TRANSPORTATION RESEARCH RECORD 807

Travel Demand Forecasting and Data Considerations

TRANSPORTATION RESEARCH BOARD

COMMISSION ON SOCIOTECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1981

Transportation Research Record 807
Price $5.80
Edited for TRB by Susan Singer-Bart

modes
1 highway transportation
2 public transit

subject areas
12 planning
13 forecasting

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Travel demand forecasting and data considerations.

(Transportation research record; 807)
1. Travel time (Traffic engineering)–Research–United States. 2. Transportation planning–Research–United States. I. National Research Council (U.S.). Transportation Research Board. II. Series.
TE7.H5 no. 807 380.5s 380.5'22 81-16927
[HE336.T76] AACR2 ISBN 0-309-03222-9 ISSN 0361-1981

Sponsorship of the Papers in This Transportation Research Record

GROUP 1–TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Leon M. Cole, Library of Congress, chairman

Transportation Forecasting Section
George V. Wickstrom, Metropolitan Washington Council of Governments, chairman

Committee on Passenger Travel Demand Forecasting
David S. Gendell, Federal Highway Administration, chairman
Moshe E. Ben-Akiva, Daniel Brand, David J. Dunlap, Robert T. Dunphy, Raymond H. Ellis, Robert E. Gall, Thomas F. Golob, Walter G. Hansen, David T. Hartgen, Thomas J. Hillegass, Joel L. Horowitz, Stephen M. Howe, Lidia P. Kostyniuk, T. Keith Lawton, Steven Richard Lerman, Eugene L. Lessieu, Frederick A. Reid, Martin G. Richards, Gordon W. Schultz, Gordon A. Shunk, Bruce D. Spear, Anthony R. Tomazinis, Edward Weiner, Yacov Zahavi

Transportation Systems and Planning Innovations Section
Richard J. Bouchard, Daniel, Mann, Johnson and Mendenhall, chairman

Committee on Transportation Planning Needs and Requirements of Small and Medium-Sized Communities
C. Michael Walton, University of Texas at Austin, chairman
Edward G. Bates, Jr., James A. Bautz, Michael J. Demetsky, David A. Doctor, Antoine G. Hobeika, R. Keith Jones, Joseph C. McBride, Carl C. McChesney, Monty C. Murphy, Edward S. Neumann, Marion R. Poole, Donald B. Rhodes, Robert R. Roberts, George E. Schoener, Kenneth W. Shiatte, Vergil G. Stover, Montie G. Wade, John H. Waggoner, Bruce B. Wilson, Eugene M. Wilson

James A. Scott, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each report. The organizational units, officers, and members are as of December 31, 1980.

Contents

Authors of the Papers in This Record

Baass, Karsten G., Département de Génie Civil, Ecole Polytechnique, Universite de Montréal, Case Postale 6079, succursale A, Montréal, Québec, Canada H3C 347

Brög, Werner, SOCIALDATA GmbH, Institute for Empirical Social Research, Hans-Grässel-Weg 1, D-8000 Munich, 70, West Germany

Ernst, Ulrich, Urban Institute, 2100 M Street NW, Washington, DC 20037

Gur, Yehuda, Urban Systems, Inc., 301 Chicago Avenue, Oak Park, IL 60302

Hartgen, David T., Transportation Data and Analysis Planning Division, New York State Department of Transportation, 1220 Washington Avenue, Albany, NY 12232

Howe, Stephen M., PAWA, Inc., 7540 LBJ Freeway, Suite 707, Dallas, TX 75251

Khajavi, Shokooh, Facility of Management Studies, University of Toronto, 246 Bloor Street West, Toronto, Ontario, Canada MSS 1V4

Kurth, David L., North Central Texas Council of Governments, P.O. Drawer COG, Arlington, TX 76011

Mekemson, James R., Department of Civil Engineering, Ohio State University, Columbus, OH 43210

Meyburg, Arnim H., School of Civil and Environmental Engineering, Hollister Hall, Cornell University, Ithaca, NY 14853

Sinha, Kumares C., School of Civil Engineering, Purdue University, West Lafayette, IN 47907

Spielberg, Frank, SG Associates, Inc., 4200 Daniels Avenue, Annandale, VA 22003

Ugolik, Wayne R., Economic and Regulatory Analysis Unit, New York State Department of Transportation, 1220 Washington Avenue, State Campus, Albany, NY 12232

Weiner, Edward, Office of the Secretary, U.S. Department of Transportation, 400 7th Street SW, Washington, DC 20590

Design of Zonal Systems for Aggregate Transportation Planning Models

KARSTEN G. BAASS

The zonal system used to represent the spatial properties of urban areas is basic to all aggregate transportation planning models and has an important impact on their results. The current approach to design of zonal systems is essentially empirical and based on experience, local knowledge, and judgment. The results of this approach are probably neither optimal nor consistent among urban areas. For this reason, a systematic computer-based procedure for the generation of aggregated zonal systems was designed. The procedure consists of three parts: data treatment, design of the zonal systems obtained, and, the central part, the grouping algorithm that allows a complete hierarchy of zonal systems to be established. The objective function for this grouping procedure contains two components: the homogeneity of the population inside each aggregated zone and the minimization of intrazonal trips. A number of constraints on the aggregated zones are introduced to ensure that good zonal systems are developed that satisfy the requirements of the transportation planning process. These constraints include an adjacency constraint, a constraint on natural and person-made barriers, a shape constraint, a constraint on equal population and equal number of member zones inside each aggregated zone, and a constraint on total trips. The aggregation methodology was applied to two initial zoning systems of 42 and 522 zones in the city of Montreal. The results of these experiments show that the procedure is able to generate good zoning systems for transportation planning purposes.

One important problem in transportation planning is the choice of an aggregated zonal system based on a detailed disaggregated system of initial spatial units, or zones, as they are available from origin-destination studies or the census. Several reasons make an aggregation necessary; for instance, depending on the planning horizon of the transportation planning study, more or less detail is needed as output of the transportation models. Consequently, the input to the models must also be at different levels of detail. Furthermore, when aggregate models are used, one needs also aggregate input and, in many cases, work at the disaggregate level may not be economically feasible.

Two problems of optimality are involved in the choice of a good aggregated zonal system. The first is to choose the optimum number of aggregated zones that would ensure a reasonable compromise between the cost of the study and the precision of the results, as illustrated in Figure 1. Once the level of aggregation is decided on, the next problem is the choice of the optimum delineation of boundaries at this level of aggregation.

A review of the literature on the problem shows that only general indications are available as to what is to be considered a good zonal system and how to derive such an aggregated zonal system. The planner is left without any guideline to achieve this and has to use knowledge of the region and judgment in designing these zonal systems. On the other hand, studies in geography and regional planning on interaction models by Openshaw (1) and others indicate that the configuration of the zonal system has an impact on the results of the planning models to a much greater extent than previously thought.

There are at least two reasons to believe that the current empirical approach will most probably give rise to nonoptimal zonal systems. First, many criteria should be considered when zones are grouped for transportation planning purposes. These criteria are based on quantitative and qualitative data of such a multitude and diversity that the human mind alone will not be able to process. These criteria can be stated in the following way:

1. Achieve a maximum of homogeneity inside the newly created zones, which is important for the trip generation and the modal split phase of the model sequence;
2. Retain a maximum of interaction between newly established zones or a minimum of intrazonal trips, which is an important requirement for the trip distribution and trip assignment models;
3. Limit the number of trip ends for the newly created spatial entity in order to avoid overloading of the adjacent street network in the assignment phase;
4. Respect physical, political, and historical boundaries as far as they are of importance from a planning point of view;
5. Avoid undesirable shapes of newly created zones;
6. Group only adjacent basic spatial units;
7. Generate only connected zones;
8. Avoid the formation of islands, which means zones that are completely contained in another zone;
9. Obtain a zonal system in which the number of households, population, area, or trips generated and attracted are nearly equal in each zone (the variation with respect to one of these variables should be kept as small as possible); and
10. Base the delineation of the zonal boundaries on the census boundaries.

The second reason is the problem of the delineation of boundaries of aggregated zones, which is a highly combinatorial problem. A simple example in

Figure 1. Problem of choice of optimum zonal boundaries.

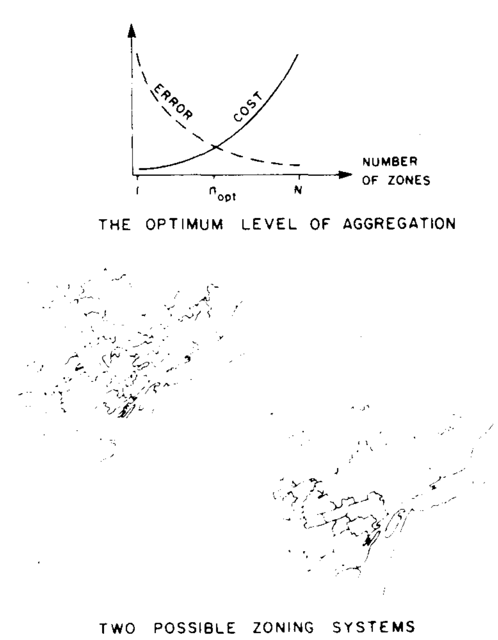
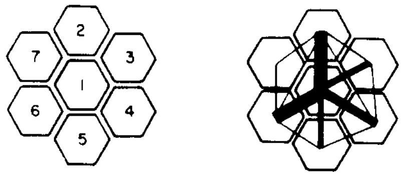


Figure 2. Enumeration of all possible combinations of basic spatial units.



NUMBER OF CLUSTERS	NUMBER OF MEMBERS	NUMBER OF GROUPINGS ADJACENT	ALL
7	1,1,1,1,1,1,1	1	1
6	1,1,1,1,1,2	12	21
5	1,1,1,1,3	21	35
	1,1,1,2,2	33	105
4	1,1,1,4	26	35
	1,1,2,3	66	210
	1,2,2,2	20	105
3	1,1,5	21	21
	1,2,4	60	105
	1,3,3	21	70
	2,2,3	21	105
2	1,6	7	7
	2,5	12	21
	3,4	12	35
Σ		334	877

Figure 3. Flowchart of the general system.

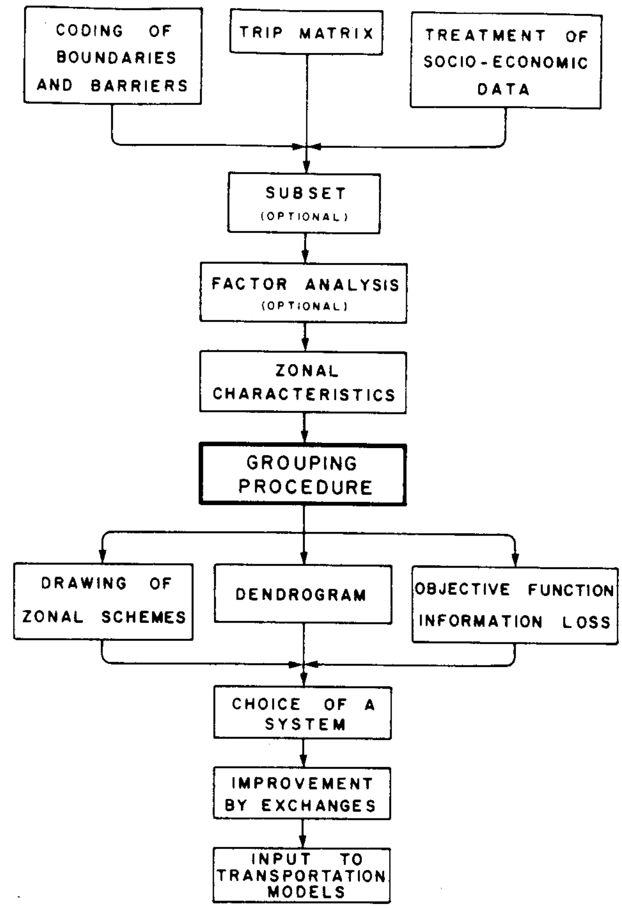


Figure 2 may illustrate this point. The optimum configuration at a six-zone level of aggregation, for example, with respect to the criterion of maximization of interaction would be given by the grouping of zone 1 and zone 7.

The conclusion that arises is that the empirical approach to zonal aggregation will not necessarily yield good zonal systems and that we need a systematic procedure for zone design that considers the important criteria and also a great number of possible combinations of basic spatial units.

GROUPING PROCEDURE

Zonal Design and Aggregation Procedure

Different transportation planning models may necessitate different zonal systems for optimal performance. In the present procedure, however, only one system (a compromise system) is derived because, for practical purposes, the planner has, in most cases, to work with one zonal system. The procedure does not consider, for the time being, the special requirements of a public transport study.

The computer-based methodology, zonal design and aggregation procedure (ZODEAG), was designed not only to derive zonal systems but also to code and treat geographical, trip data, and socioeconomic information to be fed to the grouping algorithm and to analyze the resulting zonal systems in order to help the planner to choose the system that will best suit the purpose. The general flow chart depicted in Figure 3 gives an overview of the components of the ZODEAG package.

The central part of the ZODEAG procedure is the grouping algorithm, which is essentially the algorithm to generate zonal systems, and an objective function that allows us to evaluate the zonal systems.

Grouping Algorithm

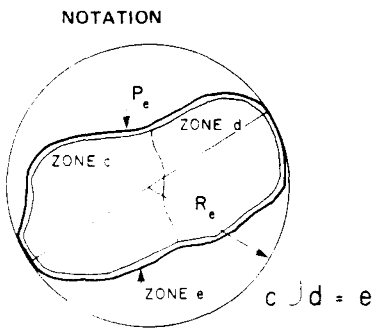
To define the optimum aggregated zonal system one would have to enumerate all possible zonal systems generated by the combination of initial spatial units into aggregated zones, and repeat this on each level of aggregation. Knowledge of all optimum systems at each level of aggregation would allow us to choose the optimum number of zones to be used in the planning study. To solve this problem by complete enumeration of all possible combinations remains impossible, even for today's fast computers.

For this reason a heuristic algorithm for generation of zonal systems was adopted, which is generally credited to Ward (2). Starting with the most disaggregated level of spatial information, one determines the value of the objective function for all possible groupings of two initial spatial units. The grouping that is the most interesting in terms of the objective function and constraints stated is then retained and the two constituting zones will remain merged at further steps of aggregation and will be considered as a new integer and indivisible zone. At the following step of the procedure all possible groupings of two zones are investigated up to the state where only one final zone remains that contains all initial spatial units. This clearly is, as all heuristics, a suboptimal algorithm, but applications of similar procedures in other contexts show that the results are of good quality.

Objective Function and the Constraints

The ZODEAG procedure allows the easy introduction of different objective functions and different constraints. The one described in Equation 1 was tested to some extent. Figure 4 illustrates the no-

Figure 4. Notation used in Equation 1.



tation used in this equation.

Objective Function

$$Z = \min\left(\alpha\left\{\sum_i\left[w_{ic}w_{id}/(w_{ic}+w_{id})\right](\bar{X}_{ic}-\bar{X}_{id})^2/SST\right\} + \beta\left[(t_{cd}+t_{dc})/\sum_{\ell}^{N}\sum_{m}^{N} t_{\ell m}\right]\right) \quad (1)$$

Constraints

Adjacency,

$$a_{cd} \neq 0 \quad (2)$$

Barrier,

$$b_{cd} = 0 \quad (3)$$

Form,

$$R_e P_e / 2 S_e \leq F \quad (4)$$

Number of members,

$$n_c + n_d \leq 2^h \quad (5)$$

where h = 1, 2 ... (ln N/ln 2).

