Reference

Report No. DOT-TSC-OST-76-10

ESTIMATING THE EFFECTS OF URBAN TRAVEL POLICIES

Charles River Associates
1050 Massachusetts Avenue
Cambridge, Ma. 02138

APRIL 1976
FINAL REPORT

Prepared for
U.S. DEPARTMENT OF TRANSPORTATION
OFFICE OF THE SECRETARY
OFFICE OF THE ASSISTANT
SECRETARY FOR POLICY,
PLANS AND INTERNATIONAL AFFAIRS
TRANSPORTATION ENERGY POLICY STAFF
WASHINGTON, D.C. 20590

Technical Report Documentation Page

1. Report No. DOT-TSC-OST-76-10	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle ESTIMATING THE EFFECTS OF URBAN TRAVEL POLICIES		5. Report Date April 1976
		6. Performing Organization Code
7. Author(s) Frederick C. Dunbar (Principal Investigator)		8. Performing Organization Report No. DOT-TSC-OST-76-10
9. Performing Organization Name and Address Charles River Associates* 1050 Massachusetts Avenue Cambridge, MA 02138		10. Work Unit No. (TRAIS) OP602/R6817
		11. Contract or Grant No. DOT-TSC-964
12. Sponsoring Agency Name and Address U.S. Department of Transportation, Office of the Secretary, Office of the Assistant Secretary for Policy, Plans and International Affairs, Transportation Energy Policy Staff, Washington, D.C. 20590		13. Type of Report and Period Covered Final Report February 1975-March 1976
		14. Sponsoring Agency Code

15. Supplementary Notes
*Under contract to: U.S. Department of Transportation, Transportation Systems Center, Kendall Square, Cambridge, MA 02142

16. Abstract

This report presents models and procedures for quick evaluation of transportation policy options on urban travel behavior. The methods described in this report can be used to estimate the travel demand effects of a wide variety of transportation policy instruments with currently available data in a matter of hours, or minutes, with the aid of a calculator.

To evaluate the effects of a transportation policy, travel is separated into work and nonwork purposes. The work travel section of the report describes procedures for applying disaggregate logit models to generally available grouped data. To analyze the effects of policies on nonwork travel, a disaggregate travel demand model is estimated which is designed to be broadly applicable to a variety of planning and data contexts.

Both the work and nonwork trip demand models and procedures are exercised on sets of policy issues which are of current interest, including gasoline taxes, parking restrictions, transit service improvements and the introduction of new modes. Where appropriate, travel demand elasticities with respect to level of service changes are computed.

17. Key Words Travel Demand, Transportation Policy, Urban Transportation, Evaluation, Aggregate Forecasting, Disaggregate Demand Model	18. Distribution Statement Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 196	22. Price

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

PREFACE

The work described in this report was supported by the Transportation Systems Center's Independent Research and Development Program and by the Transportation Energy Policies Project, sponsored by the Transportation Energy Policy Staff in the Office of the Secretary. During the course of TSC studies involving effects on urban travel of proposed policies related to energy and environmental considerations, and to research and evaluation of new transportation modes and services, a need for tools that could be used easily and quickly for prediction of travel demand effects was identified. The purpose of this project was to develop and demonstrate such methods for planning and policy evaluation use.

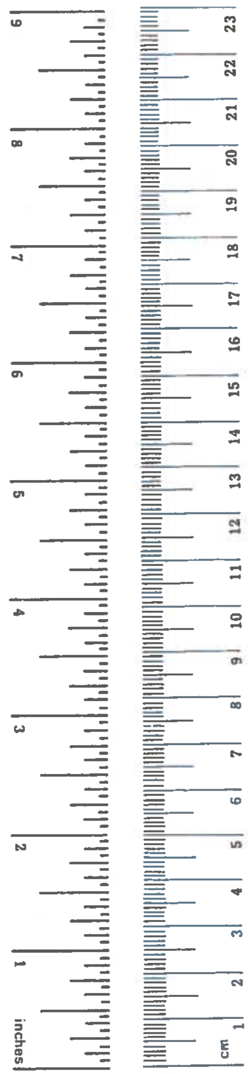
Charles River Associates gratefully acknowledges the aid it received from the Transportation Systems Center staff, particularly Donald E. Ward, in referring data and reviewing report drafts. Also, we thank the personnel of the Bureau of Census and Federal Highway Administration for their efforts in helping us understand the Nationwide Personal Transportation Survey.

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

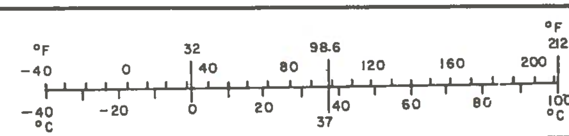
Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in = 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



CONTENTS

Section | Page

CONTENTS (CONTINUED)

CONTENTS (CONTINUED)

TABLES

TABLES (CONTINUED)

SUMMARY

Introduction

This report presents the results of a study to develop models and procedures for the quick evaluation of transportation policy options on urban travel behavior. The methods and tools for policy analysis described in this report can be used to estimate the travel demand effects of a wide variety of transportation policy instruments with currently available data in a matter of hours, or minutes, with the aid of a calculator.

To evaluate the effects of a transportation policy, travel purposes are separated into two broad groups, work and nonwork, reflecting differences between each in the structure of underlying tripmaking behavior. It was determined during the course of the study that existing mode choice models for work trips could be applied for quick policy evaluation but that currently available models of nonwork trip behavior were ill suited to the task. Consequently, the work travel section of the report describes procedures for applying probability choice models to generally available grouped data. However, to analyze the effects of policies on nonwork travel, it was necessary to estimate a simplified travel demand model which was designed to be broadly applicable to a variety of planning and data contexts.

Both the work and nonwork trip demand models and procedures were exercised on sets of policy issues which are of current interest. These included gasoline taxes, parking restrictions, transit service improvements and the introduction of new modes. Where appropriate, travel demand elasticities with respect to level of service changes are computed and validated on the results of previous research.

Models and Procedures

A review of data and source materials revealed that little was available to help the policy analyst who required quick, if approximate, estimates of the demand effects of potential policy actions. Though there is a large body of transportation related data for many separate urban areas, it varies widely in quality and is kept in cumbersome form. The only reasonably current national data base on household travel behavior is the 1969 Nationwide Personal Transportation Survey (NPTS). Additionally, the present generation of behavioral demand models -- multinomial logit models estimated on disaggregated observations -- cannot be applied directly to these data sources in order to predict accurately the effects of changes in transportation level of service. Though the journey to work is reasonably well represented by such models there is the further problem that multinomial logit models of nonwork travel are difficult to use and have, to date, failed to demonstrate their validity. The models and procedures developed in this study for work travel and nonwork travel are described separately below.

Work Travel

To apply existing probability choice models of work trip behavior to currently available data entails three major problems:

Aggregation. Quick analysis often requires data which summarizes a large amount of travel information into a relatively small number of data elements. The application of the logit model to such data usually leads to biased predictions.

Transferability. The available models have been estimated on data from specific urban localities and there is some question as to whether such models can be generalized to other urban areas.

New Modes. Typically, the most reliable models have been estimated on a limited number of mode choices (bus and auto-drive-alone) whereas the full set of mode choices available to a tripmaker includes auto passenger travel and, in some contexts, alternative forms of transit.

After a review of the performance of various existing logit models of work trip modes split, it was decided that the 1972 CRA model, estimated on Pittsburgh data, would be used for further analysis.[1] This model predicts the probability that an individual will choose either bus or driving alone for the usual journey to work. The probability is represented as a function of the level service variables (costs, line-haul and wait time, walk access time) of the alternative modes, the autos per worker in the household and a constant term. The transferability problem is solved, largely, by adjusting either the constant term or the coefficient on autos per worker. To apply the model to new modes, a heuristic approach to predicting the probability of taking each new mode is derived; the approach relies on assigning values for the level of service variables for each new mode and applying the model in such a way that the tripmaker faces a complete range of mode alternatives.

Various methods for dealing with the aggregation problem are developed; the approach to be recommended depends on the nature of the data base which is utilized in model application. For the NPTS data, the data file on individual trip records was cross tabulated into twelve market segments in such a way that aggregation bias was minimized; this produced a manageable data format for quick evaluation of national policy. To apply the model to typical urban data bases required adjustments in the model itself so that it could be used with sketch plan zone (or district level) data bases on travel behavior and transportation networks; these adjustments entailed the use

---

[1]The model is described in Thomas Domencich and Daniel McFadden, *Urban Travel Demand: A Behavioral Approach* (Amsterdam, North Holland Publishing: 1975)

of a simplified Taylor's series approximation (based on an approach to the aggregation problem originally developed by Talvitie)[1] and the use of the area of the zonal interchange as a variable to correct for aggregation bias. The procedures were validated on Los Angeles data.

Nonwork Travel

For nonwork travel, the range of choices available to a tripmaker is much more varied than in the case of work travel. An individual may choose between a number of alternative destinations including the alternative of going home after each stop on a journey and the person may choose the frequency of travel; this is in addition to the mode choice decision. Given this structure to nonwork travel behavior, it was concluded that a model which represented the range of travel choices as a continuum would be as appropriate as a model which represented travel choices as discrete entities. Such a model would be in a simplified form more suitable for quick estimation of policy impacts. Additionally, it was recognized that equations in linear form would have the decided advantage of overcoming the aggregation problem that occurs when nonlinear models are applied to aggregated and noncomparable data bases.

Two disaggregated simultaneous equation models of nonwork travel behavior were estimated from NPTS data. The unit of observation for these models is the travel record of an urban household over a four day period drawn randomly from the national population. Model I is a two equation model which predicts the number of automobile vehicle miles traveled (VMT's) and the number of transit trips by the household over a four day period; Model II is a three equation model which predicts the number of auto trips, the average length of auto trips and the number of transit trips made by

---

[1] Antti Talvitie, "Aggregate Travel Demand Analysis with Disaggregate Travel Demand Models," *Proceedings -- Transportation Research Forum, Vol. XIII* (October, 1973).

a household over a four day period. The independent variables in the models include transportation level of service characteristics, household socioeconomic variables and urban area specific descriptors of size.

The model parameters were used to compute direct and cross elasticities with respect to level of service and these elasticities were used to evaluate the effects of various policy scenarios on nonwork travel. Further study is necessary to apply the models directly to data bases which do not contain separate samples of transportation characteristics for auto and transit tripmakers. However, the travel behavior elasticities computed from application of the model to NPTS data appear to be useful for a wide range of policy applications. The independent variables in the model were chosen largely on the basis of being related to potential policy instruments to make the model a useful planning tool.

Policy Evaluations and Estimated Elasticities

To determine the effect of a policy on travel behavior, two computational approaches have been traditionally applied:

1) The policy can be represented by a percentage change in an independent variable, or set of variables, in the model and this is multiplied times the elasticity derived from the model to get the percentage change in travel behavior variables (VMT's, transit trips, etc.);

2) The policy can be represented by a new value for an independent variable, or set of variables, and the model is simulated to predict travel behavior at these new values and then compared to base case predictions of travel behavior to determine the percentage change in travel behavior variables.

In practice, it was found that the first approach was most suitable to the nonwork trip model and the second approach was used for the work trip model. Though the first approach is usually easier to apply, sometimes elasticities are not defined or they are best represented as functions of variables rather than unique numbers.

To demonstrate the models and procedures developed in the study, several policy scenarios were analyzed as examples. A brief summary of selected predicted effects is presented in Table S-1. Because changes in VMT's are presently of most interest in transportation policy objectives, only these effects are presented in the summary given in Table S-1.

The predicted effects of a policy on total travel is the weighted average of effects on work travel and nonwork travel. The weights equal the proportionate contribution to VMT's of each trip purpose category. From the NPTS survey, these weights are equal to:[1]

work travel = *.4158*
nonwork travel = *.5842*

In the case of transit, the national proportions are equal to:[2]

work travel = *.7207*
nonwork travel = *.2793*

---

[1] *Household Travel In The United States*, NPTS Report No. 7 (U.S. Department of Transportation, Federal Highway Administration, Washington, D.C.: December, 1972) Table A-2.

[2] Computed from statistics presented in Table 3-2 of this report and Table A-12 in *Home-to-Work Trips and Travel*, NPTS Report No. 8 (U.S. Department of Transportation, Federal Highway Administration, Washington, D.C.: August, 1973).

Table S-1

PREDICTED RESULTS OF SELECTED POLICY SCENARIOS USING NPTS DATA

Policy Scenario	Percent Change in VMT's: Work Travel	Nonwork Travel	Total Travel
100 Percent Gasoline Tax	-13	-15	-14
Regionwide Parking Tax[1]	-14	5	- 3
10 Percent Decrease in Transit Line haul and Wait Time	- 3	0	- 1
High Performance Dial-A-Ride[2]	- 1	-10	- 6
50 Percent Increase in Auto Fuel Economy	9	10	10
Transit within Six Blocks of All Households	- 7	-10	- 9

For specific urban areas, these proportions may well be different. Also, if the effects of policies on other variables, such as transit trips or auto trips, are predicted then the work trip versus nonwork trip weights will be different.

No brief summary of the policies being analyzed can explain all the assumptions employed in each scenario, and the reader is consequently referred to the body of the report for details. However, supplemental information is provided in Table S-2 which gives selected estimated travel demand direct

[1]Work travel results of this scenario were from simulations of a $1.00 parking tax on Los Angeles data. Nonwork travel results assume a 50 percent decline in the availability of free parking.

[2]The Dial-a-Ride service is available for all work trips with round trip distance less than 9.1 miles and for all nonwork trips regardless of distance.

and cross elasticities for various level of service variables. These elasticities, as well as the predicted policy effects will vary among urban areas. Using the procedures described in the study it is possible to refine these estimates for policy evaluation in a particular urban area using data which is specific to the planning region being analyzed.

As in the case of estimating the percentage of travel policies, the total travel demand elasticities are weighted averages of the work and nonwork elasticities. Gaps in both Tables S-1 and S-2 typically occur because some scenarios were not tested on both the work and nonwork trip modules. Additionally, some scenarios were analyzed in the body of the report which are not presented in these tables.

The results of the study indicate that it is possible to estimate the approximate effects of various transportation options quickly and with minimal commitment of computational resources. Oftentimes it is necessary to make assumptions based on ad hoc judgments and the estimates derived from exercising the models and procedures need to be interpreted with care.

Recommendations for Future Research

The implications for further research in the problem area of developing quick and accurate policy evaluation methods can be divided into the following three subject areas:

Data. Of highest priority is the processing of available data, or soon to be available disaggregate data bases, into formats which are easy to use and to manage for application of existing behavioral demand approaches. Also, to the extent that new data would be important in any particular planning application, it would be valuable to have procedures

Table S-2

SELECTED TRAVEL DEMAND ELASTICITIES

	Work Travel	Nonwork Travel	Total Travel
VMT's			
Direct Elasticity with Respect to:			
Gasoline Price	-.19	-.21	-.20
Auto Time per Mile	-----*	-.49	--- *
Cross Elasticity with Respect to:			
Transit Linehaul plus Wait Time	.32	.01**	.14**
Transit Access Time	.07-.15***	---**	---**
Transit Fare	-----*	.02	---*
Transit Trips			
Direct Elasticity with Respect to:			
Transit Linehaul plus Wait Time	-1.26	-.10**	-.91
Transit Access Time	-.25--.59***	---**	---**
Cross Elasticity with Respect to:			
Gasoline Price	.47	.00	.34
Auto Time per Mile	----*	.00	---*

*Work trip scenarios involving changes in these level of service variables were not analyzed.

**Nonwork transit trip time includes the access time conponent; access time was not isolated in the nonwork transit demand model.

***A range of elasticities for transit access time is given because the actual response depends upon the market segments where transit improvements are made.

for mounting a telephone or mailback survey with relatively few questions and a small sample (fewer than 200 households) and that would only take a matter of days to perform and code. Such surveys would be designed for application of existing behavioral demand models.

Models. More experimentation with various specifications of disaggregate nonwork travel demand models may yield more accurate predictions than the models presented in Chapter 3. One important factor in estimating such models will be the availability of better data. It is suggested that continuous choice models be tested.

System Performance. An issue which was not considered in this report is the equilibrium between travel demand and system performance. Many of the policies which were examined will have secondary or feedback effects which could be quantified with a system performance model such as a simplified network representation. Such models are currently not widely available nor has there been much effort to integrate them with shortcut travel demand prediction methods.

1. INTRODUCTION

The objective of this study is to develop procedures whereby policy effects on urban travel behavior can be estimated quickly and without a large commitment of resources. This chapter outlines the implications of these objectives and reviews the options available to planners at the national and urban level.

As discussed at length elsewhere, it is most appropriate to use behavioral demand models in order to determine the effects on travel of transportation system changes.[1] The specification of such models is to have the system attributes (e.g., travel times and costs) determine mode split, destination choice and frequency of travel. The effects of changes in these attributes (which are the result of policy implementation) are often represented by computing the percentage change in travel, either by mode or *in toto*, that would result. The percentage change in travel can be calculated in two ways:

First, the estimated demand models may provide an analytical form from which the elasticity can be calculated directly;

Second, the travel demand models may be used to simulate the effects of incremental changes in system performance and the projected demand would then be compared to base case demand estimates or data.

Neither of the two approaches is *a priori* preferable; which approach is used in any given circumstance depends upon the availability of data, the complexities of the demand model and the type and range of system changes which are being examined. Either approach, however, requires a well-specified and accurate behavioral demand model.

---

[1]See *Measurements of the Effects of Transportation Changes* (1972: CRA) for a complete discussion of this issue. Little elaboration of this point need be made, however, because it is now widely accepted in transportation research.

The models and procedures developed in this study have the characteristic that they can be utilized to perform policy evaluation in a short time (e.g., hours or minutes) with relatively limited computer and man-power resources. This requirement constrains the data format to be used in the application process. As will be discussed below, most currently available data are simply not processed so as to be easily manipulated in order to allow quick response by planners. Thus, in addition to model development, the study analyzes data base preparation issues.

The most easily applicable process to finding the effects of policies on travel behavior is to have a single number represent the elasticity of travel demand with respect to a change in a particular system attribute. Unfortunately, owing to the wide variety of transportation environments which face individual households and metropolitan areas, such an elasticity may be less precise than is desirable in policy evaluation. For example, the effects on auto behavior of a policy will be sensitive to alternate mode level of service; e.g., one would expect a gasoline tax to have a greater impact in reducing VMT's in urban areas with highly developed transit systems than in urban areas which are heavily auto dominated. This consideration leads to treating elasticities as though they were functions of other variables. However, ease of analysis and quick reaction requirements for policy evaluation indicate that the functional form for the elasticities should be relatively simple and have relatively few variables (i.e., data requirements). Moreover, the number of equations in order to determine the aggregate effects of a transportation policy should be kept to a minimum.

The rest of this chapter discusses a number of the above issues in more detail. The next section considers data commonly available for transportation policy evaluation. The following section briefly reviews existing behavioral demand models and their potential for policy evaluation in terms of specification and statistical validity. The final section reviews the applicability of current demand models to policy evaluation from the perspective of data availability and quick reaction requirements.

1.1 REVIEW OF EXISTING DATA

A short summary of available data for use by planners will help determine the characteristics a useful model should have. We make no comment on the quality of the data, but rather point out the suitability of the data in their existing forms for meeting the objectives of ease and quickness of model application. Four widely available data categories are discussed.

1.1.1. Nationwide Data

National data which has observations on both urban trip characteristics and level of service of modes is quite limited. The only available resource covering a wide range of trip purposes appears to be the Nationwide Personal Transportation Survey (NPTS). This survey of 6,000 households in 1969 is a valuable source of urban travel data.[1] Moreover, the survey will be updated in the late '70's so that a well-specified demand model can be recalibrated on and/or applied to current data at the national level at that time. Individual trip records

---

[1]See Appendix A for more description of the NPTS data base.

are available though characteristics of urban areas (aside from population categories) are not available and location of households has been purged from the data file and destroyed. Nonetheless, to the extent that the model is useful to national planners, methodologies have been developed in this study to apply to the NPTS data file and its summaries.

1.1.2 Regional Summary and Tabular Data

For most urban planning authorities, the basic source of travel related data is a one-shot household survey made, typically, during the 60's. These data tend to be updated with socioeconomic and land use information, particularly as they become avilable from the Census. Updated data on actual trips is usually estimated and lacks in reliability, even though checks are made using screenline counts. Nonetheless, tabular data, in the form of mode splits, frequency of trips by length, system inventories, etc., are sometimes available and may be relatively easy to manipulate given appropriate travel demand models.

It was found in this study that such data are not uniform across urban areas and considerable expense is entailed in locating the relevant items for even just the major cities. Tailoring a national model to accept these diverse data summaries would require that the model rely mostly on socioeconomic characteristics of urban areas and a few rude descriptors of transportation facilities. The usefulness of such a model as a policy evaluation device would be arguably suspect.

An alternative approach is to take existing, often disaggregate demand models and develop an approach to applying these to regional summaries of trip characteristics. This is tried in Sections 2 and 3 with mixed results.

1.1.3 Zonal Data

The vast majority of data available for transportation planners is processed at the traffic analysis zone level of aggregation. Some of this data, such as trip tables developed from urban travel planning packages, is spurious. Nonetheless, the best data on system attributes and socioeconomic variables is often stored in this format. Using an entire set of interzonal observations to apply a model should be avoided; there are typically 10^6 observations of zonal interchanges. One alternative is to utilize sketch plan zones but, even here, there tend to be between 2,500 and 10,000 such zonal interchanges. A final level of aggregation is the corridor; corridor data on travel are sometimes available but it is necessary to be creative about how one treats transportation system attributes at the corridor level.

The conclusions which emerged from our consideration of interzonal data are as follows: first, the demand models should accept zonal data because such data is widespread; second, a method of reducing the number of observations from a zonal interchange data set should be applied when the model is to be used for policy evaluation. With respect to the latter conclusion, two approaches to reducing the number of observations are potentially available: first, the data can be tabulated into summary formats; second, a small but representative sample of zonal interchanges may be selected from the entire population and set aside for special analysis. In Section 2, it is shown that the latter approach is generally more successful.

1.1.4 Disaggregate Data Sets

A small, but growing, source of data involve trip records and system attributes at the individual household level. Household survey records as they are currently maintained are of little value because the number of observations (in the tens of thousands) makes them too cumbersome to be applied easily. A sample of a smaller number, say less than 200, may provide a data base which can be analyzed in more detail with relatively small amounts of computer resources.

Currently, substantial research efforts into development and application of disaggregate demand models are underway by a number of diverse investigators throughout the country. Moreover, disaggregate data sets of appropriate specifications are not available in most urban areas. For these reasons it was deemed to be unproductive to focus on applying disaggregate demand models to disaggregate data sets.

The exception to this research strategy involves the NPTS data base described above. It is, operationally, a disaggregate data set of national proportions. However, as shown in Section 2, the application of existing disaggregate demand models to these data involves considerable data preparation and *ad hoc* model manipulation to reach desired ends. In Section 3, the NPTS data base is used with greater success to estimate a policy sensitive model on nonwork travel behavior.

1.2 REVIEW OF EXISTING BEHAVIORAL DEMAND MODELS

Historically, most research on urban travel demand models was motivated by the need for planners to evaluate proposed changes in urban transportation facilities. Only recently, owing to the requirements of implementing environmental and energy policy, has research been directed toward utilizing travel demand analysis for broad gauge or national issues. In either situation, behavioral travel demand models have been demonstrated to be more useful than the more conventional Urban Transportation Planning System (UTPS) approaches. Though behavioral demand models have been shown to be quite versatile in terms of evaluating a wide range of policy issues, this flexibility has the consequent disadvantage that the models tend to be rather complex and sometimes cumbersome to apply.

Within the planning context, the initial behavioral demand model was estimated by CRA in 1967 and has since become known as the direct demand model.[1] More recently, transportation research groups, principally at CRA, MIT and Berkely, have experimented with estimating and applying disaggregated logit models.[2]

This new generation of models are based on theories of rational choice making by individuals -- a major advance over UTPS in conceptualizing travel demand. The models relate travel choices (mode, destination, hour-of-day, and frequency of trips) to the costs and times spent among the various alternatives. To the extent that transportation policies can be cast in terms of changes in transportation costs and times, the models can then be used to predict the effects of these policies on travel related choices.

---

[1]Charles River Associates, *A Model of Urban Passenger Travel Demand in the San Francisco Metropolitan Area,* Cambridge, Mass., 1967.

[2]Charles River Associates, *A Disaggregated Behavioral Model of Urban Travel Demand,* Cambridge, Mass., 1972; *Policies for Controlling Automotive Air Pollution in Los Angeles,* Cambridge, Mass., forthcoming; *Disaggregate Travel Demand Models,* Cambridge, Mass., forthcoming; Ben-Akiva, M., "Structure of Passenger Travel Demand Models," Unpublished Ph.D. Dissertation, Department of Civil Engineering, M.I.T., Cambridge, Mass., 1973; Adler, T. and Ben-Akiva, M., "A Joint Frequency Destination and Mode Choice Model for Shopping Trips," MIT Department of Civil Engineering, 1974; MacFadden, D., "The Measurement of Urban Travel Demand," *Journal of Public Economics,* 1974.

As mentioned above, the behavioral demand models were developed mainly for the analysis of improved transportation facilities. For examples, the original direct demand model was part of an effort to evaluate the effects of a third Bay crossing in San Francisco, and the McFadden model (1974) of disaggregate travel demand has been developed as part of the BART impact study.

Recently, the models have been shown to be useful for the evaluation of national or regionally ubiquitous transportation controls in response to environmental and energy related goals. CRA (1975) applied its 1972 disaggregate demand model to evaluate air quality control strategies in Los Angeles. Haws, Adler and Ben-Akiva have recently attempted to determine the effects of carpool incentives on travel using disaggregate demand models (1974). The direct demand model has been applied by CRA to the problems of free transit, gasoline rationing, and fuel conservation.[1]

1.2.1 Selected Models

The models noted above have to some extent been utilized in policy evaluation studies and, consequently, it is worthwhile to describe them in more detail. In reviewing them we will first consider whether the models represent reliably travel behavior in terms of their structure and the care taken in estimation and data analysis. The issue of whether these models can be easily implemented for quick policy evaluation is discussed in a separate section.

---

[1]Kraft, G. and Domencich, T., *Free Transit*, Charles River Associates, Cambridge, Mass., 1970; Charles River Associates, *Gasoline Rationing: The Economic Effects of Gasoline Rationing on New England*, Cambridge, Mass., 1974; Charles River Associates, *Policies for Conserving Fuel*, Cambridge, Mass., forthcoming.

Table 1-1 presents summary information about the models. It can be seen that we have distinguished three categories of models: (a) disaggregate work trip; (b) disaggregate shopping trip; and (c) direct demand.

Disaggregate Work Mode Choice

Most disaggregate demand models have the generalized logit specification. For a model of n possible alternatives this is formalized as:

$$P(i,t) = \frac{e^{\theta x_{it}}}{\sum_{j=1}^{n} e^{\theta x_{jt}}} \qquad (1\text{-}1)$$

where: $P(i,t)$ ≡ probability of mode i being taken by individual t for a given origin and destination

x_{it} ≡ vector of costs and times of mode i, and socioeconomic characteristics, for individual t for a given origin and destination.

θ ≡ estimated vector of coefficients for the cost, time and socioeconomic variables.

The elasticity of demand for choice i with respect to its own travel attribute x_{it} is as follows:

$$\eta(i,x_i;t) = \theta_x x_{it}(1-P(i;t)) \qquad (1\text{-}2)$$

where: $\eta(i,x_i;t)$ ≡ the elasticity of demand for mode i with respect to attribute, x_i, for individual t for a given destination

TABLE 1-1

SELECTED TRAVEL DEMAND MODELS

Model Type	Source	Estimation Technique	Data
Disaggregate Work Trip Mode Choice	CRA [1972]	logit maximum likelihood	1967 Pittsburgh Household Interview
Disaggregate Work Trip Mode Choice	McFadden [1974]	logit maximum likelihood	1973 San Francisco Household Interview
Disaggregate Work Trip Mode Choice	Haws [1974]	logit maximum likelihood	1968 Washington, D.C. Household Interview
Disaggregate Shopping Trip: Mode, Destination Frequency Choice	CRA [1972]	logit maximum likelihood	1967 Pittsburgh Household Interview
Disaggregate Shopping Trip: Mode, Destination Frequency Choice	Adler, Ben-Akiva [1974]	logit maximum likelihood	1968 Washington, D.C. Household Interview
Direct Demand Shopping and Work Trips	CRA [1967]	constrained least squares	1963 Boston Trip Table

and given that a trip will be made[1]

$\theta_x \equiv$ estimated coefficient for attribute X

$x_{it} \equiv$ attribute (time or cost) of travel by mode i by individual t for a given destination.

The effect of the change in attribute x_{jt} on mode choice i can be determined from the cross-elasticity which is as follows:

$$\eta(i, x_j; t) = -\theta_x x_{jt} P_j(j;t) \qquad (1\text{-}3)$$

where the variables were defined as before with appropriate changes in subscripts.

The appeal of representing system effects on demand in terms of elasticities is that, conceptually at least, data requirements in applying the models would appear to be minimized. It can be seen that if there are a substantial number of variables in the estimated models then applying formulas of the type represented by (1-2) and (1-3) would conserve data resources compared to applying formulas of the type represented by (1-1). This hypothesis was tested in the research reported in Section 2 with generally discouraging results.

Each of the disaggregate work mode split models is given an abbreviated description which includes the estimated elasticity of VMT's with respect to gasoline price. Obviously, other elasticities and cross-elasticities can be computed from the models but we concentrate initially on the effects of gasoline costs because this particular elasticity has been analyzed most

---

[1] Elasticity of demand is defined as:

$$\eta(i, x_i; t) = \frac{\partial P(i,t)}{\partial x_{it}} * \frac{x_{it}}{P(i,t)}$$

commonly and is the easiest to interpret consistently across models. Also, econometric estimates of gasoline demand provide a benchmark for evaluating the performance of travel demand models. VMT elasticities with respect to gasoline prices are also calculated for each of the procedures developed for applying the demand models in Section 2 and 3; this is one of several tests used to validate these procedures.

