RP AND SP CAR OWNERSHIP MODEL IN THE CONTEXT OF DEVELOPING COUNTRIES

M. Kumar, K.V.Krishna Rao
Transportation Systems Eng., Civil Eng. Dept., Indian Inst. of Tech. Bombay, Powai, Mumbai-400 076, India
p2kumar@civil.iitb.ac.in, kvkrao@civil.iitb.ac.in

Abstract

This paper discusses the development of a car ownership decision model, which explains the decisions of households regarding their car ownership levels. The main aim of this study is to develop a disaggregate car ownership model for a case city using Revealed Preference (RP) data, Stated Preference (SP) data and Joint RP &SP data. The home interview survey data of Thane city of Mumbai Metropolitan Region (MMR), Maharashtra State, India and the work place based SP survey data of MMR is utilized for the development of car ownership model. A number of variables representing household socio-economic characteristics were included for RP model. Travel time, travel cost, projected income, car maintenance cost and car loan options are the attributes considered in constructing the SP experiment. The stated preference experiment was administered at selected places in MMR. The structure of the SP car ownership model is discussed. A multinomial logit framework is used for developing the RP car ownership model taking the household as the decision unit. The specification and the results of calibration of RP car ownership model developed are discussed. The model developed in this study exhibits satisfactory goodness-of-fit and prediction success rate. It is concluded that disaggregate modelling approach can be successfully used for modeling the car ownership decisions of households and the developed models will be useful for the transportation planners, decision makers and researchers for better transportation planning of cities in developing countries.

Keywords: Car ownership; Disaggregate modelling; Stated Preference; Revealed Preference
Topic Area: G8 Urban Transport in Developing Countries

1. Introduction

Car ownership growth has provided the momentum for the development of the transport sector for the past three decades. Car ownership modelling has received considerable attention because of the important role it plays in the overall transportation and land use planning process. Transportation planners and traffic engineers frequently require information on car ownership and utilization patterns for planning and designing the transportation infrastructure facility. It is a well-established fact that income is a primary impetus to car ownership. Increased incomes during the second half of the 20^{th} century have allowed quantum increase in car ownership in all industrialized countries, and more recently in many parts of the developing countries in the world (Dargay and Gately 1997). The car ownership is usually modeled as function of household characteristics, socio-economic variables and /or public transport services (Ben-Akiva et al. 1981; Jansson 1989; Pendyala et al. 1995). It is now well recognized that car ownership is one of the key determinants of the activity-travel behavior of individuals and households. Car ownership forecasts can be obtained using aggregate extrapolation models which model car ownership directly at the aggregate level (such as zonal, regional level) or using disaggregate car ownership models which use the household as the decision making unit and obtain zonal, regional level forecasts by aggregating over households. The disaggregate car ownership models are structurally more behavioral compared to aggregate models and are better able to capture the causal relationships and thus have become the preferred approach to model car ownership. The development of car ownership model involves collection of revealed preference (RP) and stated preference (SP) data, assumptions on the distribution of error terms of utilities, calibration of the utility functions and prediction. Many studies conducted in developed countries have proved the successful application of choice models for car ownership. In the context of developing countries, however, the applicability of disaggregate choice models for car ownership decision framework has not been explored fully.

In this context, a study has been taken up to explore the applicability of discrete choice models in arriving at a realistic decision framework for car ownership decisions in the cities of developing countries. The main aim of the study is to develop a disaggregate car ownership model for a case city using *RP* and *SP* data. The RP data set, which is mainly the home interview survey data, belongs to Thane city of Mumbai Metropolitan Region (MMR) of Maharashtra state, India. A suitably designed SP experiment was executed to synthesize the car ownership behaviour of households in MMR. The attributes considered are travel time, travel cost, type of car and car loan payment option, projected family income and servicing cost. Individuals were asked to choose the option for their car ownership level on a 5-point rating scale. In addition to the individual *RP* and *SP* models of car ownership, a joint *RP-SP* model is being developed to exploit the advantages of each type of data while mitigating the weaknesses.

A brief review of car-ownership models is provided in the next section. The details of RP car ownership model structure, data and methodology, model specification and results of calibration, model prediction successes are presented next. The subsequent part of the paper explains the structure of SP car ownership model, methodology, the SP experiment design, execution of SP experiment; SP survey experience and results. The conclusions and discussions of the RP and SP car ownership models are presented at the end.

2. Review of car-ownership models

Motorization describes the transition to higher levels of ownership and usage on the road. Over the last four decades, many studies have attempted to model the trends of motorization in various countries and periods. The models were developed mainly for understanding the determinants of motor vehicle ownership and use, and forecasting the future levels of vehicle ownership and use. With few exceptions (Hensher et al.1990; Pindyck 1979; Button et al.1993), however, most of these models focused on vehicle ownership. Good surveys of these studies can be found in Glaister (1981) and Ortuzar and Willumsen (1994). More thorough discussions of the motor vehicle ownership-forecasting problem can be found in Button, et al. (1982), Mogridge (1983), and Train (1986). Several families of models have been developed over the years. Among the earliest were the time-series extrapolation models that attempted to extrapolate the national or regional motor vehicle ownership trends to future years under an explicit assumption of a saturation level (Tanner 1974; Mogridge 1967, 1989). A popular functional form is the S-shaped logistic equation. The second family of models is the household-based car ownership model (Quarmby and Bates 1970; Bates et al. 1978), which uses disaggregate data to explain household car ownership behavior (instead of looking at the general trends). This family of models typically uses a logit model specification to estimate the probabilities of owning 0, 1, and 2 or more cars. The third family of models includes a host of econometric models that use demographic variables, income variables, and transport system variables to explain the cross-section variations or changes over time in motor vehicle ownership. Depending on the types of data used, these models can be further divided into three groups: (1) time-series models (Sweeney 1978; Chin and Smith 1997); (2) cross-section models which produce long-run income elasticity estimates (Beesley and Kain 1964; Silberston 1970; Wheaton 1982; Kain 1983; Kain and Liu 1994); and (3) panel data models (Pindyck 1979). This summary suggests that the income elasticities from time series data are typically smaller than those from cross sectional data.

In recent disaggregate car ownership models Bhat and Pulugururta (1998) pointed out that auto ownership modelling play an important role in travel demand analysis because it is a key determinant of the travel behavior of individuals and households. Discrete choice car ownership models use either an ordered-response choice mechanism or an unordered-response choice mechanism. This comparative analysis offers strong evidence that the appropriate choice mechanism is the unordered-response structure. Car ownership modelling must be pursued using the unordered-response class of models (such as the multinomial logit or probit model) and not using the ordered-response class of models (such as the ordered-response logit or probit). Brownstone and Train (1999) described and applied choice models, including generalizations of logit called 'mixed logits,' that do not exhibit the restrictive 'independence from irrelevant alternatives' property and can approximate any substitution pattern. The models are estimated on data from a stated-preference survey that elicited customers' preferences among gas, electric, methanol, and compressed natural gas (*CNG)* vehicles with various attributes. Brownstone et al (2000) compared multinomial logit and mixed logit models for data on California households' revealed and stated preferences for car ownership.

