THE TEMPORAL STABILITY OF TRIP GENERATION RELATIONSHIPS

DECEMBER 1972 - NUMBER 48

BY

E. J. KANNEL

K. W. HEATHINGTON

JHRP

JOINT HIGHWAY RESEARCH PROJECT

PURDUE UNIVERSITY AND
INDIANA STATE HIGHWAY COMMISSION

Technical Paper

THE TEMPORAL STABILITY OF TRIP GENERATION RELATIONSHIPS

TO: J. F. McLaughlin, Director
Joint Highway Research Project
December 28, 1972
Project: C-36-69I

FROM: H. L. Michael, Associate Director
Joint Highway Research Project
File: 3-7-9

The attached Technical Paper "The Temporal Stability of Trip Generation Relationships" has been authored by Edward J. Kannel, Graduate Instructor in Research, and Professor K. W. Heathington. The Paper has been accepted for presentation at the 1973 Annual Meeting of the Highway Research Board in January.

The Paper is a summary of a portion of the Final Report titled "A Disaggregate Analysis of Urban Travel Behavior" on the HPR Part I Research Study "A Micro-Analysis of the Effects of Household Socio-Economic Characteristics and Individual Attitudes on Trip Generation". Word on acceptance of that Report by the FHWA has not as yet been received.

The Paper is submitted to the Board for approval of presentation and publication. It will also be forwarded to the ISHC and FHWA for review, comment and similar approval.

Respectfully submitted,

Harold L. Michael
Associate Director

HLM:ms

cc:	W. L. Dolch	M. L. Hayes	C. F. Scholer
	R. L. Eskew	C. W. Lovell	M. B. Scott
	W. H. Goetz	G. W. Marks	J. A. Spooner
	M. J. Gutzwiller	R. D. Miles	N. W. Steinkamp
	G. K. Hallock	J. W. Miller	H. R. J. Walsh
	R. H. Harrell	G. T. Satterly	E. J. Yoder

Digitized by the Internet Archive
in 2011 with funding from
LYRASIS members and Sloan Foundation; Indiana Department of Transportation

Technical Paper

THE TEMPORAL STABILITY OF TRIP
GENERATION RELATIONSHIPS

by

Edward J. Kannel, Ph.D.
Research Assistant

and

Kenneth W. Heathington, Ph.D., P.E.
Research Engineer

Joint Highway Research Project

Project No.: C-36-69I

File No.: 3-7-9

Prepared as Part of an Investigation

Conducted by

Joint Highway Research Project
Engineering Experiment Station
Purdue University

In cooperation with the

Indiana State Highway Commission

and the

U.S. Department of Transportation
Federal Highway Administration

The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the Federal Highway Administration.

Purdue University
West Lafayette, Indiana
December 28, 1972

THE TEMPORAL STABILITY OF TRIP
GENERATION RELATIONSHIPS

Edward J. Kannel, Research Assistant, Joint Highway Research Project, Purdue University.

Kenneth W. Heathington, Research Engineer, Joint Highway Research Project, Purdue University.

ABSTRACT

Trip generation models based on household data rather than zonal aggregate data are evaluated. It has been suggested that analysis of household travel characteristics should precede aggregation so that home interview data can be used more efficiently. Relationships identified at the decision level of travel should have greater causal validity and should be more temporally and spatially stable.

The major objectives of this research are to examine the form of household travel relationships, to determine the stability of these relationships over time and to evaluate the ability of household models to estimate future travel. The potentials for reduced sample size requirements and greater applicability of disaggregate models in different urban areas are also examined.

To examine the stability of disaggregate trip generation models, household travel data were obtained from home interview surveys in 1964 and 1971. Single family households interviewed in the 1971 survey represented the identical families which were interviewed in 1964. This unique sampling design permitted the analysis of the effects which changes in the households' socio-economic characteristics during the seven year period have on trip production.

The results indicated that the household models based on the 1964 data could successfully predict household travel reported by the same households in 1971. The household models from both time periods could be expanded to adequately estimate 1964 reported zonal area travel. Parameters of the disaggregate models also appear more consistent between geographical areas and could be developed with considerably less data than comparable zonal models.

