CHAPTER IV
TRIP GENERATION[1/]

BACKGROUND

Trip generation provides the relationship between urban activity and travel and may be separated into two phases. In the first phase, an understanding and quantification of the relationship between present urban activity and present tripmaking is developed; and in the second, the results of the quantification are applied to forecasts of urban activities to develop future travel estimates.

For trip generation purposes, travel is viewed only in terms of trip ends and does not consider other characteristics such as direction and length of duration. The trips considered are usually those generated for an average weekday, but they may also be for weekend travel, for a particular trip purpose, by mode of travel or other stratification required for a specific analysis or forecasting purpose. Trip ends may be in terms of origins and destinations or in terms of productions and attractions, depending on the purpose of the forecast and the subsequent models to be used for trip distribution and modal usage.

The goal of trip generation model development is to establish a functional relationship between travel and the land use and socio-economic characteristics of the units to and from which the travel is made. A causal relationship is desired in which questions such as the following are answered:

- What is the difference in tripmaking between a family living in a high rise apartment close to the central business district and a similar family living in a single family home in the suburbs?
- What is the difference in trips to a shopping center serving a suburban area as compared to a shopping center of similar size and nature located in a central business area?

Questions of the above nature can be considered in terms of the intensity and character of the urban activities.

1/ Detailed documentation of trip generation procedures may be found in the following publications: Federal Highway Administration, Trip Generation Analysis, U.S. Government Printing Office, Washington, D.C., August 1975 and U.S. Bureau of Public Roads, Guidelines for Trip Generation Analysis, U.S. Government Printing Office, Washington, D.C., June 1967.

Intensity is the amount of activity to be found in a given areal unit (i.e. zone) and is usually stated in terms of a density measure such as the number of employees per square foot of floor area or per acre of a particular land use category, or the number of dwelling units per acre. For example, the number of trips per dwelling unit generally decreases as the number of dwelling units per residential acre increases. This is logical because dense residential developments, like high rise apartment buildings, are usually within walking distance of many services, thus alleviating the need for a vehicular trip. When residential density is low (perhaps less than 10 dwelling units per acre), trip rates are higher because most trips must be made by vehicle.

Urban activity intensity measures are usually not sufficient in themselves for trip generation relationship development. There is additional variation in travel that is accounted for by variables related to the "character" of the activity. For residential activities, character is expressed in socioeconomic terms such as family income and car ownership. With all other conditions the same, families with higher incomes generally own more automobiles and make more trips. Low income families often own no cars, rely on public transportation and walking, and thus exhibit low vehicular tripmaking potential. The higher tripmaking families usually show increases in shopping and social-recreational trips, with trips for work remaining relatively stable.

For nonresidential land uses, character reflects the type of activity--for example, industrial, retail, and commercial. The number of trips generated by a major shopping center is usually higher than the number of trips generated by a warehouse of the same size.

It is difficult to separate the individual effects of intensity and character. Each factor explains some of the variation in tripmaking, and variables reflecting these factors are used in trip generation model development regardless of the type of analysis used.

PROCEDURES FOR TRIP GENERATION

There are three basic procedures for trip generation analysis:

- Cross-classification
- Land activity rates
- Multiple linear regression

Cross-classification is a technique in which the change in one variable--trips--can be measured when the changes in other variables--land use and socioeconomic--are accounted for. Cross-classification is not heavily dependent upon assumed distributions of the underlying data and, as such, is sometimes referred to as a "nonparametric" or distribution free technique. Basically, the technique stratifies "n" independent variables into two or more appropriate groups, creating an n-dimensional matrix. Observations of the dependent variable are then allocated to the cells of the matrix, based on values of the several independent variables, and then averaged.

The land activity rate approach is based upon the development of rates in which trips are related to land use characteristics reflecting the character, location and intensity of land use. The method may also be considered a type of cross-classification analysis.

The multiple linear regression process consists of developing equations in which trips or a trip rate (i.e., trips/household) is related to independent variables which explain the variation in the dependent variables (trips or trip rate). The equations are usually developed by trip purpose and generally are based on data aggregated to the zone level as observations. Although regression is a linear technique fitting straight lines through data, transformations of variables into log functions, taking reciprocals, etc., can be made resulting in curvilinear representations.

The important statistics used in evaluating the equations developed include: (1) the multiple correlation coefficient which indicates the degree of association between the independent and dependent variables in the equation; and (2) the standard error of estimate which indicates the degree of variation on the data about the regression line established. Details of regression analysis can be found in "Guidelines for Trip Generation Analysis," FHWA, June 1967.

Regression analysis has been an important tool in trip generation analysis. A wealth of understanding of travel has resulted from application of the technique and most transportation studies undertaken in the 1960's relied on the technique. However, based upon the regression analysis of the past and current work using cross-classification and rate analysis, it appears that more efficient and straightforward trip generation procedures can now be recommended.

RECOMMENDED APPROACH TO TRIP GENERATION

The recommended approach to trip generation consists of the following components:

- Production models -- cross-classification.
- Attraction models -- trip rates based on activity units
- External trip models -- regional growth and proportions of internal trip factors
- Truck or taxi models -- separate nonhomebased attraction model or group with other nonhomebased trips.

The recommended approach is based upon the considerable amount of research and application in the area of trip generation over the last fifteen years. It is believed that this past work provides the basis for the presentation of a simple, efficient approach. The approach allows incorporation of policy sensitive factors and at the same time allows development and application in a relatively short time period and at a lesser cost than previously applied methods.

Development of Trip Production Models

In developing trip production models, several basic decisions need to be made, such as the independent variables, the trip type, and the trip purposes to be used in the analysis.

It is suggested that the variables to be used for the trip production models be income and car ownership, and that the trip type be either person trips or auto driver trips, depending on the approach to transit planning. For areas where there is minor transit use currently and no appreciable growth is expected, the trip type may be auto driver trips. The location of "0-car" households could then be utilized to define areas of high transit potential.

Car ownership is highly correlated with income, but is included in the process for several reasons--the usefulness of car ownership for modal usage models, the need to consider auto ownership saturation levels in the planning process, and the high elasticity in vehicle purchasing and travel with respect to income.

In addition to car ownership and income, most areas should consider another independent variable--density. This variable is recommended because there is a trade-off between walking and vehicular trips when density increases.

The trip purpose stratifications are usually dictated by the trip distribution and modal usage models being developed. For internal area travel, the choice of purpose will vary somewhat by size of area. The larger studies will usually consider five purposes--homebased work, homebased shop, homebased school, homebased other, and non-homebased. In smaller urban areas, three trip purposes--homebased work, homebased other, and non-homebased--are generally considered.

After these initial decisions are made, the actual trip production model development may begin. The trip production model consists of a sequence of four submodels, each developed using the origin-destination travel survey, supplemented in some cases by other sources, such as the census.

The first submodel, the income submodel, reflects the distribution of households within various income categories (e.g., high, medium, and low). This is done by relating the average zonal income to household income as shown in Table IV-1.

Table IV-1. Matrix for Income Submodel

		Average Zone Income ($)				
		0-1999	2-3999	4-5999	6-9999	etc.
Household Income Group	Low		(Percent of households in			
	Medium		each income group)			
	High					

The values in the above table may then be plotted as illustrated in Figure IV-1, with smooth curves drawn through the data to eliminate any illogical relationships.