Variation,

$$y_c y_d \leq [(\sum_j^J y_j)^2 / J(J-1)]\,\delta \quad (6)$$

Trip attraction and production,

$$\sum_m^J (t_{cm}+t_{dm}) + \sum_{\ell}^J (t_{\ell c}+t_{\ell d}) - 2(t_{cd}+t_{dc}) \leq MAXT \quad (7)$$

where

R_e = radius of the smallest circumscribing circle of zone e formed by grouping zone c and zone d,
P_e = perimeter of zone e,
S_e = area of zone e,
F = critical value of the form constraint,
y = a critical variable whose variation is to be kept as small as possible,
$\bar{X}_{ic}$ = mean of socioeconomic variable i inside zone c,
w_{ic} = weight of socioeconomic variable i in zone c,
$t_{\ell m}$ = trips between zone ℓ and zone m,
N = total number of initial spatial units,
n_c = number of member zones in an aggregated zone c,
a_{cd} = element of the adjacency matrix A,
b_{cd} = element of the barrier matrix B,
J = number of aggregated zones at the stage of aggregation considered,
SST = total sum of squares in the initial system over all socioeconomic variables i,
MAXT = the limiting value of trips attracted and produced per zone,
α, β = weight coefficients, and
δ = value to be increased when constraint has to be relaxed.

The most important criteria for good aggregated zoning systems in transportation planning are the achievement of homogeneity of population inside the new zones and the conservation of the interaction between zones. The objective function that is chosen to attain these sometimes competing objectives is a composite one. The first part reflects homogeneity, whereby homogeneity is defined as absence of difference between certain characteristic socioeconomic variables or between factors if a factor analysis has been performed previously. Mathematically, this expression represents the increase of total within-zone sum of squares when zone c and zone d are merged. This is clearly an aggregated and weighted measure of homogeneity, and the choice of the socioeconomic variables x(i) has an important impact on the results. The socioeconomic variables used in this study are those generally available from origin-destination surveys and from the census. The sum of the squared and weighted differences is divided by the total sum of squares, so that in the first step this value becomes 0.0 and is 1.0 at the end of the aggregation. The second part of the function considers the interaction between zones in terms of trip exchanges. The objective is to choose that grouping that gives rise to the smallest increase in within-zone trips, since within-zone trips are lost information for modeling purposes.

The two parts of this composite nonlinear objective function are weighted by coefficients to be chosen by the user. This allows one to put more or less importance on one or the other basic aggregation objective. To set both coefficients to the same value would imply that a one percent increase in within-zone sum of squares would be evaluated as being equal to a 1 percent increase in within-zone trips. This assumption was used as a working hypothesis in the first experiments.

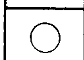
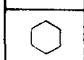
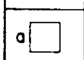
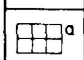
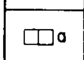
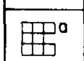
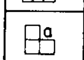
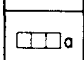
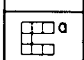
The desirable characteristics of the aggregated zones previously mentioned are obtained by application of several constraints. Most of these constraints are optional, which allows the user to achieve an aggregation in accordance with the study purpose. The only permanent constraint is the adjacency constraint, because it was found that the locational variables alone, as part of the socioeconomic variables, were not sufficient to avoid the formation of disconnected zones. This constraint is verified by inspection of an adjacency matrix that is updated after each merging. Natural, historical, and person-made barriers can be introduced in matrix form and zones on different sides of these barriers can be kept unmerged by applying the barrier constraint. The formation of zones that have unacceptable geometrical shapes and the creation of islands are avoided by using the form constraint. The criterion of form, therefore, is based on a comparison of the shape of the newly created zone and an ideal shape, in this case a circle that is defined as compact. Figure 5 shows the value of the form criterion for different zonal shapes. If F is set to a limiting value of 2.0, for instance, the formation of reasonably compact zones is guaranteed. The actual value of F used may depend on the judgment of the user.

Depending on the degree of similarity of the initial zones, aggregated zones may build up around a nucleus and thus unbalance the newly created zone system, so that heavily aggregated zones coexist with still unmerged units. This can be avoided, if desired, by applying a fourth constraint that guarantees a more uniform distribution with respect to the number of member units inside the aggregated zones. The fifth constraint tries to keep the total variation in the system with respect to a certain variable as low as possible. This may be of interest when it is desirable to achieve, for instance, a more or less equipopular zonal system.

The sixth constraint is applied in order to avoid the formation of aggregated zones that would attract or produce too many trips and thus overload the adjacent street network in the trip-assignment phase. For more detailed explanation and derivation of the formulas see Baass (3).

The aggregation procedure will be interrupted when no more mergings are possible due to the constraint or, if the user so decides, the constraints can be relaxed in the order of preference stipulated and the merging continues up to the last step. Several combinations of constraints and orders of preference in relaxing them were tried, but no formal statement can be advanced about their relative merits for the time being.

Figure 5. Value of form criterion for different zonal shapes.

SHAPE	RADIUS	PERIMETER	AREA	FORM EQUATION (1)
	r	2 π r	π r²	1.0
	r	6 r	2.6 r²	1.2
	0.7 a	4 a	a²	1.4
	1.8 a	10 a	6 a²	1.5
	1.1 a	6 a	2 a²	1.7
	2.1 a	14 a	8 a²	1.9
	1.4 a	8 a	3 a²	1.9
	1.6 a	8 a	3 a²	2.1
	2.1 a	16 a	7 a²	2.4

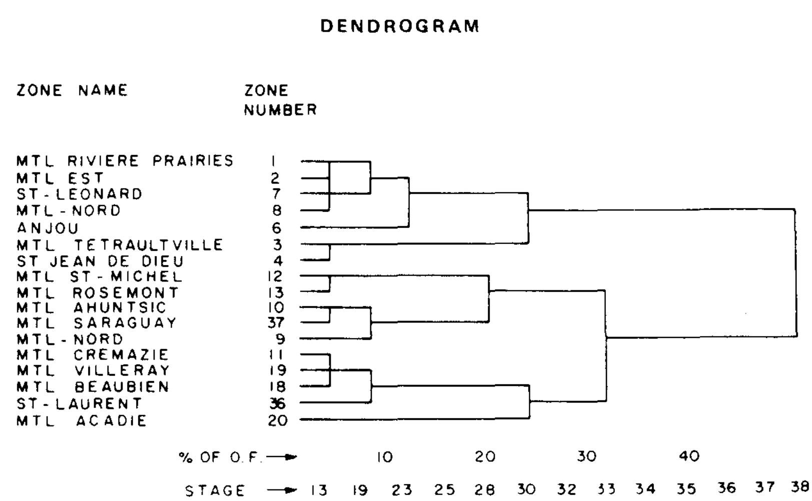
OUTPUT OF THE PROCEDURE

The optimum mergings on each level of aggregation are retained in a merge list. This is the basic input to the programs that allow an analysis of the groupings obtained. A dendrogram built by a subroutine described by Anderberg (4) and represented in Figure 6 is a useful aid in visualizing the hierarchical development of the grouping. Nevertheless, this representation becomes somewhat useless when many zones are studied. In this case a special program that prints detailed subdendrograms can be used. The geographical map of the generated zonal system is more readily understood. A ZODEAG component provides drawings of zoning maps for all or selected steps on a computer plotter. This is useful for visual analysis and is indispensable when different combinations of constraints are tested. An interactive graphic-analysis program was also included in the package, which allows rapid display of the zonal systems produced by the grouping algorithm on a graphic display terminal, and the aggregated zones can be studied in detail by using the zoom feature. Statistics about area and population, pertaining to the new aggregated zones, can also be reviewed by using this system component.

A graph that depicts the increase in objective function in relation to the level of aggregation is of great use when one must decide on the level of aggregation to be used in the planning study and also when different combinations of constraints have to be compared. A bend or a discontinuity in the curve, for instance, would suggest not to use this level of aggregation since the objective function is lower at the left of the bend.

Once a zonal system is chosen, a postoptimization program allows refinements, since it is understood that the heuristic adopted may yield suboptimal results. For the time being this step is not a fully automatic one. The planner chooses aggregated zones that could be improved, for example, for the shape characteristic. In order to achieve this, member zones can be exchanged between aggregated zones, and thus improve the shape. The user introduces into the program the exchanges thought worthwhile for the study and the program executes the exchanges if they do not violate any constraint and if the objective function is thereby improved. All data pertaining to the two aggregated zones touched by the exchange are updated. Experience showed that only very small

Figure 6. Dendrogram of hierarchical development of the grouping.



improvements could be achieved in terms of objective function, thus suggesting that the adopted algorithm achieves satisfactory groupings.

APPLICATION OF THE PROCEDURE

The ZODEAG procedure was applied to the data from the origin-destination survey in Montreal, which contained 522 initial spatial units. Twelve socioeconomic variables were reported in this survey and, in order to reduce the dimensionality of the data matrix, a factor analysis [by using biomedical computer programs (BMDP) (5)] was performed on these data. Subsequent work was done with four factors by retaining 90 percent of the initial variation in the data. These factors were readily interpreted as family status, socioeconomic status, and factors related to the trip-performing characteristics of the population. The work trip matrix was used in all experiments as information about trip interchanges between the zones, since this matrix is most significant and stable. Zonal systems for different trip purposes were demonstrated to be significantly different so that a compromise zonal system would be necessary if one wants to work with several trip purposes.

Figure 7 illustrates the results obtained when no constraints were applied to the 522 initial zones. The generated system is clearly unacceptable because of the zonal shape. Application of the form constraint alone gives acceptable results, as shown in Figure 8. The results can be improved further by introducing the barrier constraint and the constraint on shape, population variation, and number of member zones. The obtained zonal system (Figure 9) comes very close to the system actually used by the planning agency at this level of aggregation.

Graphs that depict the increase in objective function versus stage of aggregation and the increase of both of its components are shown in Figures 10-12. These graphs are useful in judging what kind of constraints to retain and where to stop aggregation in a special planning context. No limiting values can be stated at this point of the research except the rule of thumb formulated by Broadbent (6), which states that, in the case of interaction models, at least 85 percent of the interaction should be between zones. This limiting value would suggest that a fairly advanced state of aggregation would still be acceptable for interaction models, as can be seen in Figure 12. The approximate percentage of within-zone sum of squares that would still retain enough information for the trip generation and modal split models would have to be determined. The graph shown in Figure 10 also illustrates clearly that, by introducing constraints, some of the optimality must be sacrificed.

Figure 7. Stage 519 of aggregation, no constraints.

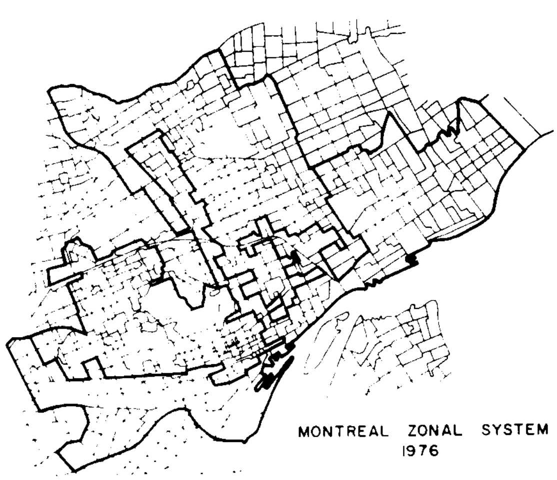


Figure 8. Stage 514 of aggregation, form constraint.

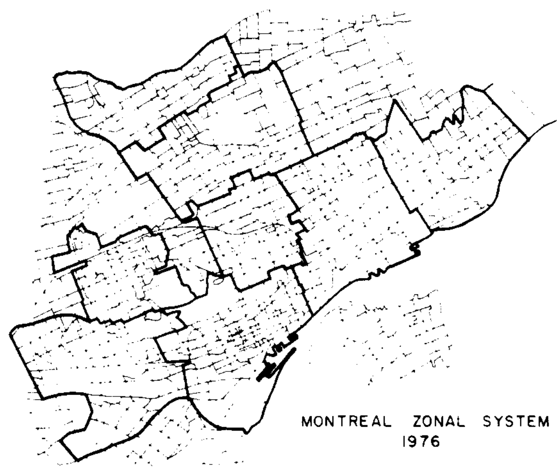


Figure 9. Stage 495 of aggregation, all constraints.

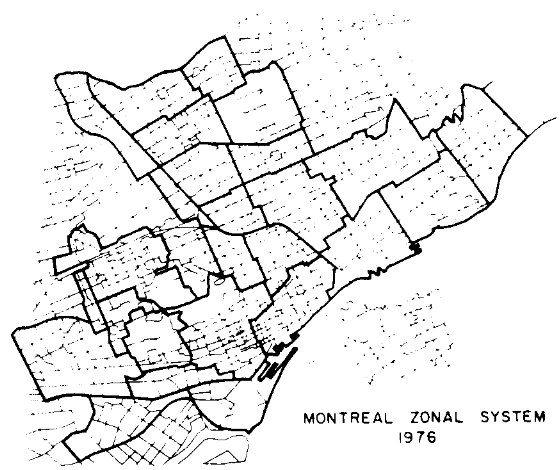


Figure 10. Increase of objective function in relation to stage of aggregation.

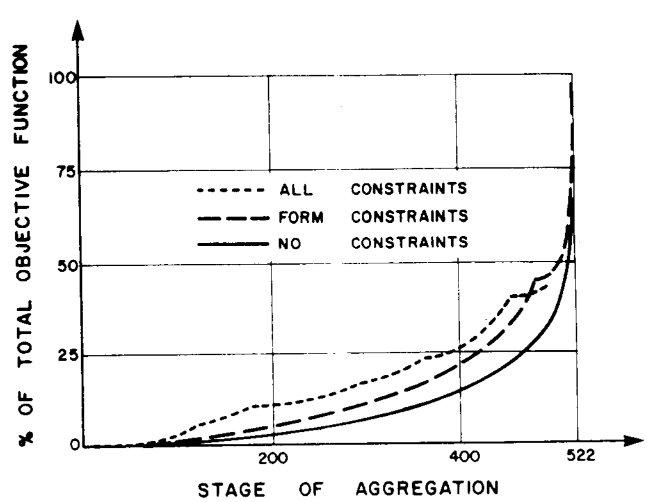


Figure 11. Increase of within-zone sum of squares in relation to stage of aggregation.

