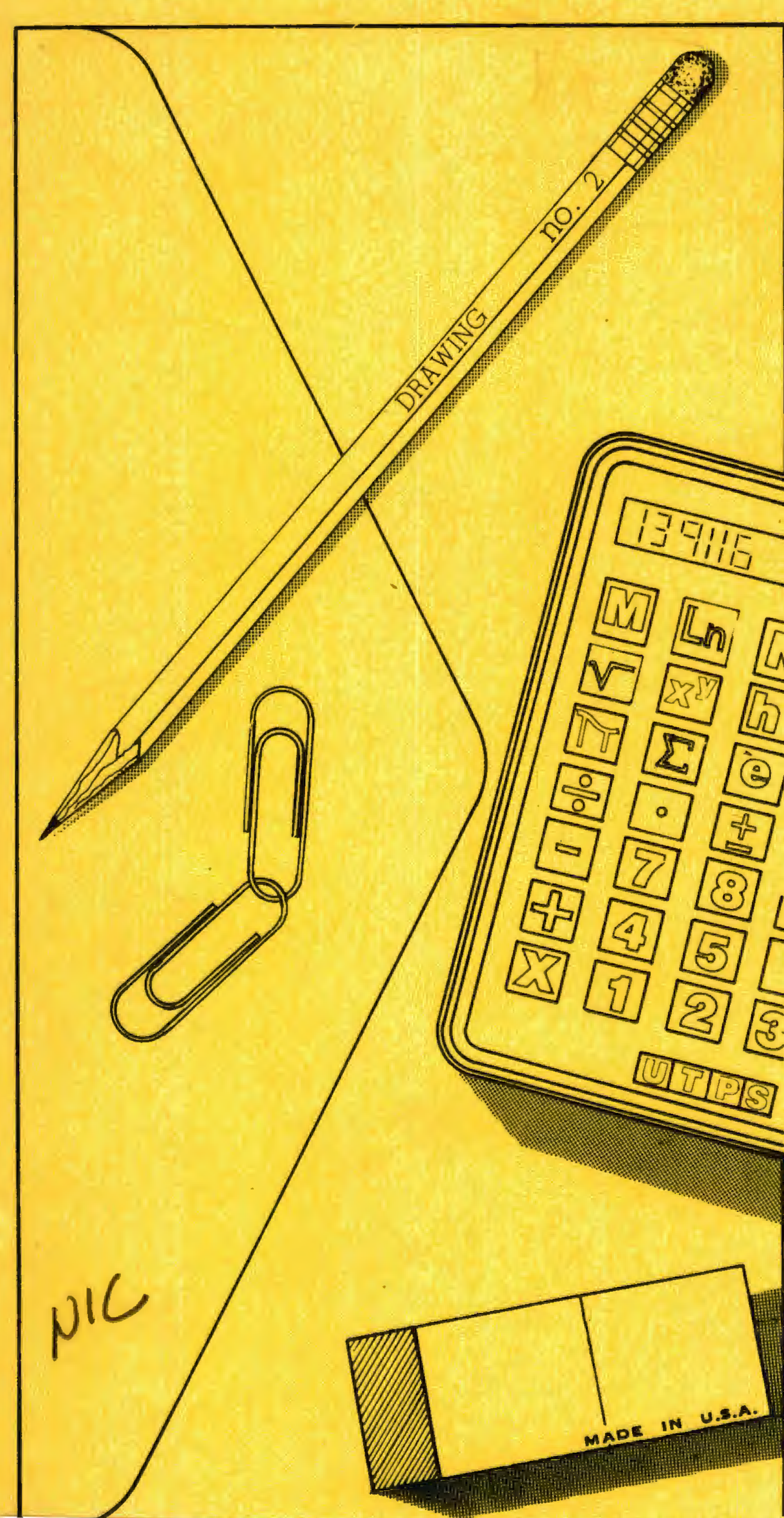


simplified aids for transportation analysis

NIC

forecasting auto availability and travel

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of technology sharing. The United States Government assumes no liability for its contents or use; neither does it endorse products or manufacturers. Trade or manufacturers' names appear only because they are considered essential to the object of this report.

26569338

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
UMTA-IT-06-9020-79-2		

4. Title and Subtitle	5. Report Date
Simplified Aids for Transportation Analysis: Forecasting Auto Availability and Travel	January, 1979
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No. UTP.PMM.77.1.1
9. Performing Organization Name and Address Peat, Marwick, Mitchell & Co. 1025 Connecticut Avenue, NW Washington, DC 20036	10. Work Unit No. (TRAIS) IT-06-9020 11. Contract or Grant No. DOT-UT-50021
12. Sponsoring Agency Name and Address U.S. Department of Transportation Urban Mass Transportation Administration 400 Seventh Street, S.W. Washington, D. C. 20590	13. Type of Report and Period Covered Final 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

In this report an analytical aid is described which provides a method for deriving and using the variable of automobile availability per household. This is an important socioeconomic variable in travel demand estimation. It is frequently not avilable, however, at the level desired, (e.g., transportation planning zone, traffic zone, or census block), or for the current year at any level. In the course of the research in which this aid was developed, a strong correlation was found between auto availability per household and the product of two other variables: the average market price of homes and the percent of home ownership. Data for these two variables are generally more readily available at the block or zone level or may at least be forecast more precisely at this level than automobile availability.

In this report, auto availability per household is described as a function of the product of these two variables. Estimates of automobile availability derived from this function can be applied to trip generation and modal split analysis in transit system studies, route patronage estimation, and sketch planning studies in urban areas of any size. Examples of trip generation and modal split applications of the variable are also presented in this report.

17. Key Words	18. Distribution Statement
Auto Ownership	Available to the Public through the National Technical Information Service, Springfield, Virginia 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	32	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Report UTP. PMM.77.1.1 Contract No. DOT-UT-50021

SIMPLIFIED AIDS FOR TRANSPORTATION ANALYSIS

Derivation and Use of Automobile Availability as a Variable for Estimating Travel Demand

Prepared by

Peat, Marwick, Mitchell & Co.
1025 Conn. Ave., N.W. Washington D.C. 20036
Based on Original Work Submitted by
Michael Holoszyc and Dale Miller

January 1979

Prepared for

U.S. Department of Transportation
Urban Mass Transportation Administration
Office of Planning Methods and Support

FOREWORD

Today's transportation planner confronts ever-changing issues within a variety of work environments. To assist him, UMTA's Planning Methods and Support Program researches, develops and distributes planning aids, including novel planning studies, and new design and forecasting techniques.

This is one of a series of six reports describing simplified aids to improve transportation decisions without resorting to computers or extensive data collection. The series, titled Simplified Aids for Transportation Analysis, presently includes the following titles:

1. Annotated Bibliography (UMTA-IT-06-9020-79-1)

2. Forecasting Auto Availability and Travel (UMTA-IT-06-9020-79-2)

3. Estimating Ridership and Cost (UMTA-IT-06-9020-79-3)

4. Transit Route Evaluation (UMTA-IT-06-9020-79-4)

5. Estimating Parking Accumulation (UMTA-IT-06-9020-79-5)

6. Fringe Parking Site Requirements (UMTA-IT-06-9020-79-6)

All are the work of recognized experts. They clearly present usable planning concepts, and add to the growing set of manual and computerized techniques comprising the UMTA/FHWA Urban Transportation Planning System (UTPS).