INTRODUCTION

The methodology of trip generation modelling used in most current urban transportation planning studies is referred to as a zonal analysis concept. The enormous body of data obtained in the home interview portion of the origin and destination study is aggregated and summarized in larger units of the total study area, the traffic zone. These zones are the smallest areas considered in all further analyses and projections. The aggregated data are used to calibrate generation models which estimate trip production occurring under present economic, social and physical conditions. Future travel is then estimated assuming that the true causal relationships have been identified and that the model parameters will remain stable over time.

This traditional trip generation modelling approach which is based on aggregated socioeconomic and land use data is subject to critical review. The modelling approach has been challenged from at least two major viewpoints. These are:

1. The modelling approach does not allow full consideration of the continuous nature of the travel decision process. Trip generation is only the first stage in the total urban transportation planning process which also consists of trip distribution, modal split, and trip assignment. In current practice each of the models is normally developed independently of the others. As a result there is no general assurance that an internally consistent network equilibrium will be achieved. The modelling process acts as though there is a given level of demand irrespective of the transportation system which is available.
2. The use of spatially aggregated data assumes that the relationships derived represent the true relationships occurring in the units which compose the aggregate total. Further, the aggregate descriptions are assumed to remain stable temporally and thus serve as a basis for prediction of future travel.

Recognizing that inconsistencies may arise because the transportation system is not explicitly allowed to affect all

stages of the model development, attempts have been made to develop explicit demand models which combine the functions of generation, distribution and modal split (2, 8, 11). Other research efforts have been directed to the development of a stochastic modelling approach which would retain the sequential nature of current planning models, but would incorporate principles of economic utility theory to include more policy sensitive variables into the model framework (12, 13, 15). These latter approaches also recognize the efficacy of using disaggregate data to estimate model parameters.

Research related to the second major point of discussion, data aggregation effects, has also shown the shortcomings in the aggregate planning model concept (1, 3, 6, 9, 10). Review of the assumptions of the aggregate models has shown that the zonal means are not adequately representative of the individual units composing the mean (9). The reasons for inadequate representivity are that:

- a) the zone sampling distributions are skewed rather than normal so that the sample mean is not the central value; and
- b) considerable heterogeneity exists within zones with respect to household travel characteristics and socioeconomic traits.

Further, aggregation of the behavioral units to a zonal description "washes out" much of the total variation which exists in the data. The aggregate data may mask the true relationships and the causal nature of the explanatory

variables. Investigations have shown that aggregation changes the strength of the associations between variables and that the model parameters are dependent upon the size of the areal unit selected in the analysis (1, 3, 9). As a result the calibrated model is applicable only at the macro-level of analysis and in the geographical area for which it was calibrated. This has further important implications in the continuing phase of the urban transportation planning study as one needs to be concerned with measurement of changing conditions. When the analyst is interested in measurement of changes, particular care must be exercised to carefully identify the explanatory variables to be used and the parameters associated with those variables. Since the aggregate models are based on large volumes of data which are averaged together, the models are not sensitive to subtle changes which occur at the basic decision level of travel. Further, the data measures habits for a single time frame; since it is financially infeasible to obtain large quantities of data which would be necessary to revise the zonal estimates, the relationships observed in the original time frame are generally assumed to be held constant throughout the planning period. Logic suggests that changes in social and cultural patterns, and changes in the physical environment will have an effect on urban travel. To be sure, comparisons of aggregated relationships within an urban area have been made for different time periods, but as noted, these aggregate relationships are dependent upon the type and level of activity

within the areal units. Since the size of some zones or the level of activity within the zones may change over the study period, it is difficult to separate the influence due to changing areal descriptions from the influence due to changing relationships of the variables in the model. The subtle changes in urban structure and individuals' status and life style cannot be detected at the macro level of analysis.

DISAGGREGATE TRIP GENERATION ANALYSIS

Although there is a recognized need for considering concepts which incorporate the interacting effects of the total travel decision process, in this paper attention is directed only towards obtaining a better understanding of travel behavior by evaluating trip generation relationships at a disaggregate, behavioral level of analysis. The household is taken as the basic decision making unit for evaluating travel behavior.

Other researchers have pointed out the shortcomings of models based on spatially aggregated data, and have indicated the desirability of identifying the more basic relationships between the socioeconomic and travel characteristics which occur at the household level of analysis. Analysis at the disaggregate level appears to provide a means of overcoming several of the shortcomings mentioned above and provides several advantages to the transportation analyst. First, since the analysis is conducted at the household level, the basic relationships are not averaged out by aggregation or

clouded by the analysts' selection of the areal boundaries. Because the parameters of the model are not tied to a particular aggregation scheme, the model can be developed and then applied to whatever aggregation scheme is necessary for the following series of models which are employed. This is of importance in the continuing phase of the transportation study as the size and shape of the planning area changes. This greater flexibility in application would allow the analyst to more effectively use data from other public records (e.g., census data) which are summarized in areal units which do not conform to the boundary scheme of the transportation study.