3. Model structure

Despite the development of a large number of aggregate automobile ownership models, many authors have worked with disaggregate models that are structurally more stable compared with aggregate models and are better able to capture the causal relationship between car ownership determinants and ownership levels. As a result, disaggregate methods have become a very popular approach to modeling car ownership levels. Further, within the class of disaggregate choice models, two general decision mechanisms have been used for car ownership modelling: the ordered-response mechanism and the unordered response mechanism. The majority of disaggregate models have used the unordered response mechanism to model ownership. These models generally fit into the random utility maximization (RUM) framework (Domenich and McFadden 1975; and Ben-Akiva and Lerman 1985). As per this framework, the utility of any alternative (car ownership) can be expressed as

$$U_{in} = V_{in} + \epsilon_{in} \qquad \textbf{(1)}$$

Where,

U_{in} = the true utility of household (or individual) n for car ownership level i

V_{in} = a deterministic component of utility and a function of exogenous variables, and it can be written as

$$V_{in} = \alpha_i + \beta X_{in} \qquad \textbf{(2)}$$

Where, α_i = constant specific to the alternative i, β = vector of parameters to be estimated, and X_{in} = vector of attributes for the individual n and the alternative i

ϵ_{in} = a random component / error term

The assumption that the error terms are independently and identically distributed as a *weibull* distribution and the application of RUM framework results in the well known multinomial logit (MNL) model. The model used in the present work is an MNL model of the form,

$$P_n(i) = exp\,(V_{in}\,)\,/\,\sum_j exp\,(V_{jn}) \qquad \textbf{(3)}$$

where, $P_n(i)$ = probability of household n choosing car ownership level i.

Using the RUM principle to estimate car ownership may be appealing especially because it does not place any restrictions on the effect of household characteristics across ownership levels. This means, for example, that it is possible for the effect of income to be highly negative for the utility of zero car ownership, positive for one and two car ownership, and zero for ownership of three or more cars.

4. Data-set and methodology for RP car ownership model

The data set for the study was mainly derived from the data of home-interview survey conducted as a part of Preparation of Detailed Project Report on the proposed Mass Rapid Transit System (MRTS) for Thane city of Mumbai Metropolitan Region (MMR), Maharashtra floated by the Maharashtra State Road Development Corporation (MSRDC). The city of Thane is located about *19 km* in the interior from main seacoast and on the northern extremity of Greater Mumbai. The physical development of Thane is circular, as the city has grown around the Central Business District (*CBD*) in the western area adjoining the Thane Railway Station. The Central Railway's main north-south corridors pass through the city. The Thane Municipal Corporation (TMC) is spread over an area of around $128.23\ km^2$ and a number of industrial estates are located in the region. The area covered under TMC was taken as the study area. It contained a population of *0.80* million in *1991* and has increased to 1.3 million as per 2001 census. The category wise motor vehicle population for the years *1996 to 1999 is* shown in Table 1. The maximum growth rate of *26.44 %* is observed for auto rickshaws and *7.17%* for cars during *1995-1999.* The car ownership was found to grow at a faster rate since the last 5 years.

Table 1. Category Wise Motor Vehicles in Thane City

Vehicle Type	**Number of Motor Vehicles**					**AAGR***
	1995	**1996**	**1997**	**1998**	**1999**	
Two-wheelers	44843	45834	48941	51018	54485	4.99 %
Three-wheelers (pass)	7657	9762	11091	12312	19571	26.44 %
Three-wheelers (goods)	2132	2215	2491	2913	3209	10.76 %
Cars& Jeeps	14209	14464	15176	16285	18746	7.17 %
Buses	1560	1560	1560	1560	1560	0.00 %
Trucks	13075	13368	14031	14709	16204	5.51 %
Tractors	198	199	201	201	202	0.50 %
Trailers	483	191	496	510	519	1.81 %
Others	389	389	389	389	389	0.00 %
Total Vehicles	84546	88282	94376	99897	114885	7.97 %

*Average Annual Growth Rate

The home interview survey data contained socioeconomic and travel information on 3500 households. This constitutes a sample size of 1.5 per cent. The *RP* data set for the development of car ownership model was extracted from this home interview survey data after applying logical checks and ensuring the availability of complete information on all the attributes. This resulted in *924* valid samples, which were used for the calibration of the car ownership model in this study. There are *641* households without any car, *241* households with one car, and *41* households with two or more cars in this valid sample. Accordingly, the choice set considered for the disaggregate car-ownership model contained three choices namely, *0 car*, *1 car* and *2 or more cars*. The methodology adopted for the development of the revealed preference car ownership model is presented as a flow chart in Figure 1

4.1 Specification and calibration of the MNL model for RP car ownership

The choice of variables for the potential inclusion in the car ownership *RP* model was guided by previous theoretical and empirical work on car ownership modelling, intuitive arguments regarding the effects of exogenous variable, statistical significance of the variables and data availability. A number of variables representing household socio-economic characteristics were included. The complete list of variables considered in the present study is listed in Table 2. The description of these variables and their categorization is also presented in Table 2. The variables considered in this study are built-up area, house ownership level, family size, number of males, number of females, number of working adults, number of non-working adults, number of school going children, number of business persons, number of service persons, number of people engaged in agriculture and labourers, number of persons of age more than forty, number of persons of age equival to or less than forty, number of retired persons, household travel expenditure, car licence holders in house hold, number of persons education more than SSC and number of persons education equal to or less than SSC.

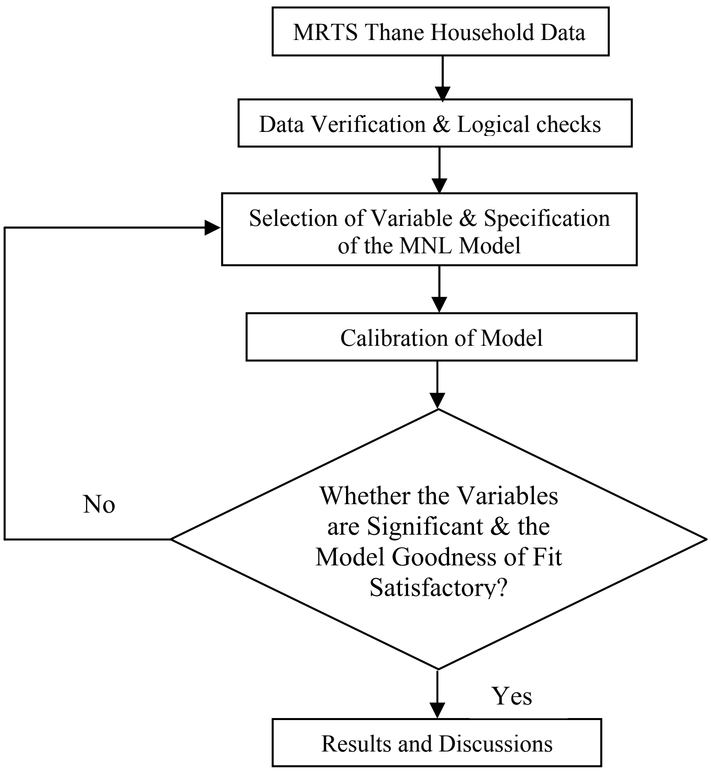

Figure 1. Flow chart of RP car ownership model methodology

With these variables, several specifications of the simple MNL model as in Equation (3) were tried before selecting the final model. As a starting point all the household and socioeconomic variables were used along with the alternative-specific constants in defining the utility of different car ownership levels. The attention was given to the use of different alternative-specific variables in utility function of different car ownership level.

