	0.000	0.52
BUS	- 0.013	- 0.047	- 0.094	- 0.000	-1.353	0.902
II Rail	- 0.011	- 0.068	- 0.087	- 0.000	0.000	0.871
I Rail	- 0.037	- 0.037	- 0.029	- 0.248	-2.041	0.543

Table 3 Calibrated Parameters of Logit model for work Trips

5.2 SP Model for a Proposed Mass Rapid Transit System

This SP experiment was designed to synthesize the mode choice behavior of travelers in horizon years when MRTS for Thane city in Mumbai Metropolitan Region would be available. The attributes considered are Travel Time, Waiting Time and Travel Cost. Waiting time for MRTS option is considered at three levels (2 minutes, 3 minutes and 5 minutes) and travel cost at three levels (Rs.5, Rs.7 and Rs.9). Travel time (one level) by MRTS is worked out based on a speed of 33 kmph. Each individual is asked to choose between the existing option (i.e., the mode used by the respondent for reaching the work place) and the MRTS option. Considering all the combinations of these three attributes one can construct 3×3=9 MRTS options. However only 7 feasible MRTS options are considered in the experiment. Respondent compares each of the 7 MRTS options with the existing option, one at a time, based on waiting time, travel time and travel cost and gives his choice as a preference on the scale given below each option. About 350 samples were collected in this survey.

Binary logit models (conditional on existing mode) were developed describing the modal shiftto MRTS from the present mode (walk, bus, auto rickshaw, car/two-wheeler). Table 4 gives the results of calibration. The subjective values of travel time and waiting time are also reported in the table. The signs of all the parameters are logical. All the variables entered the model area found to be significant as suggested by their t-values.

Variable	Walk		Bus		Autorickshaw		Car/Two wheeler	
	Coefficient	t	Coefficient	t	Coefficient	t	Coefficient	T
WT	-	-	-0.1698	-10.7	-	-	-	-
TT	-0.0669	-2.8	-0.1057	-8.0	-0.0690	-2.6	-0.3093*	-4.8
TC	-0.3310	-3.6	-0.5334	-12.7	-0.2299	-6.7	-0.3913	-4.7
Constant	-1.6050	-2.6	-	-	-0.5851	-3.5	-	-
Rho-square	0.20		0.22		0.41		0.42	
SV of WT	-		19.10		-		-	
SV of TT	12.13		11.89		18.00		47.43*	

SV: Subjective Value TT: Travel Time WT: Waiting TimeTC: Travel Cost
*waiting time included

Table 4 Calibrated Binary Choice Models Using SP Data

6. Conclusions

There is need to examine the working of an integrated land use transport model made up of behavioural submodels in the context of developing countries. In this paper the attempts made towards this in literature were reviewed. A framework for such an integrated model was suggested. The submodels could be calibrated using SP data or RP data or both. The stated preference models of mode choice developed in the context of Mumbai Metropolitan Region were presented. Other submodels like residential location, service destination, vehicle ownership, etc., can be calibrated in the similar lines using either SP or RP data or both. Based on the appropriateness of the subjective values of time derived from the two models presented in this paper it can be concluded that SP techniques can be successfully used in modelling choices relating to travel and location in the context of cities of developing countries.

REFERENCES

ALONSO, W. (1964) *Location and Land Use. Harvard Univ.* Press, Cambridge, MA.

ANAS, A. (1983) *Discrete choice theory, information theory and the multinomial logit and gravity models*. Transportation Research B, Vol. 17, 13-23.

ANAS, A. (1995) *Capitalization of urban travel improvements into residentia l and commercialreal estate: simulations of a unified model of housing, travel mode and shopping choices*. Journal of Regional Science, Vol. 35, 351-375.

ANDERSTIG, C. AND MATTSSON, L.G. (1991) *An integrated model of residential and employment location in a metropolitan region*. Papers in Regional Science 70, 167-184.

BEN-AKIVA, M. AND MORIKAWA, T. (1990) *Estimation of switching models from revealed preferences and stated intensions*. Transportation Research A, Vol. 24, No. 6, 485-495.

BOYCE, D.E. (1980) *A framework for constructing network equilibrium models of urban location*. Transportation Science, Vol. 14, 77-96.

BRADLEY, M. A. AND KROES, E. (1990) *Forecasting issues in stated preference survey research*. 69th TRB Annual Meeting, Washington DC, January 1990, USA.

BRADLEY, M. (1988) *Realism and adaptation in designing hypothetical travel choice concepts*. Journal of Transport Economics and Policy, Vol. 22, 121-137.