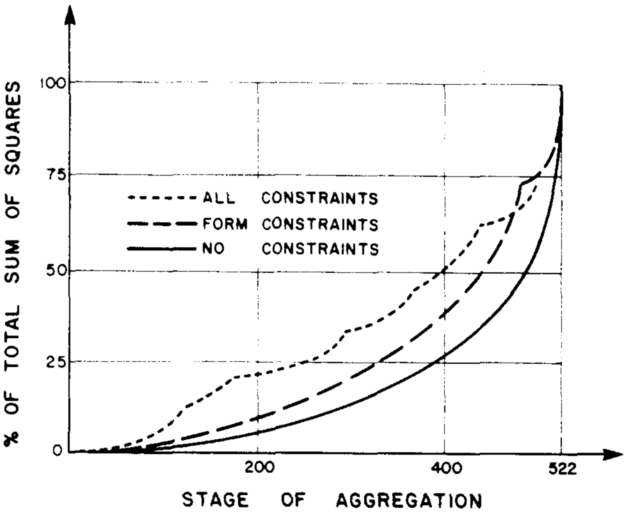
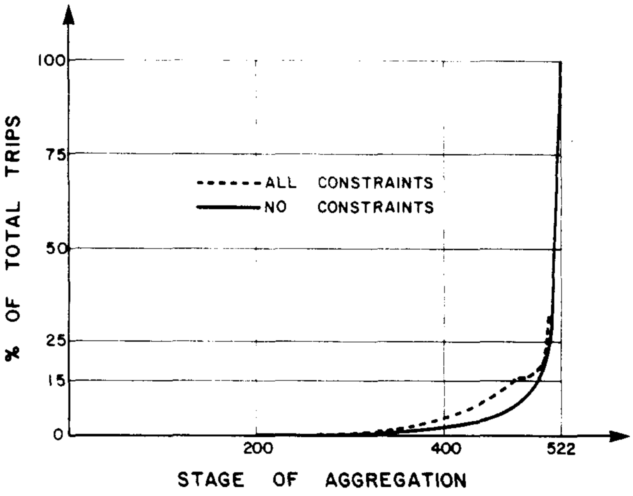


Figure 12. Increase of within-zone trips in relation to stage of aggregation.

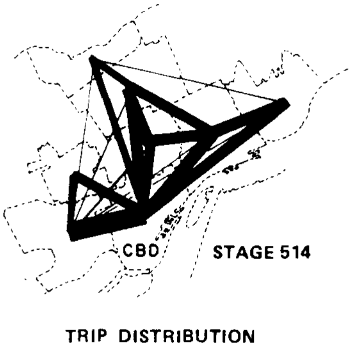


When one depicts the aggregated trip interchanges on the zonal system shown in Figure 8 (as is done in Figure 13), one realizes that 74 percent of the trips made in the 522 initial spatial units are retained in this configuration. Not only are the trips to the central business district (CBD) (zone 1) retained successfully, but also the important flows between the other zones are retained. On the other hand, 80 percent of the variation described by the socioeconomic variables are inside the aggregated zones at this level of aggregation, which makes this system clearly unacceptable for trip generation and modal split purposes.

CONCLUSION

The methodology provides a systematic approach to zonal design and zonal aggregation in transportation planning. The user can derive good zonal systems in accordance with the objectives of the planning study and with a minimum of coding effort and data processing. The procedure is also flexible enough to allow different objective functions and different constraints to be introduced and can actually be applied to at least 600 zones with 15 socioeconomic variables within reasonable computer time and memory limits. Furthermore, once the region is coded, the procedure can easily be applied to any spatial subset of it without recoding the data. Experimentation with actual data showed that the procedure gives good results. But, further research has to be done in order to determine the values of the parameters in the objective function and in the constraints and in order to determine up to what state of aggregation a transportation model still gives satisfactory results. Further research should also be devoted to the problem of the road network aggregation in relation to the zonal aggregation. Also, it becomes more evident that little is known about the interaction between the transportation planning models and the zonal systems and that the ZODEAG procedure constitutes a first step toward a better understanding of these crucial interrelationships.

Figure 13. Trip distribution at stage 514 of aggregation.



REFERENCES

1. S. Openshaw. Optimal Zoning Systems for Spatial Interaction Models. Environment and Planning A, Vol. 9, 1977, pp. 169-184.
2. J.H. Ward. Hierarchical Grouping to Optimize an Objective Function. American Statistical Association Journal, 1963, pp. 236-244.
3. K.G. Baass. A Procedure to Determine a Good Zonal System in the Transportation Planning Process. Ph.D. thesis, Univ. of Waterloo, Ontario, Canada, 1979.
4. M.R. Anderberg. Cluster Analysis for Applications. Academic Press, New York, 1973.
5. BMDP. Biomedical Computer Programs. Univ. of California Press, Berkeley, 1975.
6. G.T.A. Broadbent. Zonal Size and Singly Constrained Interaction Models. Centre for Regional Studies, Working Note CES WH 132, London, England, 1969.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

Optimal Peak-Load Pricing, Investment, and Service Levels on Urban Streets—A Numerical Example

SHOKOOH KHAJAVI

Socially optimal automobile tolls, bus fares, bus service levels, and street capacity can be determined by the use of an integrated peak-load pricing model. The objective of this paper is to develop such a model and to demonstrate the model's applicability and usefulness with regard to its implications for transportation policy. The model that is presented departs from previous ones in that it uses disaggregate travel demand models in order to derive empirically implementable pricing and investment rules for the provision of transit service. Our proposed model, as a whole, is concerned with maximizing the sum of the expected utilities derived by urban street travelers. Numerical results reveal that, given the cost and demand conditions posited, under all but the least-congested travel conditions considered, the travelers' welfare maximizing levels of automobile tolls far exceed those fees now collected by North American gasoline taxes and other automobile user charges. When the imposition of optimal automobile tolls appears impractical, the common practice of providing reserved bus lanes has much to recommend it, especially under the traffic and pricing conditions that prevail during peak hours in most North American cities. Given that automobile tolls are restricted to the gasoline tax, optimal provision of bus service implies mass transit subsidies, at least in peak hours. However, provision of reserved bus lanes would reduce substantially the travelers' welfare losses that result from subsidy reductions.

The objective of this paper is to model travel demand in a manner that derives the pricing and investment rules for a socially optimal provision of urban bus transit. The proposed model is primarily concerned with maximizing the sum of the expected utilities derived by travelers on urban streets. We use disaggregate travel demand models to derive the sum of these utilities. Disaggregate travel demand models are random utility models. They use individuals as the basic decision-making units in order to analyze travel behavior. The disaggregate demand models are used to derive the aggregate demand for different modes of transportation by the travelers on a given urban street.

Given the aggregate demand for different modes of transport, we use the functional forms for the expected utility and the equilibrium demand volumes and travel times of the modes of transport over a given urban street at peak and off-peak hours in order to express our objective function. Next, we use this objective function to derive the socially optimal bus fare, automobile toll, and bus service levels.

FRAMEWORK FOR A SIMPLE TRAFFIC ARTERY MODEL

In this numerical work, we limit our attention to the trips generated and ended along each side of a one-mile stretch of an urban two-way street. Suppose that N people per hour begin and N people per hour terminate trips along each mile of each side of this street. We assume that the origins and destinations of trips are uniformly distributed along each side. There are two modes of travel--automobile or bus. We define each trip to be M miles long. The demand for each mode is a function of the choice maker's socioeconomic characteristics and of the attributes of the alternatives, for example, trip costs and trip times. The travelers take the trip for work, shopping, or social or recreational purposes. We assume that workers take their trips in peak hours and nonworkers take their shopping or social and recreational trips in off-peak hours. Therefore, by assumption, the cross-elasticity between peak and off-peak demand travel is zero. This eases the analytical burden considerably.

Let us assume that the disaggregate demand models have the logit functional form that can be derived as a representation of utility maximization among a discrete set of alternatives under uncertainty. That is, the probability that individual s, selected randomly from the population, will choose mode i given by the multinomial logit model: $Pr_{is} = \exp(U_{is}) / \sum_{j=1}^{I} \exp(U_{js})$, where Pr_{is} is the probability that individual s will choose alternative i from the set of alternatives available, and U_{js} is the utility of alternative j to individual s. Furthermore, we assume that the utility of alternative i to individual s (U_{is}) has the following functional form: $U_{is} = \alpha Z_{is} + \xi_{is}$, where α is a vector of parameters, Z_{is} is a vector of functions of the socioeconomic characteristics of individual s and of the attributes of alternative i, and ξ_{is} is a random variable that represents an unobserved disturbance or error term.

The multinomial disaggregate travel demand models used in this numerical work are those developed by the Metro Travel Commission (MTC)-Cambridge Systematics System for the San Francisco Bay Area (1). These models are based on data from the 1965 surveys by the Bay Area Transportation Study Commission (1,2).

For the demand models that are presented in this paper and for both the peak and off-peak periods, the expected utility of a utility maximizing individual chosen at random from the population is obtained, with an approximation, by applying a combination of the Clark (3) and the Lancaster (4) methods of aggregation (5). Clark presents a set of formulas for the first two moments of the maximum of two normally distributed random variables and the covariance of the maximum with a third normally distributed variable. Then Clark proposes that his solution for the moment be used as an approximation by assuming that the maximum is itself normally distributed. This approximation then permits solution of the expected utility. Lancaster approximates logistic distribution of standard deviation $\pi/\sqrt{3}$ of ξ by a normal distribution of standard deviation $(15/16)\ (\pi/\sqrt{3})$. Following the Lancaster method of approximation, we assume that (a) ξ_i has a normal distribution with mean zero and variance $(\pi^2/6)\ (15/16)^2$ and (b) ξ_i is independent from all ξ_j. We also assume that Z_i, the explanatory variables, are distributed multivariate normal with the row vector of means $E(Z_i)$ and covariance matrix Σ_Z. Then, we have $U_i \sim N[\alpha E(Z_i), \alpha \Sigma_Z \alpha^T + (15/16)^2 \cdot (\pi^2/6)]$. By following the Clark method of aggregation, the expected utility for the peak period (period 1) $[E(\tilde{U}^P)_1]$ can be computed as an integral part of the technique:

$$E(\tilde{U}^P)_1 = \iiint_{AST} \max(U_{A1}{}^P, U_{S1}{}^P, U_{T1}{}^P)\, g_{u_1}{}^P(U_{A1}, U_{S1}, U_{T1}) \cdot dU_{A1}\, dU_{S1}\, dU_{T1} \qquad (1)$$

where

$g_{U^P}(\cdot)$ = the probability density function (pdf) of $U_1{}^P$,
A = drive alone,
S = shared ride, and
T = transit vehicle (bus) (5).

Similarly, for the off-peak period (period 2), the expected utility of a utility maximizing individual chosen at random from the population $[E(\tilde{U}^P)_{s2}]$ can be computed as follows:

$$E(\tilde{U}^P)_{s2} = \int\int_{AT} \max(U_{As2}{}^P, U_{Ts2}{}^P) g_{U_{s2}}{}^P(U_{As2}, U_{Ts2}) dU_{As2} dU_{Ts2} \quad (2)$$

where

$g_{U^P_{s2}}(\cdot)$ = the pdf of U^P_{s2},
A = automobile,
T = transit vehicle (bus), and
s = shopper or traveler who takes a social or recreational trip (4).

Given Equations 1 and 2, to complete the setting of a framework for the models that are presented in this paper, the next section focuses on the measured values of some of the independent variables that appear in the demand models and consequently in Equations 1 and 2.

MEASURED VALUES OF VARIABLES THAT APPEAR IN THE MODEL

Automobile Operating Costs

We assume that the measured automobile operating costs per mile in period t (Ca_t) represent a function of automobile travel time per mile in that period (ta_t). Specifically, the following functional form is selected for this study: $Ca_t = H_0 - H_1/ta_t$. The costs of each automobile trip are shared by A passengers. For an automobile traveler who originates a trip in period t, in a one-mile stretch of the artery total vehicle operating costs are

$$ACOST_t = (MCa_t + MFa + g)/A \quad (3)$$

where Fa is the automobile toll per mile and g is a fixed cost that is independent of the length of the trip.

Automobile Travel Time

Let us assume that buses and automobiles distribute themselves uniformly across the width of each side of each mile of the artery. Under this assumption, the measured travel time per automobile trip on each side of the artery, in period t, is a function of that side's volume to capacity ratio. Specifically, we select the following functional form for this study:

$$ta_t = ta_0 \{1 + a[(\delta X_t + M\,Na_t/A)/K]^b\} \quad (4)$$

where

ta_0 = travel time per mile at zero flow,
a and b = constants,
X_t = number of buses per hour for period t,
A = the average number of passengers per car,
M = trip length,
K = capacity of the road,
Na_t = number of automobile travelers who originate in a one-mile stretch of the road at period t, and
δ = the bus's automobile congestion equivalent.

MNa_t enters this expression because, in period t, any given point on the road is passed by automobile travelers who originated in each of the M miles that preceded that point. Therefore, the one-way in-vehicle travel time for an automobile traveler who takes trips in period t is equal to

$$ATIMET_t = M(ta_0 \{1 + a[(\delta X_t + M\,Na_t/A)/K]^b\}) \quad (5)$$

Bus Travel Time

The measured travel time per bus trip has two components. First, when not engaged in stopping and starting maneuvers, a bus is assumed to travel at the same speed as an automobile (i.e., to require ta_t min/mile in period t). In addition, in each route mile, Nt_t travelers board and Nt_t leave X_t buses at Y or fewer stops. (Y is the number of uniformly spaced bus stops per mile.) Hence, $\mu_{ST} = 2Nt_t/X_tY$ is the average number of passengers that board or leave one bus at any one stop. Suppose that bus travelers make their decisions as to when to travel independently. Then, the probability that a total of n travelers will board and alight from any one bus at any one stop is given by the Poisson distribution with parameter μ_{St}. That is, $Pr(n) = \exp(-\mu_{St})\mu_{St}{}^n/n!$. The probability that a given stop will be made, then, is $1 - \exp(-\mu_{St})$ (i.e., 1 - the probability that no one will have that stop as either origin or destination). The expected number of stops per mile is Y times this fraction. Therefore, the expected time required to travel one mile is equal to

$$tt_t = ta_t + 2N_t\epsilon/X_t + \S\, Y[1 - \exp(-\mu_{st})] \quad (6)$$

where ϵ is the time required to board or unload a passenger once a bus has stopped and § is the amount by which the time required to traverse a route segment is increased by each stop and start maneuver. Therefore, the one-way in-vehicle travel time for a bus rider who takes trips in period t is equal to

$$TIME\,T_t = M\{ta_t + 2Nt_t\epsilon/X_t + \S\, Y[1 - \exp(-\mu_{st})]\} \quad (7)$$

Access Time for Bus Riders

Most bus travelers neither live nor have their destinations on the traffic arteries traversed by the bus they use. Rather, a typical bus rider must walk to the route from an origin and from the route to a destination. Once a traveler reaches the artery, he or she must walk to the nearest bus stop. If origins and destinations are uniformly distributed between stops and there are Y uniformly spaced bus stops per mile, the one-way walking time for the traveler who uses the bus will be ht + 60/2SY min, where ht is the walking time (min) for a typical bus rider to the route from an origin and from the route to a destination and S is the average bus passenger's walking speed (5).

If the average length of a bus passenger's wait at a stop is a fraction (β) of the headway between buses ($1/X_t$), the average measured waiting time in period t is $60\beta/X_t$ min. Therefore, the average total of measured one-way access time by bus in period t is equal to (ht + 60/2SY) + ($60\beta/X_t$) min.