A second advantage proposed for the use of household level analysis is that the household provides a common base for comparing travel characteristics in different urban centers. Unlike the artificial aggregate unit such as the traffic zone, the household is basically the same size and the same internal consistency in different geographical areas. Due to the common nature of the household unit, one might expect household model parameters to be more consistent from area to area.

A third advantage is that the household relationships represent the basic relationships at the decision level and therefore are assumed to possess greater causal validity. These causal relationships are more likely to remain stable over time, thus, forming a more valid basis for the prediction of future trip generations.

Finally, since all the data which are collected in the home interview surveys are analyzed prior to aggregation, the data are used more completely and effectively. As a consequence, the possibility exists for using smaller sample surveys in the continuing study to measure the changes which occur in the basic relationships.

The purposes of the research reported in this paper were to evaluate the form of the relationships which occur in household trip generation models and to evaluate the stability of these relationships. In addition, the ability of models based on reduced sample size to estimate total area travel is examined, as is the hypothesis that the parameters of disaggregate generation models are more consistent from one geographic area to another.

STUDY DESIGN

The data used in this study to evaluate the causal validity and stability of household trip generation relationships were obtained from home interview surveys in Indianapolis, Indiana. This metropolitan area of 800,000 population conducted a basic transportation study with the home interview data collection taking place in 1964. The Indianapolis Regional Transportation and Development Study (IRTADS) was a typical example of a transportation study in which the trip generation formulations were based on aggregate zonal totals. A five percent home-interview sample was taken

representing over 10,000 interviews. The data were aggregated into 395 zones defined for the study area and factored to represent total travel volumes for the area.

In this research, a second home interview was conducted in 1971 to study changes in household socioeconomic characteristics and travel behavior, and to evaluate the stability of household trip generation models over a seven year period. The latter survey obtained measures of the socioeconomic and travel characteristics of some of the identical families which were interviewed in 1964. Earlier research had suggested several variables which could be evaluated at the household level including family size, auto ownership, stage in the family life cycle, occupational status, income, dwelling unit type, and location within the urban structure (7, 10, 16). Although simultaneous evaluation of all levels of all factors would have been desirable, such a design would have required a prohibitively large sample to obtain a sufficient number of cases for statistical stability in all possible levels and combinations of the variables. Instead, a sample was selected which, as far as possible, represented all levels of three principal socioeconomic variables: family size, auto ownership, and income. In the experimental design the confounding influence of other variables was controlled to the greatest possible extent by careful selection of the 1971 sample. To control for differences in travel behavior which may be due to differences in life style of families living in

different types of dwelling units and not due to changes in the principal variables being considered, only single family home owners were selected in the 1971 study. As indicated above, to control for differences which may be a function of social or psychological factors of the families, the 1971 survey interviewed the identical families that were interviewed in 1964. Further, only those families were selected which remained at the same dwelling unit from 1964 to 1971. In this way, differences in travel behavior which may have been due to changes in the living environment rather than changes in household characteristics could be controlled to a greater degree. Finally, travel variation which may be attributed to seasonal or daily variation was controlled by obtaining both data sets in the fall of the year and by scheduling the 1971 interview schedule such that each household recorded travel in 1971 on the same day of the week as in 1964.

Elimination of all families which were not single family, homeowners, or did not provide complete information in 1964 provided a final list of 4300 households from which the 1971 sample was selected. Table 1 provides a list of average household and travel characteristics of the 4300 households and of the 357 households from which completed interviews were obtained in 1971.

HOUSEHOLD TRAVEL CHARACTERISTICS

Several household variables such as family size, auto ownership, income, labor force, and occupational status were

Table 1. Average Household Socioeconomic and Trip Production Characteristics of Survey Sample.