More important than the production and dissemination of new tools is the experience and opinion of their user. Local issues change. Better methods evolve. Or, realistically errors may appear in the final product. We depend on you, the transportation planner, to alert us to any of the above. We need your comments and your ideas. Please let us hear them, so we can continually improve our products.

You may obtain copies of any of the above reports from the National Technical Information Service (NTIS), Springfield, VA, 22161. On your request, please include the reference number in parenthesis.

Robert B. Dial, Director
Office of Planning Methods
and Support (UPM-20)
Department of Transportation
Washington, DC 20590

ABSTRACT

In January 1976, the U.S. Department of Transportation issued a Technical Notice (DOT-1-76) requesting transportation planners, engineers, and transit operators to submit useful but not widely known manual techniques that could be developed and distributed as simplified aids for transportation analysis. Over 70 analytical aids were submitted in response to this request.

Based on an evaluation process conducted to determine the most useful, easily applied, and generally applicable techniques, several of these analytical aids have been selected and documented in sufficient detail to permit their immediate use. In addition to these techniques, three additional analytical aids were developed as part of the Short Range Transportation Planning project, and an annotated bibliography of each analytical aid reviewed was prepared. These individual analytical aids and the annotated bibliography have been prepared as separate reports and have been brought together in this manual of simplified aids for transportation analysis.

In this report an analytical aid is described which provides a method for deriving and using the variable of automobile availability per household. This is an important socioeconomic variable in travel demand estimation. It is frequently not available, however, at the level desired, (e.g., transportation planning zone, traffic zone, or census block), or for the current year at any level. In the course of the research in which this aid was developed, a strong correlation was found between auto availability per household and the product of two other variables: the average market price of homes and the percent of home ownership. Data for these two variables are generally more readily available at the block or zone level or may at least be forecast more precisely at this level than automobile availability.

In this report, auto availability per household is described as a function of the product of these two variables. Estimates of automobile availability derived from this function can be applied to trip generation and modal split analysis in transit system studies, route patronage estimation, and sketch planning studies in urban areas of any size. Examples of trip generation and modal split applications of the variable are also presented in this report.

Because the intent of this report is to provide a simplified analysis aid, modifications, embellishments, and improvements to the suggested procedures and models are encouraged should local data or previous analyses suggest a more appropriate method.

SOURCE

Michael Holoszyc, Systan, Incorporated, P.O. Box U, Los Altos, California 94022; formerly research assistant with the Princeton University Department of Civil Engineering.

REFERENCE

Michael Holoszyc and Dale Miller, "The Planning and Design of a Demand Responsive Transportation System for Morristown and Morris Township" (report prepared for the Transportation Program, Princeton University, March 1975).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

This report describes one of a collection of useful but not widely known manual techniques employed by local transportation planners, engineers, or transit operators. This particular technique provides a method for deriving the variable, automobile availability per household, as a function of two other variables for which local data at the level of detail required for transportation demand analyses is generally more readily available. Sufficient information is provided to permit the immediate use of the analytical aid; this information is presented in three sections:

- I. Introduction. This section describes the analytical aid and its applicability as a simplified aid for transportation analysis, provides an overview of the application procedure, identifies the data and information required to use the aid, and illustrates the analysis output.

- II. Deriving the Automobile Availability Function. This section presents the results of a research project which determined the functional relation used to estimate automobile availability per household and provides a method which permits an analyst to develop a similar function using local data.

- III. Using the Automobile Availability Function. This section presents two examples of travel demand models which use automobile availability per household as the principal input variable.

- IV. Shortcomings and Limitations. This section describes the shortcomings and limitations of this analytical aid to make the user aware of the limits of its applicability.

The technique reported here is oriented to the practical planner who requires a specific analytical technique but who has limited data and time to perform an indepth analysis. The soundness of the method described in this report, however, must be considered independently by the potential user for each specific application. The section on shortcomings and limitations is provided to assist the potential user in making this assessment. Modifications, embellishments, and improvements to this technique are encouraged should local data or past analyses suggest a more appropriate proceeding.

DESCRIPTION AND APPLICABILITY

Many studies have shown that automobile availability per household is a key socioeconomic variable in the trip generation and modal split phases of travel demand estimation. In general, as automobile availability per household increases, the propensity to make trips, particularly nonwork trips, increases while the probability that transit will be used for a given trip decreases. For those reasons, automobile availability is widely used as an input variable in travel demand models.

The Automobile Availability Variable

Automobile availability is defined by the Census Bureau as "the number of passenger automobiles owned or regularly used by any member of the household (including nonrelatives such as lodgers)." Each occupied housing unit is classified as having available to it one automobile, two, three or more, or none.

The Census Bureau's automobile availability data are collected from all occupied housing units in the census (100-percent sample) and is tabulated and published on the census tract level of each standard metropolitan statistical area (SMSA). The automobile availability variable, however, is generally available at no finer level of detail than the census tract, which is generally too large an area for detailed travel demand and transit demand estimation. The average census tract contains about 4,000 persons whereas the desired zonal population for demand estimation models could be as small as 100 persons depending on the location, the type of area, the level of detail, and the type of analysis necessary for a particular application. To use automobile availability as an independent variable in these models, therefore, some method must be employed to estimate the variable at finer levels of detail.

Existing and potential applications of the use of automobile availability as an input variable include the following models:

- trip generation models - regression equations;
- trip generation models - cross-classification analysis;
- modal split models;
- direct transit demand estimation models;
- bus route location models - identification of high transit potential areas; and
- models to locate route deviation pick-up points.

U.S. Bureau of the Census, 1970 Census Users' Guide, (Washington, D.C.: Government Printing Office, 1970).

Development of the Analytical Aid

In a recent study at Princeton University,[1] the census tract variable for automobile availability per household was found to be strongly related to a hybrid housing variable: the average market value of homes multiplied by the percentage of households which are owner-occupied.[2] Both components of this hybrid variable are typically available at finer levels of detail than reported values of auto availability. For census years, the components of this hybrid variable are available at the census block level and, for noncensus years, they can be estimated at this level of detail by analyzing data and information (on assessed property values and number of owner-occupied units) which can often be obtained by property from the local tax appraiser's office.

In the Princeton study, the relation between automobile availability and the hybrid housing variable was derived by plotting points for census tract data on a graph and fitting a curve to this data. Separate curves were developed for the central city, for suburban and outlying areas, and for other study area jurisdictions.

In this report, the curves derived from the work conducted by Princeton University are provided, and a method is described for calibrating similar curves using local data. To use these curves, data are collected for the hybrid variable on a small area level such as the block level. The data may then be used at this level or may be aggregated to the zonal level for the study's zone system, using weighted averages (weighted by population, for example) and an equivalency table relating blocks to zones. The hybrid variable is then computed for each zone, entered on the calibrated curve, and a value of automobile availability per household is determined at the appropriate zonal level.

The analytical aid described in this report therefore provides a method for deriving automobile availability per household for virtually any zone system using these two housing variables as input data. The

---

[1] Michael Holoszyc and Dale Miller, "The Planning and Design of a Demand Responsive Transportation System for Morristown and Morris Township" (report prepared for the Transportation Program, Princeton University, March 1975).

[2] Automobile availability per household increases as average income increases and is typically higher for homeowners, who tend to purchase either a first or second car more readily than do residents of rental units (households in rental units tend to be comprised of smaller families, couples, or unrelated individuals).

report also illustrates the use of the relation between automobile availability and the hybrid housing variable.