The variables were eliminated if found statistically not-significant or having illogical signs. The variables so eliminated were then used in the utility function of other car ownership level and again the same checks were made.

Factors considered in selecting the specification were overall goodness of fit measures, the significance of variables entered, and multicollinearity. The variables built-up area, number of people with education more than *SSC*, family size, household income, number of car license holders and number of business persons were found to be significant at *95%* confidence level and having logical signs for one car ownership utility. The variables household income, household level and family size were found to be significant at *95%* confidence level and having logical signs for two car ownership level. Whereas the variables built-up area and number of business persons were found to be significant at *90%* confidence level and having logical signs for two car ownership level. The other variables under consideration got eliminated during the process. The parameters of the logit models specified were estimated using maximum likelihood method. The software, ALOGIT was used for this purpose. The goodness-of-fit of the models was assessed with the help of the likelihood-ratio index. The results of calibration in terms of coefficient estimates and statistical validation of the *RP* model are presented in Table 2.

Table 2. Variable Definitions

Variable Code	**Description of variable**
HOL	House ownership level: categorical variable classified as own house 1, rented 2, Govt. quarter 3, company quarter 4.
BA	Built up area: categorical variable classified as less than 250sq.ft – 1; 250 to 500 – 2; 501 to 750 – 3; 751 to 1000 – 4; and more than 1000 – 5.
NM	Number of males
NF	Number of females
FS	Family size
NWA	Number of working adults in the household
NNWA	Number of non-working adults in the household
NSGC	Number of school going children in the household
NPEMSSC	Number of persons education more than SSC
NPELESSC	Number of persons education less than and equivalent SSC
NCLH	Number of car license holder in household
HHINC	Household Income: categorical variable classified as up to Rupees 2000 – 1; 2001 to 5000 – 2; 5001 to 10000 – 3; 10001 to 15000 – 4; 15001 to 20000 – 5; 20001 to 30000 – 6; 30001 to 40000 – 7; and more than 40000 – 8.
HHTE	Household travel expenditure: categorical variable & its classification is same as that of HHINC.
NBPHH	Number of business persons in household
NSPHH	Number of service persons in household
NRPHH	Number of retired persons in household
NALPHH	Number of agricultural and labour persons in household
NPAGMF	Number of persons age more than fourty
NPAGLEF	Number of persons age less than and equivalent fourty

4.2 Tests for prediction success

Though the models developed have a reasonable goodness-of-fit when measured by standard statistical tests, it is also desirable to test the performance of the models in predicting the observed choice behaviour. A prediction success table can be used for testing this condition. Prediction success tables and indices were proposed originally by McFadden (1979). Prediction success table is a cross classification between the observed and predicted choice of alternatives. The probability of choice of each alternative is estimated for each observation in the data set using the calibrated utility functions. The alternative with the highest probability is considered the chosen alternative, and this is then compared with the alternative actually chosen. The prediction success table is then obtained by cross tabulating these predicted and observed values. The prediction success table for the car ownership model developed is given in Table 4. The tabulated data indicate that the predictions are better in the case of zero car and one car ownership. In the case of zero ownership level, *88.61%* choices were predicted correctly and in the case of one car ownership level, *84.23 %* choices were predicted correctly. As the number of samples with two and more than two-car ownership level is very less, the prediction success in this ownership level is not satisfactory.

Table 3. Coefficient Estimates and Goodness-of-Fit Statistics of RP Model

Variable	Coefficient	Standard Error	t-value	Relevance of variables
BA	0.5746	0.4170	4.4	Specific to 1 car
NPEMSSC	0.2376	0.0805	3.0	Specific to 1 car
NCLH	1.4960	0.1950	7.7	Specific to 1 car
HHINC	0.5057	0.0645	7.8	Specific to 1 car
FS	-0.2870	0.0764	-3.8	Specific to 1 car
NBPHH	0.2262	0.1490	1.5*	Specific to 1 car
HHINC	0.8367	0.1050	7.9	Specific to 2 car
BA	0.2738	0.2070	1.3*	Specific to 2 car
HOL	-1.7410	0.5750	-3.0	Specific to 2 car
FS	-0.2802	0.1170	-2.4	Specific to 2 car
NBPHH	0.3427	0.1940	1.8	Specific to 2 car
0 car constant	4.4030	0.4170	10.6	Specific to 0 car
Structural Parameters				
L (0)	-1014.0191	-	-	-
L (c)	- 685.0074	-	-	-
L (θ)	- 426.2742	-	-	-
χ^2	1175.4898	-	-	-
ρ^2 (0)	0.5796	-	-	-
ρ^2 (c)	0.3777	-	-	-
Adjusted ρ^2	0.5687	-	-	-

*Not significant at 95 percent level but significant at 90 percent confidence level.

Table 4. Prediction Success Table

Car Ownership level	0 car	1 car	2 car	Observed	% Observed
0 car	568	73	0	641	69.44
1 car	38	203	0	241	26.11
2 car	6	35	0	41	4.44
Predicted	612	311	0	923	100
% Predicted	66.30	33.69	0	100	-
% Correctly predicted	88.61	84.23	-	-	-

5. SP data and methodology

A work place based Stated Preference survey of car ownership was conducted in Mumbai Metropolitan Region (MMR). The MMR covers an area of about 4355 square kilometers. Greater Mumbai a major part of MMR as per population, which is a well-developed city, covers only an area of 468 square kilometers. The Rest of MMR, which is in development stage, consists of five municipal corporations, fifteen municipal councils and a number of villages. With rapid urbanization and industrilisation during last three decades in MMR the population has increased from 9.9 million in 1981 to 17.7 million in 2001. MMR is well served by major rail and road networks. The Greater Mumbai (Mumbai city), the economic capital of India, was the launching point of modern transport in Asia. Mumbai generates abount 5% of India's GDP and contributes over one third of country's tax revenues. The city of Mumbai with its present population of over12 million generates about 11 million trips per day, with about 88% of the total trips catered by suburban railway and the public transport services provided by Brihn Mumbai Electricity and State Transport Corporation (BEST), with the average lead being 22.15 kilometers for rail and 4.67 kilometers for buses. BEST with its fleet strength of 3458 buses carries about 4.7 million passengers per day. Despite public transport system being the life line of Mumbai, the commuters are subjected to most severe over crowding in the world with 9 car rake carrying over 4000 passengers at 11 persons per square meter against normal capacity of about 1750 passengers. The Mumbai city has 63679 taxies and 101829 Auto rickshaws, which are used as intermediate public transport modes, as per 2002 statistics. The vehicle population in Mumbai city during the last 4 decades increased from 0.15 million in 1971 to 1.03 million in 2001. The population of Mumbai Metropolitan Region (MMR) grew at less than 3% per annum during 1991-2001 whereas the vehicles have grown at over 7% per annum contributing to over 50% of cars.