	1964		1964		1971	
	4300 Single Family Households	Mean	Std. Dev.	357 Households Sampled	Mean	Std. Dev.
Family Size	3.63	1.83	3.64	1.78	3.18	1.60
Persons 5 and Over	3.23	1.58	3.32	1.62	3.09	1.51
Labor Force	1.31	0.73	1.39	0.71	1.26	0.79
Mean Income	\$8,000	\$3,900	\$8,400	\$4,000	\$13,000	\$6,800
Median Income	\$7,300	--	\$7,200	--	\$14,000	--
Autos Owned	1.39	0.73	1.48	0.75	1.67	0.88
Total Trips	9.16	7.47	9.80	7.03	9.31	7.10
Home Based Trips	6.98	5.38	7.49	5.24	7.00	5.18
Home Based Work	2.00	1.49	2.12	1.34	1.99	1.51
Home Based Shop	1.33	2.03	1.43	2.13	1.04	1.73
Home Based School	1.03	1.97	1.08	1.97	0.97	1.08
Home Based Other	2.62	3.46	2.85	3.38	2.99	3.31

examined to determine their effect on household trip generation rates. All of the variables showed a statistically significant effect when considered alone, however, due to the large intercorrelation between the variables, the effect of any one variable is not independent of the others. For example, although income would be a significant explanatory variable for estimating trip production, the research indicated that income had a greater effect on auto ownership which in turn affects household travel (5). Since auto ownership appeared as a more direct cause of travel, it was selected for use in the prediction models along with the family size variable.

Graphical summaries of the relationships between family size and auto ownership and household travel are given in Figure 1 and Figure 2, respectively. Figure 1 indicates that the relationship between family size and home based trip production is nearly linear for family sizes of four or less, but as family size increases the rate of trip production increases at a decreasing rate. This overall non-linear trend agrees with the findings reported by Oi and Shuldiner (10). Since trip generation models generally assume linear relationships, large departures from linearity could have important effects on these prediction models. The analyst must recognize where the assumptions of the model are not met and the consequences of using the variable or model formulation in spite of these irregularities. This will be examined later in connection with evaluation of the predictive ability of the models.

FIGURE 1. HOUSEHOLD TRAVEL RATES FOR VARYING LEVELS OF FAMILY SIZE

FIGURE 2. HOUSEHOLD TRAVEL RATES FOR VARYING LEVELS OF AUTO OWNERSHIP

The other significant observation to be made from Figure 1 is the relatively good agreement of the curves for the two data sets. Although changes in the family composition and age structure have occurred over the years, the average trip production for families of similar size for the two periods is relatively stable. This again has important implications in developing models of travel behavior. The stability of the form of the relationship indicates the variable should be useful in forecasting models.

Figure 2 shows the corresponding curves for auto ownership. The curve exhibits strong linear trends with greater fluctuation from linearity exhibited in the 1971 data. However, the slope and intercept and thus, the effect of auto ownership on trip production, appears to have shifted in these households over the years. Such a shift in the relationships could again have special significance in the planning study. Unless a shift in the value of model parameters is detected by observation at intervals less than the planning period for which forecasts are made, the final estimate could yield considerable error. Because the disaggregate modelling approach is able to detect the subtle changes which occur, it is felt that these relationships could be monitored with smaller sample sizes and perhaps with greater frequency to detect these changes.

MODEL DEVELOPMENT

Stability of the disaggregate trip generation relationships was evaluated in three stages. First, standard linear regression models were developed from the 1964 and 1971 data sets and the parameters of these models were compared. Next, the 1964 model was used to predict the volume of travel which should be expected in 1971 if the model relationships are sufficiently stable to predict future travel in the households. The regular planning process was then essentially reversed in that the 1971 model was taken back in time to estimate the total zonal movement reported by the 4300 single family households in 1964. The 1964 and 1971 models were compared as to their ability to measure the 1964 aggregate home based trip production. Finally, a disaggregate model from all dwelling unit types was used to estimate total area travel. The consequences of using data which do not appear to meet the theoretical requirements of the model formulation are discussed.

As indicated, the relationship between family size and trip production did not appear to be linear through the entire range of the independent variable. Preliminary investigations also showed heteroscedasticity of household trip production variances for all levels of the family size and auto ownership variables. Further, the sampling distribution of the dependent variable is not a true normal distribution. Although this does not preclude the use of linear regression analysis to estimate parameters for the model, one may not be able to

make probabilistic statements about the accuracy of the model parameters with the degree of confidence that is usually associated with the statistical model. Recognizing these limitations, linear regression techniques were used to evaluate the disaggregate models.