BRADLEY, M. A. AND DALY, A. J. (1991) *Estimation of logit choice models using mixed stated preference and revealed preference information*. Preprints 6th International Conference on Travel Behaviour, Quebec, Canada.

COELHO, J.D., WILLIAMS, H.C.W.L. (1977) *On the design of land use plans through locational surplus maximization*. Presented at The 17th Congress of the European Regional Science Association.

DALY, A. (1987) *Estimating tree logit models*. Transportation Research B, Vol. 21, No. 4, 251-267.

DE LA BARRA, T. (1989) *Integrated Land Use,* Transport Modeling. Cambridge Univ. Press, Cambridge.

DOMENCICH, T.A. AND MCFADDEN, D. (1975) *Urban Travel Demand: A Behavioral Analysis*. North-Holland, Amsterdam.

ECHENIQUE, M.H. (1968) *Urban systems: towards an explorative model*. In: Land Use and Built Form Studies, Cambridge.

ERLANDER, S. (1977) *Accessibility, entropy, and the distribution and assignment of traffic*. Transportation Research, Vol. 11, 149-153.

FOWKES, A.S AND WARDMAN, M. (1988) *The design of stated preference travel choice experiments with special reference to interpersonal taste variations*. Journal of Transport Economics and policy, Vol. 22, 27-44.

HANSEN, W.G. (1959) *How accessibility shapes land use*. Journal of the American Institute of Planners 25.

HENSHER, D. A. AND TON, T. (2001) TRESIS: *A Transportation, Land Use and Environmental Strategy Impact Simulator for Urban Areas*. 9th World Conference on Transport Research, Seoul, Korea, July 2001.

KOCUR, G.T., ADLER, T., HYMAN, W. AND AUNET, B. (1982) *Guide to forecasting Travel Demand with Direct Utility Assessment*. Report NUMTA-NH-11-0001-82, Urban Mass Transport Administration, USDT, Washington, DC.

LOSCH, A. (1940) *Die raumliche Ordnung der Wirtschaft,* Jena. Translation, 1954. The Economics of Location. Yale Univ. Press, New Haven, CT.

LOUVIERE, J. J. (1988) *Conjoint analysis modeling of stated preferences – A review of theory, methods, recent developments and external validity*. Journal of Transportation Economics and Policy, Vol. 22, No. 1, 93-119.

LOWRY, I.S. (1964) *A Model of Metropolis. Rand Corporation*, Santa Monica.

LUNDQVIST, L. (1975) *Transportation analysis and activity location in land use planning- with application to the Stockholm region*. In: Karlqvist, A., Lundqvist, L. and Snickars, F. (Eds.), Dynamic Allocation of Urban Space. Saxon House, Westmead.

MARTINEZ, F. (1992) *The bid-choice land use model: an integrated economic framework*. Environment and Planning A, Vol. 24, 871-885.

MATTSSON, L.G. (1984) *Equivalence between welfare and entropy approaches to residential location*. Regional Science and Urban Economics, Vol. 14, 147-173.

MCFADDEN, D. (1974) *Conditional logit analysis of qualitative choice behavior*. In: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York.

MCFADDEN, D. (1978) *Modelling the choice of residential location*. In A. Karlquist, L. Lundquist, F. Snickars and J.W. Weibull (eds), Spatial Interaction Theory and Planning Models, North-Holland, Amsterdam

MIYAMOTO, K. (1996) *An analysis system for integrated policy measures regarding land- use, transport and the environment in a metropolis*. In: Hayashi, Y., Ro y, J. (Eds.), Transport, Land-Use and The Environment. Kluwer Academic Publishers, Dordrecht.

MIYAMOTO, K. AND VICHIENSAN, V. (2000) *Review of land-use model applications in transportation demand forecasting,* Proceedings of the International Conference CODATU IX, Mexico. Balkema, Netherlands.

MORIKAWA, T., Ben-Akiva, M. and Yamada, K. (1992) *Estimation of mode choice models with serially correlated RP and SP data*. Proceedings 6th World Conference on Transport Research, Lyon, France.

OLSEN, G. D AND SWAIT, J. (1998) *Nothing is important*. Working Paper, Faculty of Management, University of Calgary.

ORTUZAR, J.DE D AND GARRIDO, R.A. (1994b) *A practical assessment of stated preference methods*. Transportation, Vol. 21, 289-305.

ORTUZAR, J.DE D AND WILLUMSEN, L.G. (1994) *Modelling Transport*. Second Edition, John Wiley & Sons, Chichester.