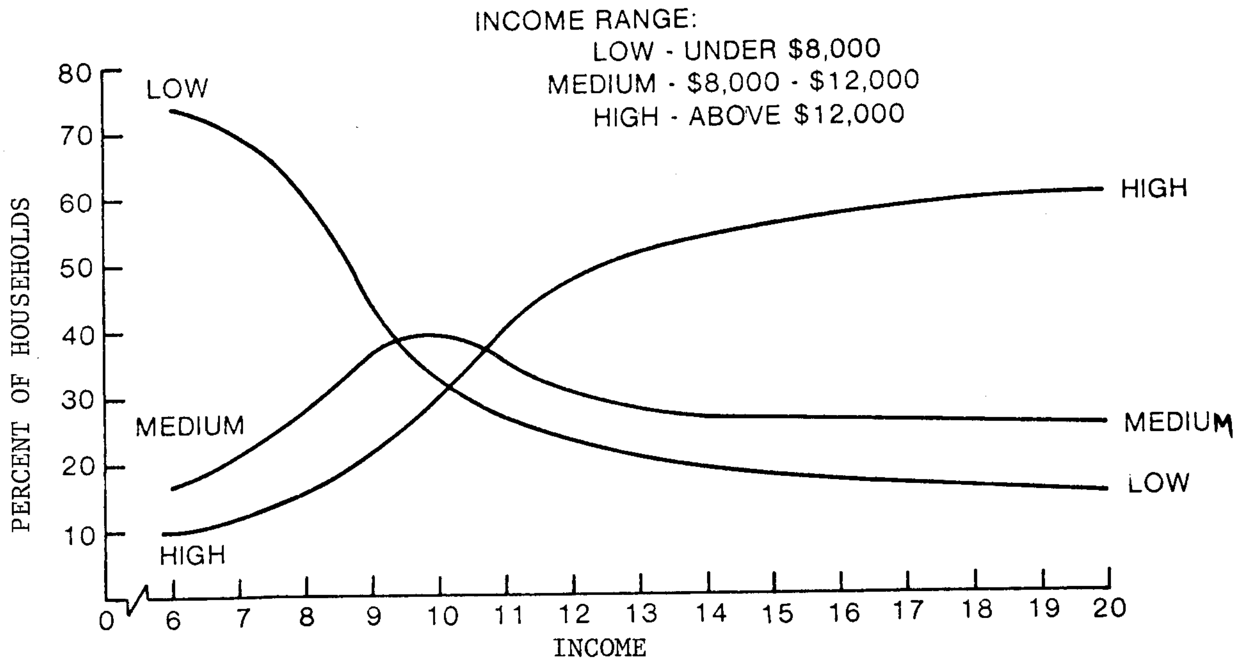

Figure IV-1. Example of Income Submodel

A good source of information for developing the income distribution submodel is the 1970 Census of Population and Housing PHC(1)-171 series.

The second submodel--the auto ownership submodel--relates the household income to auto ownership. The procedure used to develop this submodel is described in the following example.

Assume that there are twenty households in a sample and income and car ownership are being used as the independent variables. Then, for each household, information is available on the number of trips, income and car ownership as typically obtained from the home interview survey. Table IV-2 reflects the sample household data obtained from the survey.

TABLE IV-2
Example of Household Data for Cross Classification

Household	Trips	Income	Cars
1	2	4000	0
2	4	6000	0
3	10	17000	2
4	5	11000	0
5	5	4500	1
6	15	17000	3
7	7	9500	1
8	4	9000	0
9	6	7000	1
10	13	19000	3
11	8	18000	1
12	6	8000	1
13	9	7000	2
14	11	11000	2
15	10	11000	2
16	11	13000	2
17	12	15000	2
18	8	11000	1
19	8	13000	1
20	6	7000	1

From this information, a matrix is established based upon car ownership and income. Table IV-3 illustrates the matrix that results from stratifying the sample data by car ownership and income. The numbers in the various cells of the matrix represent the number of households.

TABLE IV-3
Matrix for Auto Ownership Submodel

Income ($000's)	Cars Owned 0	1	2 or More
< 6	2(67%)	1(33%)	(0%)
6-9	1(25%)	3(50%)	1(25)
9-12	1(20%)	2(40%)	2(40%)
12-15		1(33%)	2(67)
> 15		1(25%)	3(75%)

The above example contains a very small number of households in each income and car ownership category and is intended only to illustrate the development of the auto ownership submodel. In actual practice, there should be at least twenty-five observations in each cell.

The percent of households in Table IV-3 may then be plotted and the results fit with smooth curves which may be extended out past the data points based on the shape of the curves. Figure IV-2 illustrates the curves developed from the data in Table IV-3.

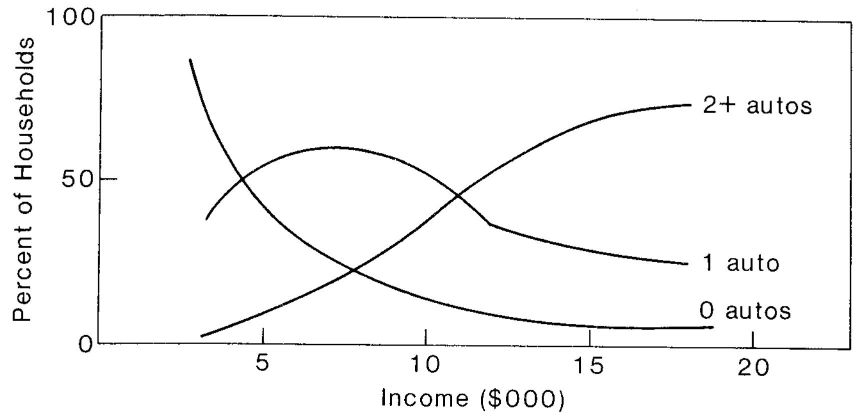

Figure IV-2. Example of Auto Ownership Submodel

From these curves, adjusted values for the number of households in each income and auto ownership category are determined and used as input in the third submodel--the trip production submodel.

The trip production submodel establishes a relationship between the trips made by each household and the independent variables. Using the trip data in Table IV-2 (number of trips for each household), as well as the adjusted number of households in each auto ownership and income category, the trip rate in each category may be determined. The trip rate is simply the number of trips in each auto ownership and income category divided by the appropriate number of households in each auto ownership and income category. Table IV-4 summarizes the trip rates resulting from this example.

TABLE IV-4

Matrix for Trip Production Submodel
(Trips/Household)

Income	Cars Owned 0	1	2+
< 6	3	5	0
6-9	4	6	9
9-12	5	7.5	10.5
12-15	0	8.5	11.5
15+	0	8.5	13.3

Again, the data from the matrix is fit with smooth curves (Figure IV-3), and the curve values are used to develop an adjusted matrix.

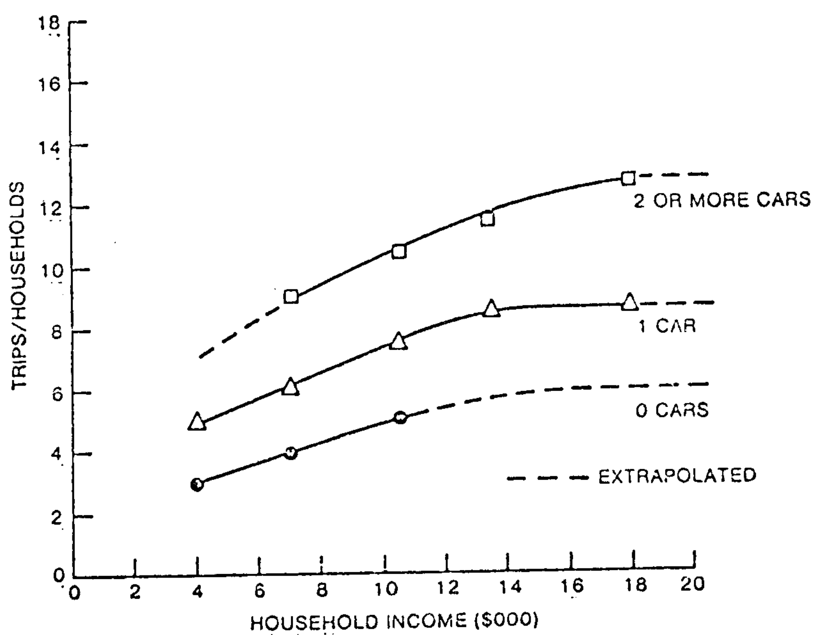

Figure IV-3. Example of Trip Production Submodel

The final submodel to be developed relates the trip purpose to income in such a manner that the trip productions can be divided among the various purposes. The trip purpose submodel is developed by accumulating the number of survey trips for each purpose within each income group and finding the percentage of the total trips each purpose constitutes within each income category. The distribution is then plotted as shown in Figure IV-4.