Table 2 presents the results of the linear models for estimating home based trip productions. As was expected, the parameters of the model have shifted somewhat over time. The degree of change is in agreement with observations made from Figures 1 and 2. That is, the parameter for family size is very similar over the period while auto ownership has greater variability. Two way analysis of variance models (ANOVA) with unequal cell sizes were evaluated to test the stability of the relationship over time (17). The time factor may be labeled simply as a years effect, but years is considered only as a surrogate for the effect of changes in other possible pertinent variables such as income and stage in the family life cycle. The statistical analysis indicated that there was not a significant change in the effect of family size over the time period, but the effect of auto ownership had changed. From Figure 2 one could speculate that the change occurred in the zero car and three car families. Indeed, when only one and two car households were considered, there was not a significant variation due to time changes.

The coefficient of determination, R^2 , and the standard error of the estimate provide other measures for comparing the two models. Both models give similar statistics for

Table 2. Household Prediction Equations for Home Based Trips.

Variables	1964		1971	
	Regression Coefficient	Std. Error	Regression Coefficient	Std. Error
Constant (a)	-0.45	--	-0.19	--
Family Size (b_1)	1.40	.13	1.46	.15
Autos (b_2)	1.92	.31	1.52	.27
$R^2 = 0.34$ $R^2 = 0.36$ Std. Error of Estimate = 4.31 Std. Error of Estimate = 4.20 $\bar{Y} = 7.46$ $\bar{Y} = 7.00$				

these measures but, for the analyst who is accustomed to observing R^2 values around 0.90 for zonal data, they are unimpressive. However, these values were not unexpected as the models are attempting to explain all of the variation in trip production - not just the variation between zones. Within any household the number of trips reported may be two to three times the average rate of trip production of all households with similar characteristics. The household model formulated here cannot hope to predict these large variations for each household. The measure of usefulness of the household model for forecasting trip production must be based on its ability to predict average travel for some higher level of aggregation. If the model is successful in accomplishing this task, then model development at the disaggregate level would be of value to the researcher as a means of evaluating causal relationships at a behavioral level and to the practitioner for developing area travel forecasts.

Estimation of 1971 Household Travel

The 1964 trip generation model given in Table 2 was first used to estimate home based trips for the 357 families in 1971. The total estimated home based travel was 2542 trips compared to the survey total of 2498 trips, i.e., an error of less than two percent.

Sufficient data were available in one and two car households, and all family size levels to statistically evaluate discrepancies in the estimated and observed trips using a

chi-square contingency analysis (14). The null hypothesis of no difference between the estimated and surveyed trips could not be rejected at the 0.01 significance level. Visual inspection of zero and three car families also did not show any major discrepancies.

The household equation was remarkably successful in estimating trip production for these households. Of course the independent variables for the prediction of 1971 trips were known exactly at each household. This is a luxury which is not available in the operational study, but it does exhibit the faithfulness of the model for estimation even though all theoretical considerations of linear regression were not met. In particular, the non-linear trend for the family size variable did not significantly reduce the effectiveness of the linear model to estimate future travel from the surveyed households.

Estimation of 1964 Single Family Zonal Trips

Since the independent variables of the household model are linear in form, zonal area trips may be efficiently estimated from the relationship:

$$Y_j = na + b_1X_{1j} + b_2X_{2j} \dots + b_mX_{mj}$$

where

- Y_j = the number of trips in zone j,
- X_{kj} = the zonal total of variable k in zone j,
- n_j = the number of households in zone j,
- a = the regression constant, and
- b_k = the regression parameter for variable k.

Table 3 shows the results of expanding the 1964 and 1971 household equations to obtain estimates of the home based trips reported by the 4300 single family households. Two prominent elements of these statistics deserve attention here. First, when the household equations are expanded to obtain zonal estimates, the percent of variation is increased from about 35 percent (Table 2) to 96 percent, while the percent standard error of the mean is reduced from approximately 60 percent to 20 percent. The adjusted values are similar to values observed in zonal regression analyses.

The second and most important point to be drawn from Table 3 is the comparability of estimates obtained from the two data sets. The 1971 model estimated the zonal trip productions reported in 1964 with the same statistical efficiency as was possible with the 1964 household data sets. This supports the basic hypothesis of this research, i.e., analysis at the household level should provide relationships which are more meaningful and these relationships should remain stable over time. In this study the disaggregate analysis did detect a shift in the effect of auto ownership for the families selected, but the overall relationship was sufficiently accurate for estimating zonal travel at a second point in time.