CRA [1972]-- The CRA mode split model was estimated on a sample of 115 individuals drawn from the 1967 Household Survey in Pittsburgh. All observations are from the same corridor which has relatively high quality bus service. Auto and transit are the only two modes represented in the model. A more complete description of the model is given in Section 2 where it is applied to Los Angeles and NPTS data.

The model was applied by CRA to estimate the effects of various pollution control policies in Los Angeles (see CRA [1975]). The model was adjusted to account for changes in auto occupancy and to be applied to data aggregated at the sketch plan zone level. By simulating the effects of a gasoline tax it was concluded that the elasticity of demand for VMT's with respect to gasoline price was -.27 in 1974.

McFadden [1974] -- The McFadden model was estimated on a specially surveyed sample of 213 households in Oakland and Berkely. The sample was designed to test the impact of Bart on commuter mode split and was accordingly stratified to overrepresent work trips to San Francisco city center. From the means of the data gathered in 1974, the estimated elasticity of VMT's with respect to auto costs was -.32. Assuming that gasoline costs are approximately two-thirds of auto commuter costs, the result implies an elasticity of VMT's with respect to gasoline price of about -.22. This figure should be caveated by noting that the elasticity was not computed by simulating the model, which would probably change the results.

Haws [1974] -- The Haws model was estimated on a sample of 448 households drawn from the 1968 Washington, D.C. Home Interview Survey. The mode choices included auto, bus and carpool. The system level of service data was of suspect quality because it was based on zonal networks rather than being developed for each of the individual observations. The carpool level of service data were based on *ad hoc* assumptions.

The Haws model was simulated on three so-called prototypical households in the Washington, D.C. area in the Adler and Ben-Akiva work [1974]. The simulated changes in VMT's with respect to a change in gasoline price implied an elasticity of -.003. This result indicates that the model forecasts travel behavior to be insensitive with respect to the costs of travel. As a planning tool, the model has limited value because its predictions of the effects of transit fare and gasoline price related policies are unreliable. Though the reasons why the model achieves these results were not presented by the authors, it can be speculated that the relatively poor quality of the data contributed to spurious parameter estimates.

Disaggregate Shopping Trip

The shopping trip category of disaggregate models is considered separately because the nature of the decision making process by individual tripmakers is more complex than in the case of work trips and shopping trip models are, consequently, structured somewhat differently. It can be assumed that an individual's options for work related travel are rather limited; in the short run the frequency, destination and time of day of work trips are presumably determined exogenously leaving only mode choice as the response to a change in transportation costs. For most other trip purposes, including shopping trips, an individual not only has considerable flexibility in choosing among modes but also can select among a wide variety of potential destinations, time of day and number of times over a given period (say, 24 hours) a trip for that purpose can be made.

Because of the increased complexity introduced into the model by the expanded number of choices, there have been relatively few complete shopping trip models estimated. The two reported on below are the only disaggregate demand models, of which we are currently aware, that consider concurrently the choices of mode, destination and frequency. Both models have been simulated for a limited range of policy scenarios.

CRA [1972] -- The CRA model for shopping trips was estimated on a sample varying in size from 73 to 140 observations drawn from the 1967 Pittsburgh Household Interview Survey. The form of the model can be depicted as follows:

$$P(d,d,f;t) = P(m;d,f,t) * P(d;f,t) * P(f;t) \tag{1-4}$$

$$P(m;d,f,t) = \frac{e^{V_{mdt}}}{\sum_{i=1}^{M_t} e^{V_{idt}}} \tag{1-5}$$

$$P(d;f,t) = \frac{e^{V_{dt}}}{\sum_{j=1}^{D_t} e^{V_{jt}}} \tag{1-6}$$

$$P(f;t) = \frac{1}{1 + e^{V_{ft}}} \qquad (1\text{-}7)$$

where

$P(m,d,f;t) \equiv$ probability of individual t choosing to make a trip to destination d by mode m.

$P(m;d,f,t) \equiv$ probability of indivdual t choosing mode m given that a trip will be made to destination d.

$P(d;f,t) \equiv$ probability of individual t making a trip to destination d by any mode given that a trip will be made.

$P(f;t) \equiv$ probability of individual t making a shopping trip to any destination by any mode over a 24 hour period.

$V_{idt} \equiv$ linear function of the costs and times of mode i from individual t's house to destination d, and of the socioeconomic characteristics of individual t.

$V_{it} \equiv$ linear function of the costs and times of all modes from individual t's house to destination j, and of the attractiveness of destination j (retail employment).

$V_{ft} \equiv$ linear function of the costs and times of all modes from t's house to all destinations and attractiveness of all destinations.

The model considered two modes, auto and bus, and varying numbers of destination choices among households with the average being four. The frequency of travel choice was limited to either one or no trips. There is a specification error in the destination and frequency choice models which, in all likelihood, does not effect the parameter estimates of this particular model but would lead to misforecasts of policy effects unless steps are taken to change the form of the V's in equations (1-6) and (1-7).[1] Also, experience in applying the frequency choice equation yielded unacceptably high elasticities, in the range of -4, for a change of trip frequency with respect to a change in trip cost.

To derive the elasticity functions analogous to Equation (1-2) and (1-3) leads to cumbersome relationships. It is simpler to simulate travel behavior in response to varying system attributes in order to compute elasticities.

The shopping trip model, excluding frequency choice, was simulated on Los Angeles data to estimate the 1974 elasticity of VMT's with respect to gasoline price [see CRA 1975]. The estimated price elasticity owing to mode and destination shifts was -.12. It was assumed that the additional VMT's conserved by reduced frequency of travel put an upper bound on the total elasticity of -.24.

Adler/Ben-Akiva [1974] -- The Adler/Ben-Akiva model was estimated on a sample of 1313 observations drawn from the 1968 Washington, D.C. Home Interview Survey. The form of the model can be depicted as follows:

$$P(m, d, f; t) = \frac{e^{V_{mdft}}}{\sum_{ij\varepsilon MDt} e^{V_{ijft}}} \qquad (1\text{-}8)$$

[1]See Ben-Akiva [1973] and William B. Tye and Leonard Sherman, *Disaggregate Travel Demand Models*, Project 8-13: Phase I Report prepared for National Highway Research Program (September 1975) for a more complete discussion of this issue.

where:

$P(m,d,f;t)$ = probability of individual t choosing to make a trip to destination d by mode m.

$V_{ijft} \equiv$ linear function of times and costs of mode i from individual t's house to destination j, and of socioeconomic characteristics of individual t.

$MD_t \equiv$ set of all possible mode, destination and frequency alternatives for individual t.

Though this model is somewhat less complex than the CRA shopping trip model, it does not yield very tractable analytical forms for the elasticities. Again, in order to compute the implied elasticities of system level of service variables, it is best to simulate the model over various scenarios.

Adler and Ben-Akiva simulated the model for three prototypical households with increased gasoline prices. The implied VMT elasticity was computed as -.06. This value is somewhat lower than would be expected based on the results of other studies of gasoline price elasticity.

Direct Demand Model [CRA 1967]

The direct demand model uses zonal interchanges as its basis of observation where their size corresponds to that of sketch plan zones. The model estimates the number of round trips by mode and purpose as a function of the times and costs of alternative modes and socioeconomic characteristics of origin and destination zones. The two relationships of interest are the auto work trip equation and the auto shopping trip equation. These are discussed separately below.

Auto Work Trip -- The auto trip was estimated on a sample of 255 zonal interchanges drawn from a trip table developed froma 1963 origin-destination survey in the Boston area. The functional form for this relationship is:

$$\frac{N_{ij}}{Y_{ij}} = \alpha X_{ij} + \beta ln X_{ij} \qquad (1\text{-}9)$$

where:

N_{ij} ≡ number of auto work round trips between zones i and j

X_{ij} ≡ vector of variables representing costs and times of alternative modes between zones i and j and representing socioeconomic characteristics in zones i and j.

Y_{ij} ≡ [employed labor force in zone of residence] * [employment in zone of work as proportion of total regional employment]

α, β ≡ estimated vectors of parameters

The functional form of the travel elasticity with respect to some level of service variable such as auto line-haul cost is:

$$\eta(X_{cij}) = \frac{\alpha_c X_{cij} + \beta_c}{N_{ij}} * Yij$$

X_{cij} ≡ linehaul auto cost for a round trip between zones i and j.

α_c, β_c ≡ estimated coefficient on linehaul costs for auto

The implied gasoline price elasticity evaluated at the mean of the observations in the estimating sample is -.25, assuming gasoline is fifty percent of auto line haul cost.

Auto shopping trip -- The auto shopping trip was estimated from a sample of 75 zonal interchanges from the above mentioned Boston 1963 trip table. The functional form for this relationship is:

$$ln N_{ij} = \alpha ln X_{ij} + \beta X_{ij} \quad (1\text{-}10)$$

where:

N_{ij} = number of auto shopping round trips between zones i and j.

Xij = vector of variables representing costs and times of alternative modes between zones i and j.

α, β = estimated vectors of parameters.

The functional form of the travel elasticity with respect to auto line-haul cost is:

$$(X_{cij}) = \alpha_c + \beta_c X_{cij}$$

where

X_{cij} = linehaul auto cost for a round trip between zones i and j.

α_c, β_c = estimated coefficients on linehaul costs for auto.

The implied gasoline price elasticity evaluated at the mean of the observations in the estimating sample is about -.44, assuming gasoline is fifty percent of auto line haul cost.

1.2.2 Brief Critique of Existing Models

In order to apply the above models, analysts should be aware of various statistical and specification problems which will affect the reliability of policy evaluations. Some problems also arise when the models are to be applied to existing data sources; this subject is covered in a separate section.

A useful starting point for critiquing existing models is provided by comparing their predictions in policy evaluation stituations. Table 1-2 summarizes the results of the above described models when applied to policy scenarios involving the price of gasoline.

It can be seen from Table 1-2 that there is reasonably uniform agreement among work trip models, excluding the Haws model. Moreover, these models compare favorably with the short run gasoline price elasticity estimates of, by now, many econometric estimates of gasoline demand functions; the results of econometric studies of gasoline demand generally indicate a short run price elasticity in the range of -.2.[1]

It can also be seen from Table 1-2 that there is little agreement among shopping trip elasticities. The lack of consistent results among shopping trip models in their application to policy issues makes them, at this stage of their development, less useful than work trip models. It also begs the question of why widely disparate elasticities among shopping trip models are obtained.

There are four principal sources of error in the models presented above which help to explain the variability in estimated elasticities. It will be seen that these estimation problems occur for both work and shopping trip models but that they are more apt to bias the results of shopping trip behavior estimates.

Cross Section and Identification Bias

Probably the major conceptual problem of the models is that they are estimated on observations which reflect location choice in addition to short run travel related choices. Thus, for example, people and businesses may have located so as to minimize travel costs or because

---

[1]See Charles River Associates, *Policies for Conserving Fuel*, forthcoming, for a review of these models.

TABLE 1-2

ESTIMATED ELASTICITY OF VEHICLE MILES TRAVELED (VMT's) WITH RESPECT TO GASOLINE PRICE FOR SELECTED MODELS

Model Type and Source:		Elasticity
Disaggregate Work Mode Split		
	CRA [1972 and 1975]	-.27
	McFadden [1974]	-.22
	Haws [1974]	-.003
Disaggregate Shopping Trip		
	CRA [1972 and 1975]	-.12, -.24
	Adler/Ben-Akiva [1974]	-.06
Direct Demand Work Trip		
	CRA [1967]	-.25
Direct Demand Shopping Trip		
	CRA [1967]	-.44

they prefer one mode over another. To the extent this represents a problem in the estimated relationships, its effect will be generally to cause over-prediction of the elasticity of travel behavior with respect to system level of service.

Recent attempts to deal with the simultaneity between auto ownership and mode choice, and among auto ownership, mode choice and housing location, have had modest success.[1] They are, in their current stage of development, not appropriate for widespread policy evaluation. Moreover, these models use a structure which is oriented toward incorporating work mode choice decisions while assuming shopping, or other, trip behvaior is not interdependent.

In fact, the robustness of work trip models indicates that the nonwork trip model is more in need of further development. The high travel demand elasticities estimated by CRA in both shopping trip models (disaggregate and direct demand) imply that, to some extent, location decisions are being merged with tripmaking decisions. That is, there is evidence that these models are picking up long run elasticities in their estimates.

For example, in the disaggregate demand model, it is appropriate to interpret the frequency choice relationship as indicating that households who prefer to shop less frequently will locate themselves so as to be further away from the constellation of alternative shopping destinations than those households who prefer to shop more frequently.

---

[1] See Lerman, S. and Ben-Akiva, M., *A Behavioral Model of Automotive Ownership and Mode of Travel*, Cambridge, Ma. 1974 for a logit model which includes work trip mode and auto ownership mode choice. See Lerman, S., *A Disaggregate Behavioral Model of Urban Mobility Decisions*, Unpublished Ph.D. Dissertation, Department of Civil Engineering, M.I.T., 1975, for a logit model of housing location, auto ownership and mode choice jointly determined. Both efforts represent technical advances in application of statistical techniques but are, at this stage, exploratory.

The same interpretation is possible when viewing the direct demand model with the additional feature that commercial establishments will also tend to locate so as to minimize distance from household markets.

There are relatively few options available to completely purge models of cross section bias when cross section data are being utilized. Gasoline demand models utilize pooled time series and cross section observations to isolate short run elasticities. Typically, time series data are not available to transportation planners.[1] Section 3 presents a non-work trip model based on cross section data which is specified in such a way so as to decrease cross section bias.

Multicollinearity

Multicollinearity occurs in equations when one independent variable can be expressed as a linear function of other independent variables in the same equation. Pure multicollinearity defined in this way is relatively uncommon but near multicollinearity is typical in transportation demand relationships. The cost of auto trips, line haul time of auto and line haul time of bus are all roughly proportional to distance traveled -- they are usually highly collinear.

The extent to which near multicollinearity is a problem depends upon sample size and the closeness of the relationship between the independent variables. Its result is to decrease the confidence with which parameter estimates can be interpreted. Several techniques are used in transportation demand research to remove the problem.

---

[1]See M. Gaudry, *The Demand for Public Transit in Montreal and Its Implications for Transportation Planning*, Unpublished Ph.D. Dissertation, Princeton, Department of Economics, 1974, for an exception.

In the effort to estimate a direct demand model, constrained least squares was employed to insure that coefficients exhibited proper signs. The result was that the stronger relationships, reflecting own elasticities, were isolated while the weaker relationships, which would have been used for computing cross elasticities, were often constrained to equal zero.

In disaggregate demand models, differences between costs and times for modes are used as the independent variables. If near multicollinearity is present, this approach will not remove it but will reduce its impact. Nonetheless, there has been a tendency for the stronger determinant of mode split, relative time between modes spent on the trip, to dominate the difference in cost between modes. Demand models estimated by Ben-Akiva, Haws, Adler and Lerman typically deal with this problem by dividing cost variables by income categories. This simultaneously reduces multicollinearity and adds an income effect into elasticity estimates.

Aggregation and Errors in Observation

One of the important attributes of disaggregate demand models is the gain in information which entails using household specific data. Observation errors typically arise of data not representing the true values of the variables used in estimation. In transportation demand modeling the use of zonal aggregates or zonal interchange data will cause errors in observation. In this connection it is interesting to note that the disaggregate demand models where the greatest care was taken to minimize errors in observation (CRA [1972] and McFadden) derived higher elasticities than the studies where the data were taken unchecked from zonal O-D tapes or were known to be *ad hoc* assumptions (Haws and Adler/Ben-Akiva). This accords with the general rule that observation errors will tend to impart downward bias to estimated coefficients.

In both disaggregate demand and direct demand models, the problem of using zonal interchange or aggregate data has several dimensions which are discussed below.

a. The meaning of "Average Level of Service" in a zone has different interpretations and different ways of being calculated. None are particularly satisfactory. The least appropriate definition of average level of service is the one which is easiest to compute -- centroid to centroid distances and times. This type of variable becomes especially inappropriate when one tries to infer an average access time to transit for, say, a sketch plan zone which is generally between 16 and 50 square miles in size.

It is obvious that the independent variables in the demand equation should be averages over some range of observations in the zone if aggregate data are used. To the extent that there are numerous auto trips in any zonal interchange, and observations on the attributes of these trips are available, the average for auto travel over trips actually made is an adequate approximation. However, this procedure is not appropriate when there are relatively few trips by a mode -- such as carpooling, transit and walking -- because the average so calculated represents only the mean attribute vector for people who actually took that mode and is typically not the vector faced by people who did not take that mode. This problem is most acute where it concerns access to transit and the opportunity costs for carpooling. In the former instance, the mean access time for the zone may be over 30 minutes, whereas the mean time for transit patrons is less than 10 minutes, both evaluated in the same zonal interchange. If average access time for zonal interchanges is computed by taking the average observed for transit patrons, then the estimates of elasticity and cross-elasticity will obviously be specious. Similarly, if the level of service for carpooling is computed from observed level of service for only those who carpool, it will, analogously, lead to zonal interchange averages which show much lower cost to carpools than are in fact the case.

b. Typically, alternate service data on trips taken from small zones do not exist for coded networks. Ignoring for the moment that the coded network misrepresents actual network times by using centroid to centroid data, there is the problem of averaging these data for zonal interchanges at the sketch plan size (there can be anywhere from 100 to 600 traffic analysis zone interchanges in a sketch plan zone interchange). A simple average is inappropriate even on the assumption that population (or land uses) are evenly distributed among small zones. To see this consider the estimation of mode split work trip equation. It is assumed that a given number of trips by any mode will be made but that there are choices to be made among modes for any given zonal interchange. However, the level of service among modes estimated as a simple average between a zonal interchange will not be the level of service, on average, faced by the individuals making the trips. Generally, the costs of travel between two zones will be less for the people making the trip than for the average costs over the area of the zones -- the people who make trips in a zonal interchange typically travel less distance than the average distance between the zones.

In the above example of mode split relationships, it is legitimate to weight the network data by total number of trips. For any zonal interchange this will give a tolerable approximation of the level of service by mode for trips actually and potentially made. The smaller the traffic analysis zone, the better the approximation. The larger the sketch plan zone, the greater the necessity of making the weighted average. However, this procedure is inappropriate if these level of service variables are to be applied to destination choice relationships. The problem is easy to see: because the level of service variables represent the alternative costs among modes of trips actually made between OD pairs, they misrepresent the costs of trips not made between these same OD pairs. Thus a different grouping of the data may have to be applied to estimation of a destination choice model, or even a direct demand model, that wouldn't be the case if only mode split relationships were being estimated.

c. Finally, a somewhat arcane point can be made about the relationship between the distribution of errors and the nature of the data. In the logit model, it is assumed that the error distribution is Weibull; in the least squares model it is assumed that the error distribution is normal. Both assumptions are jeopardized when additional errors in observation are introduced by using zonal interchange data. The distribution of level of service variables over a zonal interchange is highly truncated depending upon the distances between boundaries of the two zones. Similarly, the distribution of households about a centroid to centroid measure is also truncated. Moreover both distributions are likely to be skewed. Both features of these distributions make them unlikely to be approximated by the normal or Weibull distribution thereby incurring further error in model estimates.

Specification of Alternatives

A final problem involves specifying the appropriate range of choices for household decision making. The problem varies according to the type of choice being modeled.

a. Mode choice. The number of modes available to an individual include various forms of shared rides, being chauffeurred, transferring modes and, of more importance than is commonly recognized, walking. Mode split relationships in disaggregate demand have only satisfactorily treated auto drive alone and transit. Until better data become available, it is unlikely that more mode choices will be successfully incorporated into disaggregate demand models.

b. Destination choice. The specification of alternative destination choices in disaggregate demand models is an important area for future research. The current approach to specifying destination alternatives for shopping trip models is to develop heuristically a number of alternative destination zones. A household in an origin zone is assumed to have as its alternative destinations those destinations most often chosen by the households in its zone of residence. This representation of the true destination choice process is obviously oversimplified.[1] In the direct demand model, choice among alternative destinations is ruled out by the model's specification.

c. Frequency choice. In currently estimated shopping trip models, the frequency of trips is constrained to be zero or one. The choice of taking more than one trip or combining purposes on the same journey is not admitted.

One effect of specification error has been to make shopping trip models considerably less robust than work trip models. Clearly, alternative specifications allow for alternative courses of action when the models are simulated

---

[1]See Lerman, S. and Adler, T. "Workshop on Destination and Related Choices: Summary Report," *The Second International Conference on Behavioral Travel Demand*, 1975.

and this, to some extent, explains the variation in elasticities among the models. A final note on this subject which will be developed further in Section 3: it is quite possible that the logit approach to modeling discrete choice behavior is less appropriate for nonmode choice decisions than other model specifications.

1.3 APPLICATION OF DEMAND MODELS

Applying demand models to existing data bases involves several problems which jeopardize the accuracy of travel demand predictions. In addition to the estimation related errors described above, substantial errors in forecasting travel can occur because the models are applied to phenomena which were not originally modelled or the models are applied to data which is aggregated in various ways.

The easiest method of applying the models is to use the elasticities numerically computed in the studies cited above and apply them to policy scenarios. In this regard, it is noteworthy that the three most carefully estimated work trip models all yielded gas price elasticities between -.22 and -.27. This is a very tight cluster of estimates, especially considering that the cities represent a wide range of transit level of service (Boston, San Francisco and Los Angeles). Thus there is some merit in performing an "instant" evaluation of policies by using the elasticities from the models calculated at the means of the data upon which they were estimated. This information is, for the most part, available in the reports which present the estimation results.

Alternatively, more refinement in policy evaluation can be obtained if the models are applied to data which is also relatively refined such as trip tables or household surveys, as had been done by McFadden and CRA [1975]. Also, even though the gasoline price elasticities are similar, there is some presumption that changes in other transportation system attributes, or their combinations, would not be evaluated as uniformly across models or regions. Consequently, the disaggregate work trip mode split model being a powerful tool for policy evaluation, there is considerable merit in developing methods for applying it for the various existing transportation data bases.

Evaluating policy effects on nonwork trips presents a different problem. Existing disaggregate models of these trip purposes, other than shopping, virtually do not exist. Moreover, the review of shopping trip models presented above indicates that even these tend to be unreliable and cumbersome to apply. In the case of nonwork trips, therefore, the analysis of applying travel demand models to existing data serves a somewhat different purpose than in the case of work trips: that is, we wish to determine which type of model is most suitable for existing databases to help in further model development.

The rest of this section is devoted to discussion of two categories of problems which arise in applying travel demand models. The first of these, called generalizability problems, examines the issue of applying the models to trip purposes and mode alternatives which were not originally included in model estimates. The second issue, which has become known as the aggregation problem, involves the application of disaggregate demand models to grouped data.

1.3.1 Generalizability of Demand Models

Policy evaluation will often take place with analytical tools which are, in some sense, underdeveloped. A major issue which is as yet not completely resolved is the generalizability of disaggregate demand models across several dimensions. It will be seen by the brief descriptions given below that this issue is related to the previously discussed specification problem in estimation.

Application to New Modes

Often, logit models are estimated on a smaller choice set than the one to which they are applied. Indeed, one feature of the multinomial logit model is that it can be used to predict travel response to new modes.[1] However, it will most often be the case that *ad hoc* assumptions must be made about the most suitable adjustments to the models when they are applied to new modes.

This problem arises even when the model is being applied to existing modes. As noted in the above review, most estimated mode split models only consider bus and auto driver as the mode choice alternatives. If they are applied to only these two alternatives, and, consequently, the options of shared rides or walking are ignored, then computed elasticities will be underestimated.

There are several ways of coping with the new mode problem, though only two have actually been utilized in policy evaluation studies. The first of these,

---

[1]Considerable theoretical and empirical research has been done on the problem of potential biases which could arise when the model is used in this way. See McFadden [1974] and CRA, *Disaggregate Travel Demand...* [1975].

developed by McFadden [1974] uses *a priori* information about whether choices are made sequentially or simultaneously. The second, used by CRA in its study of pollution controls [1975], adjusts the mode specific constants based on a combination of *a priori* information and tests of the predictive ability of the models. The latter approach is pursued in Section 2 where procedures for applying work trip mode split models are developed.

Application to Other Purposes

Travel demand models are missing for a large number of trip purposes. In applying the models that do exist to determine the effects of a system change on policies, assumptions must be made about the elasticities of travel demand for purposes for which estimates are simply missing. A typical assumption is that most nonwork trips display individual elasticities with respect to times and costs similar to shopping trip elasticities. Though it can be argued that the choice structure of nonwork trips is generally more like shopping trips than work trips, the assumption obviously leads to, possibly, substantial forecasting errors. The most appropriate remedy for this problem is to estimate models for other trip purposes. In Section 3, a general nonwork travel demand model is presented.

Transferability

A final dimension to the generalizability problem is the question of whether a model estimated on one set of observations can be used to predict the travel behavior of individuals drawn from another set. The development of behavioral demand models was in large part motivated by the objective of having models which would be generalizable across regions. Models which were transferable in this way would not have to be recalibrated each time they were applied.

The operational experience of transfering models has been generally encouraging. The direct demand model estimated on Boston data was applied to San Francisco after minimal adjustment to parameter estimates to account for differences in land use density between the two urban areas. CRA also applied a disaggregate demand model to Los Angeles data, though it is not known the extent to which adjustments in the model for aggregation and new mode problems actually applied to the transferability problem. As a sobering note, the results in Section 2 indicate that McFadden's model estimated on San Francisco data is not easily transferable to Los Angeles data.[1]

1.3.2 Problems of Using Grouped Data

Within the planning context, virtually all available data is grouped whereas the most reliable and versatile travel demand models have been estimated on highly disaggregated data. Though at some future time disaggregate databases from household surveys may be in widespread use, the current and near term value of disaggregate models is limited by the lack of techniques

[1]A model not presented above, estimated on San Diego data was validated on Boston data thereby adding to the evidence that transferability can be accomplished with behavioral demand models.

to apply them to existing data. As will be shown, considerable effort has gone into developing such techniques with less than complete success. Section 2 builds on the discussion and review presented below to develop methods, relatively easy to use, which planners can draw on for applying work trip mode split models.

Theory of Applying Disaggregate Demand Models

In order to make our discussion of the aggregation problem easier to follow, we will consider only binary choice models. This also allows us to simplify the mathematical notation. For example, consider a logit model of the form:

$$P(A:t) = \frac{1}{1 + e^{Y_t}} \tag{1-11}$$

where:

$P(A:t) \equiv$ the probability that an indivdual, t, will choose option A (say, auto) given that a trip will be made between a specified origin and destination;

$Y_t \equiv$ a function of the attributes of the two modes facing t and socioeconomic characteristics of t.

If Y_t is linear in the differences between auto and transit, then the above model is the binary counterpart to the equation (1-1). Both the CRA [1972] and McFadden work trip model can be expressed in this form.

To find the demand for choice A over a population of individuals, the probabilities of each are added over the entire population:

$$N_A = \sum_t P(A:t) \tag{1-12}$$

Typically, the information needed to apply (1-12) is not available to planners. An alternative is to select a random sample of individuals for which the data for each Y_t can be obtained. It is likely that such samples could be relatively small (about 200 households) but even such modest data bases are virtually nonexistent.

Given that grouped data is all that is available, other methods must be used to apply the disaggregate demand model. These approaches build on the theory outlined below.

Suppose that over any grouping of individuals, the Y's are distributed according to a well defined mapping which can be approximated by a continuous function, $f(Y)$. Then, if $f(Y)$ is normalized to have its integral be unity over the limits of integration, the mean probability for the group is as follows:

$$\bar{P}(A,T) = \int_a^b \frac{f(Y)dY}{1 + e^Y} \tag{1-13}$$

where:

$\bar{P}(A,T) \equiv$ mean probability of choosing A for a group, T;

$a,b \equiv$ the limits of values of Y in group T.

Total demand for A is then:

$$N_A = N^*\bar{P}(A,T) \tag{1-14}$$

where:

$N \equiv$ total number of individuals in T.

Typically, $f(Y)$ will depend on the type of grouping involved. To demonstrate this we give an example. Suppose that Y is distributed uniformly over a zonal interchange. The limits of the distribution are functions of the areas of the zones. The form which $f(Y)$ takes is:

$$f(Y) = \frac{1}{b-a} \tag{1-15}$$

Then the solution to (1-13) is:

$$\bar{P}(A,t) = \frac{ln\left[\dfrac{e^Y}{1+e^Y}\right]}{b-a}\Bigg|_a^b = \frac{ln\left[\dfrac{1-P(A:b)}{1-P(A:b)}\right]}{b-a} \tag{1-16}$$

where:

$P(A:b) \equiv$ the probability of an individual choosing A when $Y = b$.

The above solution has attributes which are expected of a demand model representing probability choice. In particular, as b approaches a -- i.e., as the size of zones approach zero, representing one household traveling to one point -- the application of l'Hospital's rule to (1-10) yields:

$$\lim_{b \to a} \bar{P}(A) = P(A:a) \tag{1-17}$$

More generally, $f(Y)$ would have to be assumed using whatever information may be available about the likely distribution of access times, linehaul times and costs for the alternative modes as well as the distribution of socioeconomic characteristics across the group T. It is expected that there are a number of functional forms which are candidates for $f(Y)$.

Among the properties of *f(Y)* which would make application of equation (1-13) tractable are that its parameters be dependent, in a well defined manner, on the means of the level of service variables and the size of the zones. Also, as the zone size increases to include the whole urban area in the limit, the variance in attributes should increase at a lower rate and approach a constant.

For data aggregated by different methods, such as cross tabulations of trip and household characteristics over the region as a whole, the distribution of *f(Y)* would be different from its distribution across analysis, zones. Recently, Richard B. Westin developed an approach to applying binary choice models to such data based on the assumption that *f(Y)* is multivariate normal.[1] However, his approach is not conducive to ease of analysis and the extension to models of more than two choices has not been derived. In addition, simulations by Talvitie comparing different approaches to the aggregation problem showed no clear cut improvement in accuracy was achieved using Westin's approach when compared to other, more easily applied methods which adjust the logit model using approximative techniques.[2]

Approximation Methods: Correcting for Intrazonal Variation

Two methods have been developed to apply the disaggregate logit model to zonal data. Neither are particularly satisfactory at their current state of development, thus they will be described quite briefly.