OVERVIEW OF THE APPLICATION PROCEDURE

The use of this analytical aid should be considered whenever the transportation analyst requires an automobile availability variable at a zonal level smaller than the census tract. The aid is applied in three phases:

1. the derivation of the automobile availability function (the relation between automobile availability and the hybrid housing variable discussed above);

2. the use of the function to obtain the automobile availability variable; and

3. the use of the automobile availability variable.

Derivation of the Automobile Availability Function

There are two methods for deriving the automobile availability function. The user can either (a) calibrate the function relating automobile availability per household to the hybrid housing variable, using local census tract data, or (b) use the curves already developed in the Princeton study and presented in this report.

To calibrate a local automobile availability function, the user assembles census data on the tract level for the dependent variable, automobile availability per household, and the two components of the independent variable, average market value of homes and the percentage of homes which are owner-occupied. The user then plots each observation of automobile availability and the hybrid housing variable, one observation per census tract and fits a curve to these points.

The Princeton study suggests that separate curves should be plotted for the central city and suburban/outlying areas, and that a logical mathematical relation for these curves is of the form:

$$y^c = a + bx$$

where:

y = automobile availability per household;

x = hybrid housing variable: average market value of homes multiplied by proportion of homes that are owner-occupied;

a, b = calibration constants; and

c = a shape parameter selected to obtain the proper shape of the calibrated curve.

The procedure for deriving the automobile availability function is discussed in detail in Section II.

Use of the Automobile Availability Function

Once it is calibrated, the analyst can use the automobile availability function to derive values of the automobile availability variable at the study area zone or census block level. Zonal values of this variable can be estimated for any zone system for which zonal values of the components of the housing hybrid variable are available.

To determine zonal values of the automobile availability variable, the user first calculates a value for the hybrid housing variable for each study area zone.[1] Current housing values should be used to reflect up-to-date conditions for transportation studies, but each dollar value

---

[1] Zone structures for urban area transportation studies vary with the purpose and scope of the study, but are generally governed by the following criteria:

- Zones should be subdivided in such a way that few intrazonal trips will be made via the modes under consideration in the study.

- Zones should be defined to represent as socially and economically homogenous a group of residents as possible, since zonal travel is determined as a function of the aggregate average characteristics of its residents. Increasing uniformity within zones decreases the probability of errors when using average values for zonal values.

- Natural and artificial barriers such as rivers and expressways should serve as zonal boundaries. This tends to create homogenous zones with residents who have similar travel behavior patterns.

- Data for the independent variables must be available at the zonal level for forecasting.

should be deflated to the value at the time of calibration (1970, if census data are used as in the Princeton study).[1]

The current value of automobile availability for each zone is then estimated by entering the hybrid variable value on the appropriate calibrated curve. Zonal values of auto availability can then be used in the travel demand model the user has selected for analysis purposes.

Use of the Automobile Availability Variable

Six categories of models in which the auto availability variable can be used are mentioned above. The use of this variable is illustrated in Section III by describing the procedure for applying this analytical aid in trip generation and work-trip modal split models.

DATA AND INFORMATION NEEDED

The data input required to apply this analytical aid are the values of the two components of the hybrid housing variable on an appropriate small-area level (census block or study area zone). The most likely source of data for these two components, housing market values and the proportion of homes that are owner-occupied, is local tax assessment records.

If the analyst elects to use local data to calibrate an automobile availability function, data at the census tract level are also required for these two variables as well as for the automobile availability variable. All three variables can be obtained from information provided in the U.S. Census of Population and Housing census tract reports for each SMSA in the United States.

ANALYSIS OUTPUT

The output of this analytical aid is a set of values for the variable, automobile availability per household, at the study area zone level which can be used in travel demand models.

---

[1] The use of a ratio of consumer price indices from the two periods as a deflation factor is suggested. For 1975, the deflation factor is .72, which equals 116.3 (the 1970 national consumer price index) divided by 161.2 (the 1975 national consumer price index). The use of local consumer price indices is recommended, if they are readily available.

II. DERIVING THE AUTOMOBILE AVAILABILITY FUNCTION

This section presents the relation between automobile availability and the hybrid housing variable which was derived in the Princeton study and describes a procedure by which analysts in other areas can calibrate similar relations using locally available data.

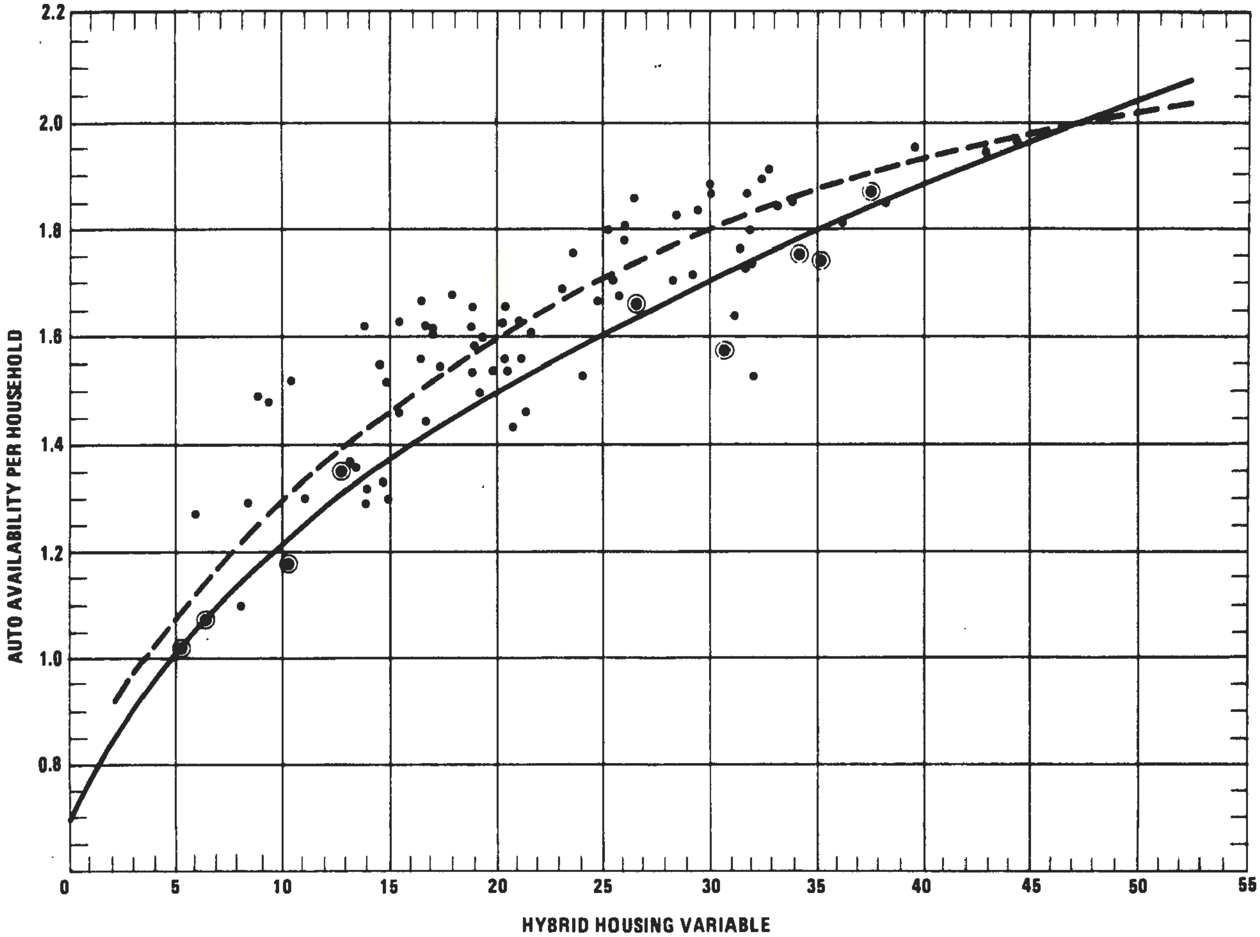
THE CALIBRATED FUNCTION

The automobile availability functions calibrated by Princeton University for Morristown and Morris County are presented in Figure 1. Along the x axis in this figure is the hybrid housing variable, the product of the average market value of homes in a zone (in $1,000s) and the proportion of households in the zone which are owner occupied. Along the y axis is the dependent variable, automobile availability per household in a zone.

