MODELING HOUSEHOLD TRIP GENERATION FOR SELECTED ZONES AT AL-KARKH SIDE OF BAGHDAD CITY

Prof. Saad Issa Sarsam  
Civil Engineering Department  
University of Baghdad  

Su’ad Saba’ Al-Hassani  
Civil Engineering Department  
University of Baghdad

ABSTRACT

Trip generation is the first phase in the travel forecasting process. It involves the estimation of the total number of trips entering or leaving a parcel of land per time period (usually on a daily basis); as a function of the socioeconomic, locational, and land-use characteristics of the parcel. The objective of this study is to develop statistical models to predict trips production volumes for a proper target year. Non-motorized trips are considered in the modeling process. Traditional method to forecast the trip generation volume according to trip rate, based on family type is proposed in this study. Families are classified by three characteristics of population social class, income, and number of vehicle ownership. The study area is divided into 10 sectors. Each sector is subdivided into number of zones so; the total number of zones is 45 zones based on the administrative divisions. The trip rate for the family is determined by sampling. A questionnaire is designed and interviews are implemented for data collection from selected zones at Al-Karkh side of Baghdad city. Two techniques have been used, full interview and home questionnaire. The questionnaire forms are distributed in many institutes, intermediate, secondary and, commercial schools. The developed models are total person trips/household, work trips/household, education trips/household, shopping and social/recreational trips/household and, person trips/person. These models are developed by using stepwise regression technique after the collected data being fed to SPSS software.

Results show that total persons trips/household are related to family size and structure variables such as number of person more than 6 year age, number of male, total number of workers, total number of students in the household, number of private vehicles. This model has coefficient of determination equal to 0.669 for the whole study area. Also the results show that the home-based work trips are related to number of worker in the household, number of male workers in the household, number of female workers in the household and number of persons of (25-60) year age; this model has coefficient of determination equal to 0.82 for the whole study area. Home-based education trips are strongly related to number of students in the household and this model has coefficient of determination equal to 0.90 for the whole study area.
Key Word: - Household, Modeling, Trip Generation, Travel Demand.
INTRODUCTION

Urban transportation planning is the process which represents the integration of many interacting characteristics of the urban environment (Brain et al, 1974). It is carried out to develop a program of high way and transit projects that should be completed in the future. As motorization is progressed, it has become increasingly evident that the traffic congestion “problem” cannot be solved by building more road infrastructure.

Forecast of demand for the system at the various levels of facility provision considered is one of the basic elements of long-term transportation planning. It is consist of the following steps (Wright, 1998):

1. Population and economic analysis.
2. Land use analysis.
3. Trip generation.
4. Trip distribution.
5. Mode choice (or model split).

Population and economic activity predictions generally are derived by specialist demographic and economic projections outside the scope of the transportation study.

There has been considerable debate on the land use. A number of models have been developed that relate changes in land use to such independent variables as: accessibility to employment, land value, intensity of land use …etc (Wright, 1998).

The population of Baghdad urban area according to the previous census is given in Table 1, whereas Fig.1 is illustrate the land uses for Baghdad city.

The four remaining steps are of considerable interest and the whole four steps are refer to as (Traditional Method) for prediction travel demand (Venigalla et al, 1999).

Trip generation involves the estimation of the total number of trips entering or leaving a parcel of land per time period (usually on a daily basis); as a function of the socioeconomic, locational, and land-use characteristics of the parcel. Usually multi-mode and purpose journeys are simplified into trip identified by a principal mode and purpose, ignoring intermediate stages and stops for secondary purposes (Hobbs, 1979). The reliability of forecasting results influences the following steps such as trip distribution, mode split, and traffic assignment. Therefore, improved trip generation models are needed to improve forecasting precision (Liya and Hongzahi, 2008).


<table>
<thead>
<tr>
<th>Census year</th>
<th>population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1947</td>
<td>666000</td>
</tr>
<tr>
<td>1957</td>
<td>785000</td>
</tr>
<tr>
<td>1977</td>
<td>2023242</td>
</tr>
<tr>
<td>1987</td>
<td>3841269</td>
</tr>
<tr>
<td>1997</td>
<td>4083258</td>
</tr>
</tbody>
</table>

The ultimate purpose of the trip generation analysis is to arrive at an estimate of the trip ends generated at each analysis unit of the study area. Traditionally, trip generation forecasts are established independently of any direct consideration of the transportation network. This assumes that trips produced at or attracted to a zone are a function only of the attributes of the zone itself and are not directly a function of the transportation network on which the trips are made (NTL, 2006).