---

[1]Westin, R.B., "Predictions from Binary Choice Models," *Journal of Econometrics 2*, (1974), pp. 1-16.

[2]See Talvitie, A., "Mathematical Theory of Travel Demand Models: A Resource Paper," The Second International Conference on Behavioral Travel Demand, forthcoming.

Probit Approximation -- One of the first investigations of the aggregation problem in travel demand assumed that the probability choice function had the probit specification.[1] The probit model rests on the assumption that unobserved "errors" in individual behavior are normally distributed. The logit model assumes that the unobservables are distributed according to the Weibull distribution which is similar to the normal. The major difference between the two specifications is that the probit model tends to approach the limits of zero and one probability somewhat more rapidly than the logit model.

McFadden and Reid developed a technique for adjusting the disaggregate probit model to zonal interchange data using the variance - covariance matrix of terms in the demand equation. By making another adjustment to the logit model, to make it perform more like the probit function, the method of applying information about variation of system attributes can be transfered from the probit to the logit model. The method is computationally cumbersome and requires extra information from zonal data bases which is not readily available.

A more important problem with the approach is that the distribution of variables within a zonal interchange is assumed to be normal about the mean of the observations. This is surely a major inaccuracy. As discussed before, variables such as linehaul time and auto cost between households in the origin zone and employment centers in the destination zone are truncated by distances between zonal boundaries. Moreover, the distribution of individuals making trips in a zonal

---

[1] Daniel McFadden and Fred Reid, "Aggregate Travel Demand Forecasting from Disaggregated Behavioral Models," unpublished, Berkeley, November 1973.

interchange with respect to level of service variables is highly skewed -- the lower the cost, access time, *et al.*, the larger the number of trips. Both the properties of truncation and skewness argue against assuming that variables in a demand model are normally distributed over a zonal interchange.

Taylor's series approximation -- Based on a suggestion by Talvitie,[1] CRA recently applied their disaggregate demand model to sketch plan zones in the Los Angeles region. The method involved taking the expected value of a Taylor's series expansion of the logit model about the mean of the data in a zonal interchange. Truncating the series after the third term yields the following expression:

$$\bar{P}(A,T) = P(A)[1 + var[Y](1-P(A))\ (1/2 - P(A)] \qquad (1\text{-}18)$$

where: $var[Y]$ ≡ variance of Y over the group T.

$P(A)$ ≡ the logit probability calculated at the mean value of Y for the zonal interchange.

One problem with the approach is that if there are n independent variables in the Y function there are, potentially, $\frac{n^2+n}{2}$ variance-covariance terms in computing $var[Y]$.[2] Even though many of these can be assumed to be zero, owing to stochastic independence or con-

---

[1]Antti, Talvitie, "Aggregate Travel Demand Analysis with Disaggregate Travel Demand Models," *Proceedings...Transportation Research Forum, Vol. XIII* (October 1973).

[2]In the CRA mode split model, n equals 7.

stancy over the zonal interchange, the resulting calculations needed for each zonal pair of data is of a large order. It was found that most terms could be made functions of the size of the zones thereby easing considerably the data requirements. It was also found that access to transit and autos per worker were, by far, the major contributors to variation of Y.

Considering that the residual terms in the Taylor's series could have been significant, the accuracy achieved in applying the above approach was somewhat surprising. Predicted auto mode split was within 10 percent of the actual and total predicted vehicle miles traveled was within 5 percent of the actual travel.

In Section 2 of this report, the Taylor's series approximation is simplified further to minimize data requirements and applied to data at varying levels of aggregation. Though some heuristic rules are necessary to utilize this approach, the results indicate that significant increases in accuracy of mode split prediction are achieved while the flexibility of applying logit models is maintained.

Aggregation Error and Area

Much of the error which occurs in applying the logit model to zonal interchange data can be related to the area of the zones. Aggregation error occurs because of variation in the variables of the logit model across individuals. Differences among level of service variables occur because individuals are located differentially with respect to transportation facilities and origin-destination points in the zonal interchange. Variation in socioeconomic characteristics tends to increase as the population of a zone increases which, under simplifying assumptions, may also be related to the area of zones.

Given a well-specified relationship between the area of zones and the distribution of Y_t, it may be possible to apply equation (1-13) in order to predict mode splits from zonal interchanges of varying sizes. Though research is needed to make this approach tractable, it does suggest the properties which a generalizable model for grouped data should have. These are summarized below:

a. It should replicate travel demand behavior. This implies two attributes: (1) elasticities and cross-elasticities have appropriate signs; and (2) demand is most elastic when the mode split is nearly equal.

b. As the variation in zonal attributes increases, the frequency with which a group chooses either mode should approach the zero and one limits less quickly. Stated another way, as within-group variation increases, and the level of service variables remain constant, the frequency with which the dominant mode is chosen should decline.

c. As the area of zones increases, the variance in mode level service variables should also increase but at a decreasing rate. Also, the variance in level of service variables should approach a constant.

d. As the variation in modal attributes declines to zero, the model should replicate individual probability choice behavior. In particular, at the zero variation level, the model may be functionally identical to a disaggregated demand model estimated on individual households.

The implication of these properties on using area to adjust disaggregate demand models for zones of varying sizes can be seen from the following formula:

$$ln \frac{N_A}{N_B} = Y(1 + f(a_i, a_j)) \tag{1-19}$$

where:

$N_A, N_B \equiv$ numbers of trips by modes A and B in a zonal interchange i, j

$f(a_i, a_j) \equiv$ scalar function of area

$a_i, a_j \equiv$ areas of zones in the zonal interchange, i, j

The function of area should have the following properties:

$$f(0,0) = 0$$

$$0 < \lim_{\substack{a_i \to \infty \\ a_j \to \infty}} f(a_i, a_j) < 1$$

$$\frac{\partial f(a_i, a_j)}{\partial a_i} \leq 0, \quad \frac{\partial f(a_i, a_j)}{\partial a_j} \leq 0 \text{ for all } a_i \text{ and } a_j$$

The results of estimating and applying a model of the form (1-19) are given in Section 2. Though the method works reasonably well when applied to zonal interchange data, one drawback is that it cannot be applied to data which is represented in the form of cross tabulations or frequencies. The problems associated with using this type of data are considered below.

Market Segmentation

Travel data is sometimes cross tabulated by distance, time, and socioeconomic characteristics of tripmakers. This format has been useful in segmenting the travel market so the impact of policies on particular socioeconomic groups can be highlighted. It will be argued below that market segmentation has the additional benefit of reducing aggregation error when such data are analyzed with disaggregate logit models.

The application of multinomial logit to market segments is actually an extension of the early development of logit analysis. Models of binary choice were originally developed from the application of statistical tools to contingency tables.[1] These models gave the probability that a "response" would occur to a "stimulus" within a specified range. For a simple, univariate model, a table giving the proportions of the sample responding at each level of stimulus will present sufficient information to estimate a model. Similarly, given a model, such as an estimated logit equation, the proportion of a sample responding to stimuli within given ranges can be predicted.

This approach can be generalized to the common specification of disaggregate mode split models. If only two modes are considered, then the response will be the proportion of trips by a given mode, say, auto. The approach is made computationally more complex as the number of different types of stimuli (independent variables such as modal attributes) increases. Instead of a column of numbers representing the sample at each level of stimulus, it would be necessary to have a multidimensional array representing the number of travelers which face alternative levels of service among modes.

---

[1]See Box, *The Analysis of Binary Data* (1970: Methuen, London).

Such an array could be potentially massive. For example, there are seven independent variables in the CRA mode split model:

auto variable cost per trip;
transit fare;
auto in-vehicle time;
transit wait, linehaul and schedule delay time;
auto access time;
transit access time;
auto availability per worker.

A categorization of each of these variables into three ranges (say, low, medium and high) would result in 3^7 or 2,187 potential cells for travelers. Obviously, in order to make the approach workable, some way to decrease the number of cells needs to be devised.

The number of cells can be reduced dramatically by making simplifying assumptions. These include the assumption that auto access time is equal for all travelers and that there are only two categories for autos per worker (0 and 1). Additionally, linehaul times and auto costs can be collapsed into the same cells by virtue of their all being correlated with distance. This approach to applying disaggregate mode split models to market cells was successfully tested in Section 2 where the number of cells was reduced to 12.

Each cell represents, initially, the number of people in the population who face the range of attributes labeling the cell. Also for each cell there is a mean probability that an individual would choose auto; this probability can be calculated, for example, by using the midpoint of the ranges of the variables and some appropriate mean figure for ranges which are open ended. The mean individual probability for each cell times the number of people included gives the number of auto trips. Total auto travel is the sum of auto trips throughout the array. Other values of interest, such as vehicle miles traveled or total transit revenues can also be computed from the array.

The reason this approach is more accurate than applying the logit model to zonal interchanges is that the amount of variation among attributes within any cell in the array can be presumed to be smaller than the variation in a zonal interchange. This is particularly true of transit access time and number of cars available: the variables which are the main contributors to biased predictions using only the means of data from zonal pairs.

The question arises as to whether existing urban data bases have readily available data in this format. In general, it can be presumed that a significant amount of tabulation would be required in order to develop a usable array. However, using a widely available prewritten cross tabulation program, the NPTS household review file was readily converted into market segments. Thus, even if new data preparation is called for, planners should seriously consider this approach to quick travel policy evaluation.

Linear Models

In transportation demand research, linear models have been eschewed in favor of the logit specification. CRA [1972] estimated a linear mode split model but made a compelling argument on theoretical grounds against it in favor of the logit model. No disaggregate linear model other than mode split has been estimated even though the linear specification can easily be applied to situations where there are many alternatives which can be ranked, such as nonwork trip generation.

On practical grounds, linear models have considerable appeal. They are easy to interpret by inspection. Multiequation systems of linear models are relatively easy to simulate, either by iterative techniques or by solving the equations simultaneously. Perhaps most importantly, the aggregation of linear models is much easier than logit models; in some situations, the model can be applied directly to means of grouped observations even when the parameter estimates are attained on individual household observations.

Models of binary choice have the following form:

$$P(A:t)\begin{cases} = \beta(Z_A - Z_B) & \text{for } 0 \leq \beta(Z_A - Z_B) \leq 1 \\ = 0 & \text{for } \beta(Z_A - Z_B) < 0 \\ = 1 & \text{for } \beta(Z_A - Z_B) > 1 \end{cases}$$

where:

Z_A, Z_B ≡ vector of attributes

β ≡ estimated vector of coefficients.

Within the probability choice framework, an implicit assumption of the disaggregate linear model is that the errors in utility are distributed uniformly.

The arguments against using linear models of binary choice tend to be theoretically convincing. Its specification is less attractive than the ogive shaped curves which result from either logit or probit analysis. Estimation of linear probability models leads to biased coefficients if specification error is minimized.

Practical experience with applying linear probability models has yielded a mixed record of success. In a recent study of cable television demand, it was found that the linear model was perferable to the logit specification.[1] The approach to determining specification error was to compute the percentage of individuals in a market whose estimated "probability" fell outside the unit interval. The result showed that less than two percent of the sample needed to be truncated. When aggregated over the market, this small error was decreased because individual errors tended to cancel.

For this study, the binary choices between auto and transit were predicted on zonal interchange data from Los Angeles with a linear model. The model performed better than binary choice logit models which were adjusted for aggregation error. However, the linear model performed poorly on higher levels of aggregation; in particular, using the means of the Los Angeles data produced a forecast of a greater than 100 percent mode split in favor of auto.

Perhaps the major drawback of the linear probability model, in its current state of development, is the binary choice limitation. Unlike the logit specification, there is no natural way of introducing new modes. Alternative modes can be nested in a binary decision tree approach, but this represents a theory of choice which, in many situations, would be highly artificial. In general, more development is needed to make linear models of more than two unranked alternatives into useful planning tools.

---

[1] CRA, *Analysis of the Demand for Cable Television* (1973).

If alternatives can be ranked, and are relatively large in number, then linear models offer distinct advantages over logit models. In Section 3, disaggregate linear models of nonwork travel behavior are estimated which predict travel frequency and average distance of trip by mode over a four day travel period for a household. The specification of this model is directed toward including as causal variables those travel attributes which can be translated into policy instruments; this allows ease of application. It can be argued that it is preferable to have continuous variable models in cases where the number of alternatives is large and ease of application has a high value.

2. WORK TRAVEL

2.1 INTRODUCTION

Recently developed disaggregate logit demand models have been successful in replicating the mode choice decision of commuters. Potentially, it is a powerful analytical device for evaluating the effects of transportation policy on work trip behavior. However, as discussed in Section 1, the available data bases on urban travel have yet to be constructed in such a way that disaggregate models can be applied with ease and accuracy. Moreover, it is unlikely that such data bases will be available on a national scale in the very near future.

This situation argues for developing techniques for capturing the policy evaluation benefits of logit models with existing data bases. In this Section, several methods are developed and tested for applying existing disaggregate demand models to, first, nationwide market segment data from the Nationwide Personal Transportation Survey and, second, to sketch plan zone data from Los Angeles. The techniques are used to evaluate the effects of a variety of transportation policy scenarios including parking restrictions, transit improvements and gasoline taxes.

2.2 NATIONWIDE POLICY EVALUATION WITH MARKET SEGMENTS

As described in the previous Section, the most useful national database on recent urban travel behavior appears to be the 1969 Nationwide Personal Transportation Study (NPTS). Applying existing disaggregate demand models to tabulations of NPTS data would be a valuable device for quick policy evaluation at the national level. The method described below gives reasonable predictions of travel behavior under a variety of policy contingencies. The effects of a policy scenario can be computed with minimal resources; calculations would be within the range of most programmable calculators, or could be performed by hand in several hours.

The approach presented below rests on tabulating the data so as to minimize the variation of mode attributes within each data grouping. As explained in our discussion of the aggregation problem, this will have the effect of mitigating the aggregation bias which occurs when logit models are applied to grouped data. A second part of the approach makes heuristic assumptions about the level of service of modes not originally included in model estimation. That is, because the model was estimated only on the alternatives of driving alone or taking transit, several adjustments in the basic model are necessary in order to predict auto passenger travel. The approach is validated by applying existing mode split models to the data and checking for consistency in predicting actual travel behavior. The CRA model (1972) is then applied to the NPTS tabulation in a variety of policy evaluation exercises.

2.2.1 Data and Model Preparation

The NPTS database is not currently stored nor tabulated in forms which allow direct application of disaggregate mode choice models.[1] In particular, much of the information about the modes not chosen by an individual was not collected in the original survey. Moreover, the cross tabulations performed by the authors of survey reports are not appropriate for application of existing logit models. In this section, we describe how the original home interview tape from the survey was cross tabulated into market segments suitable for application of the CRA work trip mode split model. This effort is essentially a three stage process: (a) the relevant variables are identified from the demand model;

[1]The NPTS survey is described in Appendix A to this report.

(b) the market segments are formed from the household interview tape; and (c) the variables representing market segments are constructed for application of the demand model. Each of these stages is discussed in separate sections below.

Mode Split Model and Variables

The general form of the logit mode split model was presented in Section 1. Here we rewrite it somewhat for the ease of presentation and model application:

$$P(a) = \frac{1}{1 + \sum\limits_{\substack{i=1 \\ i \neq a}}^{n} e^{-\alpha(x_a - x_i) - \beta y}} \tag{2-1}$$

$$P(i) = \frac{e^{-\alpha(x_a - x_i) - \beta y}}{1 + \sum\limits_{\substack{j=1 \\ j \neq a}}^{n} e^{-\alpha(x_a - x_j) - \beta y}} \tag{2-2}$$

$$ln\left(\frac{P(a)}{P(i)}\right) = \alpha(x_a - x_i) + \beta y \qquad for\ all\ i \neq a \tag{2-3}$$

where: $P(a)$ = probability of auto-drive-alone being the chosen mode;

$P(i)$ = probability of alternative i being the chosen mode;

x_a = vector of costs and times for making the trip by the auto drive alone mode;

x_i = vector of costs and times for making the trip by mode i;

y = vector of socioeconomic variables and mode specific constants;

α, β = estimated vectors of coefficients for the time, cost and socioeconomic variables and for the mode specific constants.

Each of the above relationships is derived from the general logit formulation given by Equation (1-1) in the previous chapter. For the purposes of presenting the models used in this report, we will prefer to work with Equation (2-3).

Several models were tested on the NPTS data with the result that the CRA (1972) mode split model gave the best results.[1] Using the form of Equation (2-3) the estimated model is as follows:

$$\ln\left(\frac{P(a)}{P(b)}\right) = \underset{(3.88)}{-4.77}\underset{(4.53)}{-2.24}(C_a-C_b)\underset{(1.96)}{-.0411}(T_a-T_b)$$

$$\underset{(2.69)}{-.114}(S_a-S_b)\underset{(4.06)}{+3.79}Y \qquad (2\text{-}4)$$

where: $P(a)$ = probability of auto-drive-alone being the chosen mode;

$P(b)$ = probability of transit being the chosen mode;

$C.$ = costs of making the round trip by auto (a) or transit (b), in dollars;

$T.$ = invehicle and wait time for the round trip by auto (a) or transit (b), in minutes;

---

[1]As a result of the review of existing models in Section 1, it was decided that the models to be tested included the CRA (1972) model and the battery of models estimated by McFadden (1974). The Haws and Ben-Akiva model was rejected because of unacceptably low sensitivity of travel behavior to mode costs. The CRA linear model of mode choice (1972) was rejected because new modes could be introduced only in an extremely ad hoc fashion. Tests with the McFadden models indicated the following problems: (a) for areas outside of San Francisco (where the model was estimated) the model tends to substantially overpredict transit patronage; and (b) the coefficients on transit access were so small that policies oriented toward increasing transit availability are predicted to have extremely small impacts.

$S_{.}$ = access walk time for the round trip by auto (a) or transit (b), in minutes (usually assumed to be zero for auto trips);

$Y_{.}$ = autos per worker in the household.

Because the model and its estimation is described in detail in other publications, we will not evaluate it here except to note some of the parametric test statistics.[1] For the sample size used to estimate equation (2-4), which was 115 observations, t-statistics of 1.96 and 2.33 indicate the parameter is significantly different from zero at, respectively, the 2.5 percent and 1 percent levels of significance for a one-tailed test. The t-statistics in equation (2-4) are given in parentheses under the parameter estimates and it can be seen that the estimated parameters are all highly significant. Another test of the model is whether the predicted probability of the selected mode for individuals is greater than 0.5. Equation (2-4) performed well in this respect; the model predicted the correct choice of mode for 107 of the 115 observations used in estimating the model for an accuracy level of 93 percent.

Construction of NPTS Market Segments

To construct the NPTS market segments, the work trip records from the home interview survey of urban areas was cross-tabulated across three variables: trip distance, access distance to transit, and automobile availability.[2]

---

[1]See either CRA, *A Disaggregated Behavioral Model of Urban Travel Demand* (1972) or Domencich, Thomas A. and McFadden, Daniel, *Urban Travel Demand: A Behavioral Approach* (Amsterdam: North-Holland Publishing Company, 1975).

[2]To apply the McFadden model, another cross-tabulation across household income was also performed. However, because the McFadden model was not used for scenario evaluation, for reasons given above, this data tabulation is not used.

The market segment categories are presented below.

Distance -- Trip distance was divided into two categories with the following ranges:

short trips -- zero to 9.1 miles;

long trips -- 9.1 and greater miles.

Several different methods could have been applied to determine ranges for the short and long trip category. For example, the median trip distance could have been used for the dividing line or that distance for which total VMT's in each category are equal. In the ranges actually used, the mean trip distance, equal to 9.1 miles on a round trip basis, was decided upon as the best dividing line; this figure falls between those which result from using the other two rules.

Transit accessibility -- The transit accessibility categories were determined by the distance from home to the nearest public transportation line that could be used for the journey to work. The data were originally coded in blocks and were later transformed to miles as per the instructions on the survey instrument (roughly, one block equals one-twelfth of a mile). The transit accessability categories and their ranges in distance are as follows:

high transit accessibility -- zero to two blocks;

middle transit accessibility -- three to six blocks;

low transit accessibility -- over six blocks.

These ranges, selected after an examination of the more refined breakdowns, showed which groupings would tend most to equalize the number of trips among categories.

Auto availability -- Household auto availability was was divided into two categories corresponding to the following:

- autos per worker less than or equal to .5:
- autos per worker greater than 0.5.

These categories arise naturally from the data which displays a bimodal distribution with most work tripmakers either having zero or one car per worker in the household.

The basic unit of information used to perform these cross-tabulations is the trip record for an individual's usual trip to work. For the urban areas identified in the NPTS data base, there were 1,774 such trips recorded.[1] Of these, 221 were eliminated on error checks; usually because there was insufficient data on the record. Another 101 trips were purged because they involved more than one mode of travel. The remaining 1,452 trip records form the basis of the market segments used for analysis. Note that some bias in the policy evaluation is incurred because multi-modal and walk trips are not included. This issue will be discussed along with the policy evaluation results.

Table 2-1 presents the average travel characteristics for each category of the variables used in constructing the market segments. For the distance categories, the average time and distance for each mode are given. From the survey instrument, each mode category has the following definition:

- Auto-drive- alone -- automobile - alone, truck, motorcycle;
- Transit -- bus, streetcar, commuter train, subway, elevated, etc.;
- Carpool -- automobile - with other persons.

---

[1]Appendix A contains a description of the NPTS survey and our procedures for selecting households in urban areas.

TABLE 2-1

AVERAGE ROUND TRIP CHARACTERISTICS
FOR NPTS MARKET SEGMENTS

	Distance and Time	
	Short Trips	Long Trips
Auto-Drive-Alone:		
Distance (miles)	8.63	33.26
Time (minutes)	29.04	64.53
Transit:		
Distance (miles)	8.73	36.73
Time (minutes)	50.95	109.83
Carpool:		
Distance (miles)	7.96	36.50
Time (minutes)	31.33	73.85
Percent of Trips	62.2	37.8

	Transit Accessibility		
	High Transit Access	Middle Transit Access	Low Transit Access
Distance to Transit (miles)	.068	.375	1.325
Percent of Trips	37.7	16.3	46.1

	Auto Availability	
	Less then .5 Autos/Worker	Greater than .5 Autos/Worker
Autos per Worker	.020	1.753
Percent of Trips	89.7	10.3

Table 2-2 presents the mode splits, number of trips and vehicle miles traveled (VMT's) for each of the twelve market segments. Mode split and total trips were computed directly from the data but some assumptions were necessary in order to compute the VMT's. That part of the VMT's which can be attributed to the auto-drive-alone mode is simply the sum of round trip distances for each trip made by this mode. However, the information in the data base does not allow a direct computation of VMT's incurred by carpools because, without knowing the distribution of carpool sizes, one does not know the number of passengers per vehicle and, hence, one does not know the number of vehicles used for this mode. To derive an estimate of the VMT's attributable to carpools a distribution of one-, two- and three-passenger carpools was created and each person in the carpool was credited with an equal share of the carpool's VMT's. The distribution of carpool sizes is derived from the predictions of the mode split model.[1] This distribution, of course, varies from cell to cell, but its aggregate ratio is 0.78:0.17:0.04 for one-passenger: two-passenger: three-passenger carpools respectively. The NPTS distribution, tabulated from a different part of the survey, is that, for all travel, the ratio of carpool sizes is 0.72:0.17:0.11 for one-passenger:two-passenger:three-passenger carpools respectively. Thus, the two independent estimates of passengers per auto are in reasonably close agreement.

To summarize the discussion of the construction of market segments, it should be noted that there is a large amount of flexibility in deciding the number of variables to be cross-tabulated, the number of categories to be used,

---

[1]See the following section to determine how the mode split model is used to predict carpool trips.

TABLE 2-2

MODE SPLITS, TRIPS AND VMT's FOR NPTS MARKET SEGMENTS

	Greater than .5 Autos per Worker						Less than .5 Autos per Worker					
	Short Trips			Long Trips			Short Trips			Long Trips		
	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access
Mode Split:												
Auto Drive Alone	.603	.589	.780	.644	.711	.798	.014	.083	.000	.042	.083	.400
Transit	.197	.113	.006	.178	.132	.027	.836	.917	.083	.792	.833	.000
Carpool	.200	.298	.214	.178	.158	.175	.151	.000	.917	.167	.083	.600
Total Trips	315	124	355	135	76	297	73	24	12	24	12	5
VMT's	1866	763	2663	3317	2008	8799	48	17	39	104	51	119

and the ranges to be applied. The decisions made about each of these issues reflected a desire to minimize the number of market segment cells and, at the same time, capture the essential information for application of the mode split model in data points with small associated variances. New variables and more refined breakdowns of chosen variables increase the number of cells multiplicatively rather than additively; for example, if in addition to the variables already chosen it was decided to perform a cross-tabulation using two categories of trip time, the number of market segment cells would have increased from 12 to 24. Unless one is somewhat ruthless about creating broad ranges for categories and selecting relatively few variables, the data base can easily become overly cumbersome thereby loosing the advantages of using market segments.

Though the choice of ranges and variables is, at base, rather arbitrary, there were some rules and reasons behind the decisions actually made. In addition to those already presented some of the more important of these are listed below.

- Variables were selected to conform to the independent variables in the logit model. Both autos per worker and access time to transit are direct inputs to the model. Line-haul costs and times are treated as functions of trip distance, making this variable an obvious choice on which to make a cross-tabulation.

- Though trip time data is available, and is an input into the model, it is so closely proportional to trip distance that it was deemed unnecessary to create an extra variable for the purpose of cross-tabulating by time or trip.

● Those variables which contribute most to the aggregation problem require more refined categories. Earlier research indicated that autos per worker and access to transit cause more variation in logit model log odds functions than other variables.[1] For this reason, access to transit was divided into three categories rather than two. Autos per worker naturally divides into two categories, as was mentioned above. To test whether extra accuracy would be obtained with further disaggregation, long trips were further subdivided into two distance categories thereby creating three distance categories altogether and eighteen market segments in total. It was discovered that this did not substantially increase the accuracy of model predictions.

It is apparent that the market segments created are dictated by the requirements of the model and the empirical testing of the model performance. In this sense, the market segments presented here are meant to be suggestive of what can be done for application of nonlinear disaggregate models. Because models and data bases vary, the cross-tabulations performed by other researchers for policy evaluation purposes will also vary.

Construction of Mode Specific Variables

The independent variables required for application of the mode split model need to be constructed from the variables used for creating the market segments data. The model's variables, in the two mode case of auto-drive-alone and transit, are given in equation (2-4). The variables available

---

[1]Charles River Associates, *Policies for Controlling Automotive Air Pollution in Los Angeles* (Cambridge, Mass.: forthcoming).

from the data are given in Table 2-1. In addition to the two modes represented in equation (2-4), it is useful to construct data representing auto-with-passenger modes. In describing how we have transformed the data to accept the logit model, each mode will be discussed separately.

Auto-drive-alone -- Construction of the auto specific data points is relatively straightforward. Each of the relevant variables and associated assumptions is presented below.

C_a (round trip cost) -- for each distance category, short and long trips, the cost of a trip is computed as $0.035 times the average distance (presented in Table 2-1). The cost of a trip includes gasoline, oil, tires and maintenance; the national average of these costs per mile was computed from data presented in the *Statistical Abstract of the United States: 1971* (United States Bureau of the Census; 1971), table number 854.[1]

T_a (round trip in-vehicle time) -- the data used to represent this variable is the average auto-drive-alone round trip time for each distance category presented in Table 2-1.

S_a (walk access time) -- This variable is always zero for the auto-drive-alone mode.

Transit -- Because of the limited data collected by the NPTS survey, several assumptions were made in order to

---

[1]It should be noted that it is common in transporation demand modeling to include as part of the trip costs some figures which represent cost of car purchase and insurance. However, these costs are only partially incurred by any given trip and are more appropriately costs common to all travel and a fixed cost of automobile ownership rather than trip-making. Their allocation to individual trips is, at best, arbitrary and most likely inappropriate. In the original CRA model as well as in this study, they are excluded.

construct variables consistent with the mode split model. These are described below for each of the variables.

C_b (round trip cost) -- The fare for a round trip by transit was set at the national average of $0.4928 for 1969.[1] Though it is likely that some long trip alternatives incur higher fares than short trips, there is not data which is easily available to distinguish transit trip costs between the two cost categories. It is also probable that potential transit fares are higher for individuals who took other modes but, again, this datum is missing. The assumption of a flat fare for trips of all distances and for transit trips not taken biases, somewhat, the predictions in favor of transit, all other things equal.

T_b (round trip wait plus in-vehicle time) -- the value for this variable is assumed to be the average times for long and short trips presented in Table 2-1. The average time was constructed by deducting from reported trip times an estimate of access time (presented below). By assuming that the transit time that typically occurs for transit patrons also occurs for auto drivers and passengers, the estimates of mode split will be somewhat biased in favor of transit.

S_b (walk access time) -- the access distance to transit for each trip by each mode was asked in the survey. Where responses indicated that transit was not available, an access distance of one mile was assigned.[2] Walking speed

---

[1]American Public Transit Association, *Transit Fact Book: 1974-1975 Edition* (Washington, D.C.: March 1975), p. 20.

[2]The sensitivity of mode split estimates to this assumption is relatively low because transit access distances in this range would entail a mode split prediction near zero for transit.

was assumed to be 19 minutes per mile. This speed was multiplied with the average access distances for each of the access categories, given in Table 2-1, to derive access time.

Carpool -- As discussed in Section 1, if mode choices are constrained to be auto-drive-alone and transit, the model will not typically predict the full range of responses caused by a policy. Unfortunately, the only suitable mode split models from the standpoint of accurately representing auto-drive-alone and transit choices do not include carpool as an alternative mode. However, the form of the logit model allows one to forecast the effects of a new mode if appropriate data on the times and costs of trips by the new mode is available. It was found that reasonable results could be obtained if the time and cost for carpool trips was substituted into equation (2-3) for the transit variables thereby yielding a log-odds equation for auto-drive-alone versus carpool. It is necessary to distinguish between carpools of differing sizes and treat each of these as a separate mode. In the approach used below, one-passenger, two-passenger and three-passenger carpools are considered to be three separate modes, each having a different configuration of times and costs.