GENERATION MODELS FOR ALL DWELLING UNIT TYPES

It is recognized that the 357 single family units selected for the first part of this research represent a limited inference

Table 3. Summary Statistics of Single Family Household Equations Expanded to Obtain Zonal Travel Estimates.

	1964	1971
Adjusted R^2	0.96	0.96
Adjusted Standard Error of Estimate	18.4	19.0
Mean of Zonal Trips ^a	95.8	95.8
Mean of Residuals	-4.1	-4.1
Slope: $Y_{\text{actual}}/Y_{\text{predicted}}$	0.98	0.98

^aDependent Variable is 1964 Zonal Home Based Trips

Number of Zones = 313

Household Models Based on 357 Observations

space in that they represent only a portion of the total population. The models developed for this sample can be expanded to give acceptable estimates of travel for the households from which they were selected; but would the same be true if one were to use a sample of all household and family characteristics? Further, these models have been expanded to obtain estimates of reported trips of the households from which the sample was drawn; can these models be expanded to determine the factored trip volumes which represent the trips of the total population in the study area?

These questions are evaluated here by developing a household travel model using the entire 1964 IRTADS interview data set. In addition, the models based on all dwelling unit types are used to examine the possibility of data reduction in the continuing study and the geographical transferability of the model relationships.

Estimation of IRTADS Total Urban Travel

The variables used in the home based trip production model developed by IRTADS were total zonal population and total autos in the zone. The household model developed in this study used household family size and total autos in the household to define the equivalent relationships. The relationships obtained for each of the models were:

IRTADS Model (4):

$$\text{Home based trips/zone} = 10.776 + 0.149 (\text{population}) + 1.257 (\text{autos})$$

Household Model:

$$\text{Home based trips/HH} = -.232 + 1.015 (\text{family size}) + 2.148 (\text{autos})$$

The household equation was expanded and compared with the zonal estimates made by IRTADS. The statistics for the two models appear in Table 4. A special comment is necessary when comparing the data in Table 4. The IRTADS model is based on data from 389 zones, while the household equation is expanded to represent travel from only 326 zones. As a consequence, the mean number of zonal trips is not identical for each model. The reduction in the number of zones is due to elimination in this study of all zones in which there were no reported dwelling units or labor force.

Visual inspection of the ability of household equations to estimate zonal travel is afforded by inspection of Figure 3. This is a plot of predicted home based trips against factored zonal trip estimates provided by IRTADS. If the model predicted perfectly all points should fit a 45 degree line passing through the origin. The actual regression line exhibited a slope coefficient of 1.00 and a constant term of -45. This constant is only one percent of the mean zonal trips; therefore, the model was accepted as a good fit of the data.

The residuals were examined by plotting the travel volumes against the residuals. This plot exhibited a random scatter of points. Further, Figure 4 shows a histogram of the residual distribution. This plot closely approximates the ideal normal distribution with a mean value of zero. Thus, in this study

Table 4. Trip Generation Model Statistics for Estimation of Total Home Based Trip Productions

	IRTADS Zonal Model	Household Model
R^2 ^a	0.97	0.92
Standard Error of Estimate	597	948
Mean of Zonal Trips	3287	3947
Number of Zones	389	326
Percent Standard Error	18.2	24.0
Mean of Residuals	0	-21
Slope: $Y_{\text{actual}} / Y_{\text{predicted}}$	b	1.00

^a Adjusted statistics given for household model. Original household model based on 10,532 observations.

^b Not available.

FIGURE 3 · ESTIMATED ZONAL TRIP PRODUCTIONS FROM HOUSEHOLD MODEL VERSUS ACTUAL ZONAL TRIP PRODUCTIONS FROM IRTADS

FIGURE 4 FREQUENCY DISTRIBUTION OF ZONAL RESIDUALS
DETERMINED FROM EXPANDED HOUSEHOLD
EQUATIONS

it was found that even though the household data did not meet all the assumptions for linear regression at the household level, residuals from the expanded equation did meet the criteria of independence and normality.