The Princeton study noted two significant findings during the development and use of this function. First, the automobile availability variable for a given value of the hybrid variable tended to be higher for the suburban and outlying areas than for the central city, even though the relation between auto availability per household and the hybrid housing variable was strong for data from all 88 census tracts in Morris County ($r^2 = 0.90$).[1] This difference is shown by the two curves illustrated in Figure 1. Since there was sufficient calibration data for both geographic areas, two separate curves were plotted to reflect more accurately the variations in the effect of the hybrid housing variable on automobile availability.

Second, for both the central city and the suburban/outlying cases, the relation between the automobile availability and hybrid housing variables was nonlinear. Predictably, automobile availability always increases as the value of the hybrid housing variable increases. The marginal increase in automobile availability, however, tends to decrease as the value of the hybrid housing variable increases. This tendency seems reasonable, since auto ownership increases rapidly

[1]This may be due to several factors: larger family size; dispersed work, shop, and social/recreational opportunities; and poorer transit service in the suburbs requiring more automobiles; higher incomes in the suburbs; and higher population densities and more walking trips in the central city.



KEY:

◉ Morristown or Morris Township (central city) census tracts

• Other Morris County (suburban/outlying areas) census tracts

– – Curve fitted for Morristown and Morris Township (use for central city)

— Curve fitted for Morris County (use for suburban/outlying areas)

NOTE:

Hybrid housing variable equals average market value in $1,000s multiplied by proportion of households owning homes.

SOURCE:

Michael Holoszyc and Dale Miller, "The Planning and Design of a Demand Responsive Transportation System for Morristown and Morris Township" (report prepared for the Transportation Program, Princeton University, March 1975).

FIGURE 1: CALIBRATED AUTOMOBILE AVAILABILITY FUNCTION

as affluence increases in the lower and middle income levels but levels off at just over two automobiles available per household in the higher income levels.

The curves shown in Figure 1 should be used if time, budget, or data limitations do not permit calibrating a local automobile availability function. The analyst should test these curves, however, with selected local data to determine if they reflect local experience and are therefore appropriate for local use.

If the curves are used as shown, local housing values should be converted to 1970 dollars to reflect the true socioeconomic status represented by a given value of the hybrid variable. (This is necessary because the curves were calibrated using 1970 census data. Alternatively, the hybrid housing variable shown on the x axis could be converted to current dollar values.)

CALIBRATING A LOCAL AUTOMOBILE AVAILABILITY FUNCTION

Ideally, the analyst can calibrate a local automobile availability function using local census data. To calibrate the automobile availability function, the following procedure is followed:

1. Collect and reduce census data.

2. Plot data and determine how many curves should be fit (determine if more than one relation exists between automobile availability and the hybrid housing variable for central city, suburban/outlying areas, or other study area jurisdictions).

3. Estimate the local automobile availability function(s) using statistical regression techniques.

Instructions for performing each of these steps are given below.

Collect and Reduce Census Data

The census data required to calibrate the automobile availability function is available from the U. S. Census of Population and Housing census tract reports for all SMSAs in the United States. The four data items used for calibration are recorded for each census tract in Tables H-1 and H-2 of each SMSA census tract report. These data items and their respective sources are:

- number of owner-occupied housing units in the tract (the first entry in the "Tenure, Race, and Vacancy Status" section of Table H-1);

- number of occupied housing units in the tract (the subtotal after the "Rooms" section of Table H-1);

- median market value of housing units in the tract (the last entry in the "Value" section of Table H-1); and

- number of occupied housing units with zero, one, two, and three or more automobiles available. (the "Automobiles Available" section of Table H-2).

A worksheet should be prepared for data collection and reduction such as the sample shown for Portland, Maine in Table 1. The four data items taken from the census tract reports are recorded in Section A of the form for each tract, for the total SMSA, for the central city, and for other portions of the SMSA that are of interest to the user. The summary statistics can be used to help determine whether one, two, or more curves are appropriate.

Section B in Table 1 shows how the data are reduced for calibration. Note that the Census Bureau's "occupied housing units" is used to represent the households variable used in this analysis. First, the components of the hybrid housing variable are calculated for each census tract; next the hybrid housing variable is computed for each tract and the total number of automobiles available in the tract is computed. Note that "three automobiles" is used for the category "three or more" in the weighting calculation. Finally, the dependent variable, automobiles available per household, is calculated for each tract.

Plot Data and Determine How Many Curves Should be Fit

The reduced census tract data should next be plotted on a graph, as shown previously in Figure 1. The dependent variable, plotted along the y axis, is automobile availability per household (variable B4 in Table 1). The independent variable, plotted along the x axis, is the hybrid housing variable (variable B2 in Table 1).

The analyst should plot, at least initially, the individual points for central city data and suburban data to determine whether two or

TABLE 1

EXAMPLE OF CENSUS DATA COLLECTION AND REDUCTION

DATA ITEM	IN CUMBERLAND COUNTY			PORTLAND					
	Total	Portland	Balance	Tract 001	Tract 002	Tract 003	Tract 004	Tract 005	Tract 006
A. DATA COLLECTION									
1. Number of Owner-Occupied Housing Units (from Table H-1)[1]	27,571	10,062	17,509	270	257	–	46	162	60
2. All Occupied Housing Units (from Table H-1)[1,2]	45,512	22,780	22,732	905	1,102	–	335	592	723
3. Median Housing Unit Market Value (from Table H-1)[1]	$16,700	$16,600	$16,800	$9,300	$11,100	–	$7,100	$7,300	$6,100
4. Automobiles Available[3]									
a. 1	23,356	11,347	12,009	547	567	–	109	257	139
b. 2	11,523	3,777	7,746	61	101	–	6	50	46
c. 3 or more	1,543	483	1,060	7	22	–	–	–	7
d. none (from Table H-2)	9,090	7,173	1,917	290	412	–	216	290	540
B. REDUCTION									
1. Proportion of Housing Units Owner Occupied (A1 ÷ A2)	.606	.442	.770	.298	.233	–	.137	2.74	.083
2. Hybrid Housing Variable [(A3 ÷ 1,000) x B1]	10.12	7.33	12.94	2.77	2.59	–	0.97	2.00	0.51
3. Automobiles Available per Census Tract [(A4a x 1) + (A4b x 2) + (A4c x 3)]	51,031	20,350	30,681	690	835	–	121	357	252
4. Automobiles Available per Household (B3 ÷ A2)	1.12	0.89	1.36	0.76	0.76	–	0.36	0.60	0.35
C. PLOT CENSUS TRACT DATA ON GRAPH: X Axis: Variable B2; Y Axis: Variable B4									

[1] 100% sample in 1970 census.