MODEL 1: THE BASIC MODEL

Let us suppose that government's objective is to maximize the total net benefits received by travelers, over a period of a day, from traveling on the artery. That is, given the assumptions and definitions of the last two sections, the objective function can be written as follows:

$$\text{Max } W = U - Ct - Cr \quad (8)$$

where

- U = total benefits received by travelers over a period of a day from traveling on a road (Equations 1 and 2),
- Ct = total daily operating costs of the bus company, and
- Cr = total daily rental costs of that road.

Given the equilibrium flow pattern and the travel demand volume, this model finds those values of X_t, Ft_t, Fa_t, and K that satisfy all the Kuhn-Tucker conditions for maximizing W; that is, we compute those values of X_t, Ft_t, Fa_t, and K that satisfy $\partial W/\partial X_t = \partial W/\partial Ft_t = \partial W/\partial Fa_t = \partial W/\partial K = 0$ for t = 1, 2. The model's statement about street capacity ($\partial W/\partial K = 0$) implies that arterial street capacity is expanded to the point where the value of the marginal product of the last unit of capacity produced just equals the marginal costs of providing that unit of arterial street capacity. The model's statements about bus service at each period of time ($\partial W/\partial X_t = 0$) imply that bus service at each period of time is provided to the point where the marginal benefit of the last unit of a service produced, at that period of time, just equals the marginal costs of providing that unit of service.

We choose the socially optimal automobile tolls at period t (Fa_t) equal to the congestion charge at that period [i.e., the losses an additional automobile traveler's trip imposes on (a) the existing automobile travelers (by increasing the time and vehicle operating cost of their trips), (b) bus travelers, and (c) the bus company].

Similar considerations apply to the socially optimal bus fare. We choose the price of bus in period t (Ft_t) equal to the congestion charge at that period [i.e., the losses an additional bus traveler's trip imposes on (a) the passengers already aboard the bus on which he or she travels, (b) automobile travelers, and (c) the bus company].

MODEL 2: OPTIMALITY CONDITIONS UNDER PRICING CONSTRAINTS

Mohring argues that the provision of bus service involves economies of scale (6). This argument is based on the following example: Suppose that a bus company responds to a doubling of demand along route N by doubling the number of buses that serve the route (X). Given the modal split (Nt/N), if road capacity is allowed to expand such that the arterial volume to capacity ratio remains unchanged, the average number of passengers per mile served by a bus, and hence the average travel time per trip and bus company costs per passenger served, will all remain unchanged. However, such an expansion of road capacity would cut headway between buses in half, and thereby cut the costs of waiting time per passenger in half. Nevertheless, in the framework of our simple model, the argument is not such a straightforward one. In fact, for the demand volumes that prevail under equilibrium conditions, the average number of passengers per mile served by a bus, and hence the average travel time per trip and bus company costs per passenger served, may all change. The magnitude and direction of these changes, which are sensitive to the parameter values in the model, determine whether or not the provision of bus service involves economies of scale.

In our numerical analysis, in some cases road capacity is held fixed. An increase in total travel, therefore, leads to an increase in arterial volume to capacity ratios. The resulting reduction in travel speeds tends to offset the increasing returns aspect (if there is any) of the bus system. For the combinations of parameter values studied in our numerical analysis, these decreasing returns aspects of the system largely outweigh its increasing returns features. That is why we observe that, in some cases, optimal bus fares generate revenues in excess of bus system costs.

However, for the congestion levels (ratio of trip volume to arterial capacity) that seem typical of urban areas at peak hours, model 1 yields socially optimal automobile tolls far higher than the roughly 3-9 cents/mile implied by the gasoline taxes and other excises imposed on automobile travel in North American cities. Therefore, if existing tolls on automobile travel are taken to be incapable of alteration, truly astronomical subsidies for buses would be required to maximize the net benefits of all trips.

These considerations suggest the desirability of adding the following pricing constraint to model 1: Wrong (i.e., inefficiently low) tolls are charged for automobile travel and bus operations must break even.

MODEL 3: OPTIMALITY CONDITIONS UNDER THE ASSUMPTION OF RESERVED BUS LANES

This model differs from the previous ones only in that it allows a fraction (η) of the total capacity of each side of the artery to be reserved for buses, and the remaining capacity $[(1 - \eta)K]$ is allocated to automobiles. In our numerical work, we select the values of η that comprise an integral number of lanes of the artery. For example, in the framework of our simple model, we can allocate zero, one, or two lanes to buses (i.e., η is given the value of 0, 0.5, or 1). Under this assumption, the time required for a one-mile automobile trip is

$$ta_t = ta_0 \{1 + a[(M\,Na_t/A)/(1 - \eta)K]^b\} \quad (9)$$

and in the expression for tt_t, given by Equation 6, ta_t is replaced by

$$ta_0 [1 + a(\delta X_t/\eta K)^b] \quad (10)$$

To allow reserved bus lanes is, in effect, to allow the division of an artery into two separate rights-of-way and to permit allocation of the artery's total capacity so as to equalize the value of its marginal benefit on these two rights-of-way. The policy of allowing reserved bus lanes is based on the argument that the provision of bus service involves economies of scale. The validity of this argument constitutes a strong reason for permitting reserved bus lanes. Then, the equality of the marginal products of capacity calls for a lower volume to capacity ratio on the bus than on the automobile right-of-way. The resulting increase in travel speeds for buses in comparison with travel speeds for automobiles tends to encourage bus riding and, as a result, to increase the social benefits derived by travelers from traveling on the artery.

MODEL 4: OPTIMALITY CONDITIONS UNDER THE ASSUMPTION OF RESERVED BUS LANES AND PRICING CONSTRAINTS

This model differs from model 2 in the same way that

Table 1. Parameter value combinations studied for parameters that vary between peak and off-peak periods in numerical analysis.

	Peak			Off-Peak	
Parameter	Drive Alone	Shared Ride	Bus	Automobile	Bus
A = passengers per automobile	1	2.5		2	
B = hourly bus costs ($)			15		9
β = waiting time between bus headways			_a		1/2
ha = automobile access time (min)	4.3	14.3		3.3	
g = fixed costs of automobile trip (cents)	12.5	14.2		5.3	
ht = bus access time (min)			18.5		5.6

[a]During peak periods, the waiting time between buses is ½(1 + 0.000 001 X^2).

model 3 differs from model 1. That is, model 2 involves the assumption that automobiles and buses are uniformly distributed across the artery, and model 4 allows for reserved bus lanes.

In comparison with models 1 and 3, in model 4 as well as model 2 we maximize the total net benefits to the travelers from traveling on the artery given that bus operations must break even for the case that wrong tolls are charged for automobile travel.

BENEFIT/COST IMPLICATIONS OF THE MODELS

Pseudoempirical analysis of the models described is undertaken for the parameter value combinations indicated below for parameters that have the same values in all runs and in Table 1 (5) for parameters that vary.

- M = trip length (miles) = 5,
- Y = allowable stops per mile = 8,
- K = capacity of artery (vehicles/h) = 625 number of lanes (WID),
- δ = Congestion equivalent of a bus = 3 automobiles,
- ta_0 = 2.0 min/mile,
- a = 2.62 min/mile,
- b = 5,
- Ca = cents per automobile mile = 10.565 - 5.706(1/ta),
- § = added time per bus stop made = 0.3 min,
- ε = passenger boarding or unloading time = 0.03 min, and

Cr = costs of arterial street expansion (cents) = $(100/720)\{[0.06/(1 - e^{-2.1})] + 0.06(0.342)\}WID^{1.0304}\ [\exp(12.767) + \{2917\ WID + 0.000\ 45[360(5w_1N_1 + 5w_2N_2)]\}$

The net travelers' welfare losses per trip shown in Tables 2 and 3 reflect the net welfare-maximizing levels for travelers of bus service, bus fare, automobile tolls, and road capacity, given the parameter values shown in the tables. However, these optimization problems are solved under the following two constraints: (a) the decision variables cannot have a negative value and (b) the levels of bus service and road capacity cannot be noninteger.

Benefit/Cost Implications of Model 1

Tables 2 and 3 reveal that the scale economies associated with bus operation yield only a modest reduction in system net travelers' welfare losses with increases in the scale of the system. Naturally enough, the net travelers' welfare maximizing share of bus trips (Nt/N) declines with reductions in MN/K, the volume to capacity ratio that would prevail if all trips were taken in single-occupant automobiles. With MN/K equal to 3.2, the maximization of net travelers' welfare would call for 59.4 percent of all trips to be taken by bus (Nt/N = 59.4) and 18.7 percent to be taken by shared-ride (Ns/N = 18.7) but, depending on system scale, a share of bus trips (Nt/N) in the 28.9 to 33.0 per-

Table 2. Welfare losses per trip as a function of congestion level and travel demand for peak cost conditions.

Travel Rate	Number of Lanes	Welfare Losses per Trip	Modal Split: Drive Alone	Modal Split: Shared Ride	Modal Split: Bus	Optimal Automobile Tolls (cents/mile)	Value of Marginal Product of Capacity[a] (cents)
400	4	1.74	0.775	0.225	0.000	3	0.2
400	3	2.08	0.773	0.227	0.000	3	1.0
400	2	3.17	0.420	0.250	0.330	63	7.9
800	4	3.00	0.488	0.223	0.280	40	8.1
800	2	3.36	0.219	0.187	0.594	93	26.8

Note: Length of peak period = 2 h, peak-hour bus costs = $15.
[a]Cost per number of cars/h along a given mile.

Table 3. Welfare losses per trip as a function of congestion level and travel demand for off-peak cost conditions.

Travel Rate	Number of Lanes	Welfare Losses per Trip	Modal Split: Shopping: Automobile	Modal Split: Shopping: Bus	Modal Split: Social and Recreational: Automobile	Modal Split: Social and Recreational: Bus	Optimal Automobile Tolls (cents/mile)	Value of Marginal Product of Capacity[a] (cents)
200	4	1.75	1.00	0.00	1.00	0.00	2	0.0
200	3	1.76	1.00	0.00	1.00	0.00	2	0.0
200	2	1.79	1.00	0.00	1.00	0.00	2	0.3
400	4	1.79	1.00	0.00	1.00	0.00	2	0.3
400	2	1.86	0.778	0.222	0.496	0.504	9	1.5

Note: Length of off-peak period 10 h and off-peak hour bus costs = $9.
[a]Cost per number of cars/h along a given mile.

cent range would be called for with an MN/K value of 1.6. When MN/K values are high, the provision of an optimal level of bus service would reduce net welfare losses for travelers on the system below the levels that would result with only automobile travel. When MN/K = 1.6, however, the benefits of bus service are modest. When MN/K $\leq$ 1.2, the constrained optimal level of bus service would result in net welfare losses for travelers either identical to or higher than welfare losses with only automobile travel.

The potential benefit of an optimal level of bus service when the level of congestion is high on a traffic artery could reflect the inherent superiority of mass transit vehicles for urban travel. On the other hand, these benefits could result from an inefficiently low level of capacity on that artery. Which of these possibilities is true for a particular case depends on just how costly it would be to add to the capacity of the artery. Tables 4 and 5 show that, when the congestion level is high on a traffic artery and specifically when automobile tolls are constrained to the level of gasoline taxes, the potential benefits of an optimal level of bus service result mainly from an inefficiently low level of capacity on the artery. The optimal level of arterial capacity when MN/K values are high (i.e., MN/K = 3.2) is about four 10-ft lanes. When MN/K = 3.2, an optimal level of arterial capacity would result in welfare losses of 10.7 percent less than those with the present arterial capacity levels.

However, these results are sensitive to changes in demand and, in particular, to the cost conditions that prevail in Table 4. For example, as the constant parameter (b) of the automobile travel time function (Equation 4) increases, the sensitivity with respect to a high level of MN/K increases and the sensitivity with respect to a low level of MN/K decreases. Then for a different value of b, our results may be different. In addition, in Table 4 we estimate the costs of arterial street expansion given a 6 percent interest rate and 35 years of effective lifetime for the road. As Table 5 shows, a change in the interest rate or the effective lifetime of the road changes the costs of street expansion and, as a result, the optimal level of capacity.

Benefit/Cost Implications of Model 3

Tables 6 and 7 deal with the potential benefits of reserved bus lanes when socially optimal automobile tolls and bus fares are charged and when automobile tolls are restricted to those implicit in current gasoline taxes (e.g., 45 cents/five-mile trip in peak hours and 30 cents/five-mile trip in off-peak hours) and deficit constraints.

Parameter values for Table 6 are those that seem most representative of peak-hour travel (MN/K = 1.6); off-peak values are the basis for Table 7 (MN/K = 0.8).

Perhaps the most important generalization suggested by these tables is that, even if socially optimal bus fares and automobile tolls could be charged, the reserved bus lane would result in some benefits (i.e., a reduction in welfare losses of about 31.0 percent). An optimum allocation of capacity would result in a 31.0 percent reduction in welfare losses under peak conditions (Table 6), and

Table 4. Comparison of system operation with and without reserved bus lanes under different pricing and financial constraints: base peak-hour case.

Constraint	Welfare Losses per Trip		Optimal Bus Fare (cents)		Buses per Hour		Bus Operating Cost per Passenger (cents)		Bus Share of Capacity (%)		Marginal Benefit per Subsidy Dollar		Value of Marginal Product of Capacity[a] (cents)
	Without	With	Without	With	Without	With	Without	With	Without	With	Without	With	Without
Social welfare maximizing	3.74	2.09	0	25.0	23	58	22.0	25.0	6.1	100			39.9
Pricing constraint													
Automobile toll = $0.09/mile	67.47	4.35	60	3.5	11	8	4187.6	3.5	1.1	50			3208.0
Automobile toll = $0.09/mile and bus operating subsidies = $0.0	99.43				0				0		7.2		624.7
Automobile toll = $0.12/mile	64.39	4.88	60	0	11	3	3808.5	2.0	1.1	50	7.9	14.2	3050.9
Automobile toll = $0.12/mile and bus operating subsidies = $0.0	97.25	5.02		1.4	0	3		1.4	0	50			610.4

Note: Length of peak period = 2 h, number of lanes = 2, travel rate = 800, and peak-hour bus costs = $40.

[a]Cost per number of cars/h along a given mile.

Table 5. Average daily capacity costs per mile of the arterial street as a function of number of lanes, interest rate, flow of vehicle trips over the street, and effective lifetime of the street.

Interest Rate (r) and Effective Lifetime (L)	Average Daily Capacity Cost per Mile (cents)					
	Flow of Vehicles = 400 peak and 200 nonpeak			Flow of Vehicles = 800 peak and 400 nonpeak		
	Two Lanes	Three Lanes	Four Lanes	Two Lanes	Three Lanes	Four Lanes
r = 6 percent, L = 35 years	9 964.439	14 954.143	19 991.593	10 279.439	15 269.143	20 306.593
r = 12 percent, L = 35 years	54 050.59	81 906.248	110 047.79	54 365.59	82 221.248	110 362.79
r = 6 percent, L = 30 years	10 313.775	15 484.669	20 705.194	10 628.775	15 799.669	21 020.194
r = 12 percent, L = 30 years	54 204.133	82 190.013	110 361.43	54 519.133	82 505.013	110 676.43
r = 6 percent, L = 40 years	9 727.454	14 594.243	19 507.497	10 042.454	14 909.243	19 822.497
r = 12 percent, L = 40 years	53 967.948	81 780.743	109 878.97	54 282.948	82 095.743	110 193.97

Table 6. Comparison of system operation with and without reserved bus lanes under different pricing and financial constraints: base peak-hour case.