The study of trip generation attempts to identify and quantify the trip ends related to various urban activities without describing other trip characteristics such as direction, length or duration. Usually, the interest is in trips per average weekday, but may be for weekend or special purpose travel (FHWA, 1975).
Methodology

The models start with defining the study area and dividing them into a number of zones and considering the entire transport network in the system. The database also includes the current (base year) levels of population, economic activity like employment, shopping space, educational, and leisure facilities of each zone. The following steps are convenient to be followed:

- Selection and Zoning the study area.
- Design of the questionnaire form.
- Sample sizes calculation and selection.
- Data collection and processing.

1.0 Selection and Zoning the study area

The study area is bounded between the left bank of Tigris River and the external west boundaries of Baghdad city. The population of Al-Karkh urban area reached to 1,010,408 inhabitants in 1977 census and 1,365,057 inhabitants in 1987 census and 1,797,557 inhabitants in 1997 census (Census Results, 1977), (Census Results, 1987), (Census Results, 1997).

Because of the mass of data collected relates to individual journeys, households and, centers of employment, and in its crude form, is difficult to analyze and interpret, the area being surveyed is divided into zones for the purpose of grouping the data so as to make it intelligible, amenable to analysis (Bruton, 1985).

The zoning procedure requires the division of the survey area into spatial units suitable as a basis for data collection, and later for aggregate analysis purposes. In certain cases, it may be desirable to set up two complete zoning patterns. The first of these would be the zoning pattern to be used in data collection, and would divided the area up to spatial unit best suited to be treated as individual element in sampling process. The
second set of zones would comprise the zones to be used for analysis purposes (Stopher, 1977). In large study area it is more convenient to divide the study area into sectors. The zoning system used in this analysis is based on the administration divisions shown in Fig.2. Study area zoning is shown in Fig.3

2.0 Design of the Questionnaire Form

Traditional household survey has five sections; dwelling unit information, household characteristics, car ownership characteristics, personal characteristics and, trip details (Mathew and Rao, 2007).

Dwelling unit characteristics; this section includes a set of questions designed to obtain information about the dwelling unit. Information includes:
- Dwelling unit type
- Dwelling unit ownership
- Availability of garage
- The Household characteristics; include
  - Average monthly income
  - Number of persons in the household (stratified into male and female)

Trip data; this part of the survey aims at detecting and characterizing all trips made by the household members include:
- Number of trips
- Purpose of each trip
- Mode of trips
- Origins and destinations of trips
- Time of starting and ending of trips
- Cost of trips
- Purposes of trips considered here are work, education (school or college), shopping, social/recreational, change mode and, return to
- Number of workers in the household (stratified into male and female).
- Number of students in the household (stratified into male and female).
- Number of persons less than school age
- Number of persons in each age group
- Car ownership characteristics; include:
  - Number of available cars
  - Type of each car
  - Nature of use
  - Often, cars are classified according to type and size (BCTS, 1980), therefore cars are classified into private, taxi, bus, minibus, truck, pickup and, others.
- Personal characteristics; this part includes questions designed to classify the household members (older than 6 year) according to the following aspects:
  - Gender, Age
  - Profession
  - Winning or official
  - Educational level
  - Number of residence year
  - Location of the work or education
  - Time waiting the transportation mode in minute home. It may necessary to combine some of these purposes at modeling stage (BCTS, survey report, August 1980).
- Travel modes used are:
  - Private (salon, taxi, bus, truck).
  - Government bus.
  - Taxi.
  - Mini bus.
  - Bus.
  - Motor cycle
  - Bicycle.
  - Walking.
Fig. 2 Council Municipals of Baghdad City