5.1 Stated preference survey

Stated preference techniques refer to a number of different approaches all of which use peoples' statements of how they would respond to different situations. The common feature of this particular set of stated preference techniques is their use of experimental design to construct a series of alternatives imaginary situations. Individuals are then asked to indicate how they would respond if these situations faced them in reality. This enables a wide range of situations to be investigated, which may not be measured when observations of actual behaviour are used. The statistical models would be constructed based on SP data, to measure the preferences of individuals or groups and provide a basis for forecasting behaviour. These SP methods have become popular in the estimation of human behaviour under hypothetical conditions. In the present context the respondents compare the attributes of car ownership decision like travel time, travel cost, waiting time, discomfort, projected income, car loan payment options and servicing cost of the car, etc. with those of the existing situation (mode) and state their preference. The stated preference method involves three basic aspects, viz., the design of stated preference questionnaire (experiment design), administering of the stated preference experiment, quality analysis of the responses, and model development and application.

5.2 Design of stated preference experiment

Most stated preference techniques are characterized by the use of experimental designs to construct hypothetical alternatives presented to respondents. The attributes that are used in this car ownership SP experiment are travel time, travel cost, projected family income, car loan payment option and servicing cost of car per annum. Based on the opinion of the experts it was found that these factors would play major role in taking a decision relating to car ownership. The experimental design presented in this study is known as a fully factorial design. This is because every possible of attribute levels is used. Initially the experiment is designed by taking 2 attributes (travel time and travel cost) at 1 level, 3 attributes (projected family income, car loan payment options and service cost of the car) at 3 levels. A full factorial design will yield 27 options. The experiment, however, is simplified by taking the two attributes, car loan payment and servicing cost, together as one attribute due to their dependency. The selection of attributes and levels identified in this experimental design is based on literature suggested by Kocur et al (1982), and the number of options is arrived at as per Kroes and Sheldon (1988). In the experimental design two options are dominating among other options but as per the literature at least one best/worst option should remain in the choice set, so that their logical or illogical positioning by each respondent provides some indication of the reliabilility of the responses. The SP experiment is designed as a rating experiment by constructing 9 options (1×1×3×3) by car ownership with different attribute levels. The basic structure of stated preference experiment used in the present study along with the rating scale is shown in Figure 2.

<table>
<tr><th colspan="2">Existing Work/Recreational Trip</th><th colspan="2">Car Ownership</th></tr>
<tr><td>Travel Time</td><td>Stated</td><td>Travel Time</td><td>Computed</td></tr>
<tr><td>Travel Cost</td><td>Stated</td><td>Travel Time</td><td>Computed</td></tr>
<tr><td>HH Income</td><td>Stated</td><td>Projected HH Income</td><td>3 levels</td></tr>
<tr><td>Level of Discomfort</td><td>Stated</td><td>Car Loan Payment</td><td>3 levels</td></tr>
<tr><td>Waiting Time</td><td>Stated</td><td>Servicing Cost</td><td>3 levels</td></tr>
</table>

<table>
<tr><th colspan="5">Choice Scale</th></tr>
<tr><td>Definitely Own a car</td><td>Probably Own a car</td><td>Can't Say</td><td>Definitely Stick to the Existing</td><td>Probably Stick to the Existing</td></tr>
</table>

Figure 2. Structure of Stated Preference Experiment

The information on socioeconomic characteristics and travel for work trip and recreation trip separately was collected during the survey and the required information has been transferred to the SP experiment sheet. The car ownership attributes travel time, travel cost is computed for each respondent at the time of the interview by knowing the origin and destination of his trip. The high way travel time and travel cost skims computed for 110 internal traffic Zones and the three external traffic zones viz., Thane, Kalyan and Navi Mumbai from a working transportation planning model (Mumbai Metro Study, 2003) were utilized for estimating the travel time and travel cost attributes. The travel cost was computed based on three different types of cars and their mileages per liter petrol. The three categories of cars considered in this study are Compact car (18 kms/liter petrol), Midsize car (12 kms/liter petrol) and Luxury car (10kms/liter petrol). The travel time and travel cost tables were prepared for travel by car between 10 potential residential areas and five selected work places. The potential work/industrial/business places indentified are Nariman point, Bandra-Kurla Complex, Andheri, Seepz and Thane.

The attribute, *projected husehold income* was considered at three levels (i.e., Rupees 20,000, Rupees 35,000 and Rupees 50,000). The projected income levels are worked out based on existing car owner's income levels and expert opinion. The attribute, car loan payment is considered at three levels, i.e., compact car (one time payment Rs. 56000, monthly Rs.5366), midsize car (one time payment Rs. 1 Lakh, monthly Rs.7537) and luxury car (one time payment Rs.2 Lakhs, monthly Rs.15074). The car loan payment options calculated based on present rates of bank interest on car loans, present car prices, monthly installments and margin money as shown clearly in Table 5. The attribute, servicing cost of the car per annum is considered at tree levels, i.e., Rs. 3000 (for compact car), Rs. 5000(for midsize car) and Rs. 6000 (for luxury car). These values are worked out based on the information obtained from car owners and servicing centers.

Table 5. Details of Car Loan Payment Options

Loan Details	Compact car	Mid-size car	Luxury car
Approximate cost	Rs. 2.25 Lakhs	Rs. 4 Lakhs	Rs. 8 Lakhs
Initial payment	Rs.56250	Rs. 1 Lakh	Rs. 2 Lakhs
Loan amount	Rs.168750	Rs. 3 Lakhs	Rs. 6 Lakhs
Number of installments	36	48	48
Monthly installment	Rs. 5366	Rs.7536	Rs. 15074
Rate of interest	9% per annum	9.5% per annum	Rs.9.5%
Mileage (liter petrol)	18 km	14 km	10 km

Each respondent was asked to rate 9 options (1 travel time x 1 travel cost x 3 projected income levels x 3 car loan payment levels and servicing cost levels) for work trip and 9 options for recreation/social/shopping trip on a rating scale. The respondents were told that the waiting time by car is zero and the discomfort level is 1 on a scale of 1 to 5. A typical SP option out of the 9 options floated for car ownership for work trip is shown in Figure 3. A similar SP experiment was designed for the Recreational/shopping/social trip. The respondent was asked to give the information on his/her recent family trip (maximum members of the family) for recreation/shopping/social and rate 9 Stated Preference options. The SP questionnaires were prepared to conduct the surveys at 5 work places. In addition a leaflet describing the purpose of car ownership study and showing the three types of cars and loan payment options was also prepared. The information on travel time, travel cost, HH income range, waiting time and level of discomfort for the existing trip and situation was collected and placed in the SP experiment for comparison. The level of discomfort was classified on a five-point scale (5 = standing in crowd in non air-conditioned environment; and 1= comfortable sitting in air-conditioned environment). The complete set of SP survey questionnaire contain the general description leaflet, socio-economic information sheet, work trip information sheet, recreation trip information sheet, SP experiments for work trip and recreation trip in one sheet, map of Greater Mumbai and MMR, travel time and travel cost (three types) tables.