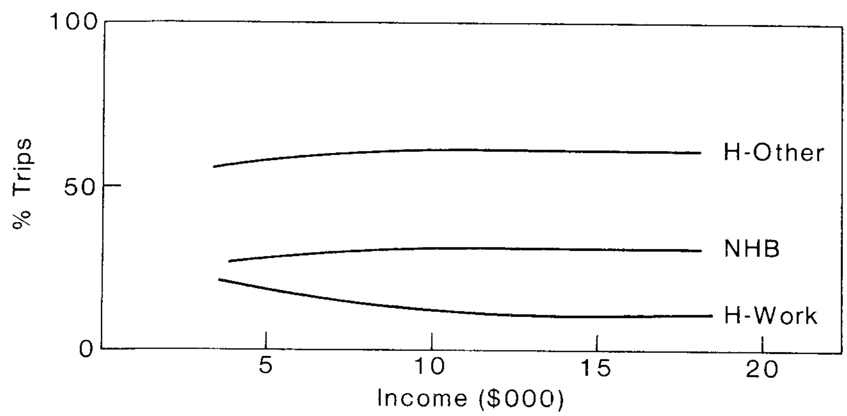

Figure IV-4. Example of Trip Purpose Submodel

It should be noted that the various submodels should be developed using an expanded sample, which simply means multiplying the survey data by the inverse of the sampling rate. By developing the submodels on the expanded survey data, much of the bias inherent in the survey data is eliminated.

Development of Trip Attraction Models

Basically, trip attraction is related to nonresidential land use for most trip purposes. For example, homebased shopping trip attractions are generally to locations where goods are sold. There may be a few shopping trips to residential land, but an amount not worth considering. On the other hand, homebased other trips usually include trips for social-recreational purposes which by nature include travel to residential land. Likewise, nonhomebased attractions would have some residential association.

There are several ways to handle trip attractions, including zonal regression, land area trip rates or cross-classification. A simplified approach is suggested based on the development of trip rates within a matrix. The matrix below is an example of that approach.

Table IV-5. Example of Procedure for Developing Trip Attraction Models

Trip Purpose	Trips per Household	Trips per Employee				Trip per Student		
		Non-Retail	Retail			Univ.	High School	Other
			CBD	Shop Center	Other			
Home Based Work	—	1.70	1.70	1.70	1.70	—	—	—
Home Based Shop	—	—	2.00	9.00	4.00	—	—	—
Home Based School	—	—	—	—	—	0.90	1.60	1.20
Home Based Other	0.70	0.60	1.10	4.00	2.30	—	—	—
Non-Home Based	0.30	0.40	1.00	4.60	2.30	—	—	—

The trip rates reflected in the above matrix are based on employees by type and/or location, number of students by type, and number of dwelling units.

As is the case with the trip production model development, the trip attraction model is developed using origin-destination travel survey data. Trips from the survey are accumulated according to the land use at the attraction end of the trip for each purpose. Usually large shopping centers are coded as separate zones to allow the trip accumulation as shown in Table IV-5.

In developing the attraction rates in the matrix, trips for the entire area are accumulated for the various cells in the matrix. For example, to obtain the trip rate per dwelling unit for homebased other trips, the total number of homebased other trips in the area is divided by the total number of households. Similarly, to obtain the rate for homebased work nonretail, the total number of homebased work trips is divided by the total areawide nonretail employment.

Special consideration should be given to major trip generators (e.g., airports, sport stadiums, hospitals, and regional shopping centers) that comprise land uses which are unique and do not show the same trip attraction characteristics that may be typical within the study area. For these cases, it may be desirable to conduct special generator studies to develop a trip rate that is characteristics of the special activity. This would involve special traffic counts and an inventory of some measure of the activity (i.e., number of enplaned passengers for an airport, or number of beds for a hospital, etc.). Various rates have been compiled from studies such as these around the country and might be a good source for a "first cut" or "borrowed" rates.*

---

* Trip Generation by Land Use Part I, A Summary of Studies Conducted, Maricopa Association of Governments, April 1974, Institute of Traffic Engineers Technical Committee, 6A6 -- "Trip Generation Rates," available from FHWA.

Application of Trip Production and Attraction Models

Trip production and attraction models are applied using estimates of the independent variables to determine the zonal productions and attractions. These estimates may be in the form of distributions (i.e., income distribution by zone) or zonal mean values. The approach discussed in this section will consider the application procedure using zonal means for the independent variables; however, this is not intended to discourage the independent forecasting of automobile ownership and income distribution. Also, although the procedure is discussed in the context of estimating future trip productions and attractions, the same procedure should be applied in the present year to validate the trip generation models. This type of analysis is discussed in more detail in the section on evaluating the trip generation results.

Figure IV-5 summarizes the trip production application process, showing how the four submodels are related. Figure IV-6 outlines the process for applying the trip attraction models.

The following hypothetical example illustrates the application process for both trip production and attraction models.

A zone located in the downtown area has the following future household and employment statistics as obtained from the urban activity forecasts:

Total number of dwelling units: 60
Zonal average income/dwelling unit = $11,000
Retail employment = 220
Nonretail employment = 650