Fig. 3 Study Area Zoning
Table 2 Sample Size for Home Interview Survey

<table>
<thead>
<tr>
<th>Study Area Population</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
</tr>
<tr>
<td>Under 50 000</td>
<td>1 in 10</td>
</tr>
<tr>
<td>50 000 –150 000</td>
<td>1 in 20</td>
</tr>
<tr>
<td>150 000 – 300 000</td>
<td>1 in 35</td>
</tr>
<tr>
<td>300 000 – 500 000</td>
<td>1 in 50</td>
</tr>
<tr>
<td>500 000 – 1 million</td>
<td>1 in 70</td>
</tr>
<tr>
<td>Over 1 million</td>
<td>1 in 100</td>
</tr>
</tbody>
</table>

Table 3 Executed Samples

<table>
<thead>
<tr>
<th>No. of sector</th>
<th>Executed size</th>
<th>Sample Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>451</td>
<td>0.112300797</td>
</tr>
<tr>
<td>2</td>
<td>116</td>
<td>0.028884462</td>
</tr>
<tr>
<td>3</td>
<td>519</td>
<td>0.129233068</td>
</tr>
<tr>
<td>4</td>
<td>451</td>
<td>0.112300797</td>
</tr>
<tr>
<td>5</td>
<td>479</td>
<td>0.119272908</td>
</tr>
<tr>
<td>6</td>
<td>490</td>
<td>0.122011952</td>
</tr>
<tr>
<td>7</td>
<td>225</td>
<td>0.056025896</td>
</tr>
<tr>
<td>8</td>
<td>550</td>
<td>0.136952191</td>
</tr>
<tr>
<td>9</td>
<td>551</td>
<td>0.137201195</td>
</tr>
<tr>
<td>10</td>
<td>300</td>
<td>0.074701195</td>
</tr>
<tr>
<td>Total</td>
<td>4016</td>
<td>1</td>
</tr>
</tbody>
</table>

3.0 Sample Sizes Calculation and Selection

The methods used to calculate the sample size for this analysis are:
1. The National Committee on Urban Transportation (NCUT) Method: The method is recommended for conducting home interviews, sample sizes is given in the table above.
2. Smith formula method: smith formula specify the use of coefficient of variation in determining the sample size as shown in the following formula:

\[ n = \frac{C.V^2 Z^2}{E^2} \]  

Where C.V = coefficient of variation (For a given variables),
E = accuracy level,
Z = normal variants, and
n = number of samples.

Table 3 shows the executed sample for each sector in the study area.

4.0 Data Collection and Processing

The collection and processing of the Household Interview Survey data will follow a set of carefully considered and approved methods and procedures. To develop accurate trip production models, a large number of data is needed. Most of the necessary data are needed to relate the amount and character of urban activity. Typically, these data come from large-scale of household travel surveys and inventories of existing land-use characteristics. Census data are also a valuable source of information for transportation planners (Wallace and Shadoff, 1998) The methods used in home interview survey are:

a-Full interview technique: Involves interviewing as many members of the household as possible in directly recording all the information.
b-Home questionnaire technique: The interviewer collects details of the household characteristics leaving forms for household residents to be completed, regarding travel information. The completed forms are collected by the interviewer after day or two (Safa-Eldeen, 2006). This study uses two methods as shown in Table 4.
Table 4 The Distribution of Types of Survey Methods

<table>
<thead>
<tr>
<th>Sector No.</th>
<th>Full Interview</th>
<th>Questionnaire</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
<td>No.</td>
</tr>
<tr>
<td>1</td>
<td>187</td>
<td>41.46</td>
<td>264</td>
</tr>
<tr>
<td>2</td>
<td>31</td>
<td>36</td>
<td>85</td>
</tr>
<tr>
<td>3</td>
<td>193</td>
<td>37.18</td>
<td>326</td>
</tr>
<tr>
<td>4</td>
<td>161</td>
<td>33.48</td>
<td>300</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>20.87</td>
<td>379</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>490</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td>230</td>
</tr>
<tr>
<td>8</td>
<td>110</td>
<td>19.09</td>
<td>440</td>
</tr>
<tr>
<td>9</td>
<td>112</td>
<td>18.32</td>
<td>400</td>
</tr>
<tr>
<td>10</td>
<td>44</td>
<td>14.62</td>
<td>206</td>
</tr>
<tr>
<td>Total</td>
<td>820</td>
<td>21.93</td>
<td>3196</td>
</tr>
</tbody>
</table>

Development of Trip Production Models
The dependent and independent variables for this analysis are as follow:

**Dependent Variables**
Y = Total Persons trips/HH
Y1 = Home-Base-Work trips
Y2 = Home-Base-Education trips
Y3 = Home-Base-Social/Recreational & Shopping trips
YP = Person trips /person

**Independent Variables**

1. **Dwelling Unit Characteristics and Income**
X101 = Dwelling unit type (House=1, other=0)
X102 = Dwelling unit Ownership (Owned=1, Rented=0)
X103 = Availability of Garage (Yes=1, No=0)
X104 = Household mean income in (ID 1000)

2. **Household Size and Structure**
X201 = Number of Males
X202 = Number of Females
X203 = Household size
X204 = Number of Male workers in the Household
X205 = Number of Female workers in the Household
X206 = Total Household workers
X207 = Number of Male students in the Household
X208 = Number of Female students in the Household
X209 = Total Household students

3. **Household by Age Groups**
X301 = No. of Person Less than (6) years
X302 = No. of Person (6 – 18) years
X303 = No. of Person (19 – 24) years
X304 = No. of Person (25 – 60) years
X305 = No. of Person more than (60) years
X306 = No. of Person more than (6) years

4. **Vehicle Ownership**
X401 = No. of Available Vehicles
X402 = No. of Private Vehicles
X403 = No. of Taxi Vehicles
X404 = No. of Other Vehicles

5. **Person Characteristics**
X501 = Age
X502 = Gender (Male =1, Female =0)
X503 = Education Level “Educated” (Yes=1, No=0)
X504 = Education Level “School” (Yes=1, No=0)
If the correlation coefficient between two variables is high, then one variable has the larger correlation coefficient with the dependent variable should be included in the models but not both the variables. For example, by inspecting Table 5 the correlation coefficient between number of female \(X_{202}\) and number of female student per household\(X_{208}\) is 0.606 and the correlation between total trips/HH and \(X_{202}\) and \(X_{208}\) are 0.507 and 0.434 respectability therefore only \(X_{202}\) should be included in the model.

**Error Tests**

The distribution of error must be normal. This assumption can be tested by normal probability plot. If 95% of points lies in the range (-2,2) then the errors are normally distributed. Fig. 5 shows the 45 line plot of errors for whole data of the study area. It can be noted that the points lies between the range of \((0,1)\) and errors are normally distributed.
Table 5 Partial Correlation Matrix Controlling for Y

<table>
<thead>
<tr>
<th>variables</th>
<th>X104</th>
<th>X202</th>
<th>X205</th>
<th>X208</th>
<th>X301</th>
<th>X303</th>
<th>X304</th>
<th>X306</th>
<th>X401</th>
<th>X402</th>
<th>X403</th>
</tr>
</thead>
<tbody>
<tr>
<td>X104</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X202</td>
<td>-0.071</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X205</td>
<td>0.303</td>
<td>0.008</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X208</td>
<td>-0.161</td>
<td>0.606</td>
<td>-0.124</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X301</td>
<td>0.006</td>
<td>0.519</td>
<td>-0.064</td>
<td>0.258</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X303</td>
<td>-0.024</td>
<td>0.324</td>
<td>-0.089</td>
<td>-0.105</td>
<td>-0.023</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X304</td>
<td>0.139</td>
<td>0.125</td>
<td>0.064</td>
<td>-0.079</td>
<td>0.362</td>
<td>-0.091</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X306</td>
<td>0.167</td>
<td>0.184</td>
<td>0.077</td>
<td>-0.188</td>
<td>0.345</td>
<td>0.158</td>
<td>0.015</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X401</td>
<td>0.284</td>
<td>-0.020</td>
<td>0.087</td>
<td>-0.058</td>
<td>0.024</td>
<td>0.029</td>
<td>0.069</td>
<td>0.127</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X402</td>
<td>0.280</td>
<td>-0.019</td>
<td>0.104</td>
<td>-0.016</td>
<td>-0.008</td>
<td>-0.004</td>
<td>0.052</td>
<td>0.097</td>
<td>0.839</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>X403</td>
<td>-0.019</td>
<td>0.010</td>
<td>-0.022</td>
<td>-0.040</td>
<td>0.032</td>
<td>0.073</td>
<td>0.029</td>
<td>0.009</td>
<td>0.134</td>
<td>-0.126</td>
<td>1</td>
</tr>
</tbody>
</table>