The major problem with including carpool as a separate set of mode choices is the lack of appropriate data. From the NPTS we have information about the distances and time associated with carpools which have actually formed. However, we have no information about the performance of carpools which would be an alternative for the individuals who take transit or drive alone. It can be presumed that people will ride in a carpool mainly if they can find tripmakers with roughly the same origin, destination and work hours. Because the origins and destinations of the passengers of existing carpools are tightly clustered, the reported trip times and distances of these carpools in the NPTS data are not representative of the average tripmaker who would probably be much farther outside of the established routes to work of potential carpool members.

The approach presented to account for carpool trips is largely heuristic. In the absence of data, assumptions and judgements must be made and those presented below reflect the subjective opinions of the project staff. In another section, with a different data base for analysis, carpooling variables are constructed with a different set of assumptions. Perhaps the main value of this exercise is the experience gained to take account of new modes when the available information about these modes is meager at best. Other researchers can use alternative judgments or better data, but the general approach to applying the model to new modes will remain unchanged.

Each of the relevant variables for application of the mode split model to carpool modes is discussed separately below.

C_{ck} (round trip costs) -- The round trip cost to a potential member of a k-passenger carpool depends on the distance traveled by the vehicle and the number of people in the carpool. For each distance category, it is assumed that the average auto-drive-alone trip distance increases by a third for picking up and dropping off any passenger. This is an ad hoc judgement of the difficulty the average driver experiences in finding passengers for a carpool. To determine the cost for each potential tripmaker by carpool, the trip distance is multiplied by the average auto operating cost per mile, \$0.035,[1] and divided by the number of carpool members. It is assumed that, over the long run, carpool members share costs equally.

T_{ck} (round trip wait plus in-vehicle time) -- The time associated with a carpool is divided into three components: the time attributable to picking up and dropping off passengers; the time associated with the linehaul journey from home to work; the schedule delay which occurs because of waiting and potential mismatches of work hours. The first step in constructing these time components is to determine average carpool speed from the data presented in Table 2-1: for short trips, the average speed is 15.25 mph; for long trips, the average speed is 29.65 mph. It was assumed that the speed of travel for picking up and dropping off passengers would be the average speed for short trips regardless of the length of the linehaul portion. As before, the distance for this portion of the trip is one-third the average distance for auto-drive-alone short or long trips, whichever is relevant, times the number of passengers in the carpool. This distance is divided by the average speed for short trips to calculate the time for picking up and dropping off passengers. The linehaul time is calculated by dividing the auto-drive-alone trip distance, for each of the short and long trip categories, by the average carpool speed for short or long trips as appropriate. Schedule delay was assumed to be twenty minutes times the number of passengers. The separate components were added together to compute the total time associated with carpool modes.

---

[1]The section of the costs of the auto-drive-alone mode discusses how the average auto operating cost per mile is derived.

S_{ck} (round trip walk access time) -- As in the case of auto-drive-alone, the access time for carpools was assumed to be zero.

Driver serve passenger -- An additional mode choice includes the option of a passenger being driven to work by another individual, usually from the same household, where the driver returns to origin after dropping off the worker (or, alternatively, uses the vehicle for another purpose before returning to the origin). Though this mode is usually not well delineated in data sources, it can be presumed that passengers reporting this trip type would have categorized it "automobile -- with other persons" in the NPTS survey. Thus this classification merges trips made by carpools with those which were chauffeured. Because the attributes of this trip are different from those of other modes, and because the response to this mode is sensitive to transportation policy, it was considered important to treat it as a separate mode.[1] The level of service variables created to represent this mode are presented below.

C_d (round trip costs) -- It was assumed that the driver serve passenger alternative entails a household member driving the tripmaker to work and then returning home for the first leg of the round trip and then driving from home to

---

[1]For a discussion of the effects of various policies on this mode choice, particularly parking taxes, see Frederick C. Dunbar, "Evaluation of the Effectiveness of Pollution Control Strategies on Travel: An Application of Disaggregated Behavioral Demand Models," in *Proceedings of the Transportation Research Forum, Vol. XVI*, (1975).

the workplace and back home with the passenger for the second leg. Because the distance associated with this mode choice is double that of driving alone (neglecting employee parking distances from place of work) the driver serve passenger mode can be calculated as twice the cost of the auto-drive-alone mode.

T_d (round trip in-vehicle time) -- The construction of the driver serve passenger mode entails triple the person time of an auto drive alone trip. A driver makes two round trips for every passenger round trip. To account for this, the dirver serve passenger time was calculated as three times the auto-drive-alone time, reported in Table 2-1, for each of the distance categories. This assumes that the extra time incurred by the driver is weighted equivalently with the time of the passenger. This assumption was modified in applications of the model to Los Angeles data, presented later in this section.

S_d (round trip walk access time) -- As with other auto oriented modes, this variable was assumed to equal zero for the driver serve passenger mode.

Autos per worker -- This variable was included in only the log odds functions comparing auto-drive-alone to transit (equation 2-3). Whenever this variable is not used, the constant term in the log odds function is set to zero. The average values are given in Table 2-1. In application of the model it was set to zero or one depending on whether the relevant category was less than or greater than 0.5 autos per worker. This construction was suggested by the data. By truncating the value this way, the coefficient on autos per worker needed to be changed. The value selected, discussed at more length later in this section, was 4.60.[1]

---

[1]There is reason to believe that adjustments in this coefficient and the constant term overcome pro-transit biases in the data and deal with the problem of transferrability.

This completes the description of the construction of independent variables for the logit model used on NPTS market segments. For quick reference, Table 2-3 is provided with summaries of the formulas representing the variables. The terms in brackets refer to data elements which appear in Table 2-1. Using these two tables, the independent variables for each market segment can be constructed.

2.2.2 Model Application

This section describes the performance of the model when applied to the NPTS market segments. To apply the model in order to predict mode splits and VMT's, the following steps are taken:

- Each of the mode specific variables for each of the twelve market segments is constructed using the formulas presented in Table 2-3 and the data presented in Table 2-1.
- For each market segment, a log odds function for auto-drive-alone verses each of the other modes is calculated using equation (2-4) with the variables constructed in the previous step and with 4.60 substituted for the coefficient on y.
- For each market segment, the probability of an individual choosing each mode, other than auto-drive-alone, is computed using equation (2-2). The auto-drive-alone probability is computed from equation (2-1).
- Mode splits for each market segment are computed as follows:
  - auto-drive-alone mode split = auto-drive alone mode choice probability;
  - transit mode split = transit mode choice probability;
  - carpool mode split = sum of one-passenger, two-passenger, three-passenger carpool mode choice probabilities plus driver serve passenger mode choice probability.

TABLE 2-3

FORMULAS FOR CONSTRUCTION OF MODE SPECIFIC VARIABLES

Auto Drive Alone

$$C_a = .035 \times [\textit{Auto-Drive-Alone Distance}]$$

$$T_a = [\textit{Auto-Drive-Alone Time}]$$

$$S_a = 0$$

Transit

$$C_b = .4928$$

$$T_b = \textit{Transit Time}$$

$$S_b = 2 \times 19 \times [\textit{Distance to Transit}]$$

Carpool (with k passengers)

$$C_{ck} = \frac{(1+k/3) \times C_a}{k+1}$$

$$T_{ck} = \frac{\frac{k}{3}\ [\textit{Auto-Drive-Alone Distance}]}{[\textit{Carpool Distance:Short Trips}]/[\textit{Carpool Time:Short Trips}]} + \frac{[\textit{Auto-Drive-Alone Distance}]}{[\textit{Carpool Distance}]/[\textit{Carpool Time}]} + 20 \times k$$

$$S_{ck} = 0$$

Driver Serve Passenger

$$C_d = 2 \times C_a$$

$$T_d = 3 \times T_a$$

$$S_d = 0$$

Autos per Worker

$$y = \begin{cases} 0 & \textit{Less than .5 Autos/Worker} \\ 1 & \textit{Greater than .5 Autos/Worker} \end{cases}$$

● VMT's for each market segment is the sum of the following VMT calculations for each mode:[1]

- auto drive alone VMT = (auto-drive-alone mode choice probability) x (auto-drive-alone distance) x (total trips)
- k-passenger carpool VMT = (k-passenger carpool mode choice probability) x $(1 + \frac{k}{3})$ x (auto-drive-alone distance) x (total trips)
- driver serve passenger VMT = (driver serve passenger mode choice probability) x 2 x (auto-drive-alone distance) x (total trips).

● Aggregate mode split is computed as the weighted average of predicted mode splits for each market segment where the weights are the proportion of the total trips in the market segment to the total trips for all market segments.

● Aggregate VMT's are computed as the sum of VMT's for each of the market segments.

With these procedures, the model was used to predict mode splits and VMT's for each of the cells in the NPTS market segment database. The aggregate predictions are given in Table 2-4. It can be seen that the predicted mode splits conform closely to the actual mode splits. Actual aggregate mode splits were calculated using the data in Table 2-2; the mode split for each market segment was weighted by the proportion of total trips in the market segment to the aggregate total and the weighted market mode splits were then summed.

Perhaps the most important single dependent variable to to the policy makers is the VMT's. The model is used to predict two VMT figures in Table 2-4. The first of these (VMT wo/DSP)

---

[1] In all model applications, only VMT's by private auto are computed. VMT's attributable to public transit vehicles are not estimated.

TABLE 2-4

PREDICTED BASE CASE VS. ACTUAL AGGREGATE MODE SPLITS AND VMT'S FOR NPTS MARKET SEGMENTS

	Actual	Predicted
Mode Split:		
Auto-Drive-Alone	.637	.635
Transit	.160	.159
Carpool	.202	.206
VMT wo/DSP	19794	19475
VMT		19913

corresponds to the VMT's which can be calculated from the data; it does not include the VMT's attributable to one-half the driver serve passenger trips (that half which is traveled by the driver without a passenger is not captured by the NPTS data). The second figure (VMT) includes all of the predicted VMT's associated with driver serve passenger trips as well as with other auto oriented trips. Using the first figure as a basis for comparing the predicted to the actual, the model predicts VMT's within 1.6 percent.[1] For most applications of the model, this error is well within predicted effects and within errors which would be associated with other causes such as data errors or parameter estimation errors. In general, the model performs well in replicating the aggregate figures from the data.

---

[1]Some of this accuracy is attributable to other effects besides the performance of the basic model. By adjusting the coefficient on autos per workers to achieve desirable results, the effect is similar to adjusting the mode specific constant so that the predictions of aggregate effects are more accurate. A complete evaluation of the model and procedures rests on its ability to predict several dependent variables owing to this single adjustment and the reasonableness of the elasticities.

To provide more information on the performance of the model, the mode split and VMT estimates for each market segment are presented in Table 2-5 in a format comparable to the actual tabulations presented in Table 2-2. A comparison of Table 2-5 with Table 2-2 indicates potential biases and the areas of greatest error. As expected, the error associated with any given market segment is greater than the aggregate error. The highest errors are associated with those market segments which have the fewest total trips (basically, the six market segments where autos per worker is less than .5). The other major item to note is that there appears to be some tendency for the model to overpredict driver serve passenger trips for short distance trips and underpredict carpool trips for long distance trips. The effects of these biases on VMT estimates under various policy scenarios will be discussed in the next section. In general, the errors associated with individual market segments tend to cancel when aggregated.

2.2.3 Predicted Policy Effects

The procedures developed above were applied to a variety of transportation policy scenarios to predict the effects of these policies on tripmaking behavior. The approach to investigating a particular policy is relatively straightforward. A policy is examined from the standpoint of how it would effect the independent variables in the logit model. This effect is quantified by changing the value of the independent variables from what they were in the base case. With the new values of the variables, the logit model is applied to the NPTS market segments data and mode splits and VMT's are predicted. The predicted mode splits and VMT's with the policy effects are then compared to the base case predictions in order to determine the impact of the policy.

TABLE 2-5

PREDICTED BASE CASE MODE SPLITS AND VMT'S FOR NPTS MARKET SEGMENTS

	Greater than .5 Autos per Worker						Less than .5 Autos per Worker					
	Short Trips			Long Trips			Short Trips			Long Trips		
	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access
Mode Split:												
Auto-Drive-Alone	.598	.667	.695	.559	.749	.851	.040	.132	.650	.016	.057	.696
Transit	.140	.041	.001	.345	.122	.002	.942	.811	.056	.981	.933	.184
Carpool	.261	.292	.303	.096	.129	.147	.018	.058	.284	.003	.010	.120
VMT wo/DSP	1890	830	2476	2737	2065	9172	30	32	78	14	25	126
VMT	2042	896	2675	2739	2066	9178	32	34	85	14	25	126

The policies to be examined include increased gasoline taxes, increased transit linehaul speeds, a variety of transit access improvements, and the introduction of a new mode (dial-a-ride).

In some cases, the aggregate effects of a policy can be summarized with a computation of the elasticity of the effect with respect to the variable which the policy changed. In calculating elasticities the following formula is applied:

$$\textit{Elasticity of X with Respect to Y} = \frac{\textit{Percentage Change in X}}{\textit{Percentage Change in Y}}$$

Most often, we will be interested in the elasticity of VMT's with respect to some other variable such as gasoline costs or transit accessibility. Sometimes the elasticity measure is not well defined because the policy control variable cannot be easily quantified; for example, improving transit access may entail making transit more available to a certain segment of the population and the aggregate effects of this policy are not well summarized by an elasticity measure.

As this example demonstrates, a complete evaluation of a policy entails more information than simply the aggregate effects. For this reason, each time a policy scenario is presented, we also give the predicted effects of the policy on each of the market segments. It is most often the case that a particular scenario has widely varying effects across different market segments.

Finally, it should be noted that analyzing the demand effects only presents one half of a policy evaluation. The cost-effectiveness of a policy option also, rather obviously, depends upon the costs involved. It will be shown, for example, that the elasticity of transit ridership with respect to changes in linehaul plus wait time is higher than the elasticity of transit ridership with respect to changes in access time. This result does not in and of itself constitute a complete policy evaluation of these two options. As will be discussed when these results are presented, the system changes implied by the two scenarios are quite different and more needs to be known about the costs associated with each policy before the implications for transit investment policy are determined.

Gasoline Tax

The model was used to predict the effects of a 100 percent gasoline tax in addition to the existing gas taxes (an assumed 7 percent state tax and 4 percent Federal tax). One of the purposes of this exercise is to compute the implied price elasticity of gasoline which the model applied to NPTS market segments would estimate. This provides a test of the approach because the result can be compared to other, independent, gasoline price elasticity estimates.

The effect of a 100 percent gasoline tax will be to increase the operating cost per mile of an auto by 50 percent. The pretax cost of gasoline was half the cost of auto operating costs in the 1969 base case. The pump price of gasoline is increased by 69 percent when a 100 percent tax rate is applied to the pretax cost of gasoline. In terms of applying the model, the new operating cost of $0.0525 is substituted for the base case figure of $0.035 per mile. The procedures for applying the model which were presented in the previous section are then followed.

The predictions of aggregate mode split and VMT's under the assumption of a 100 percent gas tax are presented in Table 2-6. In addition, the elasticity of VMT's with respect to the pump price of gas is -.184, which is somewhat lower than the results of other studies presented in the previous chapter but is within the range of statistical error. The predicted effect of the policy on VMT's is a 12.8 percent decline; transit trips are predicted to increase by a third; auto drive alone trips are predicted to decrease by 11.2 percent, and carpools are predicted to increase by 8.7 percent.

More detail on the predicted effects is given in Table 2-7 where the mode split and VMT's for each of the NPTS

TABLE 2-6

PREDICTED AGGREGATE EFFECTS
OF A 100 PERCENT GASOLINE TAX

	Base Case	100 Percent Gas Tax
Mode Split:		
Auto-Drive-Alone	.635	.564
Transit	.159	.212
Carpool	.206	.224
VMT	19913	17365

VMT Elasticity with respect to:
Auto Operating Costs = - .256;
Pump Price of Gas = - .184;
Pre-Tax Cost of Gas = - .128.

market segments is presented. In order to make a comparison with the base case predictions it is necessary to refer to Table 2-5. From such a comparison it can be seen that the gasoline tax has most impact on long trips with good to fair transit access. This is to be expected because, on a per trip basis, the gasoline tax has the highest dollar impact on long trips. At the same time, transit trips are assumed by the model to cost the same amount regardless of the length. The result is that the model predicts a higher incentive for mode switching on long trips for this scenario. It should also be remembered that there is some pro-transit bias built into the model so the predicted effects may be biased; what the model does not consider is that even if there is relatively good access to transit, those individuals who now use cars may tend to do so because the transit alternative involves long, circuitous routes or multiple

TABLE 2-7

PREDICTED MODE SPLITS AND VMT'S FOR NPTS

MARKET SEGMENTS WITH 100 PERCENT GAS TAX

	Greater than .5 Autos per Worker						Less than .5 Autos per Worker					
	Short Trips			Long Trips			Short Trips			Long Trips		
	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access
Mode Split:												
Auto-Drive-Alone	.554	.640	.677	.282	.534	.781	.029	.099	.619	.004	.016	.440
Transit	.182	.056	.001	.642	.321	.008	.957	.854	.087	.994	.979	.434
Carpool	.264	.304	.332	.077	.144	.211	.014	.047	.294	.001	.004	.120
VMT	1883	856	2592	1448	1544	8818	23	26	80	4	7	85

TABLE 2-8

PREDICTED AGGREGATE EFFECTS OF A 10 PERCENT TRANSIT SPEED INCREASE

	Base Case	10 Percent Transit Speed Increase
Mode Split		
Auto-Drive-Alone	.635	.619
Transit	.159	.179
Carpool	.206	.209
VMT	19913	19273

VMT Elasticity with Respect to Transit Speed = -.322.

Transit Ridership Elasticity with Respect to Transit Speed = 1.26.

transfers. The result of this bias is to overpredict, somewhat, the effects of the policy on mode switching and, consequently, on VMT's reduced when there is a major level of service change.[1]

Transit Speed

In this scenario, it is assumed that the combination of shorter headways and faster transit cause a uniform decrease in transit linehaul plus wait time per trip of ten percent. Access time to transit was assumed to be unchanged. This scenario was modeled by multiplying transit linehaul plus wait time by 10/11 and then applying the logit model to the NPTS market segments according to the procedures presented before.

The predicted aggregate effects of this policy are presented in Table 2-8. The predicted decline in VMT's was

[1]This effect was reduced by adjusting the coefficient on autos per worker. Policies which have small impacts on auto level of service will entail no bias in the predicted elasticities from protransit bias in the data.

3.22 percent and the predicted increase in transit trips was 12.6 percent. One of the interesting results from this exercise is the relatively high elasticity of transit mode split with respect to transit speed (1.26).

The disaggregated result of the model forecast is presented in Table 2-9. By comparing this with Table 2-5 (the base case predictions by market segment) it can be seen that the biggest impacts occur on trips with good to medium transit access and which are relatively long. As in the case of a gas tax, the effect of a uniform percentage decline in transit time will have the biggest impact in absolute terms on long trips. Consequently, those tripmakers which face the longer journeys have the most incentive to switch modes. As the figures from Table 2-1 indicate, the 10 percent decline in transit time implies a savings of about 10 minutes for long trips compared to 5 minutes for short trips. Also, as would be expected, the transit speed policy has little predicted effect on tripmakers with poor access to public transit.

Transit Access: Uniform Improvement

Because the weights that tripmakers place on access time to transit are higher than the weights placed on linehaul time, it is natural to assume that the effect of decreasing access time would be greater than the effect of decreasing linehaul plus wait time. The results of various transit access scenarios indicates that this hypothesis deserves more consideration.

The first of a series of transit access improvement scenarios involved decreasing transit access time by a uniform 10 percent for all market segments. In the base case projections, the access times to transit for high, middle and low

TABLE 2-9

PREDICTED MODE SPLITS AND VMT'S FOR NPTS MARKET SEGMENTS WITH A 10 PERCENT TRANSIT SPEED INCREASE

	Greater than .5 Autos per Worker						Less than .5 Autos per Worker					
	Short Trips			Long Trips			Short Trips			Long Trips		
	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access
Mode Split:												
Auto-Drive-Alone	.581	.661	.695	.475	.705	.850	.034	.112	.642	.011	.039	.636
Transit	.165	.050	.001	.443	.173	.003	.952	.838	.078	.987	.954	.254
Carpool	.254	.289	.304	.082	.122	.147	.015	.049	.280	.002	.007	.110
VMT	1984	889	2674	2331	1946	9168	27	29	83	9	17	116

TABLE 2-10

PREDICTED AGGREGATE EFFECTS OF A
10 PERCENT TRANSIT ACCESS TIME DECREASE

	Base Case	10 Percent Transit Access Time Decrease
Mode Split:		
Auto-Drive-Alone	.635	.631
Transit	.159	.163
Carpool	.206	.205
VMT	19913	19780

VMT Elasticity with Respect to Transit Access Time = .067.

Transit Ridership Elasticity with Respect to Transit Access Time = -.252.

access categories were 2.58, 14.25 and 50.35 minutes respectively. Thus, only in the case of short transit trips with poor access would equivalent time savings occur for equal percentage declines in access time compared to linehaul plus wait time. In all other cases, the time savings from a 10 percent reduction in linehaul plus wait time would be much greater than the time savings from a 10 percent reduction in access time. This fact helps to explain some of the results presented below.

The aggregate effects of this policy are described in Table 2-10. The decline in VMT's incurred by this policy is predicted to be 0.7 percent; the predicted increase in transit trips is 2.5 percent. Both the VMT and transit ridership elasticities are much lower for access times than for the linehaul plus wait times presented in the previous scenario.

TABLE 2-11

PREDICTED MODE SPLITS AND VMT'S FOR NPTS MARKET SEGMENT

WITH A 10 PERCENT TRANSIT ACCESS TIME DECREASE

	Greater than .5 Autos per Worker						Less than .5 Autos per Worker					
	Short Trips			Long Trips			Short Trips			Long Trips		
	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access
Mode Split:												
Auto-Drive-Alone	.596	.662	.695	.553	.733	.849	.039	.115	.619	.016	.049	.609
Transit	.144	.048	.001	.351	.141	.004	.944	.835	.110	.982	.942	.286
Carpool	.260	.289	.304	.095	.127	.147	.017	.050	.271	.003	.009	.105
VMT	2034	890	2673	2711	2023	9162	31	30	80	14	22	111

The predicted effects of the policy for each market segment is presented in Table 2-11. There it can be seen that the market segments with the greatest impact (again, comparing these with the base case predictions in Table 2-5) are those where access to transit is in the "middle" category; those with good transit access are relatively insensitive to further improvements and those with poor access would, according to the model, not find a 10 percent improvement enough of an inducement to switch modes.

This is perhaps an excellent example of a situation where a comparison between two scenario forecasts is not sufficient information to warrant making a policy decision. Before it can be said that transit speed and headway improvements are more effective for inducing new transit patrons and reducing VMT's, the implied cost of each of these scenarios should be computed. It may well be that simple extensions of route miles to increase transit access by, on average, 10 percent, are significantly less expensive than the amount of investments which would be entailed to increase linehaul speeds and reduce wait times by 10 percent. Nonetheless, the results are quite suggestive that more research on this issue may be necessary.

Transit Access: Low Availability Improvement

Based on the results of the previous section, a natural policy question to arise is whether making transit available to everybody would induce significant amounts of transit ridership. To give a rough answer, we took the low transit access market segment and assigned to it the same access time which currently obtains for the middle access group. All other variables remained unchanged although it is unlikely that any real transit service design which provided such a large change would not also affect accessibility in other market segments and linehaul and wait times in all market segments.

TABLE 2-12

PREDICTED AGGREGATE EFFECTS OF
LOW TRANSIT ACCESS IMPROVEMENT

	Base Case	Low Transit Access Improvement
Mode Split:		
Auto-Drive-Alone	.635	.600
Transit	.159	.202
Carpool	.206	.198
VMT	19913	18518

VMT Elasticity with Respect to Transit Access Time = .111.

Transit Ridership Elasticity with Respect to Transit Access Time = -.429.

The aggregate results are presented in Table 2-12. The change in average access for the whole population is 62.9 percent; the access time decline for the market segment which previously had low transit availability was 71.7 percent. This rather dramatic change entailed a decline in VMT's of only 7 percent for an elasticity of .111. Transit patronage increases by 27 percent for an elasticity of -.429. These elasticities are higher than in the previous access time scenario because the policy is directed toward those market segments where there is greater sensitivity to access time.

Table 2-13 indicates the changes that occurred in the low access category (when compared to Table 2-5). For households where the autos per worker are greater than 0.5, the predicted change in VMT's is 10.2 percent. The effect of the policy on households with low auto ownership rates is quite dramatic but because these contribute relatively little to VMT's they have a small impact on the aggregate effects.

TABLE 2-13

PREDICTED MODE SPLITS AND VMT'S FOR NPTS MARKET SEGMENTS
WITH LOW TRANSIT ACCESS IMPROVEMENT

	Greater than .5 Autos per Worker						Less than .5 Autos per Worker					
	Short Trips			Long Trips			Short Trips			Long Trips		
	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access
Mode Split:												
Auto-Drive-Alone	.598	.667	.667	.559	.744	.749	.040	.132	.132	.016	.057	.057
Transit	.140	.041	.041	.345	.122	.122	.942	.811	.881	.981	.933	.933
Carpool	.261	.292	.292	.096	.129	.129	.018	.058	.058	.003	.010	.010
VMT	2042	896	2566	2739	2066	8075	32	34	17	14	25	10

TABLE 2-14

PREDICTED AGGREGATE EFFECTS OF LOW AND MIDDLE TRANSIT ACCESS IMPROVEMENTS

	Base Case	Low and Middle Transit Access Improvement
Mode Split:		
Auto-Drive-Alone	.635	.583
Transit	.159	.225
Carpool	.206	.193
VMT	19913	17859

VMT Elasticity with Respect to Transit Access Time = .147

Transit Ridership Elasticity with Respect to Transit Access Time = -.593.

Transit Access: Low and Middle Availability Improvements

Given the above change in policy, it is interesting to determine the incremental effect of improving transit for those in the middle access category. To evaluate this effect, the low access category is again assigned the access value for walk time of the middle access group and the middle access tripmakers were assigned the value walk time of the high access group. No other variables are changed.

The aggregate results of this policy can be seen in Table 2-14. The percentage change in VMT's is 10.3 as compared to 7.0 for the previous scenario. The implied elasticities are somewhat higher for both VMT's and transit ridership. The conclusion which may be drawn from this series of scenarios is that improvements in transit access seem to have the most effect when moderate service is made better rather than when poor service is made only adequate. The effects of the policy on individual market segments is presented in Table 2-15.

TABLE 2-15

PREDICTED MODE SPLITS AND VMT'S FOR NPTS MARKET SEGMENTS
WITH LOW AND MIDDLE TRANSIT ACCESS IMPROVEMENT

	Greater than .5 Autos per Worker						Less than .5 Autos per Worker					
	Short Trips			Long Trips			Short Trips			Long Trips		
	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access
Mode Split:												
Auto-Drive-Alone	.598	.598	.667	.559	.559	.749	.040	.040	.132	.016	.016	.057
Transit	.140	.140	.041	.345	.345	.122	.942	.942	.811	.981	.981	.933
Carpool	.261	.261	.292	.096	.096	.129	.018	.018	.058	.003	.003	.010
VMT	2042	804	2566	2739	1542	8075	32	11	17	14	7	10

Introduction of a New Mode: Dial-a-Ride

One of the features of the multinomial logit model is that it allows the prediction of the probability of choosing a new mode. This attribute of the logit specification was used above in the estimation of mode splits for auto passengers; it was also used by McFadden to forecast the patronage for a new rapid transit system.[1]

To demonstrate the ability of the model applied to NPTS market segments to accept a new mode, travel behavior was simulated with an assumed widely available paratransit service. One method of modeling new service has already been demonstrated in the scenarios involving changes in transit levels of service applied to the existing mode. By embodying changes in level of service in a new mode, as is done below, the model will give somewhat different predictions of mode split effects.[2]

The new mode chosen for analysis is dial-a-ride. Relatively little suitable data on a national basis exists which can be used to give precise estimates of the expected level of service of implementation of dial-a-ride on a national scale. Indeed, the notion of national

---

[1]Daniel McFadden, "The Measurements of Urban Travel Demand," *Journal of Public Economics*, 1974.

[2]To understand the reasons why different methods of applying the model yields different predictions of the effects of equivalent system changes requires a technical discussion of the behavioral assumptions underlying the logit specification. The key assumption is known as the independence of irrelevant alternatives. There are several discussions of this assumption, its validity and methods for forecasting when the assumption is violated. For examples see McFadden, "The Measurements of Urban Travel Demand," *op. cit.*, and Tye and Sherman, *Disaggregate Travel Demand Models*, National Cooperative Highway Research Program Project 8-13: Phase I Report (CRA, Cambridge, Mass.: 1975).

implementation is rather ill-defined. Nonetheless, some rough estimates of the national potential for dial-a-ride service are developed below, using information from compendia on existing paratransit operations.

To estimate the effects of dial-a-ride requires creating a new log odds function, as represented by equation (2-3), and recomputing the mode shares for all existing modes and the new mode with equations (2-1 and 2-2). The parameters on level of service variables and autos per worker are the same in the new log odds function as they were in the others; the values for these parameters are given in equation (2-4).

Two dial-a-ride scenarios are simulated corresponding to low and high level of service. The low performance dial-a-ride scenario is based on the information provided about the characteristics of existing paratransit operations in a recent study.[1] It was concluded that the average dial-a-ride service entailed about four times more travel time (including wait time) than equivalent auto trips. The high performance scenario is based on the assumption that dial-a-ride service can be instituted which provides door to door trip times (including wait time) equivalent to existing transit, on average. In both scenarios, it is assumed that dial-a-ride is only used for short trips, as appears to be the case with actual operations.[2]

Given the above considerations, the dial-a-ride level of service variables for application of the mode split model were defined as follows:

---

[1]Ronald F. Kirby, et. al. *Working Paper: Paratransit Experience and Potential*, (The Urban Institute, Washington, D.C.: 1973).