[2] Corresponds to number of households, according to census definition.

[3] 15% sample in 1970 census.

SOURCE:
U.S. Bureau of the Census, Census of Population and Housing: 1970 Census Tracts, Final Report PHC (1)-164 Portland, Maine SMSA (Washington, D.C.: U.S. Government Printing Office, 1972).

more curves are appropriate. In the Princeton study, two distinct curves were appropriate as shown in Figure 1. If enough data points are available, the analyst may wish to further subdivide the data into smaller jurisdictions. Groups of data can always be aggregated to form larger groups if no significant trends become apparent in the smaller groups.

Estimate Local Automobile Availability Function(s)

Statistical regression techniques can be used to determine the functional relation between automobile availability and the hybrid housing variable.[1] First, the analyst must determine the form of the curve or curves which should be fit. The data must then be transformed as shown below to develop the regression equationss Finally, regression coefficients are computed and the functional relation is determined. No attempt will be made here to explain the statistical basis for regression analysis. Only the basic procedure will be given.

Figure 1 shows the characteristics of the automobile availability function developed for Morristown and Morris County. An automobile availability function calibrated for any other area should exhibit the same characteristics shown in this figure:

- the function continually increases as the value of the hybrid variable increases; and
- the amount of the increase becomes smaller; that is, it is marginally decreasing.

---

[1]A complete discussion of regression procedures and terminology can be found in a variety of texts:

- John E. Freund, Mathematical Statistics, (Englewood Cliffs: Prentice-Hall, Inc., 1971).
- Hubert M. Blalock, Social Statistics (New York: McGraw-Hill Book Co., 1960).
- Norman R. Draper, and H. Smith, Applied Regression Analysis (New York: John Wiley, 1966).

Several alternative curves which exhibit these characteristics are shown in Figure 2. Each of these curves is of the form:

$$y^c = a + bx \qquad (1)$$

where c = 2, 3, 4 for the three different curves.

As the exponent c increases, the curve bends more, representing a faster rate of marginal decrease in the automobile availability variable as a function of the hybrid housing variable.

Specific curves can be compared statistically by computing and comparing correlation coefficients and standard errors of estimate if the data do not obviously suggest a particular functional relation.

The regression equation is determined by computing the coefficients of a least-squares regression line. The first step in this process is to establish a linear regression equation of the form:

$$z = a + bx \qquad (2)$$

Nonlinear relations such as those shown in Figure 2 can be converted to linear relations by transforming the dependent variable according to the function:

$$z = f(y) \qquad (3)$$

where y = dependent variable (automobile availability per household);

z = transformed variable; and

f = function appropriate for the curve selected.

For example, if the curve selected is:

$$y^3 = a + bx, \qquad (4)$$

then the transformation to a linear form is represented by the function:

$$z = y^3. \qquad (5)$$

Each data point y value must be transformed to a z value according to equation 5 before determining the linear regression equations.

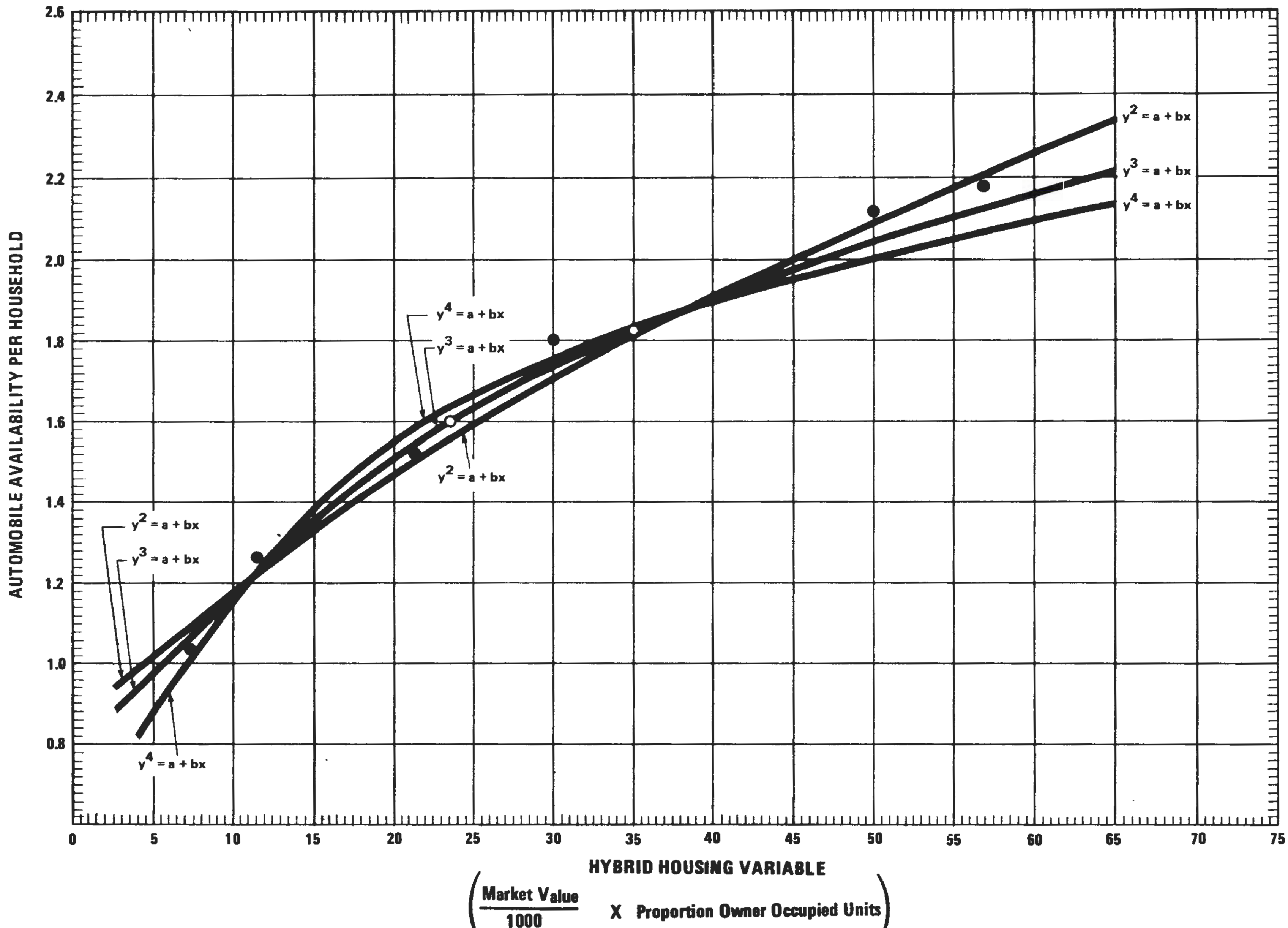


FIGURE 2: SAMPLE CURVES FOR AUTOMOBILE AVAILABILITY FUNCTION

The equations for determining the a and b coefficients which will produce the best fitting regression curve for equation 2 are the following:

$$\frac{(\Sigma z)\ (\Sigma x^2)\ -\ (\Sigma x)\ (\Sigma xz)}{n\ (\Sigma x\ )\ -\ (\Sigma x)^2} \qquad (6)$$

$$\frac{n\ (\Sigma xz)\ -\ (\Sigma x)\ (\Sigma z)}{n\ (\Sigma x^2)\ -\ (\Sigma x)^2} \qquad (7)$$

where z = transformed y value of automobile availability per household for each data point observation (for each zone);
x = x values (hybrid housing variable) for each data point observation (for each zone);
xz = the product of x and z values for each data point observation (for each zone);
Σ = summation over all data point observations; and
n = number of data point observations (number of zones).