ORTUZAR, J.DE D. (1980) *Mixed-mode demand forecasting techniques*. Transportation Planning and Technology, Vol. 6, 81-95.

ORTUZAR, J.DE D. (1996) Stated Preference data collection: from design to implementation. Proceedings Second ASC International Conference on Survey and Statistical computing, Imperial College, England.

ORTUZAR, J.DE D. (2000) *Stated preference Modelling Techniques*. PTRC Education & Research Services, London.

PEARMAIN, D AND SWANSON, J. (1990) *The use of stateted preference techniques in the quantitative analysis of travel behaviour*. Proceedings of the IMA Conference on Mathematics in Transport, University of Cardiff, September 1990, Wales.

PEARMAIN, D., SWASON, J., KROES, E AND BRADLEY, M. (1991) *Stated Preference Techniques: A Guide to Practice*. Steer Davis Gleave and Hague Consulting Group, London.

PUTMAN, S.H. (1973) *The interrelationships of transport development and land development*. Department of City and Regional Planning, University of Pennsylvania, Philadelphia.

PUTMAN, S.H. (1975a) *Calibrating a disaggregated residential allocation model DRAM: the residential sub-model of the Integrated Transportation and Land Use Model Package, ITLUP*. Presented at The Eighth Annual Conference of The British Section of the RSA, London.

PUTMAN, S.H. (1975b) *Further results from, and prospects for future research with, the Integrated Transportation and Land Use Model Package, ITLUP*. Presented at The Annual Conference of the Southern RSA, Atlanta.

SENIOR, M.L., WILLIAMS, H.C.W.L. (1977) *Model-based transport policy assessment*. Traffic Engineering and Control, Vol. 18, 402-406.

SMITH, T.E. (1978) *A general efficiency principle of spatial interaction*. In: Karlqvist, A., Lundqvist, L., Snickars, F. and Weibull, J. (Eds.), Spatial Interaction Theory and Planning Models. North-Holland, Amsterdam.

SNICKARS, F., WEIBULL, J. (1977) *A minimum information principle*. Regional Science and Urban Economics, Vol. 7, 137-168.

SWANSON, J., PEARMAIN, D. AND LOUGHEAD, K. (1992) *Stated preference sample sizes. Proceedings 20th PTRC Summer Annual Meeting,* University of Manchester Institute of Science and Technology, London.

TRAVEL MODEL IMPROVEMENT PROGRAM. (1997) *Land Use Compendium, DOT-T-99-03*.
http://www.bts.gov/tmip/papers/landuse/compendium/toc.htm

U.S. EPA. (2000) *Projecting Land-Use Change: A Summary of Models for Assessing the Effects of Community Growth and Change on Land-Use Patterns*. EPA/600/R-00/098. U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH.

VON THUNEN, J.H. (1826) *Der Isolierte Statt in Beziehung auf Landwirtschaft und National okonomie*. Hamburg. Translation, 1966. Von Thunen's Isolated State. Pergamon Press, London.

WARDROP, J.G. (1952) *Some theoretical aspects of road traffic research. In: Proceedings of the Institute of Civil Engineers,* Part II, vol. 1, 325-378.

WEGENER, M. (1985) *The Dortmund housing market model. A Monte Carlo simulation of a regional housing market*. In: Stahl, K., (Ed.), Microeconomic Models of Housing Markets. Springer, Berlin.

WEGENER, M. (1986) *Integrated forecasting models of urban and regional systems*. London Papers in Regional Science, 15.

WEGENER, M. (1994) *Operational urban models- state of the art*. Journal of the American Planning Association, Vol. 60, 17-29.

WEGENER, M. (1998) *Applied models of urban land use, transport and environment: state of the art and future developments*. In: Lundqvist, L., Mattsson, L.-G., Kim, T.J. (Eds.), Network Infrastructure and the Urban Environment. Springer, Berlin.

WEGENER, M. AND FRURST, F. (1999) *Land-use transport interaction: State of the art*. TRANSLAND Integration of Transport and Land Use Planning, Work package 2 Deliverable D2a. Institute of Spatial Planning, University of Dortmund.

WILLIAMS, H.C.W.L. (1977a) *Generation of consistent travel demand models and user-benefit measures*. In: Bonsal, P., Dalvi, W., Hills, P. (Eds.), Urban Transportation Planning. Abacus Press, London.

WILLIAMS, H.C.W.L. (1977b) *On the formulation of travel demand models and economic evaluation measures of user benefit*. Environment and Planning A, Vol. 9, 285-344.

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

WINGO, L. (1961) *Transportation and Urban Land*. Johns Hopkins Univ. Press, Baltimore, MD.