[2]See *Lea Transit Compendium: Para-transit, Vol. I, No. 8* (N.D. Lea Transportation Research Corporation, Huntsville, Alabama: 1974).

C_p (round trip cost) -- for each trip it was assumed that the dial-a-ride charged a fifty cent flat fare. This fare is relatively common among existing operations. The round trip cost of a dial-a-ride journey would consequently be $1.00.

T_p (round trip wait plus in-vehicle time) -- For low performance dial-a-ride, this variable is assumed to equal four times the in-vehicle time of the auto-drive-alone mode (T_a). For the high performance scenario, this variable is set equal to the total trip time of the existing transit mode ($T_b + S_b$).

S_b (walk access time) -- It is assumed that the dial-a-ride service is door to door and access time is, consequently, equal to zero.

The results of the model simulations are presented in Tables 2-16 through 2-19. As can be seen, the estimated response to a low performance dial-a-ride is negligible. For a high performance dial-a-ride the response is still small, but transit (including dial-a-ride patrons) ridership increases by 11.3 percent. The effect of dial-a-ride on VMT's is estimated to be 1.0 percent; it must be noted that this is an overestimate because it does not include the increased VMT's attributable to the dial-a-ride vehicles. This low effect of dial-a-ride on VMT's is the result of no trip diversion by motorists who travel long distances.

As expected, dial-a-ride compares most favorably with existing transit when the access times to fixed route systems are large. When access times are small, the effect of higher dial-a-ride fares discourages travelers even though the time penalties on conventional service are more severe owing to the disincentives of walking to transit stops. Also as expected, the predicted share of dial-a-ride passengers from the transit captive population is relatively small.

TABLE 2-16

PREDICTED AGGREGATE EFFECTS OF LOW PERFORMANCE DIAL-A-RIDE

	Base Case	Low Performance Dial-A-Ride
Mode Split		
Auto-Drive-Alone	.635	.631
Transit	.159	.156
Carpool	.206	.205
Dial-A-Ride	--	.008
VMT	19913	19857

2.2.4 Summary

The preceding results indicate that the applications of logit models to NPTS (or other) market segments holds some promise for quick, national transportation policy evaluation. It must be admitted that the examples provided are rather simple compared to some of the more complex policy issues, but the translation of a policy option into quantifiable variables in terms consistent with the logit model can yield rough estimates of demand effects in a timely fashion with few computational resources.

TABLE 2-17

PREDICTED MODE SPLITS AND VMT'S FOR NPTS MARKET SEGMENTS
WITH LOW PERFORMANCE DIAL-A-RIDE

	Greater than .5 Autos per Worker						Less than .5 Autos per Worker					
	Short Trips			Long Trips			Short Trips			Long Trips		
	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access
Mode Split												
Auto-Drive-Alone	.596	.664	.692	.559	.749	.851	.039	.121	.449	.016	.057	.696
Transit	.140	.041	.001	.345	.122	.002	.916	.744	.045	.981	.933	.184
Carpool	.260	.290	.302	.096	.129	.147	.017	.053	.196	.003	.010	.120
Dial-A-Ride	.004	.005	.005	--	--	--	.027	.083	.309	--	--	--
VMT	2034	892	2662	2739	2066	9178	31	31	58	14	25	126

TABLE 2-18

PREDICTED AGGREGATE EFFECTS OF
HIGH PERFORMANCE DIAL-A-RIDE

	Base Case	High Performance Dial-A-Ride
Mode Split		
Auto-Drive-Alone	.635	.622
Transit	.159	.148
Carpool	.206	.201
Dial-a-Ride	--	.029
VMT	19913	19718

2.3 POLICY EVALUATION WITH ZONAL DATA

The prevailing method among urban transportation planners for grouping data is to compile it into zonal formats. Most commonly, travel and level of service information is kept in the form of trip tables and networks coded on a grid system which divides an urban area into a number of zones. The relevant observation for the application of a demand model is the number of trips and level of service between two such traffic analysis zones. Because there are often more than one thousand zones in a planning region, the number of zone pairs can exceed one million. Dealing with such a database for the purpose of broad gauge policy evaluation is a time and resource consuming affair. To make urban data bases more manageable, they are often condensed into sketch plan zones or districts. The number of such zones generally number anywhere from 50 to 100. Again, the number of zonal interchanges in such a system, though much smaller than the number of traffic analysis zonal interchanges, is really too large for ease of analysis and manipulation.

TABLE 2-19

PREDICTED MODE SPLITS AND VMT'S FOR NPTS MARKET SEGMENTS
WITH HIGH PERFORMANCE DIAL-A-RIDE

	Greater than .5 Autos per Worker						Less than .5 Autos per Worker					
	Short Trips			Long Trips			Short Trips			Long Trips		
	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access	High Trans. Access	Middle Trans. Access	Low Trans. Access
Mode Split												
Auto-Drive-Alone	.586	.652	.679	.559	.749	.851	.036	.091	.203	.016	.057	.696
Transit	.137	.040	.001	.345	.122	.002	.829	.561	.020	.981	.933	.184
Carpool	.256	.285	.297	.096	.129	.147	.015	.040	.089	.003	.010	.120
Dail-A-Ride	.020	.022	.023	--	--	--	.120	.308	.687	--	--	--
VMT	2001	876	2613	2739	2066	9178	28	24	26	14	25	126

The purpose of this section is to develop methods for conditioning urban data bases and for adjusting disaggregate demand models so they can be applied for quick policy evaluation. In the sections presented below, we describe a zonal data base of manageable proportions for application of disaggregate demand models. Because the zonal interchange data presents a serious aggregation problem, methods are developed to correct for aggregation bias. The models are applied to data which has been summarized in different ways including: sketch plan zone data; the regional average of system performance; and averages of system performance among categories of trip length.

Two basic methods of adjusting the logit model to cope with aggregation bias are proposed. The first is a truncated and simplified Taylor's series approximation which was introduced in Section 1. The approach is based on the work of Antti Talvitie and CRA.[1] A second method uses the area of the zone as a substitute for the variance of the arguments of the logit function. This method was developed for this report; the area adjustment was estimated by a simple technique which is outlined later in this section.

The methods are tested on a policy scenario -- an assumed gasoline tax -- from which implied gasoline price elasticities can be estimated. From this test, only models applied to sketch plan zone data are deemed accurate enough for policy evaluation purposes. As examples of how the methods can be applied to other policy scenarios, a series of parking policy options are evaluated.

---

[1]Antti Talvitie, "Aggregate Travel Demand Analysis with Disaggregate Travel Demand Models" *Proceedings -- Transportation Research Forum Vol. XIII* (October, 1973). Charles River Associates, *Policies for Controlling Automotive Air Pollution in Los Angeles* (Cambridge, Mass: forthcoming) or Frederick C. Dunbar, "Quick Policy Evaluation with Behavioral Demand Models," presented at the Transportation Research Board Meeting, January 1976.

2.3.1 Zonal Data and Model Variables

The two methods for applying disaggregate demand models to zonal interchange data were validated on a set of sketch plan zone observations for work trips from Los Angeles. In the work reported below, only the CRA (1972) work trip model is used; the McFadden (1974) model was tested against the data but because of the problems with transit patronage overpredictions and the insensitivity of choices to transit access times, this model was not used for further analysis.[1]

As was the case in the application of the CRA model to NPTS data, it was found useful to make constructions to deal with additional modes to those on which the model was estimated. These new modes include carpooling, driver serve passenger and walking.

The Los Angeles Database

The basic unit of data for application of the disaggregate demand model is the zonal interchange. The data consist, primarily, of three types of variables: (a) number of work trips by mode as reported in the 1967 Household Survey; (b) average level of service either as reported in the 1967 Household Survey or as derived from engineering estimates for peak hour travel; and (c) relevant socioeconomic data for each of the zones tabulated from the 1970 Census of Population. The trip related data are basically the result of the 1967 Household Survey which was a 1 in 100 sample interview performed by the Los Angeles Regional Transportation Study (LARTS).[2]

---

[1]It should be noted, however, that the McFadden model was extremely accurate in predicting BART impacts.

[2]This body was recently merged with the California Department of Transportation and is now District 8 of that organization.

The sketch plan zones, which number 108, were delineated by LARTS in 1970. In a previous study, CRA selected those zones which were in Los Angeles and Orange counties for further data analysis. The exact details of the data management effort are described in that report and will only be briefly mentioned here.[1]

To construct an appropriate database, the Home Interview Survey tape of trip records was transformed into trip tables on the sketch plan level of zonal interchange. A separate time and distance file which was in the form of a traffic analysis zonal interchange matrix was also transformed to the sketch plan zone level of aggregation to get weighted averages of auto times and distances at peak hour travel. There are approximately twelve traffic analysis zones in each sketch plan zone. Finally, 1970 Census tract data was processed to derive the area and socioeconomic characteristics of each sketch plan zone. From the standpoint of aggregation error, it is worth noting that the average area of sketch plan zones is about 25 square miles and, in the sample used for analysis, the zones varied in size from 16 to 40 square miles.

Even after selecting only those zones which occurred in Los Angeles and Orange counties, there were over 4000 zonal interchanges representing separate data elements for each of the trip related variables. To bring the database down to manageable proportions, a 1 in 25 random sample of zonal interchanges was extracted for model application and policy evaluation. This yielded 172 zonal interchanges of which 89 contain a non-zero number of work trips. The descriptive statistics from these zonal interchanges corresponded to the control totals for the region indicating that the sample was representative.

---

[1] Charles River Associates, *Policies for Controlling Automotive Air Pollution in Los Angeles* (Cambridge, Mass.: forthcoming).

Work Trip Model with New Modes

The model to be used in application to the Los Angeles data is similar in many respects to the model presented for application to the NPTS market segments. There are, however, several significant changes in assumptions regarding the construction of new mode variables. These adjustments reflect the basically ad hoc nature of dealing with new modes in the absence of adequate data. It is also appropriate to use the model somewhat differently in treating new modes when the form of the database to be used is different.

Table 2-20 presents the form of the model with new modes and the construction of new mode variables. The key assumptions embodied in this model are briefly stated as follows:

Autos per worker and constant term: the autos per worker and constant term were suppressed in the logs odd functions for auto-drive-alone versus other auto oriented modes. The reason for this is that it is not clear whether the effects of a preference constant estimated on only transit and auto-drive-alone data has relevance in other auto oriented mode decisions.[1] Suppressing this constant necessitated also suppressing the autos per worker variable because the two estimated parameters are linked.[2]

---

[1]In the work of Haws and Ben-Akiva, *op. cit.*, the mode specific constant was found to vary depending upon mode choices among auto-drive-alone, transit and carpool.

[2]A more complete discussion of the relationships between mode specific variables and constants, and their effects on logit estimated relationships is contained in McFadden, D. "On Independence, Structure, and Simultaneity in Transportation Demand Analysis," Travel Demand Forecasting Project, Institute of Transportation and Traffic Engineering, University of California, Berkeley, Working Paper No. 7511 (1975). See also Tye, William and Sherman, Leonard, *Disaggregate Travel Demand Models*, National Cooperative Highway Research Program Project 8-13: Phase I Report (Charles River Associates, Cambridge, Mass.: 1975). It should be pointed out that neither the mode specific constant nor autos per worker were suppressed when the model was applied to shared rides with NPTS data. However, alternative assumptions about time and distance were also used and these may have achieved the same results as omitting the constant term and the autos per worker available.

TABLE 2-20

Summary of Work Trip Model with New Modes for Application to Los Angeles Data

1. Auto-transit log odds:

$$ln(\frac{P(a)}{P(b)}) = -4.77 - 2.24\ (C_a - C_b) - 0.0411\ (T_a - T_b)$$

$$-0.114(S_a - S_b) + 3.79Y$$

2. Auto-carpool (with k passengers) log odds:

$$ln(\frac{P(a)}{P(ck)}) = -\ 2.24\ (C_a - C_{ck}) - 0.0411\ (T_a - T_{ck})$$

3. Auto-driver serve passenger log odds:

$$ln(\frac{P(a)}{P(d)}) = -\ 2.24\ (C_a - C_d) - 0.0411\ (C_a - C_d)$$

4. Constructed cost and times variables:

$$C_{ck} = \frac{(C_a + k(.03 \times 2 \times (5/12) \times (\sqrt{a}_d + \sqrt{a}_j)))}{K+1}$$

$$T_{ck} = T_a + \left(\frac{k \times 2 \times (5/12) \times (\sqrt{a}_i + \sqrt{a}_j)}{(15/60)}\right) + 20\ (k+1)$$

$$C_d = \left(\frac{2.24 + 2 \times 4.11}{2.24}\right) \times C_a$$

$$T_d = \left(\frac{0.0411 + 2 \times .0654}{.0411}\right) \times T_a$$

where: a_i, a_j = area, in square miles, of origin and destination zones; other variables are defined as before.

Carpool cost sharing: As was the case in the previous application of the model, to NPTS data, the costs of carpool travel are assumed to be shared equally among carpool members.

Carpool distance: To form a carpool where none existed previously requires extra distance traveled to pick up and drop off passengers. For data grouped into zonal formats, it is natural to assume that the distances traveled to form a carpool are related to the size of the zones because zones of larger sizes have lower densities of households. Given certain assumptions about the distribution of origins and destinations within zones, the mean distance traveled to pick up or drop off each passenger can be computed as $(5/12) \times \sqrt{a}$, where a is the area (in square miles) of the reference zone.[1] This extra distance is traveled at both the *origin and destination*, and is doubled to represent round trips. It is multiplied by the relevant operating cost per mile ($0.03 in the case of 1967) to determine the extra cost of a vehicle used for carpooling. This cost is added to the cost of a linehaul journey between the two zones for the vehicle, which is equal to the cost of an auto-drive-alone trip (C_a), to obtain the total vehicle costs to the carpool members.

Carpool time: The extra distance traveled to pick up and drop off passengers is assumed to take time for each member of the carpool; the assumed speed for this portion of the journey is 15 miles per hour. Additionally, there is a time penalty associated with mismatched schedules and other waiting times. This time penalty was placed at 20 minutes per round trip (five minutes at the origin and destination of each link).

Driver serve passenger cost and time: As in the previous definition of driver serve passenger, it is presumed that the driver chauffeurs the passenger to work and then returns home for the first link of the trip; for the second link, the driver goes to the work place from home, picks up the

---

[1]A discussion of this distribution is presented in Frederick C. Dunbar "Quick Policy Evaluation with Behavioral Demand Models," presented at the 55th Annual Meeting of the Transportation Research Board, January 1976. See also, CRA, *Policies for Controlling Automotive Air Pollution...*, *op.cit.*

passenger and then returns home. Because the driver may have different values to place on the times and costs associated with this trip, their time and cost was weighted differently from those of the passenger. The weights were taken from the parameter estimates on cost and time of the CRA shopping trip mode split model.[1] It was thought that these parameters would be most reflective of the determinants of behavior of the chauffeur. They are equal to 4.11 for cost and .0654 for linehaul time.

For the variables not discussed above, the data provided adequate measures. Reported times and costs for auto and transit modes were checked against engineering estimates and were found to be consistent between the two sources. Transit access times for trips were derived from transit route maps; the reported access times were biased downward by the fact that these only reflect the access times of tripmakers who chose transit because, among other things, of its superior access. For zones where no transit trips were reported, linehaul and wait times were derived from system maps and route schedules. The cost of auto use was computed as the average distance in a zonal interchange times operating cost per mile of a car, $0.03, which has been defined above. Autos per workers for each origin zone was calculated from 1970 Census data as was the area for all zones.

2.3.2 Methods for Model Application

The aggregation problem is especially severe for the zone size used in the Los Angeles database. The zones are large in the sense of exhibiting substantial variation in level of service and in socioeconomic characteristics, yet they are relatively constrained in size so that methods are inappropriate

---

[1]For a presentation of this model see Domencich, Thomas A. and McFadden, Daniel, *Urban Travel Demand: A Behavioral Analysis*, (Amsterdam: North-Holland, 1975). Alternatively, see Charles River Associates, *A Disaggregated Behavioral Model of Urban Travel Demand*, (Cambridge, Mass.: 1972).

which depend upon assumptions that these variables are distributed normally.[1] To cope with the aggregation problem in applying models to urban data bases, a series of studies have developed heuristics and approximation techniques.[2] The approaches described below are in this tradition.

Direct Application

To put the aggregation problem into some perspective, and to judge the merits of various procedures for treating the aggregation problem, the work trip model described above was applied to various groupings of the Los Angeles data without any adjustments. The means of the level of service variables from the data groupings were used in the model as it appears in Table 2-20. The data groupings are described below:

Zonal Interchange: In this format, the basic unit of observation is the sketch plan zonal interchange as has been described.

Distance Segments: To provide a different, and easily manageable form of data for potential model application, the trips in the sample of zonal interchanges were grouped into three distance categories: (a) short trips where the length of the trip is less than the mean trip distance; (b) intermediate trips which fall in the range between the mean trip distance and the third quartile; and (c) long trips which are longer than the third quartile. This grouping suggests itself because the VMT elasticities with respect

---

[1]Section 1 briefly mentioned approaches which assume that the arguments in the logs odds function of a logit model are distributed normally.

[2]See, for examples, Frank S. Koppelman, *Travel Prediction with Models of Individual Choice Behavior*, unpublished Ph.D. Dissertation, Department of Civil Engineering, Massachusetts Institute of Technology (1975), and Uzi Landau, *Sketch Planning Models in Transportation Systems Analysis*, unpublished Ph.D. Dissertation, Department of Civil Engineering, Massachusetts Institute of Technology (1976).

to level of service variables tend to be different within each group. Also, the contribution of VMT's in each group is comparable (4071, 6023 and 5188 miles for short, intermediate and long trips, respectively).

Regional Means: As a final exercise, the means of the relevant variables across the sample were taken and treated as a single observation.

Owing to the way the Los Angeles data were recorded, the mode definitions are somewhat different from the mode choices offered by the model and different from the modes recorded for the NPTS data. In particular, auto drivers include the drivers in carpools (but not the drivers in the driver serve passenger mode which were ignored by the data altogether). Auto passengers include the passengers in carpools and the passengers in driver serve passenger trips. In the results reported below, the members of carpools and driver serve passenger trips were distributed according to these definitions into the relevant mode category.

The performance of the unadjusted model with the various data groupings is presented in Table 2-21. In all cases it can be seen that transit trips are underpredicted, as would be expected, and that the bias in overpredicting the dominant mode increases as the level of aggregation increases.

Predicted vehicle miles traveled are calculated from the following formula:

$$VMT = (\textit{auto driver mode split}) \times (\textit{distance}) \times (\textit{total trips}) + (\textit{auto passenger mode split}) \times (2 \times (5/12)) \times (\sqrt{a_i} \times \sqrt{a_j}) \times (\textit{total trips})$$

In the case of applying the model to zonal interchanges, the VMT's for each zonal interchange are calculated and then summed over all zonal interchanges to yield the aggregate VMT's. When distance segments are the level of observation,

TABLE 2-21

PERFORMANCE OF DIRECT APPLICATION OF WORK TRIP MODEL ON LOS ANGELES DATA

	Actual	Predicted		
		Zonal Interchange	Distance Segments	Regional Means
Mode Split:				
Auto Driver	.863	.886	.912	.927
Transit	.038	.027	.004	.001
Auto Passenger	.099	.088	.084	.072
VMT	15618	14896	15283	15979

the average distance and predicted mode split for each segment is used to compute the VMT's in each segment and the results are again summed to get the aggregate VMT's. For the regional means, there is only one observation and the above formula is applied directly to the regional means or totals of the relevant variables.

The first term in the above formula accounts for the linehaul portion of trips by auto; the second term is the assumed distance necessary to pick up and deliver auto passengers.

From Table 2-21, it can be seen that as the various applications of the model tend toward predicting the correct mode split, total VMT's become underpredicted. In the case of using zonal interchanges, the prediction error is 4.6 percent. The reason for this is that the model has some tendency to overpredict transit trips for long distance journeys relative to the actual observations. Because the long distance trips contribute proportionately more to VMT's, even if aggregate mode split estimates are accurate, the VMT estimates will be low.

From the results obtained by applying the model directly to data tabulations, there appears to be a tradeoff between ease of data handling and the degree of error in travel demand estimates owing to the aggregation problem. The error is associated with the fact that the variety of potential responses to differing level of service characteristics are suppressed as the data become merged into higher levels of aggregation. This indicates that there is some merit to investigating simplified methods for adjusting the model to account for aggregation error.

Simplified Taylor's Series

As described in Section 1, one approach to correcting for the aggregation problem is to use a truncated Taylor's series approximation. If the analyst has data from which the variances and covariances of the arguments in the logit model can be calculated, then equation (1-18) can be applied. The approach was originally developed by Talvitie[1] and applied by CRA in a study where the variance-covariance terms were estimated as functions of sketch plan zone areas.[2] However, as an examination of the CRA study will show, the calculation of the numerous variance-covariance terms can be a burdensome task.

One of the results of the CRA work was that the variance of the log odds function (*var(Y)* in equation (1-18)) must be constrained to be between zero and unity in order to meet the condition that the estimated mode split for a mode would increase as the level of service for that mode improves or

---

[1]Antti Talvitie, "Aggregate Travel Demand Analysis..."(1973)

[2]CRA, *Policies for Controlling Automotive Air Pollution...*, and Frederick C. Dunbar, "Quick Policy Evaluation..." (1976).

the level of service for other modes declines. Under this constraint, it was found that the variance of log odds functions for many zonal interchanges and mode choices in the Los Angeles database was equal to unity. In a significant number of other cases, the variance of the log odds function was relatively close to unity. This result suggests that only a limited amount of accuracy in model application would be lost if all the *var(Y)* terms in equation (1-18) were set equal to one.[1] The purpose of this assumption is to relieve the model application task of the necessity of computing a large number of variance-covariance figures and thereby simplify considerably the Taylor's series approximation.

The formula for application which this assumption implies is:

$$F(i) = P(i)\left(1 + \sum_{k=1}^{m} (P(k) - \delta)(P(k) - \tfrac{1}{2})\right) \qquad (2\text{-}5)$$

where: $F(i)$ = predicted mode share for mode i;

$P(i)$ = probability of mode i as computed by direct application of the model described in the previous section;

m = number of modes

δ = 1 if $k = i$

= 0 if $k \neq i$.

To predict the mode shares using the simplified Taylor's series, the logit model is applied directly to the means of the data (as described in the above section on the direct application of the logit model) and the resulting probabilities for each mode are used in equation (2-5) for each observation (zonal interchanges, distance segments or regional means) to compute mode splits.

---

[1] CRA, *Policies for Controlling Automotive Air Pollution...*, *op. cit.*

TABLE 2-22

PERFORMANCE OF TAYLOR'S SERIES APPROXIMATION TO WORK TRIP MODEL ON LOS ANGELES DATA

	Actual	Predicted		
		Zonal Interchange	Distance Segments	Regional Means
Mode Split				
Auto Driver	.863	.851	.884	.902
Transit	.038	.035	.006	.001
Auto Passenger	.099	.114	.110	.097
VMT	15618	14599	15107	15803

The results of applying the Taylor's series adjustment are presented in Table 2-22. In comparison with the results of direct application of the logit model (Table 2-21), it can be seen that the aggregation error has been mitigated somewhat. However, the performance of the Taylor's series approach to data grouped by distance segments and regional means is unsatisfactory. Moreover, the underprediction of VMT's is still a problem in that the error from using zonal interchange data is increased to 6.5 percent. This can be attributed to the fact that, by simplifying the Taylor's series it is presumed that the aggregation bias is distributed more uniformly across long and short trips than is really the case. As a result, transit predictions for long trips are overpredicted and this leads to low estimates of VMT's.

Area Adjustment

A separate approach to treating the aggregation problem was discussed in Section 1 which led to equation (1-19); this technique is based on the premise that the area of zones can be used to adjust log odds functions because of the relationship between the area of zones and the extent of aggregation error. A specific formula for making this adjustment is as follows:

$$ln\left(\frac{F(a)}{F(b)}\right) = \left(ln\left(\frac{P(a)}{P(b)}\right)\right)\left(\gamma + (1-\gamma)\left(\sqrt{a}_i + \sqrt{a}_j\right)^{\xi}\right) \qquad (2\text{-}6)$$

where for modal alternatives a and b, the parameters γ and ξ are estimated constants and the other terms in equation (2-6) have been defined before.

Depending upon the values of the estimated parameters, this functional form for the adjustment of the model has the desirable attributes listed in Section 1.[1] It should be emphasized, however, that this is an ad hoc method for adjusting the logit model for the aggregation problem. It relies on intuitive and observed relationships between the area of a zonal interchange and the level of variation in the log odds function. It is not rigorously derived from a well conceived theory about the relationships between area, within group variation and the logit model specification. In this sense, the area adjustment approach proposed here is more in the tradition of curve fitting rather than uncovering structural relationships. Other functional forms, consistent with the conditions imposed on adjustments to the log odds functions, were also tried but they performed less well than that represented in equation (2-6).

Because equation (2-6) is intrinsically nonlinear in parameters γ and ξ, normal linear regression methods are unsuited for parameter estimation. The approach used to estimate γ and ξ was to compare actual versus predicted

---

[1]The parameter values should satisfy the following conditions: $0<\gamma<1$
$\xi<0$

auto trips in a sample of Los Angeles zonal interchanges with various assumed values of γ and ξ. The initial sample included only auto-drive-alone and bus trips. The values of ξ which were used included either -1 or -.5; for each of these, γ was varied from .15 to .95 using increments of .05.

The best combination of values was γ equal to .6 and ξ equal to -.5. Using the predicted mode splits, a residual analysis of actual versus predicted auto trips was performed. The resulting R^2 was .997, uncorrected for degrees of freedom. Though this is a very high R^2, the residual analysis is not in itself a very robust test because simply using a weighted average of the probability across zones and using this as the mode split for every zone yielded an R^2 of .987. The inference that comes from this exercise is that the area adjusted model explains about 77 percent of the variance not attributable to simply taking the average probability.

To predict mode shares with the area adjustment approach, the logit model is directly applied to the means of the data, as described above, and these probability estimates are used in equation (2-6). When the formula is applied to zonal interchange data, the area of the origin and destination zones are used for a_i and a_j. When the formula is applied to distance segments and regional means, the values of a_i and a_j are assumed to be infinity so that the term $(\sqrt{a_i} + \sqrt{a_j})^{-.5}$ is set equal to zero.

The results of applying the approach to the full sample of zonal interchanges and the full complement of modal choices are presented in Table 2-23. The performance of the model applied to distance segments and the regional means is also given in Table 2-23. As in the case of using the Taylor's series approximation, there is a tendency for the predicted share of the dominant mode to increase as the level of aggregation increases. Unlike the Taylor's series expansion, the model applied to the distance segments gives tolerably accurate results. The performance of the model applied to regional means is unsatisfactory.

TABLE 2-23

PERFORMANCE OF AREA ADJUSTED WORK TRIP MODEL TO LOS ANGELES DATA

	Actual	Predicted		
		Zonal Interchange	Distance Segments	Regional Means
Mode Split				
Auto Driver	.863	.865	.888	.915
Transit	.038	.049	.030	.014
Auto Passenger	.099	.086	.082	.072
VMT	15,618	14,748	15,109	15,775

When applied to zonal interchanges, the approach under-predicts total VMT's by 5.6 percent; when applied to distance segments, the predicted VMT's are more accurate, with a 3.3 percent underestimate, but the mode split predictions are, on the whole, more in error. The performance of the model indicates that the approach may hold some promise, but further validation is necessary. The next section provides further tests of this and the other methods for applying the logit model.

Validation of Methods on a Gas Tax Policy Scenario

To provide an example of applying the models presented above, and to check the approaches against other results, each method was used to simulate the effects of a 70 percent increase in the pump price of gasoline in 1975. An increase in gasoline price of this magnitude represents an additional tax of about 100 percent on the pretax cost of gasoline.

The first step in the analysis is to make base case 1975 mode split and VMT predictions based on the transportation level of service obtained in 1975. It was determined that there have been only minor changes in the Los Angeles transportation environment since 1967 except for the secular increase in gasoline prices and a change in the transit fare policy. To take account of these effects in the base case forecasts, auto operating costs were changed to $0.06 per mile to reflect 1975 circumstances and transit fares between zones were set at $.25 per one way trip to represent the new flat rate fare policy. The three approaches -- direct application, Taylor's series approximation, and area adjustment approximation -- were then applied to the three data groupings -- zonal interchange, distance segments, and regional means -- to estimate a 1975 base case set of mode splits and VMT's.

The next step entailed simulating mode splits and VMT's with the assumed increase in auto operating costs caused by the change in gasoline prices. The auto operating cost under the scenario of an increase in gas price of 70 percent is $0.09 per mile. (Gasoline costs are 71 percent of auto operating costs per mile in the base case.)

The results of the two sets of simulations applied to the data indicated that the level of disaggregation of the data was the most important determinant of whether the model performed reasonably. That is to say, the method of applying the models had relatively little impact on the estimated elasticities. The estimated elasticities of VMT's with respect to the pump price of gasoline can be summarized as follows:

Zonal Interchange Data: all elasticities computed from the three approaches fall within $-.157 \pm .005$;

Distance Segments: all elasticities computed from the three approaches fall within $.105 \pm .01$;

Regional Means: all elasticities computed from the three approaches fall within $.029 \pm .01$.

Predicted mode splits also systematically varied with the level of aggregation of the data, though there was more variation in these estimates. As would be expected, the dominant mode, auto driver, had higher mode split estimates than the more aggregated the data.

Table 2-24 presents the results of applying the different methods to zonal interchange data. As can be seen, there is little to choose among the three methods without further validation. One point that does merit consideration, however, is that the direct application of the logit model tends to underrepresent transit ridership given current estimates by the Southern California Rapid Transit. In the examples of policy scenarios which follow, the two approximation methods are used instead of direct application of the logit model.

The major conclusion from this analysis seems to be that methods of disaggregating the data are much more important in dealing with the aggregation problem than the methods tried for adjusting the logit model. The elasticities derived from both the relatively aggregate data groupings, distance segments and regional means, are unacceptable given other evidence on the elasticity of gasoline demand (see Section 1). Even the application of the model to zonal interchange data yields low elasticity estimates compared to other results which are typically in the -.2 range.