Constraint	Welfare Losses per Trip		Optimal Bus Fare (cents)		Buses per Hour		Bus Operating Cost per Passenger (cents)		Bus Share of Capacity (%)		Marginal Benefit per Subsidy Dollar		Value of Marginal Product of Capacity Without (cents)[a]
	Without	With	Without	With	Without	With	Without	With	Without	With	Without	With	
Social welfare maximizing	3.46	2.35	20	0	11	23	14.7	12.7	1.7	25			4.8
Pricing constraint													
Automobile toll = $0.09/mile	3.76	2.40	0	0	11	23	115.2	17.2	1.1	25	7.0	3.3	42.6
Automobile toll = $0.09/mile and bus operating subsidies = $0.0	4.85	2.65		10	0	23		10.4	0	75			9.8
Automobile toll = $0.12/mile	3.77	2.44	10	0	11	23	112.1	17.1	1.1	25	7.7	2.8	41.5
Automobile toll = $0.12/mile and bus operating subsidies = $0.0	4.84	2.66		10	0	23		10.0	0	75			9.6

Note: Length of peak period = 2 h, number of lanes = 4, travel rate = 800, and peak-hour bus costs = $40.
[a]Cost per number of cars/h along a given mile.

Table 7. Comparison of system operation with and without reserved bus lanes under different pricing and financial constraints: base off-peak-hour case.

Constraint	Welfare Losses per Trip		Optimal Bus Fare (cents)		Buses per Hour		Bus Operating Cost per Passenger (cents)		Bus Share of Capacity (%)		Value of Marginal Product of Capacity Without (cents)[a]
	Without	With	Without	With	Without	With	Without	With	Without	With	
Social welfare maximizing	2.29	2.29			0	0	0	0	0	0	0.3
Pricing constraint											
Automobile toll = $0.06/mile	2.49	2.48	0	0	2	2	1.1	1.1	1.8	50	0.0
Automobile toll = $0.08/mile and bus operating subsidies = $0.0	2.54	2.54			0	0			0	0	0.3
Automobile toll = $0.08/mile	2.58	2.56	0	0	2	2	1.0	1.1	2.1	50	0.0

Note: Length of off-peak period = 10 h, number of lanes = 4, travel rate = 400, and off-peak hour bus costs = $25.
[a]Cost per number of cars/h along a given mile.

Table 8. Comparison of system operation with and without reserved bus lanes under different pricing and financial constraints: base peak-hour case.

Constraint	Welfare Losses per Trip		Optimal Bus Fare (cents)		Buses per Hour		Bus Operating Cost per Passenger (cents)		Bus Share of Capacity (%)		Modal Split Without Reserved Bus Lanes			Marginal Benefit per Subsidy Dollar Without
	Without	With	Without	With	Without	With	Without	With	Without	With	Drive Alone	Shared Ride	Bus	
Social welfare maximizing	3.21	2.45	30	25.0	15	29	29.7	25.0	5.0	100	0.325	0.253	0.422	
Pricing constraint														
Automobile toll = $0.09/mile	3.98		0		11		213.0		2.2		0.655	0.210	0.135	3.0
Automobile toll = $0.09/mile and bus operating subsidies = $0.0	4.85				0				0		0.757	0.243	0.00	
Automobile toll = $0.12/mile	4.06		0		11		219.8		2.2		0.640	0.203	0.157	2.4
Automobile toll = $0.12/mile and bus operating subsidies = $0.0	4.84				0				0		0.750	0.250	0.00	

Note: Length of peak period = 2 h, number of lanes = 2, travel rate = 400, and peak-hour bus costs = $40.

a 0 percent reduction in off-peak hours (Table 7). Achievement of the optimal volume to capacity ratio in peak hours requires the allocation of one lane out of four to buses, given socially optimal automobile tolls, bus fares, and bus service. In contrast, achievement of the optimal volume to capacity ratio in off-peak hours requires no allocation of lanes to buses.

As Table 8 shows, for a two-lane artery that has the same congestion level as prevails in Table 6 (i.e., MN/K = 1.6), the optimum allocation of capacity requires allocation of both lanes to buses, and this would result in a decrease in welfare losses of about 23.9 percent under peak conditions. However, the allocation of one lane out of two to buses would result in an increase in social costs of about 59 percent.

These findings are subject to a very important qualification. As Tables 2-3 indicate, under all but the least-congested travel conditions, the socially optimum automobile tolls are far higher than those implicit in gasoline taxes and other automobile user charges. A proposal to levy the tolls of 40-93 cents per vehicle mile listed in Table 2 would almost certainly generate overwhelming opposition from the public and politicians (especially in North America). If attention is restricted to automobile tolls in the neighborhood of those currently charged in North American cities (for example, 45 cents/five-mile automobile trip in peak hours and 30 cents/five-mile trip in off-peak hours), reserved bus lanes appear to have considerable merit, at least during periods of high traffic flow (peak hours). That is, reserved bus lanes would result in a reduction of 36.1-45.3 percent in social costs when automobile tolls are restricted to those implicit in current gasoline taxes, for example, 45 cents/five-mile automobile trip at peak hours.

Under the conditions given in Table 6, given that automobile tolls are set equal to gasoline taxes, the maximization of net travelers' welfare would require free bus service and would result in average welfare losses per trip of $3.76 in the absence of reserved bus lanes, 8.7 percent greater than the losses achievable with socially optimal tolls. If this pricing constraint is accompanied by the allocation of one lane out of four to buses, however, minimum average welfare losses per trip work out to $2.4--2.2 percent greater than the welfare losses attainable with reserved bus lanes when socially optimal tolls are charged.

As Table 6 shows, the maximization of net travelers' welfare requires free bus service when automobile tolls are set equal to gasoline taxes. However, for the following reasons opposition would also be likely to the provision of free bus service:

1. The setting of a zero bus fare may indeed increase the elasticity of demand for trips and
2. The optimal cost subsidies required under Table 6 conditions to maximize net travelers' welfare when automobile tolls are constrained to the level of gasoline taxes are $124.4/number of cars/h along a given mile of bus service without reserved bus lanes and $62.1 with reserved lanes. Although considerably lower than the deficits that would be required to match reductions in automobile tolls and bus fares, these subsidies, in the absence of reserved bus lanes, are still substantial. In the following section we study a deficit constraint model. This model is designed to test the welfare loss implications of a zero percent subsidy level.

Benefit/Cost Implications of Model 2

In the absence of reserved bus lanes, the elimination of bus operating subsidies would result in substantial increases in welfare losses. As Table 6 indicates, the increase in average welfare losses per trip that would result from lowering bus operating subsidies from $124.4 to 0 per number of cars/h along a given mile (i.e., from the value required to maximize net travelers' welfare subject to the gasoline tax toll constraint to zero) produces a marginal benefit per subsidy dollar of $0.01 (i.e., a marginal benefit of $7.05, when benefit is measured in terms of the decrease in total welfare losses of all trips).

Table 8 shows the effect of eliminating bus operating subsidies on welfare losses for a different level of travel demand. If we compare Tables 6 and 8, we see that, when we cut both travel demand and capacity by half, the elimination of bus operating subsidies would result in moderate welfare loss increases. (This result is consistent with the statement that provision of bus service involves economies of scale.) As Table 8 shows, the increase in average welfare losses per trip that would result from lowering subsidies from $115.0 to 0 per number of cars/h along a given mile works out to a marginal benefit per subsidy dollar of $0.01 (i.e., a marginal benefit of $3.00, when benefit is measured in terms of the decrease in total welfare losses of all trips).

Table 6 also presents the effects of an increase in the level of gasoline tax tolls on the operation of the system. It shows that increasing the gasoline tax tolls from 45 to 60 cents/five-mile automobile trip at peak hours would result in a slight reduction (0.3 percent) in welfare losses per trip. The maximization of net travelers' welfare subject to a gasoline tax toll constraint of 60 cents/five-mile trip does not require free bus service. But as Table 6 shows, the elimination of bus operating subsidies would result in substantial increases in welfare losses. When we maximize net travelers' welfare subject to the 60-cent gasoline tax toll constraint, the elimination of subsidies would result in an increase of $854.4 in total welfare losses of all trips. This increase works out to a marginal benefit per subsidy dollar of $7.7.

Comparison of Tables 4 and 6 shows that, for a higher level of congestion, increasing the gasoline tax tolls from 45 to 60 cents/five-mile automobile trip at peak hours would result in a greater reduction (2.3-4.5 percent) in welfare losses per trip.

Benefit/Cost Implications of Model 4

Subsidy restriction is less costly when reserved bus lanes are permitted. As Table 6 shows, elimination of subsidies would result in moderate increases in welfare losses. Lowering of the bus operating subsidy from $62.1 to 0 per number of cars/h along a given mile adds 25.4 cents (10.5 percent) to the average welfare losses of a trip. That is, the increase in the total welfare losses of all trips that would result from lowering subsidies from the value required to maximize net travelers' welfare subject to the gasoline tax toll constraint ($62.1) to 0 per number of cars/h along a given mile works out to a marginal benefit per subsidy dollar of $3.3.

Under both peak and off-peak conditions, when the congestion level is low, the provision of the bus system not only does not decrease but may even increase the average welfare losses per trip. As Table 7 indicates, regardless of whether reserved bus lanes are allowed, the provision of bus service itself causes some losses when MN/K is assumed to be less than one. Under the conditions given in Table 7, when reserved bus lanes are not permitted and when MN/K is set equal to 0.8, the change from the

Table 9. Comparison of system operation with and without reserved bus lanes under different pricing and financial constraints: base off-peak-hour case.

Constraint	Welfare Losses per Trip		Optimal Bus Fare (cents)		Buses per Hour		Bus Operating Cost per Passenger (cents)		Bus Share of Capacity (%)		Modal Split Without Reserved Bus Lanes		Value of Marginal Product of Capacity Without (cents)[a]
	Without	With	Without	With	Without	With	Without	With	Without	With	Automobile	Bus	
Social welfare maximizing	2.29	2.29	10.8	10.8	29	29	10.8	10.8	100	100	0.000	1.000	0.0
Pricing constraint													
Automobile toll = $0.06/mile	2.48		0		2		1.1		1.7		0.340	0.660	0.1
Automobile toll = $0.06/mile and bus operating subsidies = $0.0	3.21				0				0		1.000	0.000	11.4
Automobile toll = $0.08/mile	2.56		0		2		1.0		2.0		0.288	0.712	0.1
Automobile toll = $0.08/mile and bus operating subsidies = $0.0	3.46				0				0		1.000	0.000	11.4

Note: Length of off-peak period = 10 h, number of lanes = 2, travel rate = 400, and off-peak-hour bus costs = $25.
[a]Cost per number of cars/h along a given mile.

provision of bus service to an all-automobile travel pattern decreases average welfare losses per trip by 18.9 cents (i.e., by 7.5 percent). Therefore, under off-peak conditions, when MN/K is set equal to 0.8, the maximization of net travelers' welfare leads to minimum average losses per trip of $2.29 and requires no bus service. However, when MN/K is set equal to 1.6, as Table 9 indicates, the provision of an optimal level of bus service would reduce the average welfare losses per trip to considerably below the levels that would result with only automobile travel (i.e., a reduction of 91.8 cents--28.6 percent).

CONCLUSION

If the numerical analyses presented can be accepted as valid, they have the following policy implications:

1. Given the demand and cost conditions that these analyses are based on, the imposition of a net travelers' welfare maximizing level of decision variables (i.e., bus fares, automobile tolls, and bus service) is the best short-run solution for the traffic congestion problem. As the analyses show, the optimal level of automobile tolls at peak hours is much higher than the gasoline tax currently imposed in most North American cities.

2. The demand and cost conditions, which promise significant benefits from the imposition of optimum levels of bus fares and automobile tolls and the provision of optimum level of bus service, imply as well that road expansion would yield substantial benefits. However, the optimal level of arterial capacity changes as demand and cost conditions change.

3. Given that the imposition of optimal tolls is regarded as impracticable for technological or political reasons, the numerical analyses show that reserved bus lanes appear capable of substantially reducing current welfare losses of travel in peak hours. In other words, the numerical analyses allow us to conclude that the provision of reserved bus lanes constitutes a good solution for the traffic congestion problem.

4. Given that automobile tolls are restricted to the gasoline tax, which is much lower than the optimal level of tolls recommended for peak hours (63-93 cents/mile of automobile trip), the numerical analyses presented in this paper reveal that the optimal provision of bus service implies mass bus operating subsidies.

The above findings are based on the demand and cost conditions posited here. Our limited sensitivity analyses highlight the dependency between the results and the demand and cost conditions assumed in the study. Therefore, further sensitivity analyses are essential.

REFERENCES

1. Cambridge Systematic, Inc. The MTC Travel Demand Model Development Project: Final Report; Vol. 1: Summary Report. Metropolitan Travel Commission, Berkeley, CA, Sept 1977.
2. Cambridge Systematic, Inc. Logit Input Program: Program Documentation. Metropolitan Travel Commission, Berkeley, CA, 1976.
3. C. Clark. The Greatest of Finite Set of Random Variables. Operation Research, Vol. 9, 1961, pp. 145-162.
4. T. Lancaster. Prediction from Binary Choice Models. Univ. of Hull, Hull, England, 1979.
5. S. Khajavi. Optimal Peak-Load Pricing, Investment and Service Levels on Urban Streets. Univ. of Toronto, Toronto, Ontario, Canada, Ph.D. dissertation, 1980.
6. H. Mohring. The Benefits of Reserved Bus Lanes, Mass Transit Subsidies, and Marginal Cost Pricing in Alleviating Traffic Congestion. Univ. of Minnesota, Minneapolis, 1977.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

Development and Evaluation of a Synthetically Self-Calibrating Gravity Model

JAMES R. MEKEMSON AND KUMARES C. SINHA

The development of an alternative to the quick-response technique of using transferable parameters in trip distribution for small- and medium-sized urban areas is discussed. The proposed quick-response procedure involves an origin-zone-specific, self-calibrating gravity model in which the only input data required are the zonal productions and attractions, a zone-to-zone travel time matrix (skim tree), and the origin-zone terminal times. A travel time distribution determined from an origin-destination survey for internal trips is not needed for calibration. Tests conducted on three separate study areas indicated that the proposed model is able to reproduce trip patterns as accurately as the traditionally calibrated gravity model procedure based on origin-destination survey data. The accuracy was achieved by synthetic calibration of the model at the origin-zone level rather than at the aggregate level of the entire study area. Development of the proposed procedure was also based on the consideration that trip distribution is critically dependent on the spatial distribution of land use activities about each of the origin zones. This consideration was incorporated in the proposed procedure through the explicit measurement of the origin-zone-specific opportunity travel time distribution. The opportunity distribution for each origin zone was represented in the model by the origin-zone average travel time, computed from a gravity model trip distribution that has constant friction factors. From this initial key variable, the final model was developed, and to this the very acceptable results can be credited.