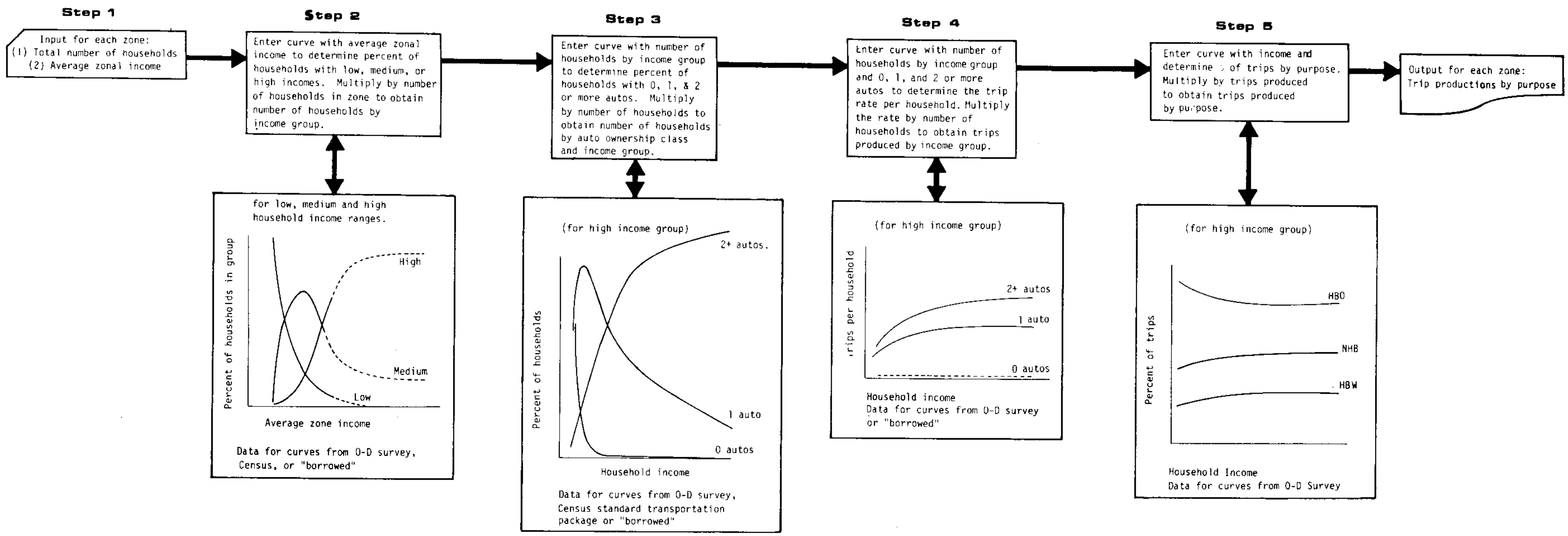

Figure IV-5. Trip Production Application Procedure

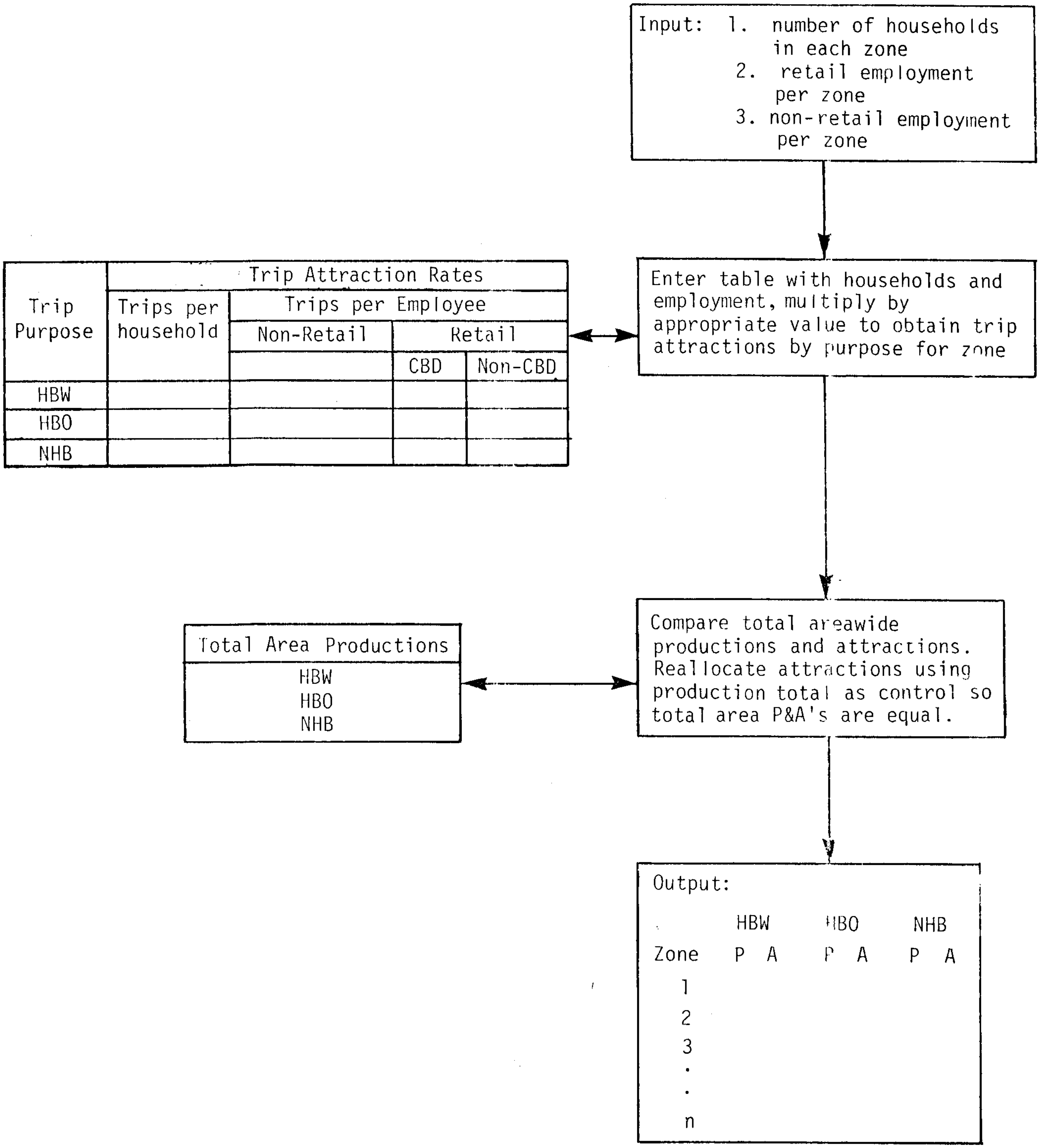

Figure IV-6 **TRIP ATTRACTION APPLICATION AND COMPARISON PROCEDURE**

Trip Production Model Application

The first step in the application process is to use the income submodel illustrated in Figure IV-7 to determine the number of households in each income group.

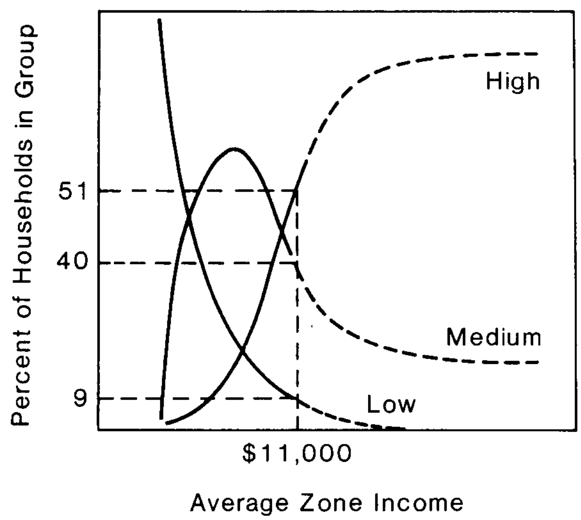

Figure IV-7. Income Submodel

The results are as follows:

Number of low income households = 60 households x 9% = 5

Number of medium income households = 60 households x 40% = 24

Number of high income households = 60 households x 51% = 31

Next, using the auto ownership submodel (Figure IV-8) and having the income submodel estimates of households by income category, the households can be further stratified by auto ownership.

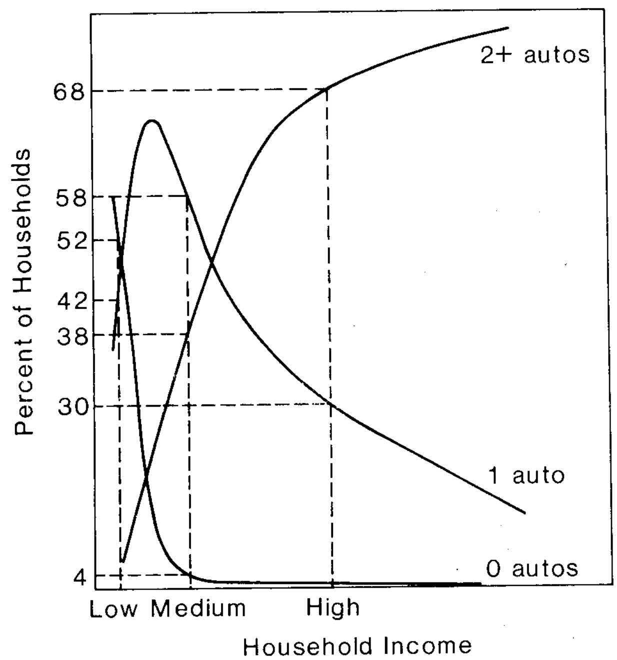

Figure IV-8. Auto Ownership Submodel

This calculation is as follows:

5 Low Income Households:	5 x 54% =	3	zero auto households
	5 x 42% =	2	1 auto households
	5 x 4% =	0	2+ auto households
24 Medium Income Households:	24 x 4% =	1	zero auto households
	24 x 58% =	14	1 auto households
	24 x 38% =	9	2+ auto households
31 High Income Households:	31 x 2% =	1	zero auto households
	31 x 30% =	9	1 auto households
	31 x 68% =	21	2+ auto households

After the households are cross-classified by income and auto ownership, the trip production submodel (Figure IV-9) can be applied to yield estimates of zonal trip productions.