5.2 SP sample size

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. The sample size requirement for stated preference experiments is very low because the data collected on each individual is used in modeling in a disaggregate manner using discrete choice theory and also because of the fact that each interviewee produces not just one observation but several on the same choice context. More recent works reported in the literature suggest that 75-100 interviews per segment would be more appropriate (Pearmain and Swanson 1990; Bradley and Kroes 1990; and Swanson et al. 1992). In the present study travelers are segmented based on their income groups like Rs.5-10 thousand, Rs.10-20 thousand, Rs. 20-30 thousand etc. It was attempted to satisfy the above sample size requirement.

Existing Work Trip		Car Ownership	
Travel Time	40 min	**Travel Time**	30 min
Travel Cost	Rs.75	**Travel Cost**	Rs.40
HH Income	Rs.25000	**Projected HH Income**	Rs.35000
Level of Discomfort	4 (Non A/C standing)	**Car Loan Payment**	One Time: Rs.1 Lakh, Monthly: Rs. 7537
Waiting Time	15 min	**Servicing Cost**	Rs. 5000
Choice:_________			

Definitely Own a car (1)	**Probably Own a car (2)**	**Can't Say (3)**	**Definitely Stick to the Existing (4)**	**Probably Stick to the Existing (5)**

Figure 3. A typical SP Option for Work Trip

5.3 Administering of SP experiment

A Team of about 8 enumerators was thoroughly trained for a week for administering the experiment on the respondents. The face-to-face work based pilot survey was conducted at Nariman Point before taking up main survey in order to arrive at a suitable survey instrument design for this study. The number of people contacted in the pilot survey was 175. The number of people satisfying the laid down criteria was 76 and those who expressed to participate in the SP interview were 25. Out of this number, 5 people discontinued in the half way. The minimum and maximum time consumed for each interview was 15 and 30 minutes respectively. Based on the experience gained in pilot survey few modifications were made to different parts of the questionnaire instruments for increasing the efficiency of survey. The main survey was administered with 8 thoroughly trained enumerators at work places, business centers and industrial areas. In majority of the cases the interviews were conducted by taking prior appointments from the concerned authorities. The enumerators would first explain the objective of the study with the help of the leaflet to the respondent and then collect his/her personal and trip information by filling the appropriate forms. The attributes travel time, travel cost, household income, discomfort level and waiting time by the existing travel obtained from trip information are then transferred to the appropriate place in the SP questionnaire. The travel time and travel cost by car is then worked out by using the travel time and travel cost tables. Each individual is then asked to compare the attributes of the existing travel pattern with those of car ownership option, and give his rating for car ownership option. The same procedure is repeated for recreational/social/shopping trip based on their recent recreational/shopping/social trip made along with maximum members of the family. The number of options that were obtained from each respondent was 18. Thus one valid sample will give 9 observations for work trip model calibration and 9 observations for recreational/shopping/social trip model calibration.

5.4 SP survey results and analysis

The number of people contacted and the final number of people who completed the SP experiment in each of the workplaces are given in Table 6. The SP survey data was cleaned thoroughly and errors were removed. The total valid samples obtained from the

SP survey are given in Table 7. Relating to the completeness of information in the SP survey sheets, 100 percent complete information in 65% samples, 90 percent information in 15% samples, 75% of information in 10% of samples, and 50% of information in the remaining 10% samples could be collected. Relating to the erroneous entries, 70% samples without any wrong entries, 10% samples with 15 percent wrong entries, 10 % samples with 30 percent wrong entries and the remaining 10% samples with more than 50 percent wrong entries were collected.

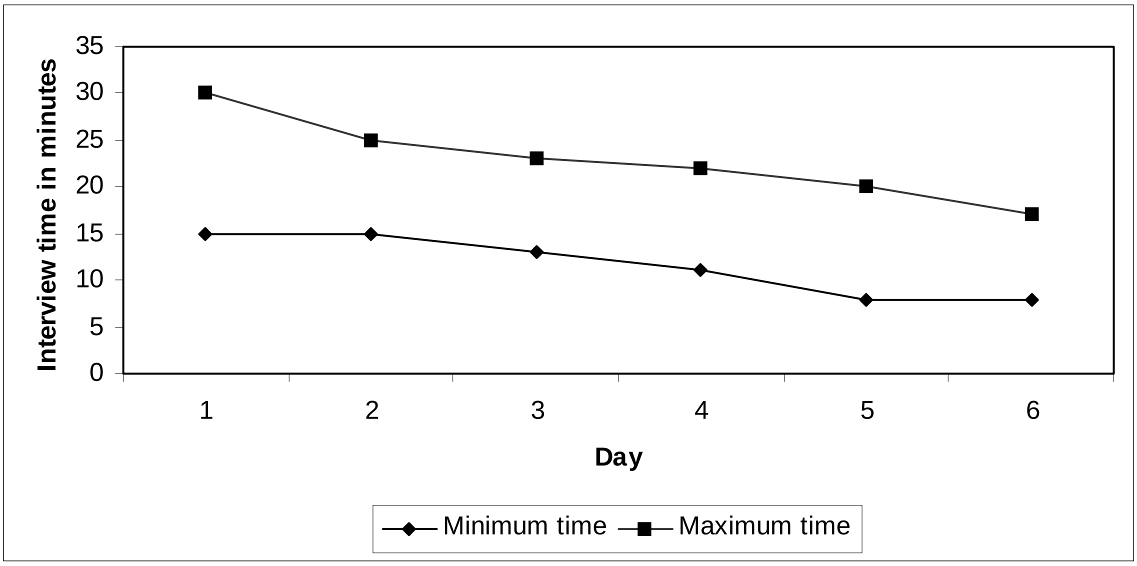

Figure 4. Variation in the duration of SP Interview from Day 1 to Day 6

The minimum and maximum durations of interview were 10 and 25 minutes, respectively, on the first day of the survey. Figure 4 shows how the duration of interview time reduced with the progress of the SP survey from first day to sixth day due to the increased efficiency of the enumerator in administering the SP experiment. Most of the respondents suggested that 10-15 minutes of interview time is acceptable to them. The respondents' understanding level of the SP experiment was found to be about 75 %.