The initiation or updating of a comprehensive transportation plan requires considerable time and money. However, the changed emphasis from traditional long-range system planning to short-range, quick-response improvement programs no longer allows for the frequently long time span between the initiation of a transportation study and the final report. In addition, the recently addressed issues of transportation impact analysis in the areas of energy conservation, air quality management, and other environmental, economic, and political issues have increased the overall scope that must be considered in a truly comprehensive transportation plan. Requests by elected and public officials for quick responses, in combination with the ever-widening scope of planning issues, necessitates that the traditional transportation analysis process be modified if it is to be relevant to the short-range planning process. Capabilities need to be developed for simplified methods in the conventional four-step estimation process (trip generation, trip distribution, mode split, and traffic assignment) and also for various impact analysis and evaluation techniques. This paper discusses only one phase of the total process, trip distribution modeling, and presents a new procedure for trip distribution gravity model calibration that is designed in the context of quick-response capabilities with limited input data requirements and without sacrificing the accuracy obtained through traditional origin and destination (O-D) survey calibration techniques (1).

CURRENT TRIP DISTRIBUTION TECHNIQUES

Six basic techniques are commonly used in trip distribution modeling. These techniques are best characterized in terms of their resource requirements of time, data, and money. A hierarchical ranking of these six techniques follows:

1. Traditional gravity model that requires a relatively large O-D survey for calibrating the travel-time-impedance function (friction-factor curve) [such a process is described in the PLANPAC/BACKPAC General Information book (2)],
2. Traditional gravity model that uses a small O-D survey for travel-time-impedance function calibration (3),
3. Traditional gravity model with a calibrated friction-factor curve from a similar urban area (4),
4. Traditional gravity model that uses standardized or default friction-factor curves or parameters based on population size (5,6),
5. Manual gravity model that has standardized or default parameters and interzonal travel times based on airline distances and other factors (6), and
6. Traditional gravity model that has constant friction factors.

The first two techniques are very similar: the only difference is the size of the O-D survey sample. Although both techniques require an O-D survey, the first technique requires a much larger commitment of time and money.

Quick-response trip-distribution techniques three and four are also similar to each other because both methods assume that similar transportation study areas exist and that these similarities can be classified. The predominant criterion for similarity has been the size of the population of the study area. Figures 1 (1,7,8) and 2 (1) illustrate the discrepancies that may occur when this assumption of similarity is used. Figure 1 shows the great variability of home-based work trip average travel time with respect to study area population size (7,8). A plot of the non-home-based trip beta (β) calibration parameter of the negative exponential friction-factor function versus the size of the population of the study area in Figure 2 also illustrates a large variability with respect to population size, especially at the lower population ranges.

The fifth technique is similar to the third and fourth in that standardized calibration parameters assume similar study areas from which these values are established. This procedure estimates travel times from airline distances and other adjustment factors. The difference in this technique is its manual instead of computerized trip distribution. It is therefore limited in terms of the number of zones and network detail because of possible time constraints.

The sixth technique, constant friction factors, assumes that trip makers do not consider travel time in their destination choice process. Such an assumption can lead to very significant errors in the transportation planning modeling process. The travel time errors that result from the use of constant friction factors in the trip distribution models of two urban areas of Indiana: Lafayette (population, 100 000) and Anderson (population, 90 000), are shown in Table 1. Average travel time for all internal trips was overestimated by 30 percent for Lafayette and 19 percent for Anderson. Total link percent-root-mean-square-error (PRMSE) for a traditionally calibrated trip distribution model and its assignment for Lafayette was 12.48 percent and the total link PRMSE by using constant friction factors was 129.83 percent. Anderson's total link PRMSE was 13.28 percent for a traditionally calibrated trip distribution model and its as-

Figure 1. Home-based work average travel time versus urban area population size.

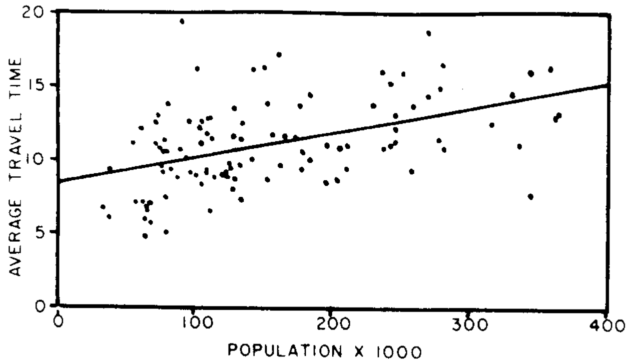


Figure 2. Variability of beta with respect to urban area population size, non-home-based trips.

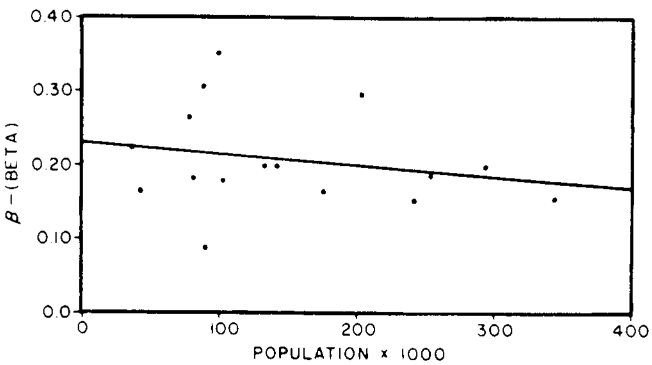


Table 1. Results of the Lafayette and Anderson gravity model with constant friction-factor values.

City	Trip Purpose	Average Travel Time Survey	Average Travel Time Model	Difference (%)
Lafayette	Home-based work	10.57	12.32	+16.56
	Home-based other	8.56	11.78	+37.62
	Nonhome based	9.12	11.42	+25.22
	Avg	9.05	11.76	+29.94
Anderson	Home-based work	11.44	12.10	+5.77
	Home-based other	10.04	12.30	+22.51
	Nonhome based	9.50	11.72	+15.58
	Avg	10.14	12.07	+19.03

signment and 86.00 percent when constant friction factors were used. The average overestimated link volumes for Lafayette and Anderson were 1760 and 771 vehicle trips, respectively. Clearly, the error introduced by using constant friction factors is unacceptable.

The above six trip distribution techniques can be further classified into three categories. Techniques one and two are sensitive to the transportation network and study area spatial land use activity distribution because of their ability to use the O-D survey data for calibration purposes. Techniques three, four, and five use borrowed or standardized friction-factor values and not derived values from O-D survey data; hence, the gravity models cannot be accurately calibrated and are therefore less sensitive to a particular study area's network and spatial land use activity distribution. Technique six is insensitive to the transportation network because it completely excludes travel time. The transportation planner thus has a choice of conducting an O-D survey for greater sensitivity and therefore greater model accuracy or using standardized, borrowed, or constant friction-factor values and probably sacrifice accuracy in return for expediency. Remember, though, that the accuracy needed is a function of the degree of precision desired by the transportation planner.

SENSITIVITY OF GRAVITY MODELS TO FRICTION-FACTOR PARAMETER ERRORS

The occurrence of some error is expected when quick-response trip distribution techniques are used; however, it would be useful to have some knowledge of the degree to which gravity models are sensitive to friction-factor parameter errors. To this end, a limited sensitivity analysis had been conducted by using the Lafayette and Anderson transportation study areas (1). The results of this analysis are summarized below.

1. The Lafayette and Anderson study areas differed greatly in terms of their sensitivity to parameter errors in friction-factor equations. Lafayette was more sensitive to parameter errors in terms of travel time estimation and traffic assignment link volumes. This difference in sensitivity was also suggested by the previously mentioned differences in errors that result from the use of constant friction factors. Sensitivity difference was probably due to underlying differences in urban structure.
2. When the single-parameter inverse travel time or negative exponential travel-time-impedance function is used, the gravity models and resulting traffic assignments are less sensitive to overestimates of the parameters than to underestimates.
3. Acceptable model results often occur when synthetic techniques are used because (a) friction-factor parameters used are approximately equal to the values that would have been determined through the traditional calibration process, (b) transportation models for the particular urban area are insensitive to errors (e.g., Anderson, Indiana), or (c) offsetting errors in trip purpose modeling lessen the total model error.

A synthetic technique that provides acceptable results due to good modeling and not to occasional model insensitivity or offsetting model errors would be greatly preferred, especially for those urban areas that are not typical or similar.

PROPOSED MODEL CONCEPT

A trip distribution technique is needed that is sensitive to both the transportation network and spatial land use activity system, but without the necessity of conducting even a limited O-D survey. The hypothesis on which the proposed model was developed is that there exists additional information within the network description, the zonal productions and attractions, and the gravity model itself by which the friction-factor relationship can be estimated at an origin-zone-specific level. Other researchers such as Voorhees (7), Wilson (9), and Fisk and Brown (10) have also suggested the possibility of calibrating the gravity model at the origin-zone-specific level.

Search for Calibration Relationships

In order to search for possible inherent relationships within the gravity model and its variables, the trip distribution gravity models for three Indiana study areas--Lafayette, Anderson, and Muncie (population range, 79 000 to 100 000)--were calibrated at the origin-zone level by using their respective O-D survey data, thereby the amount of information available for analysis was maximized. The travel time impedance function used to calibrate the model at the origin-zone level was the negative exponential function as given in Equation 1.

$$F_{ij} = e^{-\beta_i C_{ij}} \quad (1)$$

where

F_{ij} = value of the friction factor between origin zone i and destination zone j,
C_{ij} = travel time (cost) between zone i and j,
e = base of the natural logarithm, and
β_i = origin-zone-specific calibration parameter.

Thus, for each study area and trip purpose, the following origin-zone-specific information was available for further study: origin-zone productions, origin-zone β_i calibration parameter, origin-zone average travel time (AVE_i), and origin-zone travel-time variance.

One additional piece of information contained within a gravity model is the concept of the opportunity distribution as set forth by Voorhees and others (7). Opportunity distribution is the gravity model distribution that results from travel-time impedance factors of constant value at all time increments. As demonstrated by Voorhees and others, this distribution can be used as a measure of the spatial arrangement of land use activities, especially when used at the origin-zone level (7). Origin-zone-specific opportunity average travel time ($AVEO_i$) and travel-time variance were therefore computed and added to the data set.

Regression analysis procedures used in the search for additional variable relationships consisted of weighting each origin-zone observation by its respective trip productions. Another weighting procedure was used to balance the number of trip productions among the three study areas by trip purpose. These two weighting techniques were necessary to minimize bias. The first weighting scheme prevents an origin zone that has very few trip productions from having the same statistical weight as an origin zone that has many trip productions. The second weighting scheme prevents bias toward any one particular study area due to the relative number of trips per study area for any particular trip purpose.

Various researchers (10,11) have suggested that it may be possible to relate the β_i calibration parameter of the negative exponential friction-factor equation (Equation 1) to travel time statistics at the origin-zone level. To this end, extensive regression analysis by using the above-generated data set by trip purpose resulted in prediction equations of poor statistical fit. However, the correlation matrix revealed another possible approach to the problem.

A high correlation between the origin-zone-specific average travel time (AVE_i) and opportunity average travel time ($AVEO_i$) was observed for the home-based work, home-based other, and non-home-based trip purposes. This high correlation suggested that it may be possible to relate AVE_i as a function of $AVEO_i$. In other words, origin-zone average travel time AVE_i is a function of how the attractions (land use activities) are spatially distributed about the origin zone with respect to the interconnecting highway network as measured by $AVEO_i$.

Because travel time is composed of three distinct time segments (origin-zone terminal time, link travel time, and destination-zone terminal time), origin-zone-opportunity average travel time was separated into two parts before regression analysis was performed. The first part consisted only of the origin-zone terminal time ($TERM_i$), which is a constant for any particular origin-zone-based trip. The second part consists of that part of the travel time over which the trip maker can make choices as to trip length; that is, the combined link travel time and destination-zone terminal time or ($AVEO_i$ - $TERM_i$). This separation of travel time into these two parts allows the regression-analysis procedure to reveal the relative importances of these two variables with respect to trip purpose.

Regression analysis of AVE_i on ($AVEO_i$ - $TERM_i$) and $TERM_i$ was accomplished by weighting the origin-zone-specific data, as previously described, in order to properly balance the data by trip productions (origins) and by population of the study area for each trip purpose. Equations 2, 3, and 4 present the results of the regression analysis for the home-based work, home-based other, and non-home-based trips, respectively. Total R^2 and the change in R^2 for the independent variables are also given. (Because of the weighting by productions in the regression procedure, t- or F-statistics are meaningless and are therefore not given.)

Home-Based Work

$$AVE_i = -0.4328 + 1.2761\ TERM_i\ [\Delta R^2 = 0.114] + 0.8921(AVEO_i - TERM_i)\ [\Delta R^2 = 0.751] \quad (2)$$
$$R^2 = 0.865$$

Home-Based Other

$$AVE_i = -0.7499 + 1.3548\ TERM_i\ [\Delta R^2 = 0.187] + 0.7558(AVEO_i - TERM_i)\ [\Delta R^2 = 0.661] \quad (3)$$
$$R^2 = 0.848$$

Non-Home-Based

$$AVE_i = 0.5245 + 1.3416\ TERM_i\ [\Delta R^2 = 0.423] + 0.6073(AVEO_i - TERM_i)\ [\Delta R^2 = 0.296] \quad (4)$$
$$R^2 = 0.719$$

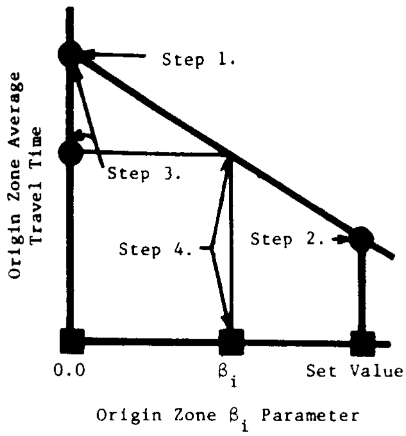
The strong statistical fit of the equations clearly shows the importance of $AVEO_i$, the origin-zone-specific opportunity average travel time, on the destination-choice process. This, in turn, demonstrates the large effect that the location of the origin zone and the distribution of land use activities about the origin zone have on the destination-choice process.

Relative importance of travel time by trip purpose in the destination-choice process is indicated by the coefficients of the $AVEO_i$ variable. In general, home-based work trips have one or few choices of destination, and the non-home-based trips have many choices of destination. Consequently, the equations show that the importance of travel time in terms of choice is the lowest for home-based work, highest for non-home-based, and of some intermediate magnitude for home-based other trips.

In addition, variable $TERM_i$, the origin-zone terminal time, becomes a greater proportion of total travel time as trip purpose changes from home-based work to home-based other or non-home-based, as is evidenced by the change in R^2. This is because $TERM_i$ is an origin-zone-specific constant and cannot be reduced by destination choice, unlike the link travel times and destination-zone terminal times.

Also of interest are the coefficients of the origin-zone terminal times. It was first expected

Figure 3. Flow chart of synthetically self-calibrated, origin-zone-specific gravity model as applied to a typical zone.



Step 1 - Average at β_i equal zero, $AVEO_i$

Step 2 - Average at β_i equal to a set value.

Step 3 - Average from regression equation, AVE_i

Step 4 - Interpolate for new β_i

Table 2. Origin-zone-specific, synthetically self-calibrating gravity model travel time statistics.