Predicted Effects of Parking Restraints

The Taylor's series approximation and area adjustment approach were used on zonal interchange data to predict mode splits and VMT's under a series of parking policy scenarios. The approach and associated scenarios are as follows:

- Taylor's Series: a $1.00 parking tax in the Central Business District (CBD) only; and a $1.00 parking tax throughout the region.

TABLE 2-24
PREDICTED 1975 EFFECTS OF 70 PERCENT
GASOLINE PRICE INCREASE USING WORK TRIP MODELS
APPLIED TO LOS ANGELES ZONAL INTERCHANGE DATA

	Base Case		70 Percent Increase In Gas Price
Direct Application			
Mode Split			
Auto Driver	.842		.795
Transit	.053		.085
Auto Passenger	.105		.120
VMT	13,340		11,819
VMT Elasticity W.R.T. Gas Price		-.162	
Taylor's Series			
Mode Split			
Auto Driver	.807		.764
Transit	.065		.096
Auto Passenger	.127		.139
VMT	13,057		11,638
VMT Elasticity W.R.T. Gas Price		-.154	
Area Adjustment			
Mode Split			
Auto Driver	.816		.771
Transit	.080		.110
Auto Passenger	.103		.120
VMT	13,123		11,714
VMT Elasticity W.R.T. Gas Price		-.153	

- **Area Adjustment:** Decreased parking availability in the CBD; and decreased parking availability and associated increase in parking cost by $0.50 per trip.

To evaluate each scenario, the 1975 Los Angeles base case is first presented. Each scenario is then represented by an appropriate change in the relevant level of service variables. Elasticities, per se, were not estimated because they were not definable in these particular scenarios.[1]

Taylor's Series Applied to Parking Taxes

The two tax scenarios are represented by increasing the auto cost variable by $1.00 per auto trip when the tax applies. It is to be noted that one of the advantages of using zonal interchange data is that origins and destinations of trips can be precisely defined. Thus, only CBD oriented trips from the sample would have their auto costs changed in the CBD parking tax scenario. When the model simulates the effect of a tax policy, it is assumed that carpool members share the tax equally. The results of this simulation are presented in Table 2-25.

It can be seen from Table 2-25 that the predicted impact of a regionwide tax is much greater than that of a CBD only tax. This is to be expected, especially in a city such as Los Angeles which exhibits spread development and a diffuse pattern of employment centers.

---

[1]In the base case, both parking costs and auto access time are zero. To compare an elasticity for these scenarios entails dividing by zero.

TABLE 2-25

PREDICTION OF WORK TRIP MODE SPLITS AND VMT's USING TAYLOR'S SERIES APPROACH TO LOS ANGELES ZONAL INTERCHANGES WITH PARKING TAX SCENARIOS

	Base Case	$1.00 Parking Tax: CBD Only	$1.00 Parking Tax: Regionwide
Mode Split			
Auto Driver	.807	.792	.582
Transit	.065	.079	.194
Auto Passenger	.127	.129	.224
VMT	13,057	12,566	11,213
Percent Change in VMT		-3.8	-14.1

Area Adjustment Applied to Parking Availability

It is reasonable to expect that a policy of decreased parking availability in the CBD will increase the walk access time of trips with a destination in the CBD. The extra walk time from parking spot to place of work was assumed, for the purpose of this policy simulation, to be 7.5 minutes, or 15 minutes for a round trip. For non-CBD trips the auto access variable remains zero. It is further assumed that carpool trips entail dropping off passengers at the place of work, involving no access time, whereas the driver must bear the burden of the 15 minute access walk. The average access time for carpool members is represented by 15 minutes divided by the number of carpool members; the access time penalty is shared equally by increased payments or carpool driving rotation.

It may also be the case that a policy which rations parking availability will cause parking costs to increase. In the 1975 base case, these costs are negligible and are consequently assumed to be zero. It may also be the case that parking restrictions include both decreased availability and a tax. In either event, two scenarios were represented; the first involves increased access time in the CBD for auto trips without an auto cost increase; the second involves both an increase in access time and an assumed $.50 per trip auto cost increase for work journeys to the CBD. The results of these policy simulations are given in Table 2-26.

TABLE 2-26

PREDICTION OF WORK TRIP MODE SPLITS AND VMT's USING AREA ADJUSTMENT APPROACH TO LOS ANGELES ZONAL INTERCHANGES WITH PARKING AVAILABILITY SCENARIOS

		15 Minute Auto Access Time in CBD	
	Base Case	No Parking Cost Increase	$.50 Parking Cost Increase
Mode Split			
Auto Driver	.816	.806	.800
Transit	.080	.087	.092
Auto Passenger	.103	.107	.108
VMT	13,123	12,812	12,630
Percent Change in VMT		-2.4	-3.8

It is to be noted that evaluation of the parking restrictions without accounting for potential parking charge increases would underrepresent the impacts of the policy. In general, the predicted effects of this type of parking restriction are quite similar to the predicted effects of a parking tax.

2.3.3 Summary

In summarizing the application of disaggregate models to zonal interchange data it is helpful to compare the above results to those obtained from applying the models to NPTS data. Using market segments is computationally less cumbersome than using zonal data because there appears to be no need to adjust the model for aggregation error if market segments are created to minimize variation; also, the eventual number of observations to be used with market segments is much smaller than the number used with zonal data and this can mean the difference between having to use a calculator or a computer. In terms of predictive accuracy, there is little evidence offered that one data source is preferable to the other. Zonal data has two advantages: first, for any particular urban area it may already be available in a form suitable for model application; second, if policies are to be differentiated by corridors or other parts of the region (such as CBD vs. non CBD) then zonal data is more appropriate because it is geographically specific. For both types of data formats, the problems of aggregation, transferrability and new modes can be solved with a reasonable degree of success to estimate quickly the aggregate effects of alternative transportation policies.

3. NONWORK TRAVEL

3.1 INTRODUCTION

For trips other than the journey to work or shopping, there is little in the way of available models and evidence which can be used for transportation policy evaluation. Even shopping trip models, as described in the review of existing models in Section 1, are cumbersome to apply and, at present, not robust in terms of the implied elasticities of travel demand with respect to level of service variables. Current attempts at modeling shopping trip behavior are focusing on the application of the generalized logit specification and there is some question as to whether this is an appropriate structure for representing nonwork travel decisions.

Given these considerations, there appears to be some basis for using alternative modeling strategies to estimate nonwork travel demand. If the purpose of the development of such models is to provide relatively simple and workable tools for policy analysts, linear specifications should be tested. The linear structure minimizes aggregation bias when applied to data grouped at levels different than that used for the estimation sample. It is also, typically, easier to work with than nonlinear forms. The major drawback to linear models is that they may introduce specification error into the representation of travel behavior. In this sense, the development and application of linear models is dictated by determining through empirical exercises the relevant tradeoffs between specification error and the additional complexities which arise from using more accurate functional forms. In either case, for nonwork tripmaking there is little *a priori* to choose from in comparing linear versus logit specifications. The former assumes that decisions are made among a continuum of alternatives whereas the latter assumes that decision makers must decide among a relatively small number of discrete options. Though neither assumption is probably entirely correct, the development of a model explicitly for the purposes of quick planning applications indicates that a relatively simple linear specification may be suitable.

This section presents the estimates of several linear relationships which can be combined in various ways to form models of nonwork trip behavior. The purpose of the model is to find the effects of certain key variables on nonwork travel. As such, the models represent substantial simplification of the process of nonwork travel behavior. However, to the extent that policy instruments can be translated into effects on the exogenous variables in the models, they can be used for policy and planning analysis. Each model is a simultaneous system of equations representing the demand for auto and transit services. The relationships were estimated on disaggregate household data from the Nationwide Personal Transportation Survey. In a final part of the section, the models are used to analyze policy scenarios involving auto trip costs, transit service improvements and parking restrictions.

3.2 DISAGGREGATE LINEAR MODEL OF NONWORK TRIP BEHAVIOR

To develop a model of nonwork travel behavior, tripmaking was conceptualized from the standpoint of consumer theory in order to isolate the key variables and to correctly specify the relevant equations. It was also necessary to work within rather severe data constraints which ultimately had strong impacts on the model specification and statistical techniques used. The estimated equations reflect, in large degree, the unique qualities of the NPTS household data and the attempt to concentrate on those variables which are of current interest in national transportation policy and planning. A rather severe validation of the model was attempted on independent Los Angeles data with mixed results. Each of the above issues is discussed in more detail in the following sections.

3.2.1 Modeling Issues

The underlying theory of behavior implied by a model determines the model specification and the estimation techniques which are to be employed. Thus, theorizing about nonwork travel behavior suggests that care should be used to construct the relevant independent variables in order to avoid simultaneous equations bias as well as to develop a model consistent with our assumptions about the underlying structure of travel decisions. Moreover, the extent of simultaneity in the phenomena represented by the model requires consideration be given to estimation methods which remove simultaneous equations bias. A final issue, partially dictated by the nature of the data, is the existence of unobserved variables in the equations and the estimation procedures necessary to deal with this problem.

Conceptualization of Nonwork Tripmaking Behavior

Transportation services are an intermediate good to households. They are typically not consumed for their own pleasure but rather they are the means to an end. One result of viewing transportation in this manner is to conclude that a household is ultimately interested in a set of origins and destinations rather than trips per se.

The origins and destinations, for most trips, are not fixed but rather tend to be substitutable. In most urban areas there are many alternative destinations for shopping, recreation etc., and even home based trips can be minimized by linking several trip purposes into a single journey or by the simple expedient of making fewer trips.

With regard to auto trips, and overview or simplification of the system condenses the representation of the origin/destination/frequency choice decision making process to relatively few variables explaining miles traveled. There are as a consequence two ways of viewing the household's travel behavior:

VMT decisions -- the most simplified approach is to consider that the household consumes a number of VMT's over an appropriate time period to satisfy its demand for transportation services. How it distributes these VMT's over the urban area is of little concern in the model. Neither do we consider the frequency with which auto trips are made. Obviously, these factors affect the number of VMT's, but the orientation of the model is to abstract from the more detailed considerations of destination and frequency choice. The variables which affect destination and frequency choice are used to model the demand for VMT's directly.

Joint frequency and average trip distance decisions -- at a somewhat more disaggregated level of decision making, it can be presumed that the household consumes a certain number of trips with a codetermined average length. In this model form, the frequency of travel is predicted but the exact origins and destinations are not specified. It is expected that, all other things equal, the larger the number of trips, the shorter the average distance traveled (because the two are substitutes). Again, the variables which affect destination choice are included in the model but they are used to influence trip frequency and distance directly.

With respect to transit trips, there are fewer options available to tripmakers because of the fixed route constraint on transit trips. Distance traveled is of secondary importance in the consumption of transit trips; that is, it affects the transit choice decision but, unlike the case of autos where there is a nearly ubiquitous road network, the trip itself tends to be tantamount to the origin/destination decision. Viewed from this perspective, the number of transit trips made is the relevant item consumed by households.

To specify each of the relationships discussed above requires three types of variables which determine the travel decision. These include:

Cost and Time -- Each mile traveled or trip made incurs a penalty in terms of cost and time. In the cases of VMT's and average auto distance, these penalties include the marginal cost per mile of operating the vehicle and the time per mile of the journey. For transit and auto trips, the relevant variables are the cost and time of the entire trip. To the extent possible, it is useful to separate cost and time into components such as out of pocket charges vs. gasoline costs or walk access time vs. wait and linehaul time. It should also be noted that in the data available it is often the case that the reported cost and time related variables are simultaneously determined with the dependent variables of trips and distance; the problems introduced by this simultaneity will be discussed in sections on model specification and estimation technique.

Urban Form -- The nature and size of the urban area and of the place of residence will present various opportunities and constraints to travel. Large agglomerations of activity may cause less distance to be traveled than would otherwise be expected; households in low density locations will tend to travel farther, all other things equal; large urban areas offer more opportunities for trip destinations than small areas. The effects of the various urban variables are often difficult to judge *a priori* but it is likely they have a significant impact on travel behavior. To the extent data is available which describes the salient features of urban areas, appropriate variables should be included in the models of travel demand because they contain the information about alternative destinations that would be used in more disaggregated models of decision making. If the model is estimated on a sample of households drawn from a large number of urban areas, excluding variables describing urban form can bias the estimates of travel demand parameters. This issue is considered further in the discussion of estimation technique.

Socioeconomic Characteristics -- There are also a large number of household specific factors which will affect travel decisions. Among the most important of these is income and household size. Income effects also interact with time and cost effects because it can be presumed that the value of time and cost elasticities are sensitive to income budget constraints and income related opportunity costs. Other important socioeconomic factors include the number of employed persons and the age distribution of household members. It can be expected that households with greater numbers of employed members will have higher opportunity costs associated with nonwork travel. The age distribution of a household affects the nature of transportation demands.

Model Specification

Model specification is determined primarily by the underlying theory of behavior which is assumed for the travel decision making process. The simplifications described above do not, in and of themselves, necessarily imply that the estimated relationships will be less cumbersome for planning purposes than models based on a more complex or detailed set of assumptions about travel behavior. In order to realize the advantage of abstracting from specific origin/destination choices, it is necessary to introduce further simplifications in terms of the functional forms used to describe the particular relationships. In addition, there are a number of other specification problems which must be considered in order to estimate valid relationships. The most important of these involves the simultaneous determination of travel choices and reported level of service. For the purposes of discussion, the specification issues can be grouped into three categories presented below.

Linear Forms -- As mentioned previously, the purpose of developing the nonwork travel model is to have a versatile yet easy to use tool for policy analysis. The linear specification adds flexibility to model application because it can be applied to data of widely varying groupings with minimal loss of accuracy owing to aggregation error. It is also easy to interpret through examining the parameters. In some circumstances this can lead to a nearly instant evaluation of policy options. However, these advantages of the linear specification are not without their costs in terms of the specification error involved in representing nonwork travel behavior and options.

The important elements of specification error introduced by the linear form are listed below:

--The implicit assumptions about the utility of travel as a function of trips and distance by mode are dubious. In particular, the utility of travel is presumed to be linear with respect to trips and distance if the demand for trips and distance is specified as a linear function of level of service variabl

--Some dependent variables are assumed to be continuous rather than divisible into discrete alternatives. With respect to number of auto trips or length of auto trips, this is probably a reasonable approximation, especially if the time period being represented is a week or longer. However, because relatively few transit trips are taken, the use of a continuous variable to approximate a discrete number of transit trips leads to some bias in the estimated effects of level of service variables on transit demand.

--The implicit assumptions about the distribution of travel opportunities and system constraints are also very approximate. The linear specification simplifies some of the complexities of urban form and its interactions with travel demand. The result is that policies which affect land use to a significant extent will be evaluated with less accuracy than policies which affect transportation level of service.

--Finally, many of the interactions among independent variables which affect travel decisions can only be modeled in the linear specification in a way which is vulnerable to aggregation error. For example, the relationship between income and travel time is multiplicative rather than additive. Unless income and time spent traveling are stochastically independent, which is unlikely, the estimated parameters of value of time from disaggregated data cannot be applied easily and accurately to aggregate data.

Understanding these limitations on the linear models enables analysts to make judgements about the confidence with which policy evaluations are made. At the very least, the model can be used to test hypotheses about travel behavior and the effects of various policy instruments.[1] However, we can expect that linear models can do more than this in that, within the constraints discussed above, policy contingent forecasts with the linear specification are possible.

[1]An example of this type of analysis is given in Joel Horowitz, "Effects of Travel Time and Cost on the Frequency and Structure of Automobile Travel" unpublished (Environmental Protection Agency, Washington, D.C.: May, 1975).

Disaggregate Level of Observations -- The benefits to estimating models with disaggregate data were described briefly in Section 1. The main advantage to using disaggregate data is that there is minimal information loss of the sort which occurs when data is grouped. Some of the extra information gained is more detail on the range of options available to tripmakers; this helps mitigate the effects of cross section bias which were a problem with the direct demand model and led, potentially, to overestimates of level of service elasticities.

Simultaneous Determination of Variables -- There are two types of simultaneity problems which arise in the models developed for this study. We discuss these in turn:

--In order to determine the cross elasticities among modes, it was necessary to include the trips made by transit among the arguments of the auto travel demand equations. However, it is likely that auto and transit travel are simultaneously determined.

--Some of the level of service variables, particularly time spent for the trip, were only reported when the mode was chosen. This means that alternate mode data, to the extent it exists, does not include the true level of service confronting the household for all trips but only for the trips actually made. For example, trips made by transit were selected to be the ones with most favorable level of service when compared to auto options. Thus transit time and transit availability for trips actually made are jointly determined with mode choice.

To deal with these problems, statistical methods were employed to remove simultaneous equations bias from the estimated relationships. The result of not using these techniques is that the elasticities and cross elasticities would have been substantially underestimated.

Estimation Techniques

In the event that a relationship includes jointly dependent variables as arguments in the function, ordinary least squares is inappropriate as an estimation technique. Estimating coefficients for endogenous variables violates the assumption that the variables are exogenous. There are several methods available for dealing with this problem.[1] The method employed in this study is two stage least squares. The essence of this approach is to use exogenous variables to create expected values of the endogenous variables which are used as arguments of estimated relationships. This removes the stochastic dependence with occurs when two or more endogenous variables are used in the same equation.

Another problem which occurs with the data base employed for estimating the demand models is the existence of unobserved variables which are city specific. It can be expected that the observations of household trip records from different cities will be affected by system characteristics and urban area characteristics for which the available data in the trip record provide no information. At a minimum, this implies that the error terms associated with household observations have distributions which vary from city to city. This problem was similar to earlier econometric studies which had to use similar types of data and for which an estimation technique called error components was developed.[2] Estimation with error components is used in this study to test the sensitivity of the parameter estimates to this potential source of error.

---

[1]Most standard econometric textbooks have relatively complete discussions of these methods and they are not repeated here. See, for examples, Henri Theil, *Principles of Econometrics* (John Wiley & Sons, New York City: 1971), Franklyn M. Fisher, *The Identification Problem in Econometrics* (McGraw-Hill, New York: 1966), and Carl F. Christ, *Econometric Models and Methods* (John Wiley & Sons, New York: 1966).

[2]See Zvi Griliches, "Errors in Variables and Other Unobservables," *Econometrica*, 42 (1974) and Marc Nerlove, "A Note of Error Component Models," *Econometrica*, 40 (1972).

The estimation results for the travel demand models are presented in the section after the following discussion of the data and variables used in model development.

3.2.2 Data Used for Model Estimation

The basic data source for estimation of the nonwork travel behavior models is the Nationwide Personal Transportation Survey (NPTS). Individual travel records from the trip day reports over a four day period were processed to give observations on travel by households.[1] Table 3-1 presents the variables taken from the trip records that appear in the estimated models. Other socioeconomic and geographic data were also processed and were used as instrumental variables in the two stage least squares estimates of the models.

Though Table 3-1 and the associated footnotes give a relatively complete description of the data elements, certain features are worthy of further elaboration. These are discussed below.

--Level of service data is only reported for trips actually made. The consequence of this shortcoming on estimation techniques was discussed above. It has the additional consequence that only separate (though overlapping) samples are available to estimate auto and transit tripmaking behavior. Out of a total sample of 765 urban households for which complete records were available, 638 took one or more auto trips and 108 took one or more transit trips.[1] The effect of this is to introduce some bias, especially in the transit trip demand equation, in the estimated results. The bias arises

---

[1]See Appendix A for a discussion of the NPTS database.

[2]In both the auto and transit samples some additional observations were lost when consistency checks were made on survey responses. Thus, trips which were unlikely distances, families with no members over the age of four, etc., were purged from the sample.

TABLE 3-1

VARIABLE DEFINITIONS
FOR NONWORK TRAVEL MODELS

Variable Name	Variable Definition
VMT	Vehicle miles traveled by a household for nonwork trips over a four day period[1]
#D.TRP	Number of nonwork automobile trips by a household over a four day period[1]
D.DIST	Average distance of each nonwork automobile trip by a household over a four day period, in miles[1]
#T.TRP	Number of nonwork transit trips by a household over a four day period[2]
D.CO/MI.V.HH	Average gasoline price per mile of a nonwork auto trip for a household divided by the household wage per minute in minutes/miles[2,3]
D.CO/TRP.V.HH	Average gasoline cost per nonwork auto trip for a household divided by the household wage per minute, in minutes[2,3]
D. TIME	Average travel time for an auto nonwork trip by a household. in minutes[4]
D.TM/MI	Average travel time per mile for an auto nonwork trip by a household, in minutes/mile[4]
D.V.HH.TIME	Average travel time for an auto nonwork trip by a household multiplied by the household wage per minute, in cents[3,4]
D.V.HH.TM/MI	Average travel time per mile for an auto nonwork trip by a household multiplied by the household wage per minute, in cents/mile[3,4]
T.TIME	Average travel time for a transit nonwork trip by a household, in minutes/mile[4]
T.V.HH.TIME	Average travel time for a transit nonwork trip by a household multiplied by the household wage per minute, in cents[3,4]
D.PKAV	Fraction of household's nonwork auto trips for which free parking was available.[5]
TAVL.D.T.TRIP	Fraction of household's nonwork auto and transit trips for which transit was available within 6 blocks

TABLE 3-1 (CONTINUED)

Variable Name	Variable Definition
H.H.$	Household income, in dollars/year[6]
#CARS	Total number of cars owned by the household
#LIC.D	Total number of licensed drivers in the household
HH.SZE	Total number of household members
#PPL>4	Number of household members aged 5 or older
#EMPLY	Number of employed persons in household
F.H>HH	Dummy variable equal to one if a female head of household and zero otherwise
URBAN	Coded variable indicating population of urban area ranging from 1, for largest area, to 8, for smallest area[7]
SM.SZE	Coded variable indicating population of SMSA ranging from 2, for smallest area, to 7, for largest area[8]
PLCSZE	Coded variable indicating population of household residence place ranging from 0, for smallest, to 15, for largest[9]

[1]Nonwork trips exclude trips to school and church for those individuals less than 26 years old.

[2]Each automobile in a household was given a code corresponding to the size of car if domestic (three size classes) and a separate code if foreign. A dictionary of auto makes and models for each code was obtained from the Bureau of Census. This was used to compute the average miles per gallon for cars in each code category. The data on miles per gallon were obtained from back issues of Consumer Reports. The automobile used for each trip in the survey was given by the respondent. The amount of gasoline consumed for nonwork trips was estimated based on the size category of car and the associated miles per gallon. The cost of gasoline to each urban area's households for each interview month was obtained from 1969 data on gasoline price by city from the Oil and Gas Journal, Vol 67 (Petroleum Publishing Co., Tulsa, weekly); this data was used to compute the gasoline cost per mile of nonwork auto trips.

TABLE 3-1 (CONTINUED)

[3]Household wage per minute is equal to household annual income in dollars divided by 120000 minutes and transformed to cents.

[4]Time for a trip has the following codes:

15 min. or less
16-30 min.
31-45 min.
46 min.-1 hr.
Bet. 1 and 2 hrs.
2 hrs. or more

This variable was decoded by using the midpoints of intervals and 150 minutes for trips over 2 hours.

[5]For each trip, a variable was assigned a value of 1 if free parking was available; it was assigned a value of 0 if parking was not free, if the tripmaker did not park or if the respondent did not know. D.PKAV is the average value of this variable over all trips. Because parking at home is typically free, this variable will take on values between 0.5 and 1.0.

[6]Income was coded into 11 categories in the NPTS data tape. The household income associated with each code is the midpoint of the range of incomes in the relevant category. No incomes in the highest, open ended, category were recorded in the sample used for estimation.

[7]The code urban population correspondence is as follows:

urban in urbanized area:

1	3,000,000 or more
2	1,000,000 - 2,999,999
3	250,000 - 999,999
4	under 250,000

urban not in urbanized area:

5	25,000 or more
6	10,000 - 24,999
7	2,500 - 9,999
8	rural

[8]The code SMSA population correspondence is as follows:

2	100,000 - 249,999
3	250,000 - 499,999
4	500,000 - 999,999
5	1,000,000 - 1,999,999
6	2,000,000 - 2,999,999
7	3,000,000 and over

TABLE 3-1 (CONTINUED)

[9]The code place population correspondence is as follows:

0	under 200
1	200 - 499
2	500 - 999
3	1,000 - 1,499
4	1,500 - 1,999
5	2,000 - 2,499
6	2,500 - 4,999
7	5,000 - 9,999
8	10,000 - 19,999
9	20,000 - 24,999
10	25,000 - 49,999
11	50,000 - 99,999
12	100,000 - 249,999
13	250,000 - 499,999
14	500,000 - 999,999
15	1,000,000 or more

because there are no values of the dependent variable equal to zero. For dependent variables which typically deviate from zero, such as VMT's, this is not a major source of error. However, because so few transit trips are actually made, there is significant potential for misforecasts with the transit equation in policy evaluation contexts. This issue will be discussed in more detail in the model validation section.

--Many of the important variables are available only as internal categories. For example, time spent for the trip is divided into six time intervals. Income data is also coded into classes, as are urban descriptors. In the cases of trip time and income, variables were decoded by selecting the mid-point of the intervals for the corresponding codes. The result of data represented this way is to increase considerably the error of the estimated parameters and the equations. It does not, however, lead to biased estimates unless the decoding results in values for the variables which are different from the mean of the actual values in the interval. Assuming the mean value of each interval was actually selected, there is a probability that the actual time or income is a different value than the mean though it falls into a range within the interval. This is a random error associated with the variable which has the consequence of increasing the standard error of the estimated equation. Thus, R^2 measures of goodness of fit and t-tests of statistical significance lose some of their explanatory value. In particular we would expect to see low R^2 measures associated with the equations because of the randomness associated with using coded variables.[1]

--Automobile operating costs were constructed from other sources. In order to get the necessary variation in auto costs per mile among households, the information on the size class of cars from the survey was translated into miles per gallon for each trip. Because the urban location of each household had been identified, it was possible to use a gasoline price, quoted in the trade press, for each urban area and date of travel. Thus the auto costs per mile variable reflects city and date specific gasoline prices as well as the individual gas consumption rates of the car used.

---

[1]To see how this effect works, consider a simple linear stochastic model:

$$y = a + bx + u,$$

where y is the dependent variable, x is the true value of the independent variable, u is the stochastic component of the equation, and a and b are parameters to be estimated. Now suppose that instead of observing x directly, we observe the mean of an interval in which x appears. The relationship between x and its observed value is:

$$x = m + e,$$

where m is the observed value and e is the random and unobserved component. If m is used, the estimating equation is as follows:

$$y = a + bm + (u + be).$$

(Footnote continued on following page.)

Table 3-2 presents the salient statistics for the samples used in estimation and for the complete data base before consistency checks. It can be seen that the auto and transit samples have substantial differences. Additionally, it should be noted that the average auto distance, ten miles, is rather long compared to results which typically obtain in urban planning data. Because the NPTS observations are perceived data rather than engineering estimates, there is some reason to suspect that the reported distances were longer than actual trip distances.[2]

---

[1] (Continued from previous page.)

The term in brackets is the new stochastic error attributable to both the error of the original model and the randomness associated with the unobserved component of the independent variable. Unless e is identically equal to zero, the standard error of the latter estimated equation will obviously be larger than the standard error of the first estimated equation. Consequently, the R^2 will be lower for the latter equation and the t-statistic for b will be lower when m instead of x is used for the independent variable. It will not necessarily be the case, however, that the two estimates of b will differ systematically. The latter equation can still yield unbiased parameter estimates.

[2] K. Pat Burnett has reported systematic variations between perceived and engineering estimates of level of service. (Second International Conference on Behavioral Demand, Asheville, 1975). Estimated elasticities from perceived data of VMT's with respect to other variables correlated with distance, such as time and auto cost, tend to cancel the differences between perceived and engineering data, thereby making such elasticities estimated with either type of data more comparable. However, elasticities of VMT's with respect to variables not related to distance, such as parking charges or transit fares, may be biased upward as a result of using perceived data.

TABLE 3-2

SAMPLE STATISTICS

	Auto Sample Mean	Transit Sample Mean	General Sample (before consistency checks) Mean	Max	Min	Standard Deviation	Not Available
VMT	79.32	25.17	68.66	98.4	0.0	95.48	1
#D.TRP	10.01	3.194	8.505	60.0	0.0	7.617	0
D.DIST	9.448	4.926	9.448	155.0	0.0	16.36	115
#T.TRP	.2241	3.065	.4601	16	0.0	1.397	0
D.CO/MI.V.HH	.3269	--	.3257	1.589	.07777	.2195	123
D.CO/TRP.V.HH	3.008	--	3.315	68.55	.09930	5.616	115
D.TIME	18.93	11.15	18.93	115.0	0.0	15.01	115
D.TM/MI	3.234	--	3.233	30.0	.3556	2.653	123
D.V.HH.TIME	168.7	--	168.7	11.98	0.0	147.9	123
D.V.HH.TM/MI	29.18	--	29.17	325.0	.2607	27.91	123
T.TIME	--	30.43	30.43	80.0	0.0	13.16	654
T.V.HH.TIME	--	215.1	215.1	883.3	0.0	101.8	654
D.PKAV	.8983	--	.8983	1.0	0.0	.3452	115
TAVL.D.T.T.TRIP	.3057	.8310	.3706	1.0	0.0	.3949	41
H.H.$	10940.	8832	10422	20000	2500	5127	0
#CARS	1.511	.7407	1.353	5.0	0.0	.7943	0
#LIC.D	1.925	1.139	1.752	5.0	0.0	.8653	0
HH.SZE	3.497	3.231	3.426	15.0	1.0	1.900	2
#PPL>4	3.124	2.870	3.064	9.0	1.0	1.666	1
#EMPLY	1.500	1.370	1.441	6.0	0.0	.8359	0
F.H>HH	.1317	.2685	.1660	1.0	0.0	.3721	0
URBAN	2.878	1.630	2.706	8.00	1.0	2.339	0
SM.SZE	5.694	6.324	5.750	7.0	2.0	1.423	0
PLCSZE	11.21	13.26	11.52	19.0	1.0	3.357	0

3.2.3 Estimation Results

Four separate equations, each with alternative specifications, were estimated using two stage least squares. These relationships can be combined into two distinct simultaneous equation models: the first determines VMT's and transit trips; the second determines number of auto trips, average distance of an auto trip and number of transit trips. All equations represent travel behavior over a four day period for an entire household. Each of the estimated equations and models is discussed below.