Table 6. Details of SP Survey Efficiency at Different Locations

Location	Number of people contacted	Number satisfying the criteria	Number willing to participate	Number discontinued (half-way)	Number completed
Nariman Point*	523	243	85	12	73
B-K Complex	369	156	103	12	91
Andheri	401	195	88	13	75
SEEPZ	402	174	106	11	95
Thane	368	162	71	7	64
Total	2063	930	453	55	398

* including pilot survey samples

Significant number of respondents were not ready to disclose their home address and their household income. The non-response rate is found to be 42 %.

Table 7. Valid Samples obtained at Different Locations

Location	Total samples	Valid samples
Nariman Point	73	65
B-K Complex	91	82
Andheri	75	67
SEEPZ	95	85
Thane	64	58
Total	398	357

The number of times each choice option was chosen is shown as a bar chart in Figure 5. The frequency of response for the choice "definitely own a car" is 29%, and that for "probably own a car" is 14.44% for work trip; where the frequency of response for the choice "definitely own a car" is 39%, and that for "probably own a car" is 17% for recreational trip. Further, at different projected income levels the option "definitely own a car" was chosen 6.47%, 34.57% and 51.91% times respectively at Rs.20000, Rs.35000 and Rs.50000 income levels for work trip. Similarly for recreation trip, these percentages were 7.45%, 39.09% and 53.46% at Rs.20000, Rs.35000 and Rs.50000 respectively. It was observed from the data analysis that travelers are giving priority for recreational trip rather than work trip in owning a car. At these projected income levels the number of times each choice option was chosen for work trip and recreation/shopping/social trip is shown as a bar chart in Figures 6 & 7 respectively.

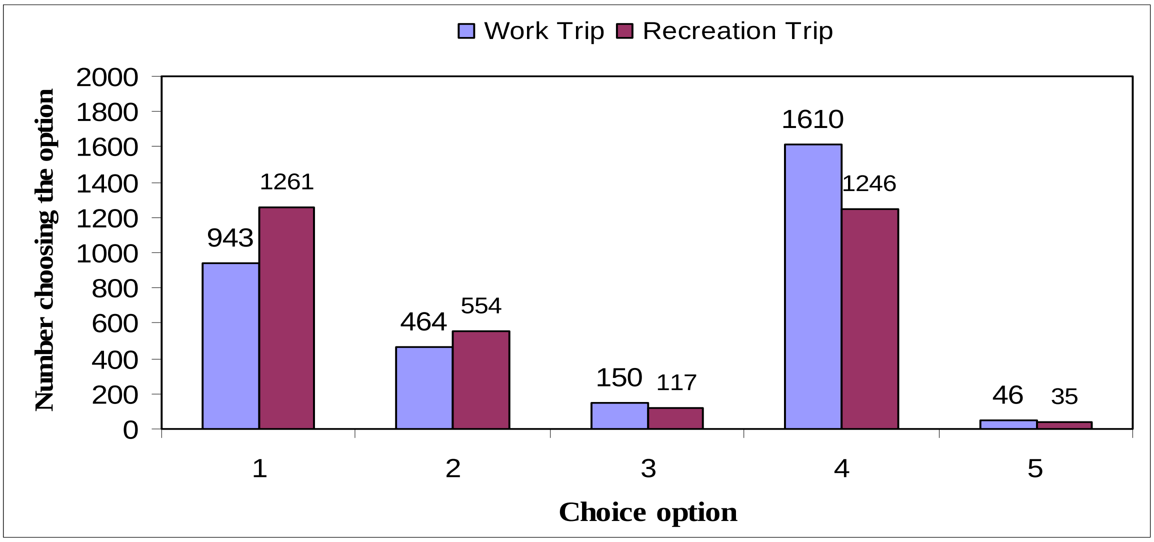

Figure 5. Number Choosing the Option for Work Trip and Recreation Trip

It was observed from Figure 6 that the frequency for the option "definitely own a car" is increasing considerably when the income levels are increasing and the similar phenomenon can be observed in Figure 7. The frequency of choice of the options ("definitely own a car" and "probably owning a car") for recreation /shopping/social trip were always more than that of work trip at all income levels. The behaviour of the travelers, however, is to be modeled by developing a car ownership model using the SP data collected.