City	Trip Purpose	Average Travel Time O-D Survey	Average Travel Time Model	Difference (%)
Lafayette	Home-based work	10.57	10.97	+3.78
	Home-based other	8.56	8.56	+1.05
	Nonhome based	9.12	8.93	-2.08
	Avg	9.05	9.09	+0.71
Anderson	Home-based work	11.44	11.27	-1.49
	Home-based other	10.04	10.02	-0.20
	Nonhome based	9.50	9.64	+1.47
	Avg	10.14	10.14	0.00
Muncie	Home-based work	9.07	8.82	-2.79
	Home-based other	7.48	7.37	-1.46
	Nonhome based	6.93	7.04	+1.63
	Avg	7.52	7.46	-0.80

that the coefficients would have a magnitude very close to one. The coefficients were found to be greater than one. No explanation could be rationalized. Further investigation would be necessary to account for this result.

After estimating equations were developed by trip purpose for the origin-zone-specific average travel time (AVE_i) based on the origin-zone-specific opportunity average travel time ($AVEO_i$), a procedure was then formulated to incorporate these equations into a gravity model.

Synthetic Gravity Model Formulation

One method of calibrating the gravity model would be to first distribute all trips with a constant friction factor, thereby allowing for the computation of $AVEO_i$. AVE_i could then be computed by using the regression equations discussed in the previous section. These estimated origin-zone-specific average travel times then could be used to repeatedly adjust the origin-zone-specific β_i parameters until the gravity-model origin-zone-specific average travel times match the estimated values. However, this type of calibration procedure would require many iterations before convergence of β_i for all the origin zones would occur. The many iterations required for convergence would consume a large amount of computer time. This would be an undesirable characteristic for a synthetic model to possess.

Consequently, another method was developed for eventual adoption for synthetic calibration and was based on three assumptions regarding the interaction of origin-zone average travel time and the origin-zone β_i parameter. The first assumption was that the origin-zone average travel time varies linearly with the origin-zone β_i parameters (12). The second was that origin-zone average travel time is independent of the β_i parameters of other zones. Third, origin-zone average travel time is independent of the attraction-constraint iteration procedure. These assumptions, however, do not always hold true for some of the individual zones when the origin-zone β_i parameters are allowed to vary independently, as in the proposed model. This aspect was not demonstrated until after the model was developed and tested, but as the final results indicated, aggregation of the origin-zone results up to the entire study area, as evidenced by the acceptable traffic assignment, nullified these minor errors. This is typical of other transportation planning procedures.

The estimating equations of AVE_i and the three assumptions listed above were incorporated into a gravity model procedure that can best be described in conjunction with Figure 3. First, trips are distributed with all origin-zone-specific β_i parameters equal to 0.0 (constant friction factor) and then the resulting origin-zone-opportunity average times ($AVEO_i$) are computed. Second, trips are distributed with all origin-zone-specific β_i parameters equal to a single preselected value and the resulting origin-zone average travel times are computed. This step provides a linear relationship between each of the origin-zone-specific β_i's and average travel times. Third, by using the estimating equation for the appropriate trip purpose and the $AVEO_i$ values computed in step one, AVE_i is determined. Fourth, by using the data points from steps one and three and then estimating AVE_i, the origin-zone-specific β_i parameters are obtained through interpolation. All trips are then distributed by using the interpolated origin-zone-specific β_i parameters.

EVALUATION OF THE PROPOSED MODEL

The gravity model procedure was used to distribute the trips for Lafayette, Anderson, and Muncie, Indiana. The resulting average travel time statistics by trip purpose and for all internal trips are given in Table 2. Comparison of these statistics to the corresponding O-D survey values is also shown. Lafayette's total internal trip average travel time was only 0.71 percent greater than the O-D survey value; Anderson's total internal average travel time was identical to the O-D survey value. Muncie's total internal average travel time was 0.80 percent lower than the O-D survey results.

The average travel times for each trip purpose also indicate very acceptable results; the largest percentage error among the area's individual trip purposes was 3.78 percent, this being the only one outside of Federal Highway Administration's (FHWA) suggested ±3 percent error range when survey data

Figure 4. Volume group link comparison for Lafayette all or nothing traffic assignment: origin-zone-specific, synthetically self-calibrating gravity model versus O-D survey.

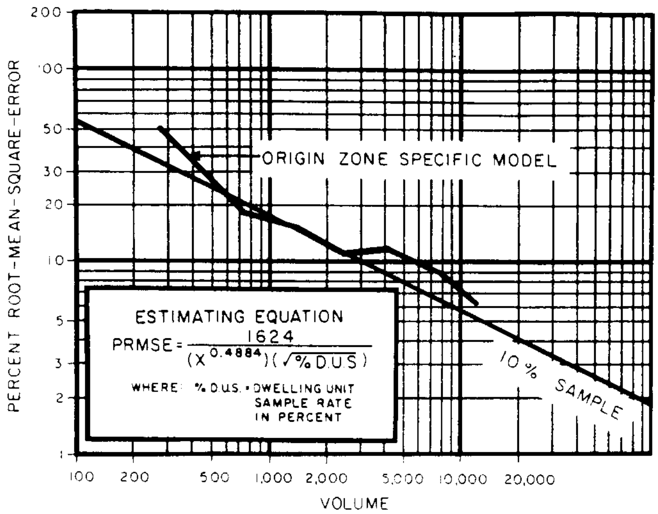
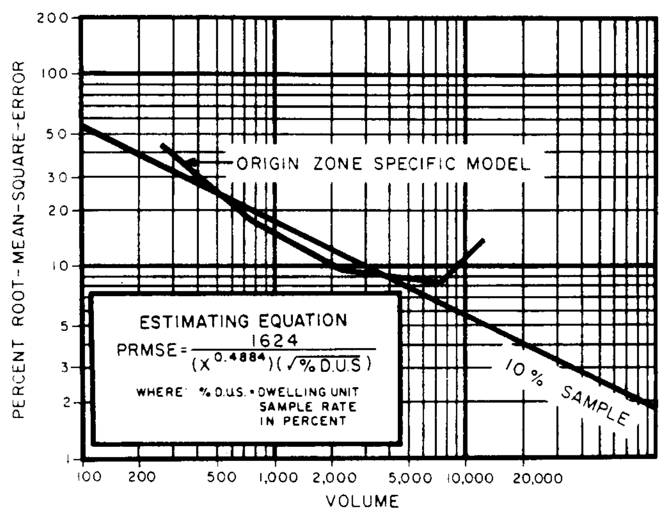


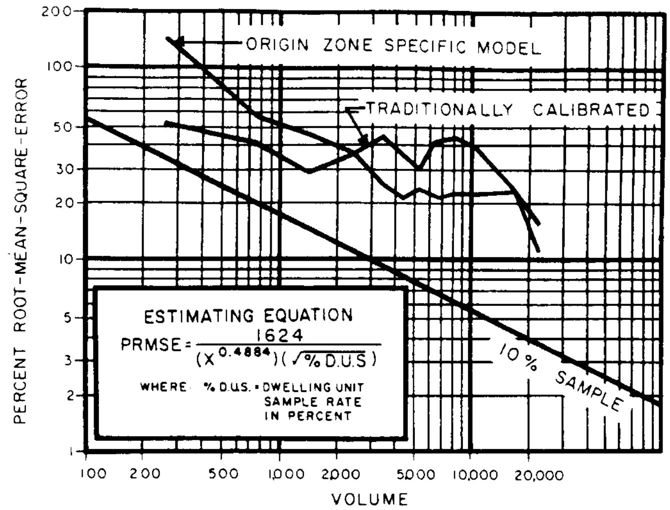
Figure 5. Volume group link comparison for Anderson all or nothing traffic assignment: origin-zone-specific, synthetically self-calibrating gravity model versus O-D survey.



are used for calibration (13).

Volume group link comparisons of the three synthetic gravity models' all-or-nothing traffic assignments were performed by using their corresponding O-D survey traffic assignments as the base for comparison. An all-or-nothing traffic assignment was used because only an analysis of gravity model accuracy was desired; errors due to capacity-restraint differences and ground count errors are eliminated. Total link trip PRMSE for Lafayette was computed to be 12.67 percent. This value was practically identical to the corresponding value (12.48) obtained from the gravity model calibrated by the traditional procedure. Total link trip PRMSE was calculated to be 14.55 percent as compared with the traditionally calibrated gravity model value of 13.28 percent for Anderson. A surprising result occurred for the Muncie study area, where total link trip PRMSE was 47.7 percent--a major improvement compared with the traditional calibrated gravity model result of 67.8 percent. Volume group link comparisons are shown graphically in Figures 4, 5, and 6. Also shown on the figures are the expected

Figure 6. Volume group link comparison for Muncie all or nothing traffic assignment: origin-zone-specific, synthetically self-calibrating gravity model versus O-D survey.



PRMSE values associated with a 10 percent dwelling unit sampling rate for the O-D home interview survey. In addition, Figure 6 also shows the PRMSE statistics for the traditionally calibrated gravity model. The PRMSE values for Lafayette, Anderson, and Muncie were all observed to be acceptable.

Of special interest in Figure 6 is the result that PRMSE values for the proposed model for the higher volume links are significantly lower than results obtained with the traditional model. Also, note that a probable explanation for the high PRMSE values for Muncie by using both the traditional and proposed Muncie gravity models may be that the Muncie study area included a relatively high number of zones and links for an urban area of its population size. The zones and links are proportionally much greater than in either Lafayette or Anderson, which suggests that the Muncie zone and network system were probably too detailed for the number of trips produced. For example, the Muncie data included 1249 out of 2698 links that had an average daily traffic of less than 500 vehicles/day; peak-hour traffic on these links would probably not exceed 50-60 vehicles/h.

Hence, the trip distribution and traffic assignment results of the three test areas indicate that the proposed origin-zone-specific, synthetically self-calibrating gravity model is capable of distributing trips at least as accurately as the traditionally calibrated gravity model without the need for an O-D survey.

Revision of the Calibration Equation

Because of the inability to give an interpretation to the constant terms in Equations 2, 3, and 4 and the fact that travel time is composed of only terminal times and link travel time, a regression analysis was performed to force the regression through the origin and therefore eliminate the constant term in the equations. The result of this forced regression is given in Equations 5, 6, and 7.

For home-based work trips,

$$AVE_i = 1.1910\,TERM_i + 0.8638(AVEO_i - TERM_i) \quad (5)$$

For home-based other trips,

$$AVE_i = 1.122\,34\,TERM_i + 0.7033(AVEO_i - TERM_i) \quad (6)$$

For non-home-based trips,

$$AVE_i = 1.2524\,TERM_i + 0.6856(AVEO_i - TERM_i) \tag{7}$$

Trial runs of the model with the above equations indicated no significant change in the model results compared with the equations that have the constant term.

Computer Availability and Program Requirements

The model was initially coded as optional user code in the Urban Transportation Planning System (UTPS) UMODEL program. A similar version was also coded for the PLANPAC package. A FORTRAN IV program is also available for incorporation into other packages.

For the 271 internal zone Muncie study area, 180 s of central processing unit (CPU) time were required to execute the model for the home-based other trip purpose on a CDC 6500. The same Muncie home-based other trips required 14 min of CPU time on an IBM 370/148 by using the UTPS UMODEL program. CPU time for the UMODEL version is highly variable, depending on UMODEL options and reports selected.

Test Case Application

The Indiana State Highway Commission will use the origin-zone-specific, synthetically self-calibrating gravity model during the initial transportation plan development study for Bloomington, Indiana, during the summer and fall of 1981. This test case application will provide a real-world test of the model.

FUTURE RESEARCH

It is hoped that this model can be tested by using the data from a much larger urban area in the 750 000 to 1 000 000 population range. Such an urban area would provide a large data base for testing the model and for comparing the characteristics of a large urban area with those of the three smaller study areas used in the development of the model.

SUMMARY AND CONCLUSIONS

Because of the need to reduce time, money, and effort involved in the transportation planning process, as well as to address emerging planning issues, many new techniques need to be developed and evaluated. One such technique is the proposed synthetic trip-distribution procedure, which consists of an origin-zone-specific, synthetically self-calibrating gravity model in which the only input data required are the zonal productions and attractions, a zone-to-zone travel-time matrix (skim tree), and the origin-zone terminal times. A travel-time distribution is no longer needed for calibration, thereby eliminating the necessity for an O-D survey for internal-internal trips.

Tests conducted on three separate study areas indicated that the proposed model is able to reproduce trip patterns as accurately as the traditionally calibrated gravity model procedure that is calibrated on O-D survey data. The accuracy was achieved by synthetically calibrating the model at the origin-zone level rather than at the aggregate level of the entire study area. Development of the proposed procedure was also based on the consideration that the trip distribution is critically dependent on the special distribution of land use activities about the origin zone. This consideration was incorporated in the proposed procedure through the explicit measurement of the origin-zone-specific opportunity travel-time distribution.

ACKNOWLEDGMENT

This paper summarizes the results described in a report entitled Evaluation and Development of Synthetic Trip Distribution Modeling Techniques as Applied to Small Urban Areas, which documents a research project conducted by the Joint Highway Research Project staff at Purdue University (1). This project was funded by FHWA through the Highway Planning and Research funds administered by the Indiana State Highway Commission. We accept sole responsibility for the contents of this paper.

REFERENCES

1. J.R. Mekemson. Evaluation and Development of Synthetic Trip Distribution Modeling Techniques as Applied to Small Urban Areas. Joint Highway Research Project, Purdue Univ., West Lafayette, IN, JHRP-79-5, July 1979.
2. Computer Programs for Urban Transportation Planning: PLANPAC/BACKPAC General Information Manual. Federal Highway Administration, April 1977.
3. R.E. Foster and others; Texas Transportation Institute, Texas A&M University. Consequences of Small Sample O-D Data Collection in the Transportation Planning Process. Federal Highway Administration, FHWA-RD-78-56, Jan. 1976.
4. W.L. Grecco and others. Transportation Planning for Small Urban Areas. National Cooperative Highway Research Program, Rept. 167, 1976.
5. UTPS--Urban Transportation Planning System. Planning Methodology and Technical Support Division, Urban Mass Transportation Administration; Urban Planning Division, Federal Highway Administration, HHP-20, 1977.
6. Comsis Corporation. Quick-Response Urban Travel Estimation Manual Techniques and Transferable Parameters--A Users Guide. National Cooperative Highway Research Project, Rept. 187, 1977.
7. Alan M. Voorhees and Associates. Factor and Trends in Trip Lengths. National Cooperative Highway Research Program, Rept. 48, 1968.
8. Y. Chan and others. Review and Compilation of Demand Forecasting Experiences: An Aggregation of Estimation Procedures. Pennsylvania Transportation Institute, Pennsylvania State Univ., University Park, June 1977.
9. A.G. Wilson. Further Developments of Entropy Maximizing Transport Models. Transportation Planning Technology, Vol. 1, 1973, pp. 183-193.
10. C. Fish and G.R. Brown. The Role of Model Parameters in Trip Distribution Models. Transportation Research, Vol. 9, 1975, pp. 143-148.
11. J.A. Tomlin and S.G. Tomlin. Traffic Distribution and Entropy. Nature, Vol. 220, 1968, pp. 974-976.
12. G.M. Hyman. The Calibration of Trip Distribution Models. Environment and Planning, Vol. 1, 1969, pp. 105-112.
13. Calibration and Testing a Gravity Model for Any Size Urban Area. Federal Highway Administration, March 1975.