Comparison of the predictive ability of the zonal totals model and the household model indicates that the latter produces estimates with somewhat greater variation. It must be noted, however, that parameters of the IRTADS zonal equations are estimated to produce the minimum error in the zonal productions. By definition, the sum of the residuals must be zero. On the other hand, the parameters of the household model are estimated to produce minimum error at the household level. The mean of the residuals at the household level must be zero, but generally there can be no assurance that the residual sum will be zero when the model is used to estimate larger area travel. The degree to which the mean residual approaches zero provides another measure of the applicability of the expanded equation. The mean residual represents less than one percent of the zonal mean.

Potential for Data Reduction in Continuing Study

It was demonstrated above that the disaggregate household model could be expanded to produce total area travel. However, since the aggregate and disaggregate models were both formulated from a data base which includes over 10,000 home interviews, there has been no indication that the household modelling approach would save data collection expenditures. To estimate

the savings possible it would be necessary to conduct a full scale analysis of sampling variability and expected errors. From this analysis the ideal sample size necessary to obtain estimates within desired confidence limits could be determined. In this research a single sub-sample was drawn to determine the order of magnitude of sample size reduction which might be possible. This sub-sample was equivalent to a one percent sampling rate, whereas the IRTADS sample was designed as a five percent sample. Table 5 repeats the adjusted household equations given in Table 4 for the five percent sample and provides the comparable statistics for the one percent sample (2240 cases). The ability of the two household equations to predict total travel is very similar. The standard error of the estimate is actually somewhat smaller for the smaller sample size but, on the other hand, the mean residual is larger. Additional research is required to obtain more complete knowledge concerning the full extent of possible data reductions for trip generation as well as the other phases of travel forecasting. Certainly though, the contention that sample size requirements may be reduced for estimation of household trip generation appears to be substantiated.

Geographical Variation in Model Parameters

The final advantage proposed for disaggregate analysis was that observed relationships should be more consistent from area to area because the analysis unit is not tied to an artificial

Tab . 5. Comparison of Predictive Ability of Household Models Based on a Five Percent Sample and a One Percent Sample.

	Sample Size (Households)	
	10532	2240
Adjusted R^2	0.92	0.92
Adjusted Standard Error of the Estimate	948	922
Mean of Zonal Trips	3947	3947
Mean Residual	-21	+56
Slope: $Y_{\text{actual}}/Y_{\text{predicted}}$	1.00	1.02

^a Household equations expanded to estimate home based travel in 326 zones.

areal description and the household unit is of the same basic internal consistency in different geographical areas. A limited examination of this aspect is provided in Table 6.

The parameters given in Table 6 provide a measure of the degree to which household and zonal model parameters are comparable for two study areas, i.e., Indianapolis and the Tri-State area which includes New York City. The Tri-State area equations were developed in the research by Kassoff and Deutschman (6).

The magnitude of the household model parameters for the independent variables are strikingly similar for the two study areas, even though the areas themselves would not be considered as comparable in nature. The largest variation is in the magnitude of the constant term. One might reflect that the constant term of the model is the geographic factor which explains differences in household travel in the two areas. Of course, other differences in average trip rates in the areas would be reflected by differences in the average value of the independent variables.

On the other hand, there are substantial differences in the parameters of the zonal based models. Although this comparison is only for two study areas, the basic premise that household parameters measure a more stable, basic relationship appears to be substantiated.

Table 6. Aggregation Effects on Trip Generation Model Parameters for Two Urban Areas.

Level of Analysis	Dependent Variable (Y)	Independent Variable (X _i)	Indianapolis		Tri-State Area	
			Number of Observations	Model Parameters	Number of Observations	Model Parameters
Household	Trips per Household	X ₁ = Persons per Household	10,532	Y = 1.146X ₁	5,032	Y = 1.064X ₁
				+ 3.169X ₂		+ 3.169X ₂
		X ₂ = Autos per Household		- 0.192		+ 0.292
<hr/>						
Zone	Avg. Trips per Household per Zone	X ₁ = Avg. Persons per Household per Zone	299	Y = 1.092X ₁	305	Y = 2.054X ₁
				+ 5.139X ₂		+ 3.458X ₂
		X ₂ = Avg. Autos per Household per Zone		- 2.37		- 2.94

SUMMARY AND CONCLUSIONS

Analysis of travel behavior using the household as the basic unit provides a method of evaluating the changing relationships which occur over time. A disaggregate trip generation model developed from data obtained from 357 single family households in 1964 was able to predict the home based trips produced by the same families in 1971 with an average error of less than two percent. The household models from 1964 and 1971 also exhibited the same degree of statistical efficiency when expanded to estimate total zonal trips reported in 1964 by the single family households from which the 1971 sample was selected.