Vehicle Miles Traveled

Table 3-3 presents two estimated equations for VMT demand. The first specification, equation (3-1) is preferred. Most variables have appropriate signs with adequate tests of confidence except for cases noted below. Though computed elasticities are in a separate table, some effects should be described here. In particular, each transit trip substitutes for about 24 VMT's, indicating that transit trips serve more purposes than auto trips,which average about 10 miles. Also, each incremental licensed driver contributes about 15 miles over a four day period. (There are nearly two licensed drivers on average in the auto trip sample.)

The geographic variables indicate that as the size of the SMSA (*SM.SZE*) and the place where the household lives (*PLCSZE*) both increase, the total distance traveled tends to decline. However, as the size of the urban center (*URBAN*) increases, more miles are traveled. This effect can be explained because urban centers attract auto trips from surrounding areas and such trips tend to be longer than neighborhood journeys. Increased size in the SMSA and place of residence indicate more opportunities closer to home for travel destinations.

TABLE 3-3

ESTIMATED VMT EQUATIONS
(Two Stage Least Squares)

Variable	Equation (3-1)	Equation (3-2)
	Coefficients (t-statistics)	
VMT	Dependent	Dependent
Constant	162.5 (4.842)	151.3 (3.830)
#T.TRP	-24.24** (2.209)	-23.83** (2.114)
D.TM/MI	-7.838 (2.609)	-6.616 (1.769)
D.V.HH.TM/MI	- .4751 (1.572)	- .6158 (1.547)
D.CO/MI.V.HH	-51.01 (2.244)	-44.91 (1.753)
#PPL>4	4.966 (1.855)	4.948 (1.838)
URBAN	- 3.394 (2.042)	- 3.364 (2.018)
SM.SZE	- 2.897 (1.0468)	- 3.004 (1.0828)
PLCSZE	- 1.979 (1.671)	- 1.956 (1.650)
#LIC.D	15.14 (2.738)	15.52 (2.439)
D.PKAV	-20.04 (0.921)	-17.65 (0.795)
HH$	-----	.0008462 (0.5642)
#CARS	-----	-1.699 (0.252)
R^2 (corrected)	.1059	.1052

**indicates jointly dependent variable

Number of observations = 638

The estimates also indicate that as free parking becomes more available, the number of VMT's decline. Though the t-test on this coefficient is not significant at the 90 percent level of confidence, it was included in the equation because the parameter value persisted through many alternative specifications indicating a higher degree of robustness than the t-test would signify. The result suggests that driving more is an alternative to paid parking, or lack of parking. Extra VMT's are added to journeys when people search for free parking; drivers serving passengers to avoid parking tariffs use the car more than would be the case if it were parked. In any event, small declines in the availability of free parking do not seem to decrease journeys. If free parking availability, other than at home, decreased by 10 percent then the number of VMT's would increase by about 0.8 miles.

Equation (3-2) was included in Table 3-3 to demonstrate the effects of household income and automobile availability. Changes in income have little direct impact on the number of VMT's. Besides being statistically insignificant, the estimated coefficient indicates that as income for the household increases by $1000, then miles traveled increases by 0.8. The effects of income in the demand equation are more important through interaction with the time and price variables than when it is included as a separate variable. The number of cars was included in one specification because some policy instruments may affect auto ownership and thereby affect auto travel. As can be seen in equation (3-2) the inclusion of an auto ownership variable gave counterintuitive and statistically insignificant results. Alternative models of auto travel demand, presented below, achieved different results.

Number of Auto Trips

Table 3-4 presents the estimated demand equations for number of driver trips. With exceptions noted below, parameters are reasonably significant and have the proper signs. Equations (3-3) and (3-4) have the same specification but differ in that the latter used instrumental variables to estimate the cost and time parameters. The reason instrumental variables were used in equation (3-4) on cost and time is that these variables are both directly related to distance which is simultaneously determined with auto trip frequency. Consequently, there is some reason to believe that the interdependence between cost, time and number of auto trips would cause biased estimates if ordinary least squares were used. As can be seen from the estimated equations, the effect on cost and time coefficients of using instrumental variables was significant. The time coefficient was doubled and the cost coefficient was decreased by one-third. Though equation (3-4) is preferred on theoretical grounds, in order to choose between these two equations requires additional information about the elasticities which can only be computed in conjunction with the average auto distance relationship presented later.

The other variables have the expected effects. Larger household income increases the average number of trips directly and by decreasing the cost elasticity. An additional licensed driver in the household contributes nearly two auto trips over a four day period. Additional employed members of the household decrease the number of trips suggesting that the opportunity cost of travel is higher if more people work. Also, the addition of a car increments the number of trips by 1.3 over a four day period. The result of having fewer destinations with free parking is consistent with the evidence presented in the VMT equation -- this would increase the number of auto trips.

TABLE 3-4

ESTIMATED AUTO TRIP EQUATIONS

(Two Stage Least Squares)

Variable	Equation (3-3)	Equation (3-4)	(Equation (3-5)
	Coefficients (t-statistics)		
#D.TRP	Dependent	Dependent	Dependent
Constant	11.12 (5.612)	11.44 (5.702)	11.63 (4.672)
#T.TRP	-1.622** (4.003)	-1.589** (3.894)	-1.457** (3.609)
D.DIST	-----	-----	-0.8258** (3.584)
D.TIME	-.02476 (0.993)	-.05269** (1.063)	-----
D.TM/MI	-----	-----	-.02386 (0.221)
D.CO/TRP.V.HH	-0.1546 (1.893)	-0.1085** (0.7509)	-----
D.CO/MI.V.HH	-----	-----	-1.352 (0.728)
H.H.$	.00008514 (1.360)	.00009403 (1.380)	.00008279 (1.004)
SM.SZE	-.4137 (2.184)	-.4051 (2.136)	-.4398 (2.322)
#PPL>4	.8279 (4.500)	.8122 (4.385)	.8254 (4.477)
#LIC.D	1.725 (3.552)	1.711 (3.514)	1.716 (3.526)
#EMPLY	-.7852 (2.012)	-.7778 (1.992)	-.7953 (2.014)
#CARS	1.308 (2.715)	1.313 (2.713)	1.330 (2.764)
D.PKAV	-5.671 (3.586)	-5.695 (3.597)	-5.667 (3.580)
R^2 (corrected)	.1716	.1710	.1729

** indicates jointly dependent variable

Number of observations = 638

It is to be noted that the t-statistics for cost and time terms are not significant. The estimated parameter values remain relatively unchanged through alternative specifications, however, and the results presented in equation (3-3) appear to be robust though possibly biased through simultaneous equation error.

Equation (3-5) is presented because it is easier to use for policy analysis than the other equations. By separating the cost and time terms form the distance variable, the amount of interdependence with distance is lowered for purposes of exercising a simultaneous equation system. As expected, there is a negative relationship between distance per trip and the number of trips.

Average Auto Distance

Two equations for predicting the average distance in miles for a one-way auto trip are presented in Table 3-5. Instrumental variables were used for estimation of the coefficient on the number of driver trips. Cost and time variables are represented on a per mile basis which minimizes the interdependence between these variables and the dependent variable. Though other variables which appear in the demand equation for number of auto trips could have been included in the average distance equation, it would have been inappropriate to include them in addition to number of driver trips.

Equation (3-6) is preferred. All signs are as expected and the t-tests are significant with the exception of coefficient for household income. Equation (3-7) demonstrates the effects of not including income as a separate variable; that is, the elasticity of distance with respect to auto operating costs per mile declines in value and significance. Because of this effect, and because the coefficient value on income

TABLE 3-5

ESTIMATED AVERAGE AUTO TRIP DISTANCE EQUATIONS
(Two Stage Least Squares)

Variable	Coefficients (t-statistics) Equation (3-6)	Equation (3-7)
D.DIST	Dependent	Dependent
Constant	28.65 (7.703)	26.37 (9.436)
#D.TRP	- .5890** (5.323)	- .6077** (5.576)
D.TM/MI	-1.582 (8.019)	-1.583 (8.010)
D.CO/MI.V.HH	-6.116 (1.718)	-3.704 (1.516)
H.H.$	- .0001472 (0.9301)	-----
URBAN	- .3713 (1.650)	- .3630 (1.611)
PLCSZE	- .3136 (1.932)	- .3085 (1.898)
R^2 (corrected)	.0928	.0896

**indicates jointly dependent variable

Number of observations = 638

was robust under alternative specifications, it was deemed advisable to include income in the final equation.

As can be seen from equation (3-6), extra auto trips substitute for longer journeys; the average distance per trip decreases by about half a mile for each extra auto trip. The urban center size and place of residence size variables have effects which are consistent with the result of the VMT demand equations. That is, as the size of the urban core increases, the average distance traveled per auto trip also increases but as the place size increases the average distance declines.

Number of Transit Trips

Table 3-6 presents three alternative specifications of the estimated transit trip demand equations. For reasons described before, coefficients on transit time and transit availability were estimated with instrumental variables. It should be noted that transit time is the entire time of the trip and is not disaggregated into access, wait and linehaul components. Owing in large part to the poor quality of the data, the test statistics for the transit demand equation indicate substantial randomness in estimates of individual parameters. However, coefficients generally have reasonable values.

The affects of household size and number of people over four must be considered in unison. The parameter estimates on these variables suggest that additional household members will contribute to additional transit trips if they are over the age of four but will cause fewer transit trips if they are aged four or younger. This reflects the burden of making transit trips with young children. Another socioeconomic variable, whether there is a female head of household, is more difficult to interpret. Equation (3-10) included this variable but the apparent strengthening of the estimation results which occurs may simply be a statistical anomaly.

TABLE 3-6

ESTIMATED TRANSIT TRIP EQUATIONS
(Two Stage Least Squares)

Variable	Coefficients (t-statistics) Equation (3-8)	Equation (3-9)	(Equation (3-10)
#T.TRP	Dependent	Dependent	Dependent
Constant	1.444 (1.158)	.9555 (0.910)	1.495 (1.217)
T.TIME	-0.009959** (0.711)	-----	-.01170** (0.846)
T.V.HH.TIME	-----	-.002212** (1.526)	-----
TAVL.D.T.TRP	1.707** (1.666)	1.872** (1.898)	2.391** (2.258)
H.H.$	-.00006254 (1.365)	-----	-.00008049 (1.753)
HH.SZE	-.3722 (1.377)	-.3706 (1.371)	-.4544 (1.688)
#PPL>4	.7877 (2.357)	.7757 (2.324)	.8450 (2.558)
F.H>HH	-----	-----	-1.144 (2.138)
R^2 (Corrected)	.0942	.0963	.1206

** indicates jointly dependent variable

Number of observations = 108

The time spent in making transit trips was specified in two different ways in equations (3-8 and 3-9). In the first specification, income was separated from the time variable with the result that the statistical significance of time was practically nil. In the second specification, income was combined with time by multiplying the trip time with the household wage per minute. The resulting coefficient appears to be stronger but it is probably picking up the negative influence of income more than the time penalty associated with transit trips.

Given these considerations, equation (3-8) is the preferred estimate, though there is little basis for rejecting the others. Alternative specifications which included the effects of auto trips, auto travel level of service, and geographic variables were attempted but with poor results. These estimates indicated that there is negligible cross elasticity of transit trips with respect to auto trip level of service variables in the ranges represented by the data.

Model I Elasticities

Model I is a two equation system using relationships which represent the demand for VMT's and the demand for transit trips. Though several combinations of equations from Tables 3-4 and 3-5 are possible to form Model I, the preferred relationships are equations (3-1) and (3-8). These were used to compute the elasticities represented in Table 3-7.

Computation of own elasticities using the linear Model I is relatively straightforward. To see this consider the general linear equation:

$$Y = \alpha_0 + \alpha_1 X_1 + \dots + \alpha_n X_n$$

Then the elasticity of Y with respect to X_i is:

$$\frac{\frac{dY}{dX_i}}{\frac{Y}{X_i}} = \frac{\alpha_i X_i}{Y}$$

TABLE 3-7

Model I

LEVEL OF SERVICE ELASTICITIES

Own Elasticity of VMT's	
With respect to:	
Auto travel time per mile	-.4944
Gasoline cost per mile	-.2102
Free parking availability	-.1006
Own Elasticity of Transit Trips	
With respect to:	
Transit trip time	-.0989
Transit availability	.4628
Cross Elasticity of VMT's*	
With respect to:	
Transit trip time	.0068
Transit availability	-.0317

*Percentage change in VMT's divided by percentage change in transit trips = -.0685.

Each of the elasticities presented in Table 3-7 were computed at the means of dependent and independent variables used in the estimation sample, given in Table 3-2, except for parking availability, which is discussed further below. The elasticity of VMT's with respect to auto time per mile is the sum of the elasticity of VMT with respect to D.TM/MI and D.V.HH.TM/MI. The elasticity of VMT's with respect to auto gasoline cost per mile is the elasticity of VMT with respect to D.CO/MI.V.HH.

To compute the elasticity of VMT's with respect to free parking availability it was necessary to adjust the mean for residential parking. On the assumption that half of all trips ended at home, .5 was subtracted from the mean of this variable to yield .3983.

The cross elasticity of VMT's with respect to transit trip level of service variables is equal to the product of VMT elasticity with respect to transit trips times the transit trip elasticity with respect to transit level of service.

The VMT elasticities with respect to time and cost are within the ranges validated by the other studies reviewed in Section 1. The free parking availability elasticity does not have supporting evidence from other studies because no research has been done on the overall effects of parking restrictions.[1]

The transit trip time elasticity is low compared to the results presented in Section 2 on the work trip. This may be a reflection on the average quality of transit service, and preferences against transit, for nonwork trips.

---

[1]For a review of the existing research in this area see CRA, *Policies for Controlling Air Pollution*... (forthcoming).

Over a four day period only .224 transit trips were made by households which also made auto trips. This also accounts for the low cross elasticity of VMT's with respect to transit level of service; a large percentage change in transit trips has a small effect on VMT's.

Model II Elasticities

To form Model II, three equations are used to represent the demand for auto trips, average auto trip distance and the demand for transit trips. The equations for these relationships are given in Tables 3-4, 3-5 and 3-6. There are potentially a large number of combinations of equations from these tables to create Model II and the selection of a preferred set is somewhat subjective. For the purposes of computing Model II elasticities, equations (3-4, 3-6, and 3-8) were used. The elasticities are given in Table 3-8.

Model II is more complicated to use than Model I. One problem which arises is that average trip length per household times average trip distance per household is not equal to average VMT's per household. Thus, simply multiplying the predictions of equations (3-4 and 3-6) will not yield suitable predictions for VMT's. Fortunately, a simple method is available to make an approximate adjustment.[1]

To demonstrate the problem we note the following statistical identity:

$$\overline{VMT} = (\overline{\#D.TRP})(\overline{D.DIST}) + COVARIANCE\ (\#D.TRP,\ D.DIST)$$

where variables with bars over them indicate means over the

---

[1]The most accurate method for computing elasticities or other policy effects would be to simulate the model over all the household observations in the sample. Though this approach would not be difficult if the requisite software and data base are available, it is not pursued here.

TABLE 3-8

Model II
LEVEL OF SERVICE ELASTICITIES

Own Elasticity of VMT's

With respect to:	
Auto travel time per mile	-.4968
Gasoline cost per mile	-.2189
Free parking availability	-.0854
Auto ownership	.0745

Own Elasticity of Transit Trips

With respect to:	
Transit trip time	-.0989
Transit availability	.4628

Cross Elasticity of VMT's*

With respect to:	
Transit trip time	.0013
Transit availability	-.0062

*Percentage change in VMT's divided by percentage change in transit trips = -.0134.

sample of households. It can be seen that the mean of VMT's differs from the product of mean auto trips and distance by the covariance between auto trips and distance. This covariance is negative and can be computed for the auto sample from statistics presented in Table 3-2:

$$COVARIANCE(\#D.TRP,\ D.DIST) = \overline{VMT} - (\overline{\#D.TRP})(\overline{D.DIST}) = -15.25$$

To compute elasticities, it will be necessary to take account of this covariance term. The elasticity of VMT with respect to some variable X will be:

$$\frac{\dfrac{d\ \overline{VMT}}{dX}}{\dfrac{\overline{VMT}}{X}} = \frac{\dfrac{d[(\#D.TRP)(D.DIST)]}{dX} - \dfrac{d[COVARIANCE(\#D.TRP,\ D.DIST)]}{dX}}{\dfrac{\overline{VMT}}{X}} \tag{3-11}$$

Concentrating only on the numerator, on the right hand side of the equation, the derivative of the covariance term is not tractible unless simplifying assumptions are made. One reasonable assumption is that this derivative is proportional to the other derivative in (3-11) where the constant of proportionality equals the following:

$$\frac{COVARIANCE(\#D.TRP,\ D.\ DIST)}{(\overline{\#D.TRP})(\overline{D.DIST})} = -.1612$$

Thus, as an approximation, the change in VMT caused by a unit change in X can be computed as follows:

$$\frac{d\overline{VMT}}{dX} = .8388 \left((\frac{d\#\overline{D.TRP}}{dX}) \ \overline{D.DIST} + \#\overline{D.TRP} \ (\frac{d \ \overline{D.DIST}}{dX}) \right)$$

Further complications are introduced because $\#\overline{D.TRP}$ $\overline{D.\ DIST}$ are simultaneously determined. One method of dealing with this is to solve the two equations simultaneously at the old and changed values of X and compute the resultant change in VMT's. Another method, which is applied below, is to use the chain rule in computing the appropriate derivatives. This will allow us to calculate the change in VMT's which result from the change in X by using the partial derivatives in the following equation:

$$\frac{d\ \overline{VMT}}{dX} = .8388 \left[\left(\frac{\partial\#\overline{D.TRP}}{\partial X} + (\frac{\partial\#\overline{D.TRP}}{\partial\overline{D.DIST}})(\frac{\partial\overline{D.TRP}}{\partial X}) \right) \overline{D.DIST} + \#\overline{D.TRP} \left(\frac{\partial\overline{D.DIST}}{\partial X} + (\frac{\partial\overline{D.DIST}}{\partial\#\ \overline{D.TRP}})(\frac{\partial\#\overline{D.TRP}}{\partial X}) \right) \right] \quad (3\text{-}12)$$

Equation (3-11) is the formula used to compute the derivatives for elasticities presented in Table 3-8. To give an example of its application, we will show how the change in VMT with respect to a change in time per mile ($D.TM/MI$) is calculated. All the parameters and data necessary for computation of the derivatives are given in equations (3-4, 3-6, and 3-8), and Table 3-2.

First we note that the partial derivative of number of auto trips with respect to time per mile is equal to the coefficient on trip time in equation (3-4) times the average distance of the trip:

$$\frac{\partial\#\overline{D.TRP}}{\partial\overline{D.TM/MI}} = -.05269\ (\overline{D.DIST}) = -.4978$$

Next we compute the partial derivative of trip frequency with respect to average trip distance as follows:

$$\frac{\partial \#\overline{D.TRP}}{\partial \overline{D.\ DIST}} = \left(\frac{\partial \#\overline{D.TRP}}{\partial \overline{D.TIME}}\right)\left(\frac{\partial \overline{D.TIME}}{\partial \overline{D.DIST}}\right) + \left(\frac{\partial \#\overline{D.TRP}}{\partial \overline{D.CO/TRF.V.HH}}\right)\left(\frac{\partial \overline{D.CO/TRP.V.}}{\partial \overline{D.DIST}}\right)$$

$$= -.05269\ (\overline{D.TM/MI}) - .1085\ (\overline{D.CO/MI.V.HH})$$

$$= -.2059$$

The partial derivatives of average distance with respect to time per mile and number of auto trips can be calculated directly from the relevant coefficients in equation (3-6):

$$\frac{\partial \overline{D.DIST}}{\partial \overline{D.TM/MI}} = -1.582$$

$$\frac{\partial \overline{D.DIST}}{\partial \#\overline{D.TRP}} = -.589$$

This completes the computation of partial derivatives which appear in equation (3-12). Using the values for $\#\overline{D.TRP}$ and $\overline{D.DIST}$ given in Table 3-2, the change in average household VMT's over a four day period caused by a unit change in time per mile can be calculated using formula (3-12):

$$\frac{d\ \overline{VMT}}{d\ \overline{D.TM/MI}} = -12.1854$$

The elasticity of VMT's with respect to time per mile is then computed as follows:

$$\frac{\dfrac{d\ \overline{VMT}}{d\ \overline{D.TM/MI}}}{\dfrac{\overline{VMT}}{\overline{D.TM/MI}}} = -.4968$$

which is the value given in Table 3-8.

The other elasticities are computed similarly using formula (3-12). For variables which do not appear in equation (3-6), the partial of $\overline{D.DIST}$ is identically zero. This eases, for example, the computation of auto ownership and free parking availability elasticities.

It can be seen from Tables 3-7 and 3-8 that there is little difference between VMT own elasticities for Model I and Model II. However, the VMT cross-elasticities with respect to transit service are much smaller for Model II than for Model I. This can be attributed to some misspecification in Model II where number of transit trips was not included in the average auto distance equation. In comparing the results of Model II with Model I, it is apparent that transit trips substitute for longer driver trips.

3.2.4 Model Validation

Two exercises were performed with Model I (equations (3-1 and 3-8)) to determine the problems involved with transferring the equation to alternative sources of data. The results indicate that the model is not instantly generalizable except, possibly, for the elasticities.

The first attempt at validation was simply to apply the model to the aggregate sample of 765 households which includes 115 households not taking auto trips. As mentioned before, the estimation of the model on a truncated sample excluding observations with zero VMT's could bias the parameter estimates. By applying the model to the aggregate sample, some indication of the extent of this bias can be determined. The results are as follows:

-- actual nonwork VMT per household over a four day period = 68.66: predicted nonwork VMT per household over a four day period = 70.63;

-- actual nonwork transit trips per household over a four day period = .4601: predicted nonwork transit trips per household over a four day period = 2.260.

As can be seen from the above results, the use of aggregate data which includes households who did not take transit trips, especially when level of service data is missing, overpredicts the number of trips taken. This implies that the preferred method for using the transit equation is to use the elasticities which are purged, to a large extent, of the scale effects which would cause overprediction if the equation were applied directly. If a data base can be reconstituted into overlapping auto samples and transit samples, as was done for estimation, then the equations may be applied directly.

As a further test of the model, it was applied to an aggregated sample of 992 households from a Los Angeles sketch plan zone. This area is a low income region south of Los Angeles city center with better than average bus service for the Los Angeles region. Level of service data are engineering estimates based on time and distance matrices for traffic analysis zones. The results of applying Model I to this data are as follows:

-- actual nonwork VMT per household over a four day period = 17.95: predicted nonwork VMT per household over a four day period = 45.00;

-- actual nonwork transit trips per household over a four day period = .4609: predicted nonwork transit trips per household over a four day period = 2.301.

Again the effect of not having separate samples for auto and transit users produces the result that travel by each mode is overpredicted. In the case of transit, the percentage error is of the same proportion obtained when Model I is applied to national data. When used to predict VMT's, the model is less accurate when applied to Los Angeles data than when applied to the national sample, though the direction of change is correctly predicted.

This result created the suspicion that the reported distances in the NPTS were grossly out of line with actual distances. Yet when the results of the NPTS sample are expanded to a full year of travel, the VMT's per auto attributable to nonwork travel are equal to 4,630 miles per year. Given the rule of thumb that about half of all VMT's occur on nonwork trips, the reported distances are consistent with the accepted datum that the average car accumulates 10,000 miles per year. This throws some doubt on the validity of the Los Angeles data.

One reason for the difference in estimates could be the presence of unobserved variables, the effects of which could not be estimated. To test for this problem, several equations of various specifications were estimated using error components regressions. The parameters obtained from these estimates typically did not vary significantly from the parameters presented in Tables 3-3 through 3-6. The regressions did indicate that the constant terms for auto related equations should be somewhat smaller than those given by standard two stage least squares estimates if the equations were to be applied to Los Angeles; however, when the error components estimates were applied to the Los Angeles data, they also significantly overpredicted VMT's and transit trips. It was concluded from these results that the absence of city specific unobserved variables in and of themselves did not significantly bias the results.

Given the conflicting evidence, there are relatively few claims which can be made about direct application of the models until procedures are developed to deal with the problem of separate samples used for estimation. It is likely that such procedures can be developed routinely. In the meantime, it is more appropriate to apply the estimated elasticities.

3.3 POLICY EVALUATION WITH NONWORK TRAVEL MODEL ELASTICITIES

The specification of the nonwork travel model makes it somewhat less cumbersome to use in policy evaluation applications than the work trável models presented in Section 2. However, in applying the nonwork model, it is important to translate the policy instruments in such a way that they effect the independent variables accurately. This poses some problems in model application which are discussed by example below. Four types of policy scenarios are considered dealing with gasoline taxes, parking restrictions, fuel economy measures and transit level of service improvements.

Each of the policy scenarios represent broad gauge options available for national urban transportation policy. The exercise of the model under various policy contingencies points up the strengths and weaknesses of performing instant policy evaluation with known elasticities. Often the policy instruments are not represented directly by the variables in the model and assumptions must be made about the appropriate correspondence. For some scenarios, supply effects need to be known for a complete evaluation and, in the absence of knowledge about such effects, further assumptions must be made. A consequence of this situation is that separate analysts using the same model can arrive at different conclusion about the results of a particular policy because of the differences in judgement on key assumptions. The examples of model application presented below indicate how judgements and assumptions affect the outcome of policy evaluation.

3.3.1 Gasoline Tax

One of the more routine applications of the model is to changes in gasoline price. This would occur if an additional tax was imposed on the sale of gasoline to motorists. As in the case of the work trip gasoline tax scenario, it is assumed that a nationwide tax of 100 percent is placed on the pretax cost of gasoline. The result of such a tax on the pump price of gas is to increase its value by, on average, 70 percent. Using either Model I or Model II, the predicted effect of this action would be to decrease nonwork trip VMT's by about 15 percent or about 12 miles per four day period for the average household. This is computed by multiplying the relevant elasticity presented in either Table 3-7 or Table 3-8 by the percentage change in the pump price of gasoline.

3.3.2 Parking Restrictions

Because the nonwork trip models do not have well defined parking cost elasticities, it is only possible to make very approximate estimates of the effects of various parking policy options. In order to evaluate the effects of parking restrictions they must be translated into the resultant change in free parking availability (*D.PKAV*). Even then, not all of the impacts may be predicted because of the absence of parking price effects in the models.

In the work trip section of this report, the policies analyzed included parking taxes and the rationing of parking places through government regulations. Even if such policies were applied locally to those areas where paid parking is typically the rule, the regionwide effect would be a reduction in the availability of free parking. The extent to which free parking availability would decline cannot be known with accuracy until the supply response of parking sites with respect to cost changes is analyzed. Nonetheless, some inferences about the impacts of such policies on nonwork VMT's can be determined from the line of reasoning pursued below.

In the absence of parking restrictive policies, it can be presumed that the pattern of parking charges over an urban area is related to the demand for parking and the cost of land for alternative uses. Parking rates charged by entrepreneurs represent, in large part, a pure rent to the urban land. If such land is inelastic in supply, then parking taxes would cause little increase in the price charged to motorists but would, instead, be passed backwards to decrease rents. In the longer run, parking sites would be converted to uses which offer higher rents decreasing the supply of parking around premium locations. This would increase the demand for parking in other locations where, perhaps, parking was free before the imposition of the policy.

Consequently, the result of a parking restriction policy has two effects on free parking availability: the first is to create parking charges where none previously existed by the imposition of a direct tax; the second, and more difficult to model, is to increase the proportion of paid parking sites and the proportion of travel destinations where parking is not convenient by artificially restricting supply or by using taxes to create such a supply response. In either case, the areas where free parking would change is on the fringes of activity centers where paid parking already obtains or in activity centers where free parking obtains but the parking restrictions and taxes could be easily monitored. Many potential trip destinations for nonwork travel, such as private homes or small semiurban commercial establishments, would not be effected nor could any parking restriction policy be inexpensively enforced.

To determine the implications of parking restrictions consistent with those imposed in work trip scenarios, it was assumed that a broad based parking control plan would change free parking availability at potential trip destinations on the order of 50 percent. Using the elasticity on free parking availability from Model I, given in Table 3-7, the result of such an action would be to increase the nonwork VMT's by 5 percent to about 4 miles per four day period per household.

Given that the increased restrictions on parking tend to increase nonwork VMT's, there is some doubt about the efficacy of such a policy to curb emissions or to reduce energy consumption. All effects of such a policy need to be considered and these include the induced changes in work trip VMT's and the impact of higher parking prices on nonwork travel, before such a policy could be realistically proposed as a solution to environmental and energy problems. Even if this is done, there is reason to suspect that parking rationing by itself will not have the desired result in a number of urban contexts.

3.3.3 Fuel Economy of Automobiles

It has been asked whether mandated fuel economy regulations on new cars might be counterproductive in efforts to conserve energy. The issue arises as a result of lower fuel consuming automobiles being cheaper to operate and thereby inducing more travel. The nonwork travel model can be used to give reasonably precise estimates of fuel saved with miles per gallon regulations including the effects of induced travel.

To apply the model, the average miles per gallon for travel in the 1969 sample needs to be known. Using the Consumer Reports figures for new cars, the average gasoline consumption rate was 16 miles per gallon. This is somewhat higher than rates typically given for 1969 (around 13.5 miles per gallon in the Statistical Abstract) and the difference can probably be attributed to the decay in fuel efficiency as cars are used over time and to the differences in data reporting. However, the Consumer Reports data are well suited to regulations which apply to cars as they come off the assembly line.

For the purposes of illustration, suppose that the effects of a federal restriction on the fuel economy of new cars is such that the average miles per gallon increases by 50 percent to 24 miles per gallon. The result of this regulation would be to decrease auto gas costs per mile by 50 percent. Application of either the Model I or the Model II elasticity to this change in auto costs indicates that the resulting increase in VMT's is on the order of 10.5 percent or about 8 miles per household per four day period. The net fuel consumption decline would be about 40 percent.