To plot the resulting equation, the calculated values must be transformed back to their original form by applying the inverse of the function in equation 3. For example, for the transformation expressed in equations 4 and 5, the calculated z values must be substituted into equation 8:

$$y = \sqrt[3]{z} \qquad (8)$$

The resulting equation can be compared with the original data to determine the degree of fit using the correlation coefficient (r). The correlation between the estimated values of y and the observed value of y (y_o) is determined by the following equation:

$$r = \frac{n\Sigma yy_o\ -\ \Sigma y\Sigma y_o}{\left\{\left[(n\Sigma y^2\ -\ (\Sigma y)^2\right]\left[n\Sigma y_o^2\ -\ (\Sigma y_o)^2\right]\right\}^{1/2}} \qquad (9)$$

A correlation coefficient closest to the value 1 indicates the strongest correlation or functional relationship.

A brief example of the application of the procedure described above is illustrated in Table 2. Suppose the x and y values shown in Table 2 and illustrated in Figure 2 have been computed from census data. The array of data point observations, as shown in Figure 2, suggests a relation represented by equation 1:

$$y^c = a + bx$$

with the shape parameter $c = 2$. The automobile availability variable must therefore be transformed according to equation 3 as shown in the last column in Table 2 before applying the regression formula (equations 6 and 7):

$$z = y^2$$

Next the regression coefficients are computed according to equations 6 and 7 as shown in Table 2. The correlation coefficient between the estimated values of y (as determined from the regression equation) and the observed values of y are then determined as shown in Table 2.

TABLE 2

EXAMPLE OF APPLYING REGRESSION TECHNIQUE

CENSUS TRACT	HYBRID HOUSING VARIABLE (x)	AUTO AVAILABILITY PER HOUSEHOLD (y)	TRANSFORMED AUTO AVAILABILITY ($z = y^2$)
1	7.0	1.04	1.082
2	21.0	1.52	2.310
3	49.0	2.06	4.244
4	35.0	1.84	3.386
5	50.0	2.12	4.494
6	30.0	1.80	3.240
7	12.5	1.28	1.638
8	23.5	1.60	2.560
9	56.5	2.18	4.752
	284.5		27.706

$\Sigma\ x =$ 284.5

$\Sigma\ z =$ 27.706

$\Sigma\ x^2 = (7.0)^2 + (21.0)^2 + (49.0)^2 + \ldots + (56.5)^2 = 11{,}416.75$

$\Sigma\ xz = (7.0)\ (1.082) + (21.0)\ (2.310) + (49.0)\ (4.244) + \ldots + (56.5)\ (4.752) = 1{,}053.573$

$n = 9$

$$a = \frac{(\Sigma z)(\Sigma x^2) - (\Sigma x)(\Sigma xz)}{n(\Sigma x^2) - (\Sigma x)^2} = \frac{(27.706)\ (11{,}416.75) - (284.5)\ (1{,}053.573)}{9(11{,}416.75) - (284.5)^2} = 0.760$$

$$b = \frac{n(\Sigma xz) - (\Sigma x)(\Sigma z)}{n(\Sigma x^2) - (\Sigma x)^2} = \frac{9(1{,}053.573) - (284.5)(27.706)}{9(11{,}416.75) - (284.5)^2} = 0.073$$

Regression equation:

$z = y^2 = 0.760 + 0.073x$

To plot the curve, calculate sample points:

x	z (computed from regression equation)	y ($\sqrt{z}$)
5	1.125	1.061
10	1.490	1.221
20	2.220	1.490
30	2.950	1.718
40	3.680	1.918
50	4.410	2.100
60	5.140	2.267

Plot x, y pairs.

The correlation coefficient is calculated as follows:

CENSUS TRACT	AUTO AVAILABILITY PER HOUSEHOLD (y)	
	OBSERVED (y_0)	ESTIMATED (y)
1	1.04	1.127
2	1.52	1.514
3	2.06	2.083
4	1.84	1.821
5	2.12	2.157
6	1.80	1.718
7	1.28	1.293
8	1.60	1.573
9	2.18	2.210

$\Sigma\ yy_0 = (1.04)(1.127) + \ldots + (2.18)(2.210) = 27.770$

$\Sigma\ y = 1.127 + \ldots + 2.210 = 15.496$

$\Sigma\ y_0 = 1.04 + \ldots + 2.18 = 15.44$

$\Sigma\ y^2 = (1.127)^2 + \ldots + (2.210)^2 = 27.852$

$\Sigma\ y_0^2 = (1.04)^2 + \ldots + (2.18)^2 = 27.706$

$$r = \frac{9(27.770) - (15.496)(15.44)}{\left\{[(9)(27.852) - (15.496)^2]\ [(9)(27.706) - (15.44)^2]\right\}^{1/2}} = 0.981$$

III. USING THE AUTOMOBILE AVAILABILITY FUNCTION

Once an automobile availability function is calibrated, automobile availability per household on a zonal level may be determined for use in an appropriate travel demand forecasting model by:

- collecting or assembling data for the two hybrid variable components at the appropriate zonal level for the forecast year;
- computing the value of the hybrid variable for each zone; and
- entering the value on the calibrated curve to obtain the estimated zonal value of automobile availability per household.

In some cases, the independent variables may be only available on a level smaller than the transportation planning or traffic zone. This data must then be aggregated to the study area zone level. To do this, the analyst must determine the value of the hybrid variable for each study area zone by calculating an average over all of the smaller zones in the study area zone.[1]

In this section, two examples are presented to illustrate the procedure for using the automobile availability function in trip generation models and work-trip modal split models.

TRIP GENERATION MODELS

Automobile availability per household has traditionally correlated well with trip-making propensity. Trip generation models, therefore, often include automobile availability as an independent variable.

Trip generation models have been calibrated locally for most urbanized areas as a part of each area's ongoing transportation

---

[1] The average value for the study area zone would in most cases be determined by calculating the weighted average of each of the smaller zones in the study area zone. Weighting should, if possible, be by number of households, the population unit for the independent variables.

planning program. These models are generally based on data from local home interview or telephone surveys. Typically, each model was developed and used by the local metropolitan planning organization or the city or county planning agency. The particular model used in this example, originally calibrated for a 1956 Modesto, California, transportation study, was taken from a general transportation planning text and modified for use in the Princeton study.[1]

The model, shown in Figure 3, is a graphical cross-classification model in which the number of daily trips (from home for all trip purposes for a particular zone) is estimated as a function of two independent variables, zonal values of automobile availability per household and household size. As shown on Figure 3, automobile availability per household is represented on the x axis. Individual curves are plotted on the graph for different values of the household size variable.

Trip generation increases as both household size and automobile availability increase. Furthermore, when either of these variables remains constant, an increase in the other variable results in increases in travel but at a diminishing rate. This suggests, in the case of increasing household size, that some trips serve the purposes of the entire household; thus, larger households do not make proportionately more trips per person. In the case of increasing auto availability, the most significant increase in trip-making occurs when the first car is acquired; trip generation increases are smaller with the acquisition of additional automobiles.