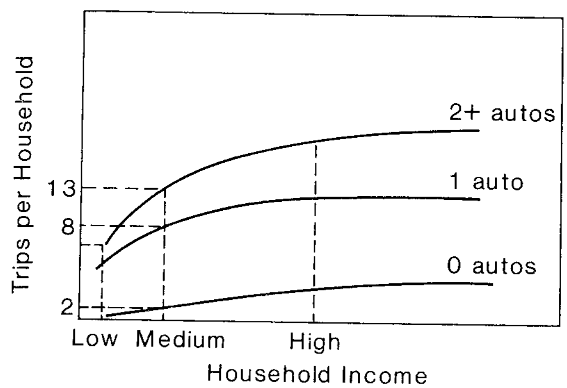

Figure IV-9. Trip Production Submodel

The calculation is as follows:

	Auto Ownership			
Income Group	0	1	2+	Total
Low	3HH x 1 trip/HH = 3 trip ends	2HH x 6 trip/HH = 12 trip ends	0HH x 7 trip/HH = 0 trip ends	15
Medium	1HH x 2 trip/HH = 2 trip ends	14HH x 8 trip/HH = 112 trip ends	9HH x 13 trip/HH = 117 trip ends	231
High	1HH x 3 trip/HH = 3 trip ends	9HH x 11 trip/HH = 99 trip ends	21HH x 15 trip/HH = 315 trip ends	417

Finally, with estimates of trip productions by zone and household income, the fourth submodel (Figure IV-10) can be entered to split the trip by trip purpose.

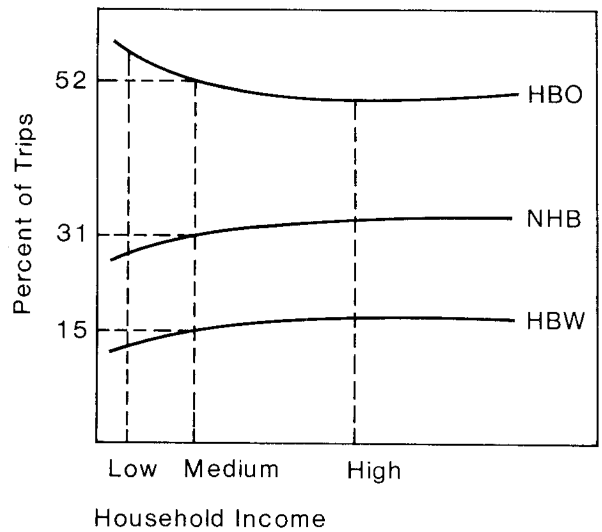

Figure IV-10. Trip Purpose Submodel

The calculations are as follows:

Zone 1:

Household Income Group	Trip Productions	Percent by Purpose	Trip Productions by Purpose
Low	15	15% HBW	2
		57% HBO	9
		28% NHB	4
Medium	231	17% HBW	39
		52% HBO	120
		31% NHB	72
High	417	18% HBW	75
		50% HBO	209
		32% NHB	133

Summary: 116 Homebased Work Productions

338 Homebased Other Productions

209 Nonhomebased Productions

Trip Attraction Model Application

With estimates of the number of households and employment levels in each zone, and with the trip attraction rates (Table IV-6), forecasts of trip attractions can be made.

Table IV-6. Trip Attraction Rates

Trip Purpose		Attractions per Household	Attractions per Non-Retail Employee	Attractions per Downtown Retail Employee	Attractions per Other Retail Employee
Trip Purpose	Home based work	negligible	1.7	1.7	1.7
	Home based other	1.0	2.0	5.0	10.0
	Non-home based	1.0	1.0	3.0	5.0

The calculations are as follows:

Home Based Other Attractions

60 households x 1 trip end/HH	=	60 attractions
220 downtown retail employees x 5 trip ends/emp.	=	1100 attractions
650 non retail employees x 2 trip ends/emp.	=	1300 attractions
Zone 1 home based other attractions	=	2460 attractions

Similarly, calculations for other purposes yield:

Home based work attractions = 1479

Non-home based attractions = 1370

The models are applied in the above manner for all the traffic analysis zones in the study area to determine estimates of the productions and attractions.

For trip distribution purposes, trip production and trip attraction estimates should have an areawide balance by purpose. For each home-based purpose the areawide summation of attractions should equal the areawide summation of productions as estimated by the trip generation models. If the two are not in agreement, the trip production estimates usually are taken as the control since characteristics of the home, such as auto ownership and income, more adequately reflect changing travel characteristics than do nonresidential variables. However, if there are large discrepancies it is important that more than a cursory evaluation and adjustment be undertaken. Both the production and attraction relationships should be examined to determine if adjustments are required in future application due to unforeseen situations. For example, there may be some overbuilding of shopping facilities resulting in higher than expected shopping trips attraction generation rates. An evaluation of this might indicate a lowering of the shopping trip rate for shopping center locations and leaving the rates constant for CBD and other nonshopping center locations.* Likewise, the percentage of trips in each purpose category should be examined for consistency with known trends. For example, most repeat origin-destination studies show that work trips are declining as a percentage of total trips.

---

* It is also quite possible that increases in trips to a particular land use type activity (e.g., shopping centers) would be reasonable because of increases of a general rise in socioeconomic levels at the production (or residential) end of the trip. Reducing the trip attractions to balance with the productions could result in an underestimation of future total travel.

Forecasting Considerations

The application of the trip generation models in the future year requires the forecasting of a few key land use and socioeconomic variables. It is important that these variables be forecast with care since subsequent transportation systems analysis will be dependent on such forecasts.

A key variable in the trip generation model development and application is income; therefore, a procedure for forecasting income should be developed. One procedure for income forecasting is described in appendix A of "Trip Generation Analysis."

This procedure described is based upon an examination of income distributions for several past years on a constant dollar base. The technique allows for the extrapolation of the historical income distributions to the forecast year based upon the knowledge that the proportion of families in the lower income ranges is decreasing and that the proportion of families in the higher income ranges is increasing. The procedure is iterative in nature, requiring the application of differential growth factors to income ranges and the subsequent plotting of the resulting distribution to determine if the distribution appears to fit the historical changes in the income distribution.

The assumption of the stability of the relationships between trips and land use and socioeconomic variables over time is basic to forecasting. No matter how carefully the trip generation procedure is developed or how accurately the estimating relationships mirror observed data, considerable forecasting error may result unless the variables are forecast within a reasonable degree of accuracy and the relationship in fact does remain constant over time. It is often easy to forget that the quality of the trip generation estimating procedure is only as good as the quality of the future estimates of the land use and socioeconomic forecasts. Also, since the assumption of time invariance is generally made when forecasting, it is extremely important that relationships be chosen which are expected to exhibit a high level of stability over time. The transportation planner must not become so involved in the mechanics of model development that he loses sight of the goal of providing meaningful travel forecasts.

External Trip Generation Model

Up to this point, the procedures described for development and application of trip generation models have been based on internal trips--trips that begin and end within the study area. The remaining trips are either through trips--trips that have neither end in the study area--or internal-external trips--trips that have one end in the study area.

Through trips are generally handled by the application of growth factors to the inventoried through trips. These growth factors are developed based on forecasts of transportation facility development which may add or subtract from the attractiveness of travel through the study area. Population and economic activity in the region from which the through trips began are also analyzed and forecast.