Dependent Variable: total trips /HH

Fig. 5 45 line Plot of Errors

Models Estimation
Many methods of linear regression are available, they are:
Enter, Stepwise, Remove, Backward, Forward
The best and commonly method used to determine parameter of prediction model is
stepwise method (Abed, 2010). This method computes the simple regression model for each independent variable. The independent variable that has the largest F-statistic is chosen as the first entering variables. If at least one variable exceeds the standard, the procedure continues. The procedure considers whether the model will be improved by adding a second independent variable and so on. It examines all variables to determine which has the F-statistic which suite the selected F-statistic to enter criteria (Al-Zaidy, 2005). Either F value or probability of F are used as enter criteria. Probability of F equal to 0.05 is used in the analysis, this correspond to 3.84 F value.

Table 6 shows the resulted models with coefficient of determination $R^2$ for whole data of the study area.

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Models</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y$</td>
<td>$3.019+1.362X_{206}+0.758X_{209}+0.493X_{206}+0.356X_{201}+0.243X_{402}$</td>
<td>0.669</td>
</tr>
<tr>
<td>$Y1$</td>
<td>$0.182+0.478X_{206}+1.194X_{204}+1.26X_{205}+0.08X_{304}+0.035X_{303}+0.07X_{305}$</td>
<td>0.82</td>
</tr>
<tr>
<td>$Y2$</td>
<td>$0.378+1.792X_{209}$</td>
<td>0.900</td>
</tr>
<tr>
<td>$Y3$</td>
<td>$2.789+0.13X_{303}+0.243X_{201}–0.533X_{205}+0.326X_{103}+0.23X_{303}+0.362X_{304}$</td>
<td>0.089</td>
</tr>
<tr>
<td>$Yp$</td>
<td>$0.43+1.034X_{502}+0.283X_{505}+1.297X_{514}+1.141X_{513}+1.128X_{511}$</td>
<td>0.478</td>
</tr>
</tbody>
</table>

**Conclusions**

Major conclusions found in the trip production analysis are that the first model ($Y$) is the preferred from the rest models because of the preferred statistical results and the simplicity of the application. On the other hand, purposes trip models ($Y1$, $Y2$, and $Y3$) cannot be use in application because all purposes trips models must be applicable to use. Work and education trip models are good from the statistical point of view, while shopping and social/recreational trip models are poor from the statistical point of view. Whereas person trip models ($Yp$) are weak. It may be due to the non linear relationship between $Yp$ and person’s variables.

Other conclusions found are:

1. Most of the household variables have statistically significant relationships to each trip purpose, apparently due to multiple collinearities between a household's numbers of person’s workers, and vehicles.
2. In general, the strongest single predictor of trips in a given purpose is the number of persons eligible or most expected to make such trips. For work trips, this is the number of employed persons; for education trips, it is the number of school-age children and persons in a college age group. For shop and other trips, this is simply the number of persons counted as trip makers such as number of male, number of female workers per household.
3. The most effective independent variables on total trips/HH are household size variables, from these variables the most notable effect is number of person more than 6 year. The most effective independent variables on person trips/person are gender and modes of travel.
4. In many models the constant and income parameters have (t) value less than 2 and hence these parameters are insignificant. Therefore, the constant and income are eliminating from these models.

**References**


Brain; Memmott; and Bone, "Principles And Technique of Predicting Future Demand for Urban Area Transportation", M.I.T. Report No. 3, 1974.
Bruton, M. J. , "Introduction to transportation planning", Hutchinson, London, 1985

Baghdad Comprehensive Transportation Study (B.C.T.S), Scott Wilson Kirkpatrick & Partners, Amanat-Al-Assima "Survey Report , August 1980".

Federal Highway Administration , "Trip Generation Analysis " , U.S. Department of Transportation, Washington, D.C. August 1975


Liya Y.; Hongzahi G ; and Hai Y. , "Trip Generation Model Based on Destination Attractiveness " , TSINGHUA Science And Technology ISSN 1007 -0214 , Vol.13 , No.5, October 20


Quade, P.B. and Douglas, " Travel Demand Model Development and Application Guidelines ", Oregon Department Of Transportation , Planning Section, Transportation Planning Analysis Unit , September 1994