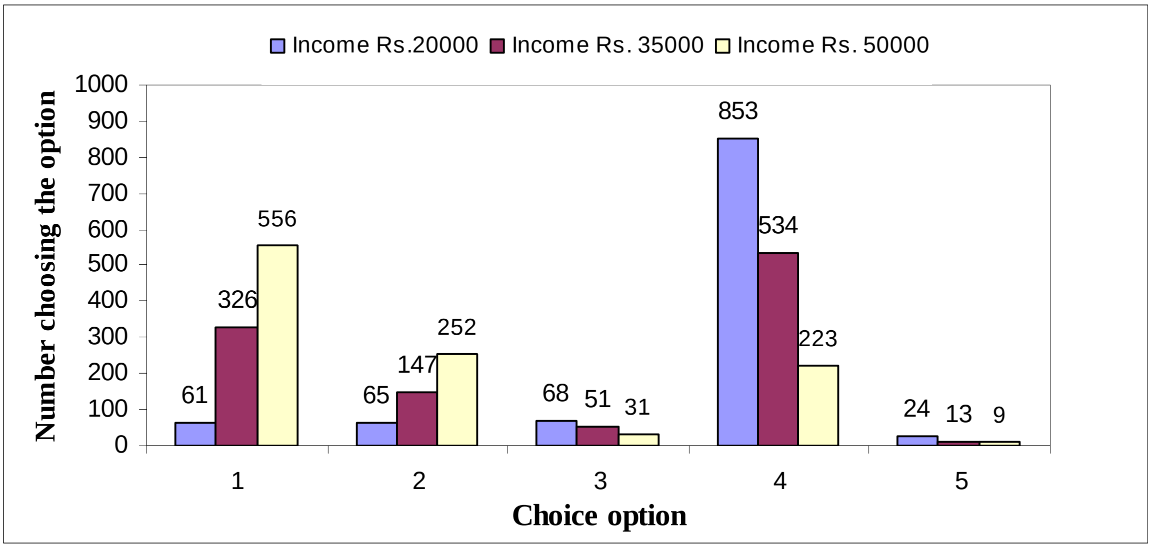

Figure 6. Number Choosing the Option at Different Income Levels for Work Trip

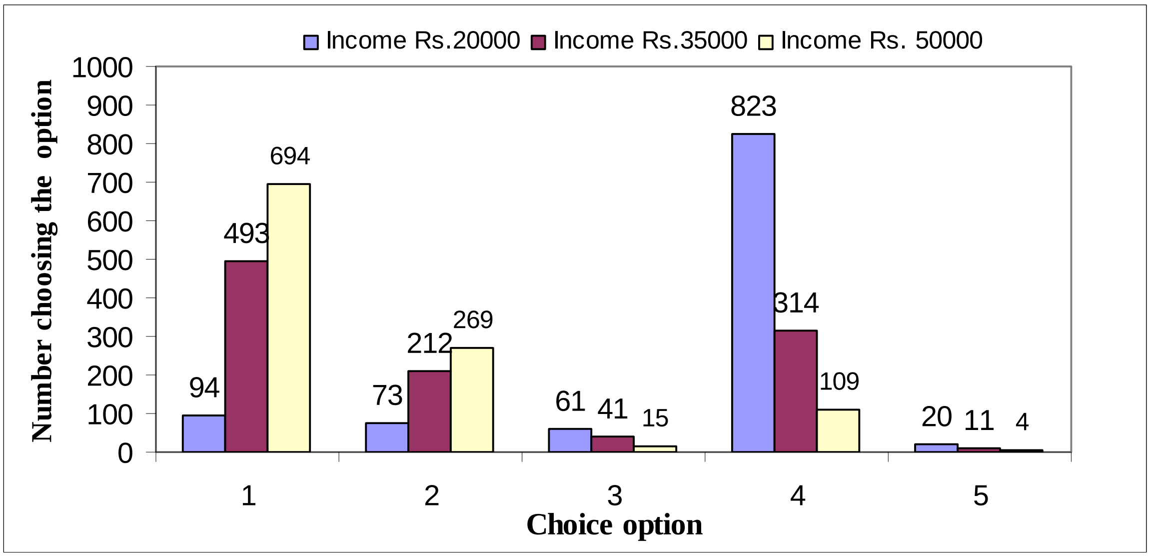

Figure 7. Number Choosing the Option at Different Income Levels for Recreation Trip

5.5 SP car ownership model

With suitable specification using the socioeconomic characteristics and attributes floated in the SP experiment, several specifications of a simple MNL model as in Equation (3) will be calibrated and tested before selecting the final SP car ownership model. The factors that will be considered in selecting the specification are overall goodness of fit measures, the significance of variables entered, and multicollinearity. The methodology adopted for the development of stated preference car ownership model shown in the Figure 8.

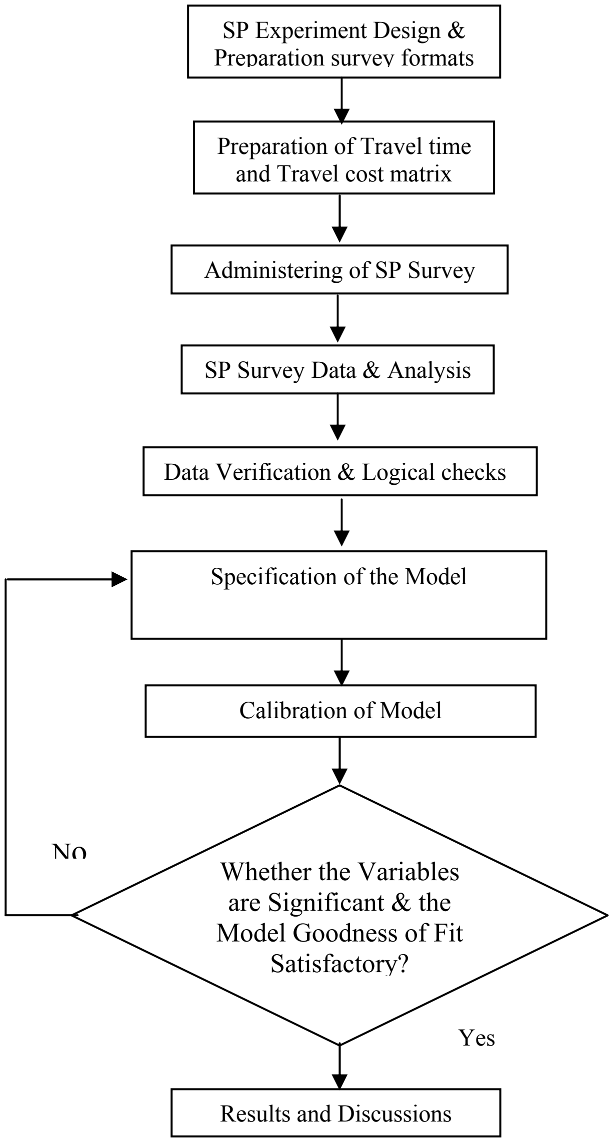

Figure 8. Flow chart of SP car ownership model methodology

6. RP and SP joint modelling

In addition to the individual RP and SP models, a joint RP-SP model will be developed to exploit the advantages of each type of data while mitigating the weaknesses. The usual approach to joining RP and SP data has been to test whether the utility of RP and SP choice models are proportional. That is, both data sets are pooled to test the hypothesis

$$H1: (\gamma_{rp}, \beta_{rp}) = \mu\, (\gamma_{rp}, \beta_{rp}),\ \mu > 0 \qquad (4)$$

Where γ and β are parameter vectors corresponding, respectively, to the alternative-specific constants and the design attributes, for the RP and SP data sets. In the multinomial logit (MNL) model, μ is the scale of the Gumbel distribution assumed to describe the stochastic component of the utility function (Ben-Akiva and Lerman 1985), and is inversely related to the unobserved error distribution of the RP data. Hypothesis H1 represents a very stringent requirement for the respective coefficients to satisfy: it demands not only that the response to the design variables be proportional in the data sources. The data set of RP original and SP work trip information consisted of 923(22.27

percent) and 3213 (77.72 percent), respectively. A joint RP and SP model will be developed utilizing both these data sets.

7. Conclusions

The RP car ownership model developed and calibrated in this study exhibits good results in terms of goodness-of-fit measures. The ρ^2 value was found to be *0.58*. In many trials this value varied between *0.50 and 0.59*. This value of ρ^2 could be considered as an indication of a good fit according to Hensher and Johnson (1981). The hypothesis related to nest structure could not be supported for car ownership model, as the value of the structural coefficient of the nested tree was not found significant. Therefore, only the results of MNL model had been presented. The log-likelihood ratio (χ^2-statistic) of the best model indicated that all the variables relating to household socioeconomic characteristics that entered the model were collectively significant at *95* percent confidence level in explaining the car-ownership levels. In the prediction success test, the model has predicted the observed ownership levels closely. A suitably designed SP experiment was executed and the responses of individuals to various alternative options were recorded on a 5-point rating scale. This SP experiment was designed to synthesize the car ownership behavior of households. The attributes considered were travel time, travel cost, type of car and loan payment options, and servicing cost of car and projected family income. A total of 398 samples were collected and after cleaning and thorough checking 357 valid samples were obtained. The work trip SP options of 3582 and recreational/shopping/social SP options of 3582 were obtained from these 357 valid samples. The non-response rate found from the eligible samples was 42%. The minimum and maximum time for the each interview decreased considerably from first day of the survey to the last day of the survey. The level of understanding of SP experiment by the respondent was found to be 75%. The frequency of response for the choice "definitely own a car" is 29%, and that for "probably own a car" is 14.44% for work trip; where the frequency of response for the choice "definitely own a car" is 39%, and that for "probably own a car" is 17% for recreational trip. Further, at different projected income levels the option "definitely own a car" was chosen 6.47%, 34.57% and 51.91% times respectively at Rs.20000, Rs.35000 and Rs.50000 income levels for work trip. Similarly for recreation trip, these percentages were 7.45%, 39.09% and 53.46% at Rs.20000, Rs.35000 and Rs.50000 respectively. It was observed from the data analysis that travelers are giving priority for recreational/shopping/social trip rather than work trip in owning a car. The behaviour of the travelers, however, is to be modeled by developing a car ownership model using the SP data collected. Based on the RP model developed in this study, it can be concluded that the model specification developed for the case city is acceptable. The results clearly indicate that the disaggregate modelling approach can be successfully used for modeling the car ownership decisions of households in the context of Developing countries also.