The disaggregate model for estimating total home based travel from all dwelling units was judged to be nearly comparable with the zonal model for estimating present travel. However, since the disaggregate model is sensitive to measurement of change in the behavioral unit, the household model is preferred. Indications are that the data set may be reduced by as much as 80 percent for estimating trip generation parameters at the disaggregate analysis level. Further, since the household is the basic unit in all urban areas, analysis at the household level can help the planner understand true travel variation between geographical areas, rather than apparent differences which are a function of the size of areal unit selected within the study area.

Certainly additional research is required to determine the limits of the sample size necessary for estimating travel

and the degree of geographical biases which exist. Also, consideration must be given to the data requirements of other aspects of travel forecasting, i.e., trip attraction, distribution, modal split and assignment. In the continuing study, the analyst must determine the degree to which the existing calibrated models can simulate changing travel patterns. Will the sample size which provides adequate information about changes in trip generation rates also provide sufficient data to evaluate changing attitudes and patterns of spatial distribution? Behavioral model research for the other planning models may also indicate increased efficiency. Careful planning of the survey design may provide information adequate for development of all disaggregate models. If knowledge of the complexities of travel behavior can be attained at this disaggregate level, the analysis could be conducted at this level and then aggregation may proceed to whatever level is necessary. The important item to emphasize is that the disaggregate model approach is sensitive to changes which occur at the behavioral level and therefore provides a means to measure changes. This is an essential consideration as the transportation analyst considers the changing conditions which occur during the continuing planning process. After evaluation of these changes at the behavioral level, aggregation may proceed to whatever analysis unit is necessary.

REFERENCES

1. Bureau of Public Roads, U. S. Department of Transportation, Guidelines for Trip Generation Analysis, U. S. Government Printing Office, 1967.

2. Domencich, T. A., G. Kraft, and J. P. Vallette, "Estimation of Urban Passenger Travel Behavior: An Economic Demand Model", Highway Research Record 238, Highway Research Board, Washington, D. C., 1968.
3. Fleet, C. R., and S. R. Robertson, "Trip Generation in the Transportation Planning Process", Highway Research Record 240, Highway Research Board, Washington, D. C., 1968.
4. Indianapolis Regional Transportation and Development Study, "Trip Generation Analysis", Technical Work Paper, 1966.
5. Kannel, Edward J., A Disaggregate Analysis of Urban Travel Behavior, An Unpublished Ph.D. Dissertation, Purdue University, June, 1972.
6. Kassoff, H. and H. Deutschman, "Trip Generation: A Critical Appraisal", Highway Research Record 297, Highway Research Board, Washington, D. C., 1969.
7. Lansing, J. B., and G. Hendricks, Automobile Ownership and Residential Density, Institute for Social Research, University of Michigan, Ann Arbor, 1967.
8. Manheim, M., "Practical Implications of Some Fundamental Properties of Travel Demand Models", Paper presented to Highway Research Board, January, 1972.
9. McCarthy, G. M., "Multiple Regression Analysis of Household Trip Generation - A Critique", Highway Research Record 297, Highway Research Board, Washington, D. C., 1969.
10. Oi, W. and P. Shuldiner, Analysis of Urban Travel Demand, Northwestern University Press, 1962.
11. Quandt, R. E., and W. Baumol, "Abstract Mode Model: Theory and Measurement", Journal of Regional Science, Vol. 6, No. 2, 1966.
12. Reichman, S. and P. Stopher, "Disaggregate Stochastic Models of Travel Mode Choice", Highway Research Record 369, Highway Research Board, Washington, D. C., 1971
13. Charles River Associates, Inc., "A Disaggregate Behavioral Model of Urban Travel Demand", Final Report on FHWA Contract FH-11-7566, March, 1972.
14. Siegel, S., Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill Book Company, New York, 1956.

15. Stopher, P. R. and T. Lisco, "Modelling Travel Demand: A Disaggregate Behavioral Approach - Issues and Applications", Transportation Research Forum, 1970.
16. Stowers, J. and E. Kanwait, "The Use of Behavioral Surveys in Forecasting Transportation Requirements", Highway Research Record 106, Highway Research Board, Washington, D. C., 1966.
17. UNEQUAL, Program Available in Purdue University Computer Library, 1970.