There are several ways to handle internal-external travel. One way is to group external trips so that they are "produced" at the external stations on the cordon line and "attracted" to the internal zones. The number of attractions is a function of the character of the internal zone and the distance between the zone and the cordon line. Research has indicated a consistency in the pattern of external trip ends in a study area (see Figure IV-11). Generally, more trips are attracted to the central business district than to the outlying zones of the study area.

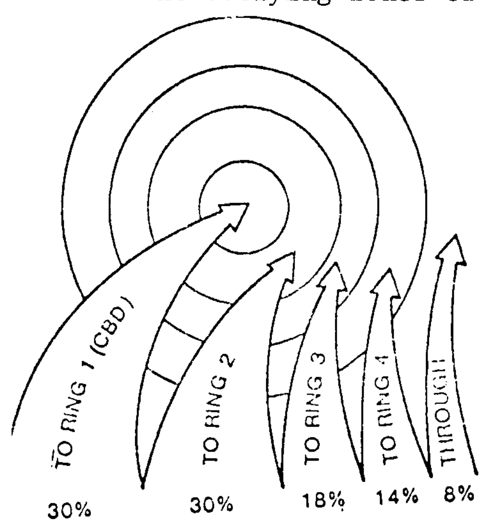

Figure IV-11. Distribution of Traffic Approaching a Typical Metropolitan Area of One Million Population

Usually, there are not enough external trips for them to be analyzed alone; therefore, it is suggested that the internal ends of external trips be treated as a proportion of all other trips.

The following ratio, calculated by zone, forms the basis for this approach:

$$\frac{\text{Number of internal ends of external trips}}{\text{Number of all other internal trip ends}}$$

After these ratios have been determined, the averages of the zonal ratios can be calculated and analyzed, by rings, from the central business district.

If no pattern is apparent, an alternative is to average the zonal ratios by districts, then examine those ratios for a pattern. If a large traffic generator attracts a significant amount of external traffic, it should be analyzed separately.

At the external station--the production end of the trip--the forecast of future trip ends should be based on a growth factor that reflects the growth in travel within the travel corridor of the external station, including the area beyond the cordon.

In forecasting it is necessary to balance the growth in trips, as determined from the external factors, with the growth determined from the analysis of the zonal ratios by ring. The more reasonable and logical value should be used as a control.

It is also necessary to use judgement in selecting the ratio to use in forecasting. Over time, the outer rings may take on the character of the inner rings in the base year, thus affecting their ratios. Also growth of urban activity up to and beyond the cordon may lead to a change in the ratio for the ring immediately inside the cordon.

Truck and Taxi Models

Depending on the volumes of truck and taxi trips, the generation analysis can be done either separately or in combination with other travel analysis in the nonhomebased category.

In the case of an unusually large truck terminal or a manufacturing site with a high rate of truck activity, separate growth factors may be required--growth factors that reflect the potential growth of the individual site.

Taxi trips generally have a rather definite pattern. For example, taxi trip generation rates usually decrease in direct proportion to distance from the central business district, and in direct proportion to increases in car ownership.

If the volume of taxi or truck trips is high enough to warrant separate analyses, they can be developed based on the number of households and employment.

For the above approach, trip productions and trip attractions would be set equal. In some analyses, truck and taxi trips are combined and treated as one trip "purpose."

EVALUATING TRIP GENERATION RESULTS

The cross-classification approach is basically a common sense approach which minimizes the amount of statistical evaluation required; therefore, most of the evaluation is focused on checks of the reasonableness of the results.

Reasonableness Checks

The design of the cross classification matrix is based on the choices of the independent variables on which to stratify the trip rates, and on the choices of categories on which to stratify the independent variables.

It is recommended that the chosen trip rate (i.e. trips per household) be plotted against the possible choices of independent variables for stratifying the trip rate in the cross classification matrix to develop a "feel" for the data and relationships at hand. While plots of the data are very useful, they are only two dimensional; for this reason, too much reliance should not be placed on the scatter diagrams produced because the relationships may change when a third variable is added.

Regression and correlation also can provide useful information in determining the independent variables for the cross-classification matrix. A simple correlation matrix will provide the interrelationships between the trip rate and possible independent variables. The coefficient of simple correlation (r) is a measure of association between two variables, and the matrix gives the correlation coefficients for all possible pairs of variables. An examination of this matrix will provide information about relationships between the independent and dependent variables. Strong relationships may then be evaluated for logic.

Although the above tools are important, variables which appear to be the most logically related to the trip rate variable should receive the most attention. Variables that reflect a causal relationship that has some likelihood of remaining stable over time should receive the highest priority.

Rules of Thumb for Developing and Evaluating a Cross-Classification Matrix

1. Number of Observations

The number of observations for any "cell" of the matrix should be large enough so that the mean rate developed for the cell can be reflective of travel for future application. It is suggested that at least twenty-five observations be accumulated in each cell. Where there are less observations, consideration should be given to combining the cell with another. The following are some guidelines to follow in combining cells:

1. The final result should involve the fewest possible combinations.

2. Zero auto ownership is considered more significant in tripmaking than income category. Therefore, necessary combinations for the zero auto ownership category should be made across income categories.

3. Where one or more automobiles are owned, income level is considered more significant in tripmaking than the number of autos owned. Therefore, combination of adjacent cells should generally be within an income category where possible.

4. When more than two cells are combined, the shape of the combination should be rectangular rather than "L" shaped.

5. A small cell should not be combined with a much larger cell if it might be combined with another deficient cell in an adjoining category, thereby producing one cell which satisfies the 25-observation minimum requirement.

Another alternative where there are too few observations is to ignore the cell when plotting the data to form a curve. Then, after a smooth curve has been formed, the value from the curve can replace the value of the cell.

2. <u>Cells With High Standard Deviations</u>

The cell values should not have too wide a dispersion as reflected by the standard deviation of the observations about the mean.

Criteria should be established to determine the point at which the standard deviation is "too high." Where many cells of the chosen cross classification matrix have a large standard deviation, consideration should be given to either stratifying on an additional variable or reevaluating the initial choice of variables. Where only a few of the cells of a matrix have a large standard deviation as a percent of the mean, the ranges established should be reevaluated to determine if new ranges should be used or if certain cells should be subdivided. For example, if a matrix utilizes income stratified into $4,000 increments (i.e., 0-$4,000, 4-$8,000, 8-$12,000, etc.) and results in a high standard deviation in the trip rates, then perhaps a stratification on $2,000 income increments should be tested (i.e., 0-$2,000, 2-$4,000, etc.), or perhaps only the lowest cell may need to be subdivided.

3. Plotting Cell Values

The cell values in cross-classification matrices should always be plotted.

Plotting the cell values will indicate the results of the matrix design are logical. For example, if a plot shows dips and rises instead of a smooth functional relationship, investigation should be made of the ranges established and of the distribution of values within the cells.

If a very flat relationship is discovered--if there is very little variation in the trip rate with variation in the stratifying variable--the planner should consider using an alternative matrix--stratifying variable.

If the curve shows a relationship that is clearly contrary to logic (if, for example, the trip rate decreases with increasing car ownership), the source data should be evaluated for soundness and accuracy.

4. Plotting Observed Versus Estimated Values

The trip generation relationships should be applied to base year data to develop zonal trip productions and attractions. These productions and attractions can then be compared to the base year productions and attractions from the origin-destination surveys. A plot of observed against estimated values would be useful in this evaluation. This is an excellent way to identify analysis areas that have unique travel characteristics. An example plot is shown below.

The points in the figure below represent individual analysis area. Points that fall well away from the 45 degree line may be located on a map of the study area. Separate analyses may be done on these areas to examine them for unique characteristics of geographical bias.

This type of plot is suggested for all of the cross-classification tables, including trip productions, trip attractions, and car ownership and purpose stratifications. The results of applying the cross-classification matrix approach can be improved if unique traffic generators are removed from the matrix. A major shopping center or an air terminal are typical examples which may warrant deletion from the cross-classification analysis. Separate analyses are then required for these areas.