References

Bates, J.J., Gunn, H.F., Roberts, M., 1978. A Model of Household Car Ownership, Traffic Engineering and Control 19(11/12) 486-491& 562-566.

Beesley, M.E., Kain. J.F., 1964. Urban Form, Car Ownership and Public Policy: An Appraisal of Traffic in Towns, Urban Studies 1(2).

Ben-Akiva, M., Lerman, S. R., 1985. Discrete choice Analysis: theory and Application to Travel Demand. Cambridge, MA: MIT Press.

Ben-Akiva, M., Manski, C.F., Sherman, L. A., 1981. Behavioral approach to modeling household motor vehicle ownership and applications to aggregate policy analysis, Environment and Planning 13(A) 399-411.

Bhat, C.R., Pulugurta, V., 1998. A Comparison of Two Alternative Behavioral Choice Mechanisms for Household Auto Ownership Decisions, Transportation Research B 32(1) 61-75.

Bradley, M.A., Kroes, E., 1990. Forecasting issues in stated preference survey research. 69th TRB Annual Meeting, Washington DC, January, USA.

Brownstone, D., Train, K., 1999. Forecasting new product penetration with flexible substitution patterns, Journal of Econometrics 89, 109-129.

Brownstone, D., Bunch, D. S., Train, K., 2000. Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles, Transportation Research 34B, 315- 338.

Button, K., Ndoh, N., John H., 1993. Modeling Vehicle Ownership and Use in Low Income Countries, Journal of Transport Economics and Policy 27(1) January.

Button, K., Pearman, A.D., Fowkes, A.S., 1982. Car Ownership Modeling and Forecasting, Gower Publishing Company Ltd.

Chin, A., Smith, P., 1997. Automobile Ownership and Government Policy: the Economics of Singapore's Vehicle Quota Scheme, Transportation Research 31A (2) 129-140.

Dargay, J., Gately, D., 1997. The demand for transportation fuels: imperfect price reversibility?, Transportation Research 31B(1) 71-82.

Domenich, T.A., McFadden, D., 1975. Urban Travel Demand a Behavioral Analysis. North-Holland Publishing Company, Amsterdam, The Netherlands.

Glaister, S. 1981. Fundamentals of Transport Economics, New York: St. Martins Press.

Hensher, D. A., Johnson, L., 1981. Applied Discrete Choice Modelling, Croom-Helm, London, United Kingdom.

Hensher, D.A., Frank, W. M., Nariida, C. S., 1990. The Demand for Vehicle Use in the Urban Household Sector, Journal of Transp.Economics and Policy 17(3) 119-137.

Jansson, J. O., 1989. Car demand modeling and forecasting: a new approach, Journal of Transport Economics and Policy 23,125-140.

Kain, J. F., 1983. Impacts of Higher Petroleum Prices on Transportation Patterns and Urban Development. in Theodore E. Keeler (ed.). Research in Transportation Economics 1, JAI Press Inc.

Kain, J.F., Liu. Z., 1994. Efficiency and Locational Consequences of Government Transport Policies and Spending in Chile, Harvard Project on Urbanization in Chile, Harvard University.

Kocur, G. T., Adler, T., Hyman, W., Aunet, B., 1982. Guide to forecasting travel demand with direct utility assessment. Report No. UMTA-NH-11-0001-82, Urban Mass Transportation Administration, US Department of Transportation, Washington, DC.

Kros, E. P., Sheldon, R. J., 1988. Stated preference methods, Journal of transport Economics and Policy 22, 11-25.

McFadden, D., 1979. Quantitative models for analyzing travel behavior of individuals, in Behavioral Travel Modeling, Edited by D.Hensher and P.R. Stopher, Croom Helm, London.

Mogridge, M. J. H., 1967. The Prediction of Car Ownership. Journal of Transport Economics and Policy 1(1) January.

Mogridge, M. J. H., 1983. The Car Market, Pion, London.

Mogridge, M. J. H., 1989. The Prediction of Car Ownership and Use Revisited, Journal of Transport Economics and Policy 23(1) January.

Mumbai Metro., 2003. Master Plan for Mumbai Metro: A feasibility report, IIT Bombay, Mumbai, India.

Ortuzar, J.de D., Willumsen, L.G., 1994. Modelling Transport, John Wiley & Sons.

Pearman, D., Swanson, J., 1990. The use of stated preference techniques in the quantitative analysis of travel behaviour, Proceedings of the IMA Conference on Mathematics in Transport, University of Cardiff, September, Wales.

Pendyala, R.M., Kostyniuk, L.P., Goulias, K.G., 1995. A repeated cross-sectional evaluation of car ownership, Transportation 22, 165-184.

Pindyck, R., 1979. The Structure of World Energy Demand. Cambridge: MIT Press.

Quarmby, D. A., Bates, J. J., 1970. An econometric method of car ownership forecasting in discrete areas. MAU Note 219, Department of the Environment London.

Silberston, A., 1970. Automobile Use and the Standard of Living in East and West. Journal of Transport Economics and Policy 4(1) January.

Swanson, J., Pearmain, D., Loughead, K., 1992. Stated preference sample sizes. Proceedings 20th PTRC Summer Annual Meeting, University of Manchester Institute of Science and Technology, September, England.

Sweeney, J.L., 1978. Energy Policy and Automobile Use of Gasoline. Stanford University.

Tanner, J.C., 1974. Forecasts of Vehicles and Traffic in Great Britain: 1974 Revision, TRRL Report LR 650, Transport and Road Research Laboratory, Crowthorne, England.

Train, K.E., 1986. Qualitative Choice Analysis: Theory, Econometric and an Application to Automobile Demand, The MIT Press.

Wheaton, W. C., 1982. The Long-Run Structure of Transportation and Gasoline Demand, The Bell Journal of Economics 13.