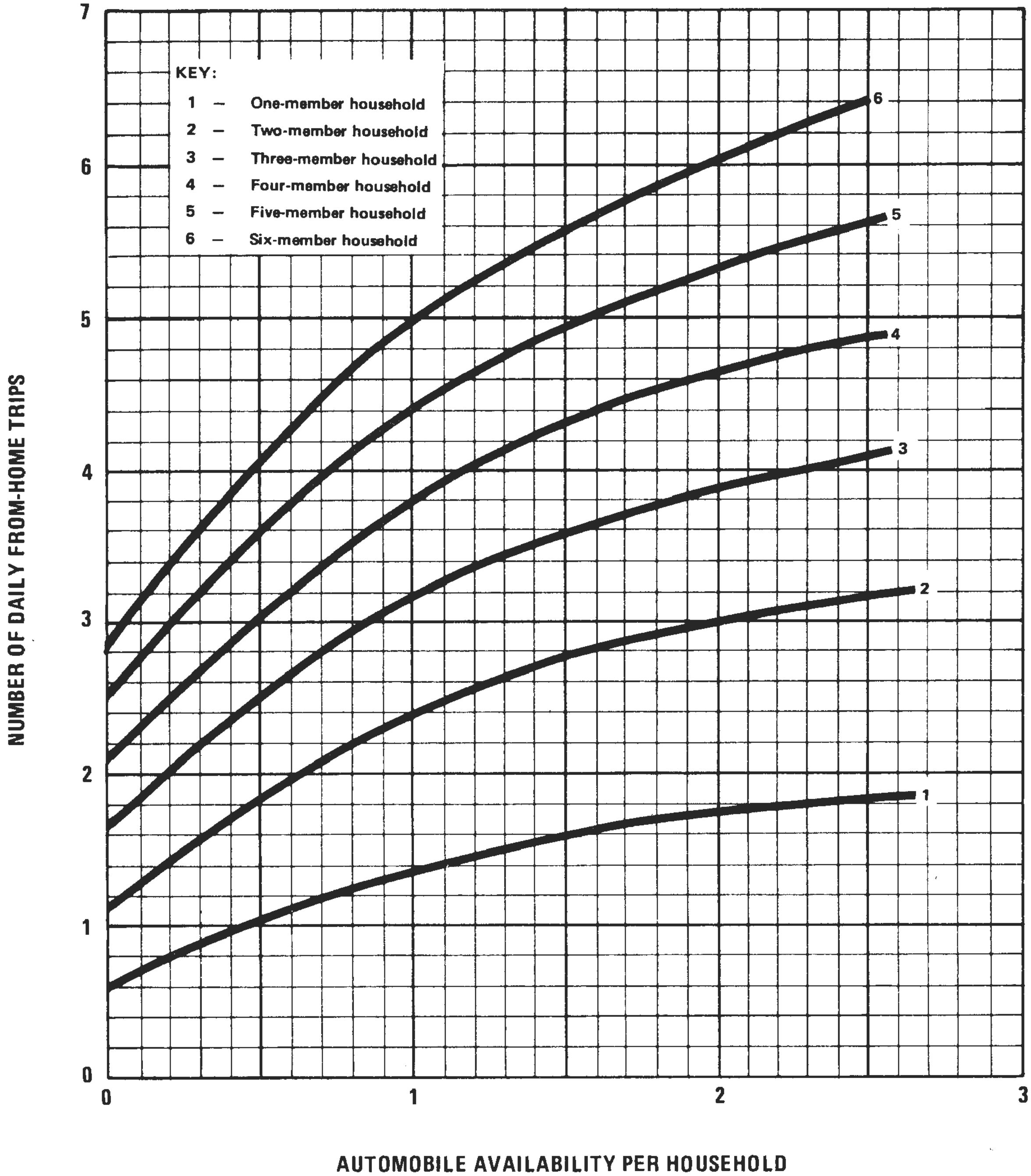
Figure 3 is used by entering the x axis at the value of the automobile availability per household for a particular zone (as determined from the auto availability function). Moving vertically up from this point, the user finds the point at or between curves which represents the average household size for the particular zone. The number of daily from-home trips for the zone can then be read on the y axis.

WORK-TRIP MODAL SPLIT MODELS

The second example of the use of the automobile availability function is illustrated by a work-trip modal split model. This type of

---

[1] Walter Y. Oi and Paul W. Shuldiner, An Analysis of Urban Travel Demand (Evanston: Northwestern University Press, 1962).



SOURCE: Michael Holoszyc and Dale Miller, "The Planning and Design of a Damand Responsive Transportation System for Morristown and Morris Township" (report prepared for the Transportation Program, Princeton University, March 1975).

FIGURE 3: EXAMPLE OF A TRIP GENERATION MODEL – TRIPS PER HOUSEHOLD VERSUS AUTOMOBILE AVAILABILITY AND HOUSEHOLD SIZE

model estimates modal split, the percent of work trips made via auto, transit, and other modes on a trip interchange basis. The total number of work person trips made between zones are split between transit, walking, and auto, plus other modes. To use this model, the user must generate or obtain a work trip table and zonal values for each of the input variables used in the model.[1]

Automobile availability per household correlates well with modal split. Tripmakers in households with two or more automobiles available have a greater propensity for making a given trip by automobile than those in households wih one automobile or no automobile. This relationship is generally true for all trip purposes. For work trips, for example, significant modal split differences occur in households with zero, one, and two or more automobiles. Generally, the second car in a household is purchased primarily for transportation to and from work.