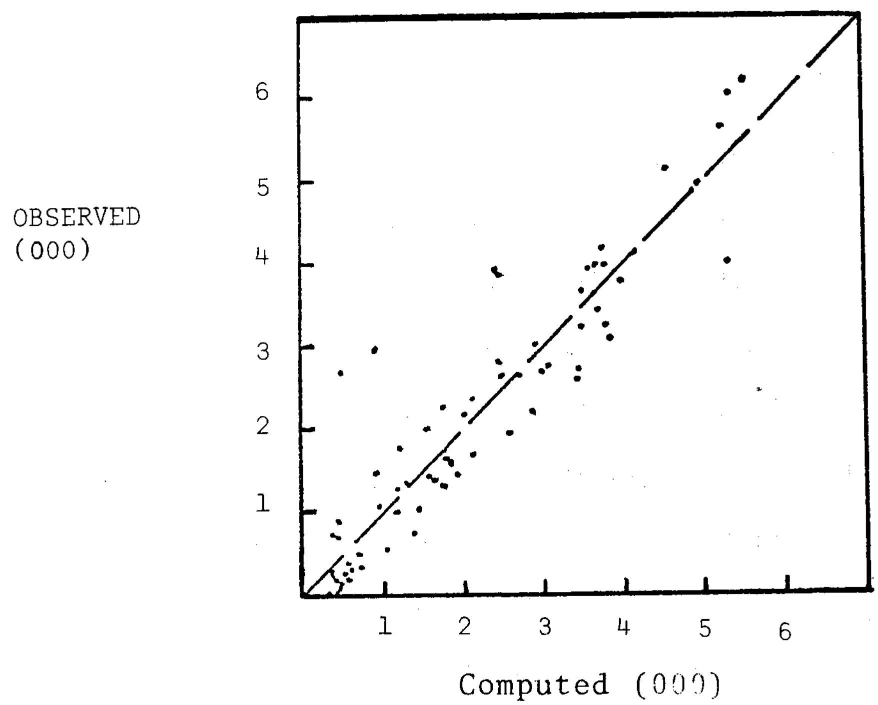

Figure IV-13. Plot of Observed vs. Estimated Values of the Dependent Variable--Total Trip Productions by Zone

COMPUTER PROGRAMS FOR TRIP GENERATION

The following descriptions outline general program operation and application philosophy. More detailed documentation necessary for program utilization is available in the publications "FHWA Computer Programs for Urban Transportation Planning," July 1974 and "BMD, Biomedical Computer Programs."1/ These tools will provide for more efficient and effective trip generation analyses. This is important for both initial studies and studies performing surveillance activities or updating existing trip generation relationships.

Cross Classification

Figure IV-14 illustrates the process for applying the computer programs in PLANPAC using cross classification. (See Figures I-4 and I-6, also). As this figure indicates, the program DUSUM merges the relevant data from the dwelling unit records (number 1 cards) and the trip records (number 2 cards). The resulting output is a dwelling unit record summarizing, by sampled unit, any trip and socioeconomic information desired.

Program DUSUM assumes that both input data sets (dwelling unit and trip cards) are in sort by sample number. Edits are made, however, in both sets for proper sort and the user is warned (1) if either data set is out of sort, (2) if there is no dwelling unit record for a particular trip record, or (3) if there is more than one dwelling unit record for a given trip record, the program will not abort, but will generate error messages which can be quite numerous.

The user specifies the "from" and "to" purpose codes and the "mode" codes of the trip purposes and modes desired in the dwelling unit summary output. Trips may be summarized by purpose in either origin-destination or production-attraction format, and the output data set may be in either binary (normal) or EBCDIC format, at the option of the user. If combined purposes are desired, the output of DUSUM may be input to program REFORM or BPR09S (BMD09S) and summed before entering program XCLASS. Although DUSUM is designed to merge No. 1 and No. 2 cards, other socioeconomic and trip data sets may be merged as long as certain constraints are met in describing these data.

The output of DUSUM, or REFORM/BPR09S if trip purpose combinations are desired, is then used as input to program XCLASS. This program forms the core of the cross classification analysis procedures. It provides cross-tabulations, in various forms, of input data based on stratifications of up to three independent variables. Observations in the input data are assigned to cells of a matrix created by user-specified categories of the

---

1/ BMD Biomedical Computer Programs, W. J. Dixon, editor. Available at a cost of $8.25 from the University of California Press, 2223 Fulton, Berkeley, California 94720.

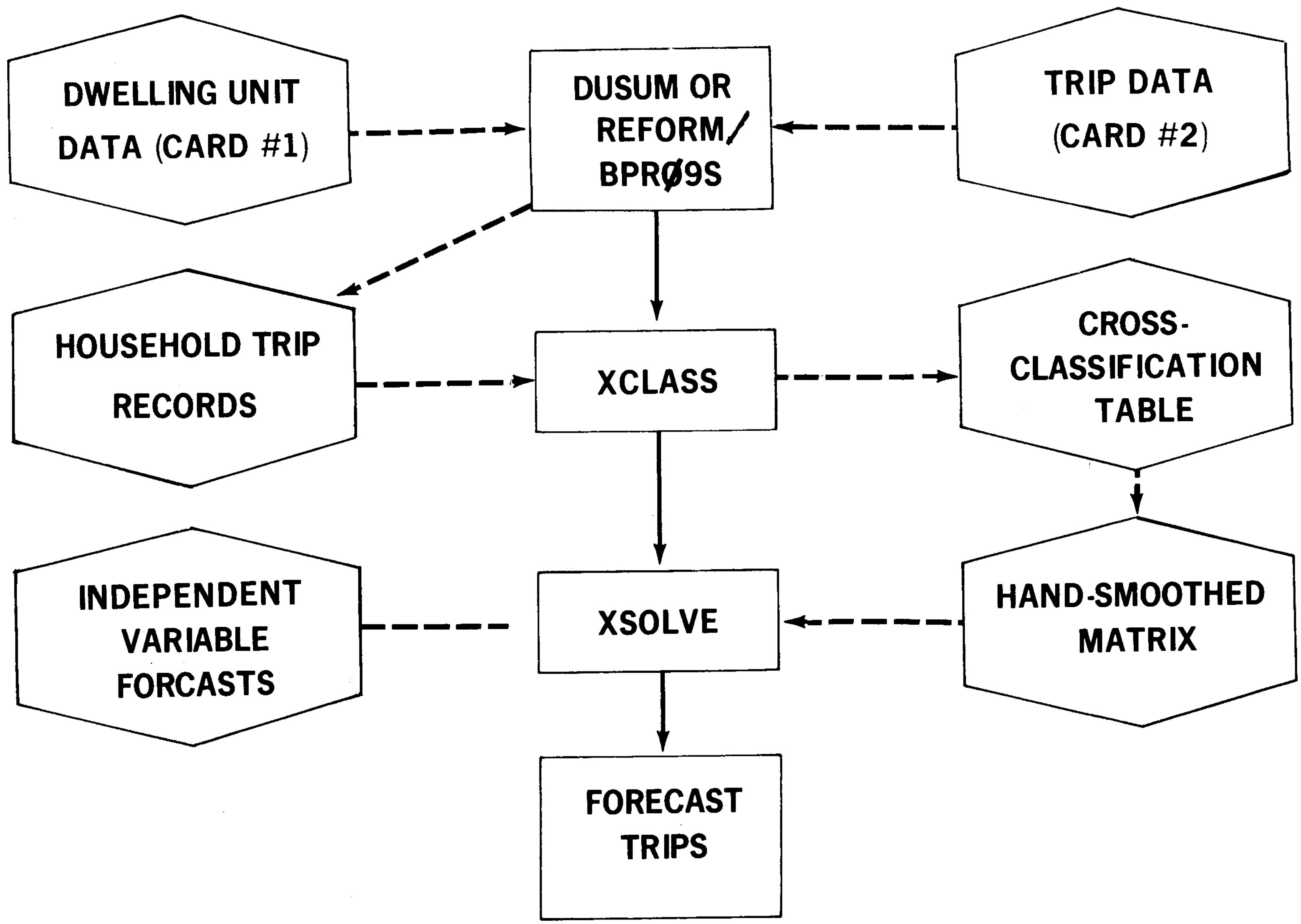

FIGURE IV-14 CROSS-CLASSIFICATION PROGRAMS IN PLANPAC two or three independent variables. Cell totals, row and column totals and a grand total of the number of observations are printed. Also, a dependent variable is summed and averaged by cell and printed with row and column totals and averages. Cell standard deviations are also provided. The cell means of the dependent variable represent the trip rate by dwelling unit types (income and auto ownership) and are used in the application of the cross classification procedure.