The implications from the elasticity of VMT's with respect to automobile gas costs per mile are that net percentage fuel reductions owing to increased fuel economy by automobiles are .8 times the percentage change in miles per gallon.

3.3.4 Transit Availability

The effects of various transit policy options on VMT's and transit policy options on VMT's and transit trips cannot be very refined using the nonwork travel models because transit variables are highly aggregated. Also transit costs were excluded from the model because of the lack of data on this in the NPTS survey.[1] Moreover, trip time by transit was not divided into access, wait and linehaul components. Finally, the model does not readily accept new modes as a separate travel choice from either auto or transit. The result of these shortcomings is that the analysis of any particular

---

[1]Conceptually it would have been possible to include the affects of the transit fares on transit ridership by adding data from other sources to the household data set used for estimation. Other information on levels of transit service which were city and even place specific also could have been merged with the data base. This effort would be similar to the approach used in computing auto travel costs which used city specific data on gasoline prices from sources other than the NPTS. It is suggested that if further model estimation is pursued using the NPTS data base that these adjustments be made.

policy scenario involving transit must be judgemental in translating policy instruments into model inputs.

Some implications about transit policy are readily apparent from simple inspection of model parameters and elasticities. For example, it is unlikely that anything except for the most major of auto disincentives will have any significant impact on diverting motorists to transit for nonwork travel. Moreover, because the cross elasticity between transit trip time and VMT's is quite small, it is improbable that simple adjustments in linehaul and wait times for transit will divert motorists to transit. For example, a 10 percent increase in transit speeds, a scenario analyzed in the work trip section, would reduce VMT's by at most 0.068 percent.

With regard to transit fare policy, it is interesting to use external estimates of the transit fare elasticities in association with the estimated VMT elasticities from Model I. Assume, for example, that the transit fare own elasticity is -.33 as is the usually accepted figure. If this is multiplied times the VMT elasticity with respect to transit trips from equation (3-1), given in Tables 3-7, the resulting estimate of VMT cross elasticity with respect to transit fares is .022. Thus if transit fares were uniformly lowered by 50 percent, the consequent decline in nonwork VMT's would be only 1 percent or about .8 VMT's per household per four day period.

To analyze the effects of increased transit availability, the models can be applied more directly. The definition of transit availability, consistent with the definition of the variable in the transit demand equation, is the proportion of trips for which a transit mode was available within one half mile (six blocks). A scenario consistent with the ones applied to the work trip model is to consider the effects of making transit available, within six blocks, for all nonwork trips. One method for achieving this would be to institute a regionwide door to door dial-a-ride service. Another approach may be simply to extend conventional bus service by the necessary route miles to cover all the relevant origins and destinations with some ill-defined but minimal number of transfers.

The result of such a policy on nonwork trips is to increase the availability of transit, as defined above, by 327 percent. Applying the relevant cross elasticity on VMT's given in Table 3-7 to this scenario yields the result that such a policy would decrease VMT's by about 10 percent or by about 8 miles per household per four day period. This is obviously a rather rough estimate because different methods of changing transit access will have different effects: for example, dial-a-ride clearly minimizes the walk time to transit while conventional bus systems still may entail substantial walk access, yet both changes are predicted to have the same effects. In order to predict the differential impact of various types of transit improvements the transit demand equation should be refined with more information specific to the service characteristics of particular urban systems.

Beyond the short run effects of transit service improvements, proponents of transit system investments often make the claim that the long-run changes in urban form caused by transit offer substantial efficiencies in transportation when compared to spread development in auto dominated urban forms. To our knowledge, this argument has never been validated by an examination of the data on comparative travel patterns among cities of alternative development patterns. As shown below, the results of the estimated VMT equations cast some doubt on the effects of increased density on reducing nonwork VMT's.

To analyze this issue, we use equation (3-1) and consider the assumed impacts of a new system on the urban area specific variables of SMSA population, urban center population and place size population. Of these, it can be presumed that SMSA population would remain unchanged because the effect of the new system would be to redistribute population within the region rather than to change the level of population. The effects of the system may be considered to be on the order of changing the average place size and urban center size from that which obtains in the auto trip sample to the averages in the transit trip sample. Not surprisingly, the average of both place size and urban center size would increase. By increasing the size of place of residence, on average, there is a consequent decline in VMT's per household per four day period equal to 3.9 miles. However, the increase in the attractiveness of the urban center causes an increase in average VMT's of about 4.2 miles. Thus the net effect on nonwork VMT's of increasing densities and the importance cf the downtown through major transit systems may well be negligible.

3.3.5 Summary

Even though the specification of the nonwork trip models were somewhat restrictive in terms of the number of variables available to test policy options, the estimated elasticities were applied to a relatively diverse set of policy scenarios. These scenarios and the predicted effects are summarized in Table 3-9.

The results of estimating the nonwork travel demand changes caused by various policy instruments can be summarized as follows:

TABLE 3-9

SUMMARY OF POLICY SCENARIO

Predictions Using Nonwork Travel Model Elasticities

Policy Scenario	Percentage Change in VMT's
100% Gasoline Tax	-15%
50% Decline in Free Parking	5%
50% Increase in Auto Fuel Economy	10%
10% Increase in Transit Speed	0%
50% Decline in Transit Fare	- 1%
Transit Available for All Trips	-10%
Higher Density Urban Form	0%

Auto Travel Controls

To achieve objectives related to reducing VMT's, emissions or energy consumed, policy instruments should correspond as closely as possible to these objectives. Thus, for examples, taxes on VMT's through a gas tax or regulations on energy efficiency have more impact than parking regulations. In fact, if parking regulations are applied to off-peak travel, they may well have counterproductive effects.

Transit Improvements

The response of travelers to transit for nonwork trips appear to be significantly different from their response for work trips. It can be presumed that mode switching for nonwork trips is more sensitive to transit access than to transit fares and linehaul times. Consequently, transit innovations for non-work trips should probably be oriented toward serving a dispersed set of origins and destinations with relatively low level of service compared to peak transit level of service which serves work trips. This strategy will not reduce auto travel by a large amount but it appears to have a larger impact than more inflexible transit systems, such as transit facilities with their own guideways, on off-peak trips in both the short and long run.

The inadequacies with the nonwork travel demand estimates given in this chapter are largely the result of gaps in the NPTS data base. The lack of information on alternate mode level of service and, perhaps, city specific variables certainly reduces somewhat the confidence with which model predictions can be made. This problem was most severe in the estimation of transit demand and, possibly, caused the low values for transit direct elasticities. However, even higher values for transit own direct elasticities would not have had much effect on increasing the VMT cross elasticities with respect to transit level of service because so few transit trips are made for nonwork purposes. Future attempts to estimate disaggregate travel demand models with national databases would do well to merge the household trip data with publicly available external data on the level of service for transit in specific urban areas.

The other major shortcomings of the model are the result of a conscious tradeoff in model specification to make it easy to use. The linear form of the equations and the relatively high degree of aggregation of the relevant variables undoubtedly misrepresents to some degree the actual behavioral process in something as complex as nonwork travel demand. However, by representing the range of travel choices with continuous variables rather than a relatively small number of discrete options, the linear specification may be an improvement over some probability choice specifications. Future work in estimation of nonwork travel demand may productively consider model forms which wed the strengths of both specifications such as probability choice models which apply to a continuous range of options.

Finally, the direct application of the linear model is now limited by the use of two separate samples for model application. This was done because of the lack of data on modes not chosen. Insofar as alternate mode data are among the missing elements from many disaggregate data sets, some research into how to apply models estimated on separate (though overlapping) samples would potentially have a high payoff. As mentioned before, from such an effort it is likely that routine methods could be developed for direct application of the models to widely available data.

APPENDIX A
THE NPTS SURVEY

The Nationwide Personal Transportation Survey (NPTS) is a collection of household interviews which were conducted between April 1969 and January 1970. Approximately 6000 households were surveyed (half of them four times) and every state and the District of Columbia is represented. Each individual in a given household is interviewed; three types of information are obtained: (1) general socioeconomic information about the individual and the household (income, auto ownership, etc.); (2) a categorization of the usual work-trip, shopping-trip, and travel-to-school patterns; and (3) a record of each trip taken by each individual in the household on a designated day. The FHWA Office of Highway Planning (Division of Program Management) has published a series of pamphlets based on the survey, reporting tabulations of the aggregated survey data projected to nationwide or annual estimates.

To conduct the NPTS survey, the country was divided into approximately 1900 primary sampling units (PSU's). These PSU's were grouped into 235 strata of one or more and are roughly homogeneous according to some socioeconomic criteria. One PSU was selected from each of the 235 strata. From each chosen PSU, a sample of households was randomly selected. In total approximately 6000 households were selected, and they were divided into two groups of approximately 3000 households each. One group was first interviewed in April 1969, and then three subsequent times: July 1969, October 1969, and January 1970. It appears there is a slight variation in the number of households interviewed in each of the four months; this is probably due to the difficulty of finding people at home, etc. In the first interview the three types of information outlined above are obtained: (1) general socioeconomic, (2) usual travel patterns, and (3) specific trip records. The followup interviews were concerned only with trip records. The second group of households was interviewed only once, in August 1969; all three types of information were obtained in this interview.

The questionnaire used in the survey consists of seven parts:

- control card, aimed at obtaining basic geographic and socioeconomic information about the household;
- Section I, Automobile Record, with an entry for each car owned/used by the household, asking ownership, miles traveled, parking and passengers for the work trip;
- Section II, Shopping, asking about shopping trips to the CBD;
- Section III, Travel to Work, asking about the usual trip to work;
- Section IV, Driver Information, asking the miles an individual has driven in the past year;
- Section V, Travel to School, asking about the usual trip to school;
- Section VI, Travel Day Report, recording all trips taken by an individual on a designated day; and
- Section VII, Overnight Travel, recording all overnight trips taken during the week before the designated travel day.

A copy of the questionnaire begins on the following page. The top section is a transcribed portion of the control card. On the first visit to each household, the control card was filled out, and each individual in the household completed

NOTICE – All information which would permit identification of the individual will be held in strict confidence, will be used only by persons engaged in and for the purposes of the survey, and will not be disclosed or released to others for any purposes.

FORM NPT-2
(7-10-69)

U.S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS
ACTING AS COLLECTING AGENT FOR THE
U.S. DEPARTMENT OF TRANSPORTATION

HOUSEHOLD QUESTIONNAIRE – AUGUST 1969

NATIONWIDE PERSONAL TRANSPORTATION SURVEY

BUDGET BUREAU NO 41-S69011
APPROVAL EXPIRES DECEMBER 1970

1

a. Ident. Code	b. Household No.	c. Control No. PSU	Rot.	Segment	Serial	Str.

d. Type of structure	e. Race	f. SMSA	g. Place	h. State

i. Subsample	j. Designated travel day: Day of week	Mo./day	k. No. of hhld. members (all ages)	l. Number of automobiles
.				

m. Automobile

Auto No.	Year	Make	Office use	n. Principal user Line No.

o. *(If no automobile)*
1 ☐ Auto available
2 ☐ Not available

p. Income

q. Interviewer's code

r. OFFICE USE

s. Date of interview

t. Noninterview reason
1 ☐ NOH
2 ☐ TA
3 ☐ Ref.
4 ☐ Other Type A
5 ☐ Other type – *Specify*

(Fill c, b, c, f, g, h, i, j, q.)

Section I – AUTOMOBILE RECORD

Now I have some questions about your – – (first, second, etc., automobile)	Auto No.	Auto No. 2	Auto No.
1. Is it owned by somebody living here?	1 ☐ Yes 2 ☐ No *(Go to Q. 3)*	1 ☐ Yes 2 ☐ No *(Go to Q. 3)*	1 ☐ Yes 2 ☐ No *(Go to Q. 3)*
2a. Was it purchased new or used?	1 ☐ New 2 ☐ Used	1 ☐ New 2 ☐ Used	1 ☐ New 2 ☐ Used
b. In what month and year was it bought? (Examples 10/67, 04/68)	Month / Year	Month / Year	Month / Year
3. About how many thousand miles was it driven during the past 12 months?	Miles *(Thousands)*	Miles *(Thousands)*	Miles *(Thousands)*
4. Is it used at least once a week in going from home to work?	1 ☐ Yes – Entire trip 2 ☐ Yes – Part-way 3 ☐ No *(Go to next auto or Sec. II)*	1 ☐ Yes – Entire trip 2 ☐ Yes – Part-way 3 ☐ No *(Go to next auto or Sec. II)*	1 ☐ Yes – Entire trip 2 ☐ Yes – Part-way 3 ☐ No *(Go to next auto or Sec. II)*
5. How many people are usually in the automobile going to work, including the driver?	Number	Number	Number

CODE KEY →
1 – Commercial parking garage or lot
2 – Employer provided space
3 – Fringe parking
4 – Other lot or garage
5 – On the street
6 – No all day parking used
7 – Other

6a. What type of parking facility is usually used for the trip to work – the employer's lot, a commercial lot, on the street, or what?			
	If code 6 go to next auto or Sec. II		
b. Is there a cost for parking?	1 ☐ Yes 2 ☐ No *(Go to next auto or Sec. II)*	1 ☐ Yes 2 ☐ No *(Go to next auto or Sec. II)*	1 ☐ Yes 2 ☐ No *(Go to next auto or Sec. II)*
c. How much?	\$ ____.____ 1 ☐ Day 2 ☐ Week 3 ☐ Month	\$ ____.____ 1 ☐ Day 2 ☐ Week 3 ☐ Month	\$ ____.____ 1 ☐ Day 2 ☐ Week 3 ☐ Month
d. Does . . . pay by putting coins into a meter?	1 ☐ Yes 2 ☐ No	1 ☐ Yes 2 ☐ No	1 ☐ Yes 2 ☐ No

Section II – SHOPPING

ASK for SMSA residents only – 1 or 2 as second digit of identification code

Now we are interested in where people shop –
(Ask 1 and 2 for (1) wife or (2) female head or (3) male head)

1. During the past 3 months has . . . gone to the main business district of ______ principally to shop?
   - 1 ☐ Yes → How many times?__________ *(Go to Q. 3)*
   - 2 ☐ No

2. What were the reasons for not shopping there?
   (Mark all boxes that apply)
   - 1 ☐ Goods available locally
   - 2 ☐ Too far away
   - 3 ☐ Difficulty of parking
   - 4 ☐ Difficulty of driving in congested area
   - 5 ☐ No automobile
   - 6 ☐ Other – *Specify*

3. How far is it from home to the nearest public transportation line to go to the main business district of __________?
   - 1 ☐ Less than one block
   - 2 ☐ 1–2 blocks (less than ¼ mile)
   - 3 ☐ 3–6 blocks (¼ – ½ mile)
   - 4 ☐ Over 6 blocks (over ½ mile)
   - 5 ☐ No public transportation available
   - 6 ☐ Lives in main business district

Note Fill remaining pages for household members 5 years old or over.

Section III – TRAVEL TO WORK

(3)

1. Line No.	2. CHECK ITEM 1 [ ] This person is 16 years old or older and has an entry in Control Card question 16b. *(Fill in Sec. III, IV, and V as applicable)* x [ ] All others *(Fill in Sec. IV and V as applicable)*

We are interested in where people work and how they get to work. 3. Is the place where . . . works located in a city?	1 [ ] Yes → What city? ______ 2 [ ] No 3 [ ] Don't know — State? ______
4. How far is it from home to the place where . . . works? (Actual travel distance)	Miles ______ *(Enter nearest full mile)* 1x [ ] No fixed place 2x [ ] At home } *(Go to Sec. IV)* 3x [ ] Less than ½ mile (5 blocks)
5. How much time is usually required for . . . to get to work from the time he leaves until he arrives at work?	Minutes
6. How does . . . usually get to work? *(Mark all appropriate boxes)*	1 [ ] Bus or street car 2 [ ] Commuter train, subway, elevated, etc. 3 [ ] Automobile – with other persons 4 [ ] Automobile – alone 5 [ ] Truck 6 [ ] Motorcycle 7 [ ] Walk only *(Go to Q. 10a)* 8 [ ] Other – including bicycle – Specify ______
7. How far is it from home to the nearest public transportation line that . . . uses (could use) to get to his place of work?	1 [ ] Less than 1 block 2 [ ] 1 to 2 blocks (less than ¼ mile) 3 [ ] 3 to 6 blocks (¼ to ½ mile) 4 [ ] Over 6 blocks (over ½ mile) 5 [ ] None available } *(Go to Q. 10a)*
(Ask if boxes 1 and/or 2 – is not marked in Q. 6) 8. What is the reason . . . does not use public transportation to go to work? Anything else? *(Mark all boxes that apply)*	1 [ ] None available 2 [ ] Not convenient to get to 3 [ ] Not convenient to place of work 4 [ ] Too many transfers 5 [ ] Too expensive 6 [ ] Too crowded or uncomfortable 7 [ ] Takes too long 8 [ ] Need auto for work 9 [ ] Other – Specify ______ *(Go to 10a)*
(Ask if either box 1 or 2 – is marked in Q. 6) 9. What is the reason . . . uses public transportation to get to work? Anything else? *(Mark all boxes that apply)*	1 [ ] No driver's license 2 [ ] No car available 3 [ ] No car pool available 4 [ ] Cheaper than auto 5 [ ] Safer than auto 6 [ ] No parking problems 7 [ ] No driving strain 8 [ ] Faster 9 [ ] Other – Specify ______
(Ask for persons 21 years old or older) 10a. Does . . . work at same location as 5 years ago?	1 [ ] Yes 2 [ ] No 3 [ ] Not working 5 years ago *(Go to Sec. IV)*
b. Does . . . live at same location as 5 years ago?	1 [ ] Yes 2 [ ] No
c. Compared with the time it took . . . to get to work 5 years ago, is the time to work	1 [ ] About the same as 5 years ago 2 [ ] At least 10 minutes more 3 [ ] At least 10 minutes less

Section IV – DRIVER INFORMATION

(Ask for licensed drivers only) 1. About how many thousands of miles did . . . drive during the past 12 months, including driving as part of work?	1 [ ] None 2 [ ] Under 5,000 3 [ ] 5,000 – 9,999 4 [ ] 10,000 – 14,999 5 [ ] 15,000 – 19,999 6 [ ] 20,000 – 24,999 7 [ ] 25,000 – 29,999 8 [ ] 30,000 and over

Section V – TRAVEL TO SCHOOL

(Ask Sec. V for persons 5–18 years old) Now I would like to ask some questions about transportation to school. 1. Last May was . . . attending or enrolled in school?	1 [ ] Yes 2 [ ] No *(Go to Sec. VI)*
2. Was it a public or private school?	1 [ ] Public 2 [ ] Private
3. What grade was . . . attending?	Grade ______ *(Enter "0" for kindergarten or 1–12, 13+)*
4. About how many miles was it from home to . . .'s school? *(If less than one mile, enter "0")*	Miles
5. About how long did it take . . . to get from home to school?	Minutes
6. How did . . . usually get to school? *(Mark only one box)*	1 [ ] School bus – No charge 2 [ ] Public transportation – No charge } *(Go to Sec. VI)* 3 [ ] School bus – Charge 4 [ ] Public transportation – Charge 5 [ ] Walk, bicycle 6 [ ] Automobile – Driver 7 [ ] Automobile – Passenger 8 [ ] Motorcycle 9 [ ] Other } *(Go to Q. 7)*
7. Was free school bus or free public transportation available?	1 [ ] Yes 2 [ ] No

4

Section VI – TRAVEL DAY REPORT

a. Line No.	b. Age	c. Sex	d. Employment status (C.C. 16a)	e. Occupation (C.C. 16b)	f. Retired Code (C.C. 17)	g. Licensed driver (C.C. 18)
		1 [ ] Male 2 [ ] Female	1 [ ] Yes 2 [ ] No			1 [ ] Yes 2 [ ] No

Now I have some questions about the trips taken on ________. A trip is anytime you went from one place to another by motor vehicle or some form of public transportation. For example, going to work by automobile would be one trip, going to lunch by automobile would be a second trip, returning to work from lunch would be a third trip.

Reference day is from 4:00 a.m. to 3:59 a.m. the following day

1. Did . . . go any place at anytime on __________?

1 [ ] Yes – One or more trips not previously reported (*Fill columns*)
2 [ ] Yes – All previously reported } (*Go to Q. 14a*)
3 [ ] No }

	Trip 1 (5)	Trip 2 (6)	Trip 3 (7)	Trip 4 (8)
2. At what time did . . . start the (1st, next) trip he took on __________?	___:___ 1 [ ] a.m. 2 [ ] p.m.	___:___ 1 [ ] a.m. 2 [ ] p.m.	___:___ 1 [ ] a.m. 2 [ ] p.m.	___:___ 1 [ ] a.m. 2 [ ] p.m.
3. How far is it from where . . . started to where he went?	______ Miles 0 [ ] Less than ½ mile (5 blocks)	______ Miles 0 [ ] Less than ½ mile (5 blocks)	______ Miles 0 [ ] Less than ½ mile (5 blocks)	______ Miles 0 [ ] Less than ½ mile (5 blocks)
4. How long did it take to get there?	1 [ ] 15 min. or less 2 [ ] 16–30 min. 3 [ ] 31–45 min. 4 [ ] 46 min.–1 hr. 5 [ ] Bet. 1 and 2 hrs. 6 [ ] 2 hrs. or more	1 [ ] 15 min. or less 2 [ ] 16–30 min. 3 [ ] 31–45 min. 4 [ ] 46 min.–1 hr. 5 [ ] Bet. 1 and 2 hrs. 6 [ ] 2 hrs. or more	1 [ ] 15 min. or less 2 [ ] 16–30 min. 3 [ ] 31–45 min. 4 [ ] 46 min.–1 hr. 5 [ ] Bet. 1 and 2 hrs. 6 [ ] 2 hrs. or more	1 [ ] 15 min. or less 2 [ ] 16–30 min. 3 [ ] 31–45 min. 4 [ ] 46 min.–1 hr. 5 [ ] Bet. 1 and 2 hrs. 6 [ ] 2 hrs. or more

CODE KEY →
1. To work
2. Business, other than to work
3. Shopping
4. Other family or personal business
5. To school or church
6. To doctor or dentist
7. Vacation
8. Visit friends or relatives
9. Pleasure driving
10. Other social or recreational
11. Other
Return home (reclassification required)

	Trip 1	Trip 2	Trip 3	Trip 4
5. What was the main reason for this trip? (*If "return home" enter the main purpose of the outgoing trip(s), plus "R.H." (Enter one code.)*)	Code	Code	Code	Code
6. In addition to . . . did anyone else living here go on this trip? (*List line numbers of other household members* **5 years old or older** *who went on this trip.*)	0 [ ] No others Line numbers	0 [ ] No others Line numbers	0 [ ] No others Line numbers	0 [ ] No others Line numbers

CODE KEY →
1. School bus
2. Other bus and/or street car
3. Elevated or subway
4. Other train
5. Airplane
6. Taxi
7. Automobile – Driver
8. Automobile – Passenger
9. Motorcycle or motor bike
10. Truck (including pick-up)
11. Other

	Trip 1	Trip 2	Trip 3	Trip 4
7. What means of transportation were used for this trip? (*If more than one, circle major means.*)	Code (*If code 1–5 only go to Q. 13*)	Code (*If code 1–5 only go to Q. 13*)	Code (*If code 1–5 only go to Q. 13*)	Code (*If code 1–5 only go to Q. 13*)
8. Was public transportation for this trip available within 6 blocks (½ mile)?	1 [ ] Yes 2 [ ] No 3 [ ] Don't know	1 [ ] Yes 2 [ ] No 3 [ ] Don't know	1 [ ] Yes 2 [ ] No 3 [ ] Don't know	1 [ ] Yes 2 [ ] No 3 [ ] Don't know
(*Complete questions 9–12 if code 7 or 8 was entered in Q. 7*) **9. What automobile was used?** (*Transcribe automobile number from C.C.*)	Automobile No. or 9 [ ] Not an auto listed on the C.C.	Automobile No. or 9 [ ] Not an auto listed on the C.C.	Automobile No. or 9 [ ] Not an auto listed on the C.C.	Automobile No. or 9 [ ] Not an auto listed on the C.C.
10. Who drove the automobile for this trip?	Line No. ______ 99 [ ] Not a household member	Line No. ______ 99 [ ] Not a household member	Line No. ______ 99 [ ] Not a household member	Line No. ______ 99 [ ] Not a household member
11. Was parking free for this trip?	1 [ ] Yes 2 [ ] No 3 [ ] Did not park 4 [ ] Don't know	1 [ ] Yes 2 [ ] No 3 [ ] Did not park 4 [ ] Don't know	1 [ ] Yes 2 [ ] No 3 [ ] Did not park 4 [ ] Don't know	1 [ ] Yes 2 [ ] No 3 [ ] Did not park 4 [ ] Don't know
12. How many people were in the automobile including the driver? (*Include children under 5 and non-household members.*)	______ Number 0 [ ] Don't know	______ Number 0 [ ] Don't know	______ Number 0 [ ] Don't know	______ Number 0 [ ] Don't know
13. Did . . . go anywhere else on __________?	1 [ ] Yes – One or more trips not recorded (*Go to next column*) 2 [ ] Yes – All trips recorded } Go to Q. 14a 3 [ ] No }	1 [ ] Yes – One or more trips not recorded (*Go to next column*) 2 [ ] Yes – All trips recorded } Go to Q. 14a 3 [ ] No }	1 [ ] Yes – One or more trips not recorded (*Go to next column*) 2 [ ] Yes – All trips recorded } Go to Q. 14a 3 [ ] No }	1 [ ] Yes – One or more trips not recorded (*Go to next column*) 2 [ ] Yes – All trips recorded } Go to Q. 14a 3 [ ] No }

14a. During the 7 days ending (the day before travel day) ________ did . . . return home from a trip after being away from home one or more nights?

1 [ ] Yes – One or more trips not previously reported (*Go to 14b*)
2 [ ] Yes – All trips previously reported } *Fill Sections III–VI for next person 5 years old or older*
3 [ ] No }

b. How many such trips ended during the 7 days? Number ______ (*Go to Sec. VII*)

Sections I-VII of the questionnaire, skipping the sections which did not apply. On subsequent trips to Group I households, those first interviewed in April, only Sections VI and VII, which record specific trips, were completed.

For work trip data, the relevant information is on the control card and in Sections I, III, IV and VI of the questionnaire. The information from Section III, Travel to Work, and Section VI, Travel Day Report, are stored in separate computer files, and it is relatively expensive to merge the two files. The Travel Day Reports form a statistically proper sample, for they record actual trips rather than general patterns of behavior. Given the same sampling process, a collection of different individuals' impressions of their usual trip-to-work, as is done in the Travel to Work section, is likely to be less representative of their aggregate behavior than the collection of records of their actual worktrips in the Travel Day Report. Alternatively, the Travel to Work questionnaire gives more detailed information about access to public transportation and parking costs. Transit access time is one of the most important variables in making a mode-split prediction and changing the cost of parking is potentially one of the most important transportation controls. It was decided to use the Travel to Work section, reasoning that most people do have fairly accurate impression of their trip to work, and that the bias introduced into the sample would not be significant.

The disaggregate work trip mode split models require information about the various modes available for the work trip (access time, linehaul time, cost) and socioeconomic variables (income, auto ownership). To construct a work trip mode split data base, the following information was obtained from the NPTS files: (1) distance to work in miles for the mode actually chosen (Section III, question 4); (2)time to work in minutes for the mode actually chosen (III, 5); (3) usual choice of mode for the work-trip (III,6); (4)distance to public transportation in city blocks (III,7); (5) parking costs, if any (I, 6b); (6) auto ownership (control card, p); and (7) income (control card, p).

Before the Bureau of the Census released the data from the survey, the geographic identity of each PSU was masked. Only the region of the country and certain population characteristics of the PSU are reported. A reasonably successful effort was made to decode the demographic data and identify the urban areas of the various PSU's. Forty-three distinct Standard Metropolitan Statistical Areas (SMSA's) containing 56 PSU's have been identified, including 22 of the 24 SMSA's with more than a million inhabitants -- although some identifications are somewhat tentative. The PSU/SMSA identifications are made in detail for the households identified in April (Group I) and August (Group II). There is not a perfect overlap of PSU/SMSA's identified in each of these two groups. Of the approximately 3000 households in (Group I) 961 are contained in identified PSU's; in (Group II) 971 are in identified PSU's.

The method of locating a PSU is straightforward. The data from the survey is stored in two computer files: one file contains the information about travel patterns; the other contains specific trip records; the socioeconomic data is in both files. Each household and every individual is represented in both files. To decode the PSU locations, the file containing the trip records was used. In the file, six population/location items appear for each household:

1) geographic location (the country is divided into 9 regions);

2) urbanized area (if the household is or is not in an urban area and 1960 size);
3) SMSA (if the household is in the center city of the SMSA, not in the center city but in the SMSA, or not in the SMSA);
4) 1960 SMSA size;
5) type of place (if the place which contains the household is an incorporated area, or the center city of both an urbanized area and/or SMSA, etc.); and
6) 1960 place size.

These six pieces of information are used to determine the possible SMSA's that the household, and thus the PSU, are contained in. If only one SMSA fits the appropriate criteria, this identifies the location of the PSU. The entire PSU is contained in this SMSA, for PSU's do not cross SMSA borders.

Additional assumptions are made to identify some SMSA's. On the basis of the criteria outlined above, Chicago/Detroit and New York City/Philadelphia are two sets of indistinguishable pairs. Chicago was separated from Detroit by noting that Chicago has a rapid transit system and Detroit does not; a distinction between the two cities was made by comparing the choice of mode for the trip records in the appropriate PSU's over the 5 interview months. In Pennsylvania and New Jersey the definition of some types of suburban communities is different than the definition in New York State. This distinction was used to separate the New York City and Philadelphia SMSA's. Other less significant pairs were separated by other, sometimes more tenuous, techniques.

APPENDIX B

REPORT OF INVENTIONS

As a result of the work performed under this contract, improvements in models and procedures for quick evaluation of transportation policy options for urban travel behavior were achieved, as described particularly in Sections 2.1, 2.2, 3.1, and 3.2