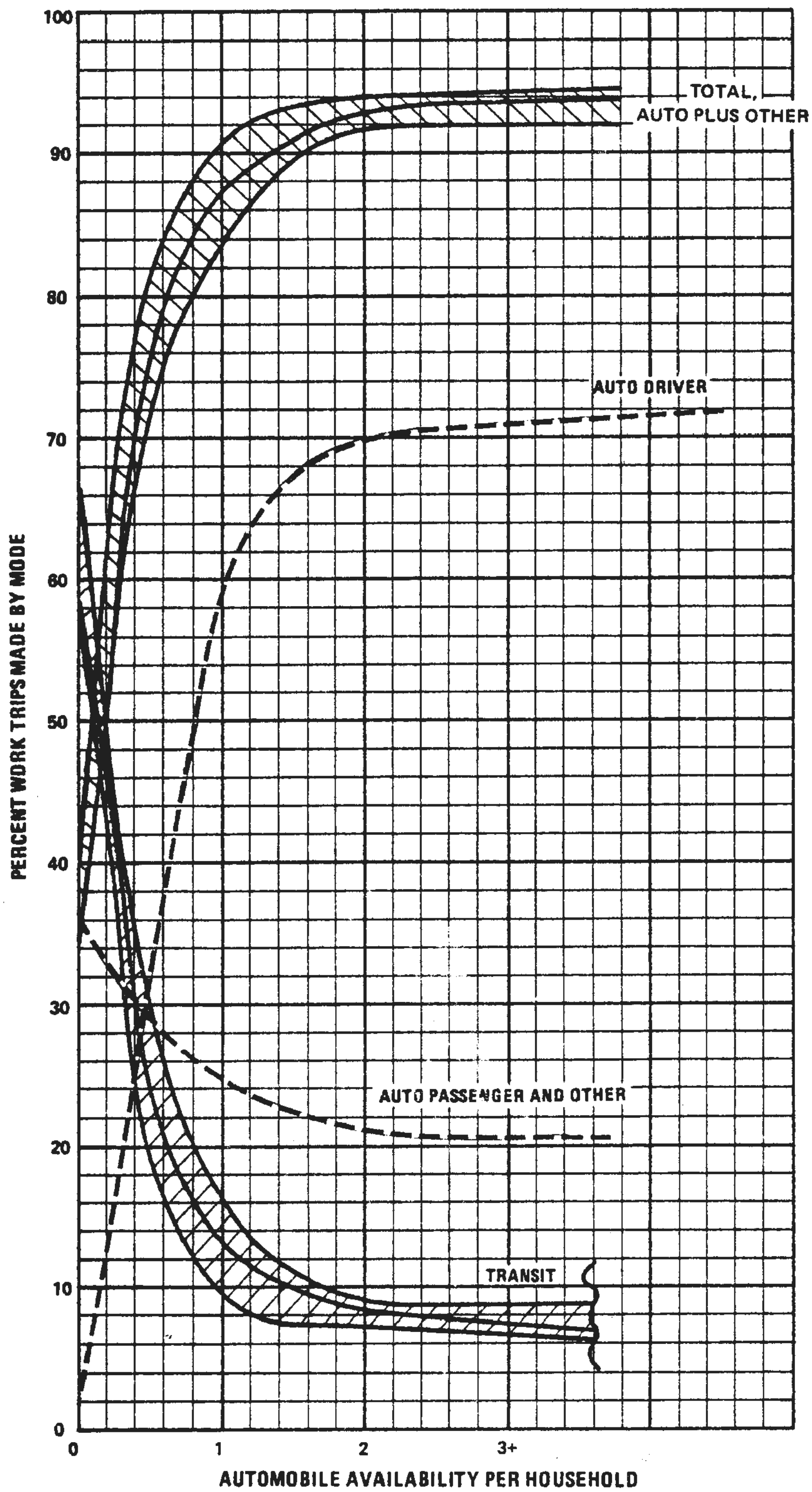
Modal split models have been calibrated locally for many urbanized areas. They are typically based on data from local home interview or telephone surveys, but 1970 census journey-to-work data can also be used for calibration.

The model used in this example was developed from small urban area data collected in the Nationwide Personal Transportation Study.[2] The model was developed to estimate work-trip modal split in small urban areas, as a part of the process for planning transit in urban areas with less than 500,000 in population. For this model, tabulations were made for both suburban and small nonsuburban communities to derive modal share ranges for each level of the independent variable, automobile availability per household. The resulting range in values has been preserved in the multimodal plots shown in Figure 4 and can be assumed appropriate for determining high and low estimates of modal split.

---

[1] For small urban areas, an abbreviated work trip table is sufficient for the travel information requirements. This would consist of daily trip interchanges between each residence zone and only the dominant employment zones, notably the central business district, industrial parks, large factories, large office complexes, universities, hospitals, and shopping centers where transit potential exists.

[2] Nationwide Personal Transportation Study, data collected by U.S. Bureau of the Census for the Federal Highway Administration, 1969.



NOTES:

1. Relationship is for small urban areas only.
2. Transit service for the trip to work must be available within 6 blocks of residence.
3. Transit service must be fixed-route, fixed-schedule service only.
4. Other mode includes truck, motorcycle, bicycle, and taxi and excludes walk.
5. Curves are derived from small urban area data from the Nationwide Personal Transportation Study.

SOURCE:

Analyzing Transit Options for Small Urban Communities (Urban Mass Transportation Administration, January 1978).

FIGURE 4: EXAMPLE OF WORK TRIP MODAL SPLIT MODEL – MODAL SHARE AS A FUNCTION OF AUTOMOBILE AVAILABILITY

As illustrated in Figure 4, the work-trip modal split for two modes (automobile and transit) and for three modes (automobile driver, automobile passenger, and transit passenger) is expressed as a function of automobile availability per household. The curve shown in Figure 4 is based on the assumption that fixed-route, fixed-schedule transit service to a worker's place of employment is available within six blocks of the traveler's residence.

The model is applied by entering the x axis at the value of the automobile availability variable for the study area zone. High, low, and average estimates of transit and auto modal split for a particular trip interchange can then be read on the y axis. The model is applied for only those trip interchanges served by transit during peak periods. Zones may be split if the service area within six blocks of the transit routes serving a zone does not cover the entire zone.

Note that bus service availability for the work trip is the only level of service variable in this model. Because other variables, such as travel times and service frequency, are not included, the model should be used primarily for transit feasibility studies and evaluation of major system changes. It is not sufficiently sensitive to permit evaluation of system improvements or alternatives other than service coverage.

IV. SHORTCOMINGS AND LIMITATIONS

When deciding if the method described here for deriving automobile availability is applicable in a particular situation, the user should be aware of the potential limitations of the technique. These limitations fall in the following six areas:

- theoretical validity;
- transferability of calibrated curves;
- validity of census data;
- variations in market values;
- variations in data; and
- curve selection.

VALIDITY OF THE THEORY

The automobile availability per household variable correlated well with the hybrid housing variable in the Princeton study. The relationship is plausible, but there is no assurance that the correlation would be as strong if calibrated for another urban area. Consequently, it should be calibrated locally, if possible, or at least validated with a limited number of data points.

TRANSFERABILITY OF CALIBRATED CURVES

No attempt has as yet been made to transfer the calibrated curves to another urban area. Distribution of the variables involved tend to vary widely from one area to another, and the specific relationship developed in the Princeton study, therefore, may not be transferable. A local calibration would correct this limitation. A validation with a limited number of data points is recommended as a minimal checking procedure.

VALIDITY OF CENSUS DATA

Census data may require updating before being used for either calibration or forecasting if the analysis is performed five to nine years after the census year. Generally, however, the relationship determined by a curve calibrated for a census year is reliable as long as the calibration year values for the housing market value variable are converted to forecast year prices, or the forecast year prices are converted back to calibration year values. The conversion can be made using local consumer price index ratios.

VARIATION IN MARKET VALUES

It is possible that the conversion described above might not accurately project the automobile availability function from the calibration year to the forecast year. Market value growth rates could vary for different neighborhoods in the urban area. A given data point, for example, could rise or fall on the curves when its hybrid variable value for the calibration year is converted to a forecast year, whereas both the real values of the hybrid variable and automobile availability per household variable remained at the same point on the curve. This reflects the fact that, in general, the reliability of any given calibrated relation diminishes over time.

VARIATIONS IN DATA

As with most analytical models, selected data observations show wide variations from a calibrated relationship. One could find zones, for example, which have low values for the hybrid housing variable, but high values for automobile availability per household--perhaps a university campus, or a residential neighborhood near a military installation in which a large percentage of housing units are rented. A plot of data points should reveal such anomalies which can be eliminated for calibration and adjusted for estimating purposes.

CURVE SELECTION

Curves may be selected from graphical or visual "best fit" judgements or from linear regression statistical curve fitting techniques. In any of these methods, the analyst should be aware that the range of values represented by the selected curve may be quite wide. In situations where the automobile availability measure is critical to the development of travel demand estimates, a "high" and "low" curve may be used to determine the range of possible forecasts due to variability in the selected curve.

☆ U. S. GOVERNMENT PRINTING OFFICE : 1979 281-568/59

17353