PRKTAB can also cross-tabulate dwelling unit summary records, and its various options make it more versatile than XCLASS. In many cases, one run of PRKTAB is equivalent to six or seven runs of XCLASS.

Some advantages of PRKTAB:

1. 20 tables may be built per run.

2. 6 independent variables (one or two columns and 5 rows) can be used for stratification.

3. Mathematical operations (add, subtract, divide and multiply) can be performed between two tables to build a third table.

4. A scale can be applied to each cell of a table before printing.

5. Row and/or column variables can be weighted.

6. Row and/or column categories can overlap.

7. Row and/or column percentages can be determined.

8. Percentages of total observations can be developed.

9. Any internal table can be saved.

10. Define cards can be used to create a format description of the input data.

11. Row and/or column categories can be given literal values (i.e., income/2000 = 1/).

12. Nonadjacent codes can be grouped together by use of item 11, above, (i.e., STRUCTYP/other = 2,4-6,8/) would cause the values 2,4,5,6,8 to be grouped together to form the category "other."

Some Disadvantages of PRKTAB:

1. Initial set up requires more control cards.

2. Each table (output automatically by XCLASS) must be built separately with PRKTAB (observations, total trips, trip rates).

3. Cell standard deviations are not available.

4. Any table that is saved is not suitable input to program XSOLVE directly; it must first be punched onto cards in the proper format. (Note, however, that this is recommended for XCLASS output, also, since handsmoothing of the curves should be done before running program XSOLVE).

Also see the discussion of PRKTAB in Chapter XII, Section G.

As Figure IV-14 indicates, the cross classification matrix of trip rates is used (after hand smoothing) as input to program XSOLVE, along with the forecast of the independent variables. These independent variable forecasts are then matched with the appropriate cell of the desired trip purpose cross classification matrix, with the program interpolating between cells if necessary. The total number of dwelling units for each analysis unit and the proportion of those dwelling units for which the input independent variable forecasts are representative must be specified. If desired, an average value for the independent variables may be used and a proportional representation of 100 percent would be specified. The output of XSOLVE is an estimate of trip ends by trip purpose and geographic area designation (e.g., zone), suitable as input to trip distribution programs.

Regression Analysis

The regression analysis programs that have been prepared are generally sophisticated in their operation and are often extremely powerful in their nature. That is, they employ a method of successively adding variables to a regression equation with the objective of obtaining the "best" final equation. In this method one variable is added at each step and statistical tests are conducted to determine the "improvement" in the equation. Variables continue to be added until the maximum step specified is reached, there are no more variables or there are no more variables which satisfy certain statistical limitations specified by the user. One such program, BPRO2R (a modified BMDO2R in PLANPAC) or BMDO2R (in BACKPAC), Stepwise Regression, is a part of a battery of statistical and data analysis programs originally developed for medical research at UCLA. The complete package of the BIMED series and the BIMED-P series are distributed as part of the BACKPAC dataset. These programs are extremely versatile and many are suited for use in the transportation planning process. The BPRO2R (BMDO2R) program computes a sequence of multiple linear regression equations in a stepwise manner by adding one variable to the regression equation in each step. The variable added is the one which makes the greatest reduction in the error sum of squares. In addition, variables can be forced into (or out of) the regression equation. Nonforced variables are automatically removed when their F values become too low.2/ Equations with or without regression intercept may be selected.

2/ Further information about what statistics the planner should look for when developing equations by this method can be found in the Trip Generation Analysis manual referenced on the first page of this chapter.

Once the planner has established equations using stepwise regression or some other method, program PRATT may be used to solve those equations. It computes trip productions and attractions (dependent variables) from given zone data (independent variables) using the conversion equations. It will also adjust to user-supplied control totals and do other checking under user-controlled options. As in cross-classification, the output is an estimate of the trip ends by zone suitable for input to trip distribution programs.

As mentioned, all of the BIMED series (BMD) and the BIMED-P series (BMDP) are distributed in BACKPAC. Those listed below have been found to be particularly useful to planners:

a. BMD01D - Simple Data Description. This program computes simple averages and measures of dispersion of variables, omitting those values which the user specifies for exclusion from the computations.

b. BMD05D or BPR05D - General Plot Including Histogram. This program produces graphs and histograms.

c. BMD09S or BPR09S - Transgeneration. This program performs selected transgenerations on specified variables in the data. Input may be from punched cards, from BCD tape, or from binary tape.

d. BMD01V - Analysis of Variance for One-Way Design. This program computes an analysis-of-variance table for one variable of classification, with unequal group sample sizes. Data may be input from punched cards, from BCD tape, or from binary type.

e. BPR02R - Stepwise Regression Analysis. This program is an expanded version of BMD02R. It provides additional statistics and plots plus extended capability for selecting observations. Better run identification is also available through user supplied ID cards.

MONITORING AND SURVEILLANCE FOR TRIP GENERATION

The continuing transportation planning process emphasizes the need to monitor and, if needed, update trip volume estimates in light of changing land use and socioeconomic characteristics. Since trip generation estimating relationships are usually derived from cross-sectional data for one period in time, and are subject to change with time, it is also extremely important that the relationships be evaluated periodically.

Since trip generation supplies the direct link between travel and changes in the land use pattern it is necessary to periodically evaluate the relationships for stability. Additionally, the changing character and intensity of land use must be accounted for.

The intensity and character of land use in a study area are continually undergoing transformation. The most dramatic example can be seen in the central business district of most any city in the country. Small, old office buildings give way to parking lots sandwiched between two surviving buildings which eventually yield to the forces of time in the same fashion. After a brief existence, the parking lot becomes the site of a large modern office building. Changing character and intensity of residential land is just as evident. Small apartment houses are replaced with higher ones, resulting in a considerable increase in residential density. In the newer areas vacant land is utilized in developing commercial and residential land uses.

The procedure for trip generation described in this Chapter is efficient concerning data requirements for monitoring and surveillance. For residential trip generation, the land use and socio-economic forecasts include the number of dwelling units and income. For non-residential generation, the forecasts required are dwelling units, employment and school enrollment. The need for updating this information is primarily a function of the age and growth pattern of a metropolitan area. In older cities actual updating may not be as critical as in rapidly growing urban areas. This does not however alleviate the need for adequate and timely evaluations in all areas. To stay abreast of travel demands and changing land use activity in a dynamic and rapidly growing area, evaluation of the forecast annually may not be too often.

In addition to application of trip generation rates to changing land use patterns, it is important to evaluate the trip generation relationships. For example, for some unknown reason the trip making rate of a one car household with a $10,000 income may be increasing. The analyst should periodically evaluate the stability of the developed relationships over time. For the recommended procedure the data requirements for monitoring change in the trip generation relations may be limited to a small sample travel survey coupled with site surveys of selected non-residential land uses.