Publication of this paper sponsored by Committee on Transportation Planning Needs and Requirements of Small- and Medium-Sized Communities.

What Will Happen to Travel in the Next 20 Years?

DAVID T. HARTGEN

This paper develops a baseline projection of travel and energy use for New York State for the period 1975-1995. The projection is developed from an equilibrium forecast of gasoline price, supply, improvements in average car efficiency, and population. This projection is then adjusted by using straightforward elasticity approaches to account for major trends in household and population characteristics, the economy and inflation, urbanization, and automobile ownership. Results show that upward pressure on travel will be caused by increased car efficiency, growth in population, suburbanization, and automobile ownership, in that order. Downward pressure will be exerted by energy price, supply embargoes, inflation, and employment, in that order. The projection of travel into the 1990s suggests that, in spite of higher-than-historical inflation, rapid increases in energy price, and a slowly growing economy, travel will continue to grow slowly and rise 40-50 percent over 1975 rates by 1995. In the same period, gasoline use will fall 10-20 percent, spurred downward by price rises and the rapid increases in car efficiency. Periodic supply shortfalls are likely and will slow the growth in travel but will not reverse it.

Forecasts of traffic are a basic element in decisions concerning transportation investments. Based on such forecasts, analyses can be made of benefits to users and nonusers; impacts on the economy, safety, and accidents; operating costs; energy; air quality; noise; and congestion.

Traffic forecasts are generally prepared by two basic approaches. In major urban areas, forecasts are generally based on the four-step travel-simulation process, in which travel on street links is an output of the assignment process. In nonurban areas, where assignment-based modeling does not exist or may not be appropriate, estimates are generally made by growing present traffic into the future by using projections of population, households, cars, employment, county or town vehicle miles of travel (VMT), or other parameters. Whenever such forecasts are made, whether in rural or urban contexts, they are likely to contain errors. These errors result from misspecification (wrong variables), misestimation (wrong assumptions about input level), or wrong coefficients.

These problems have always been with us, but have recently been highlighted by national and international events, the effects of which have not explicitly been accounted for in prior forecasts. Among these are rapid increases in energy price and periodic shortfalls in gasoline; high inflation, possibly coupled with recession in the economy; changes in employment and unemployment rates; slow economic growth; changes in life-style, family structure, and population age; automobile ownership; and automobile efficiency. Some of these factors may have been accounted for in the particulars of a traffic forecast for a given project, but the need to incorporate such factors, in general, is real. Indeed, forecasts prepared without consideration of such factors are, at best, likely to be received with disbelief; at worst, dismissed as unreasonable.

Ideally, we need to develop new tools for traffic forecasting that can handle various assumptions concerning these and other factors. Some of the research necessary to do this is now under way. But we need not wait for the results; much of the knowledge is available now, in disparate studies and reports. The purpose of this paper is to pull together this information and focus on the findings of the studies, not the methods used. A number of studies have attempted such forecasts at the national level. The Federal Highway Administration (FHWA) (1) prepared joint forecasts of travel, population, employment, and fuel use after the 1973-1974 energy crisis. This effort has recently been updated by the U.S. Department of Transportation (DOT) (see Spielberg and others in this Record). Both documents, however, do not lend themselves to use in making direct traffic forecasts or in adjusting previously developed forecasts because they are primarily overviews of travel and economic growth. The approach taken in this paper is to develop adjustment factors based on empirical evidence and travel elasticities that are applied to baseline forecasts to obtain adjusted estimates.

BASELINE PROJECTION FOR TRAVEL

It is often suggested that, now that energy prices are rising rapidly and supplies are tightening, travel growth will slow, perhaps, even reverse, thus the need for new projects will be short circuited. Prices will probably rise and curtailments may occur, but it does not follow that VMT will decline (more likely, its growth may slow). To see this, consider jointly the relationships among travel, gasoline price, gasoline consumption, and, most importantly, average passenger car efficiency.

The following analysis is based on extensive modeling of energy futures of the New York State Department of Transportation, documented elsewhere (2,3). The analysis assumes an equilibrium model of travel, gasoline supply, and price, expressed as follows:

$$VMT_F = VMT_{75}(POP_F/POP_{75})[1 + e_1(\Delta X_1/X_1) + e_2(\Delta X_2/X_2) + \ldots] \quad (1)$$

$$GD_F = (VMT_F)(1/EFF_F) \quad (2)$$

where

- VMT = vehicle miles of travel in 1975 and future (F) year;
- POP = population;
- GD = gasoline demand;
- X = independent variable, including gasoline supply, price, unemployment, and labor force; and
- EFF_F = automobile efficiency, over the road, miles/gal (actual vehicle in-use efficiency).

The model is operated by balancing gasoline supply and demand scenarios against price and efficiency. The forecast reflects a baseline set of assumptions as follows:

1. Real prices increase by 2 percent/year,
2. Average new-car efficiency follows federal efficiency standards to 1985 and is constant thereafter,
3. Population growth is moderate, and
4. Fuel supplies are adeauate (at higher prices).

The analysis (Figure 1) shows that the most significant effect on travel is the effect of improved efficiency of new model year vehicles and over-the-road New York State fleet efficiency. Federal law mandates that new cars increase in corporate average fuel economy (CAFE) from 1978 to 1985. New York's CAFE has exceeded the standard for the 1978, 1979, and 1980 model years, as can be seen in the following table, which gives New York State CAFE values through November 1980.

Model Year	Standard	New York State Domestic New Car CAFE
1978	18	19.2
1979	19	19.7
1980	20	21.4
1981	22	23.8
1982	24	
1983	26	
1984	27	
1985	27.5	

As a result of such trends, the average over-the-road efficiency of cars in New York will increase from 13.26 miles/gal in 1976 to 23.78 miles/gal in 1995, about 79 percent (4.0 percent/year). This means that, all things being equal, gasoline demand could fall 4 percent/year, or travel could grow 4 percent/year, solely because of fleet turnover and increases in car efficiency. But, not all of this potential gasoline savings is actually saved; a portion of it is reinvested in additional travel that otherwise would have been curtailed in the face of rising gasoline prices. To say it another way, improvement of vehicle efficiencies means that more miles can be driven on the same amount of gasoline. This encourages continued growth in travel, without a subsequent increase in gasoline fuel use. Table 1 and Figure 1 show that travel is projected to grow 46 percent (2.3 percent/year) during the 1976-1995 period, but gasoline consumption will actually decline 7.6 percent (0.4 percent/year) during the same period. These projections essentially parallel those of DOT (4), a recent National Cooperative Highway Research Program (NCHRP) analysis of energy scenarios (5), and a national assessment of energy use (Spielberg and others in this Record).

Federal decontrol of domestic crude oil, continued world pressure on supplies, and political factors are likely to result in continued increases in the real price of gasoline. Prices are expected to rise to about $2.00/gal (1978 dollars) by 1995. The projections incorporate estimates of the elasticity of gasoline sales with respect to price, estimated to be about -0.15 in current years and rising uniformly to -0.50 by 1995. This implies an increased public sensitivity to gasoline price increases. (An assumption about a lower absolute elasticity would lead to even higher prices.)

Figure 1. New York State baseline forecasts of travel, energy, gasoline price, and fleet efficiency.

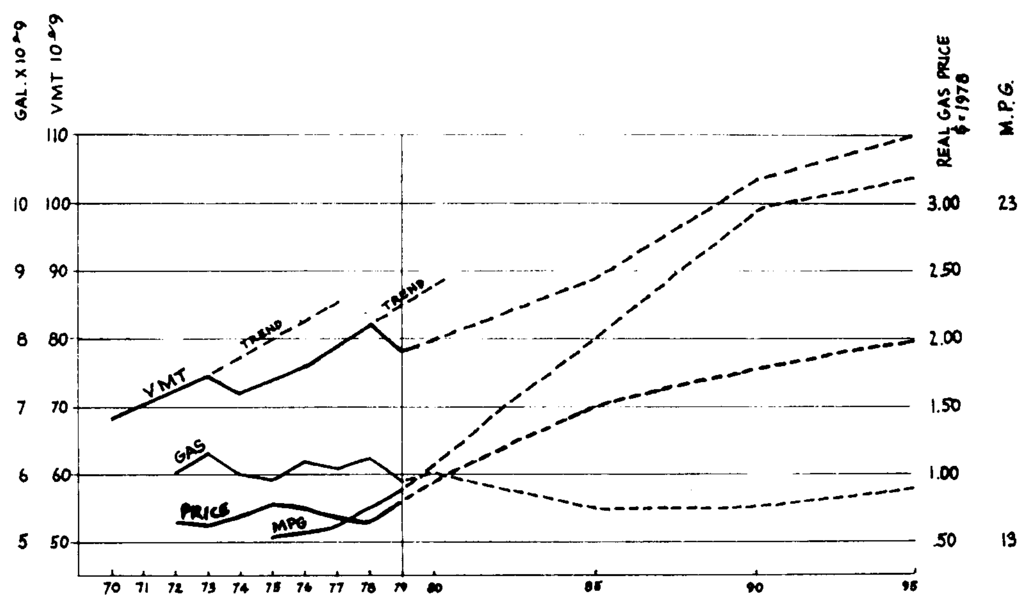


Table 1. New York State baseline forecasts of travel, energy use, gasoline price, and fleet efficiency.

Year	Vehicle Miles of Travel[a] (000 000 000s)	Index[b]	Total Gasoline Use (000 000 000s)	Index[b]	Real Gasoline Price (1978 dollars)	Index[b]	Car Efficiency	Index[b]
1968	62.326	0.831						
1969	64.641	0.862						
1970	67.042	0.894						
1971	69.532	0.927						
1972	72.115	0.961	6.063	0.978	0.625	0.839		
1973	74.794	0.997	6.321	1.020	0.624	0.838		
1974	72.222	0.963	5.998	0.967	0.746	1.001		
1975	73.621	0.981	5.985	0.965	0.757	1.016	13.110	0.989
1976	75.020	1.000	6.200	1.000	0.745	1.000	13.262	1.000
1977	78.260	1.043	6.122	0.987	0.744	0.999	13.529	1.020
1978	81.50	1.086	6.202	1.000	0.693	0.930	13.955	1.052
1979	79.50	1.059	5.871	0.947	0.798	1.071	14.418	1.087
1980			6.017	0.970	0.998	1.340	14.853	1.120
1985	89.073	1.187	5.485	0.885	1.495	2.007	19.156	1.444
1990	103.049	1.374	5.486	0.885	1.775	2.383	23.063	1.739
1995	109.786	1.463	5.730	0.924	1.970	2.644	23.776	1.793

Notes: Figures for 1980-1995 are forecasts.
Annual growth from 1976 to 1995 is projected to be +2.3 percent for vehicle miles of travel, -0.4 percent for gasoline, +8.2 percent for the real price of gasoline, and +4.0 percent for car efficiency.

[a]Vehicle miles of travel for 1968-1975 are adjusted estimates. [b]Index in 1976 = 1.0.

Table 2. Changes in efficiency and travel.

Year	Vehicle Miles of Travel	Car Efficiency (miles/gal)	Vehicle Miles of Travel (%Δ)	Efficiency (%Δ)	Ratio
1976	75.020	13.262			
1979	77.810	14.418	3.7	8.7	0.425
1985	89.073	19.156	18.7	44.4	0.421
1990	103.049	23.063	37.4	73.9	0.506
1995	109.786	23.776	46.3	79.3	0.584

Table 3. Vehicle miles of travel adjustment factors for average car efficiency.

Difference in Over-the-Road Efficiency (%)	1980, Ratio = 0.42	1985, Ratio = 0.42	1990, Ratio = 0.50	1995, Ratio = 0.58
-10	0.958	0.958	0.950	0.942
-5	0.979	0.979	0.975	0.971
0	1.000	1.000	1.000	1.000
5	1.021	1.021	1.025	1.029
10	1.042	1.042	1.050	1.058
20	1.084	1.084	1.100	1.116
30	1.126	1.126	1.150	1.170
40	1.168	1.168	1.200	1.232
50	1.210	1.210	1.250	1.290
60	1.252	1.252	1.300	1.348

Notes: Ratio values are taken from Table 2.
Formula: Factor = 1 + (ratio) (percentage difference)

Table 4. Example adjustment for revised fuel efficiency.

Year	Base Data Traffic Forecast	Miles/Gal Old	Miles/Gal Revised	Change (%)	Vehicle Miles of Travel Adjustment Factor[a]	New Traffic Forecasts
1975	1000	14.0	14.0	0	1.000	1000
1980	1150	14.0	14.8	+6	1.024	1178
1985	1294	14.0	19.1	+36	1.151	1489
1990	1423	14.0	23.0	+64	1.320	1878

[a]Interpolated from Table 3.

Table 5. Adjustment factor for price increases with no shortfalls.

Time Frame	Elasticity	Percentage Above Baseline Real Prices: 10	25	50	75	100
1980-1982	-0.15	0.9850	0.9625	0.9250	0.8875	0.8500
1983-1987	-0.35	0.9650	0.9125	0.8250	0.7375	0.6500
1988-1992	-0.50	0.9500	0.8750	0.7500	0.6250	0.5000

Notes: To use, multiply future year vehicle miles of travel by factor in table that corresponds to elasticity and price change.
Formula: F = 1 + (elasticity) (percentage increase).

VARYING THE ASSUMPTION

The above (or some other) forecast can be adjusted to account for new projections or additional assumptions.

Car Efficiency

If the analyst has a forecast of traffic that assumes a constant efficiency (compared with a base year), it should be adjusted to account for the greater mobility provided by the increased efficiency. Figure 1 provides a clue to the magnitude of this adjustment. Table 2 summarizes the percentage growth in travel and car efficiency over 1976, from Figure 1.

These data suggest that the proportion of the gain in average car efficiency, which shows up as VMT, is about 42 percent at present and rises to 60 percent by 1995. This information can be expressed as adjustment factors, such as in Table 3. As an example, if an analyst has estimates of traffic volume for a facility (based on an assumed over-the-road efficiency of 14.0 miles/gal), this can be adjusted as shown in Table 4.

This example shows that traffic forecasts prepared in the early 1970s (hence increased car efficiencies are not assumed) are likely to substantially underestimate VMT, particularly if they included rapid price increases. Ironically, forecasts that include both increasing efficiency and price rises are likely to require little net adjustment because, as will be shown, these two tend to cancel each other out.

Gasoline Price Increases

The baseline projection assumes an equilibrium price profile that rises more rapidly than inflation. This projection is our best assessment of the price trends at this time. Note that it shows real prices of gasoline doubling during the 1980-1990 period. In the event that the actual or implied price increases in a traffic forecast do not follow this track, some adjustments must be made in the values of estimated vehicle miles of travel.

Table 5 shows adjustment factors to apply to forecasts to account for price increases above or below estimated values. Previous estimates do not necessarily need to be those contained in Table 1; any prior estimate can be adjusted. To use these factors, the analyst selects the appropriate elasticities of travel and gasoline price, then estimates the percentage difference between the price forecast and the actual price (both are real prices in 1978 dollars).

The table further assumes that elasticities are likely to increase over time; this is consistent with our observations of increased public sensitivity to higher real prices and generally higher estimates of long-term elasticity in the literature, compared with short-term elasticity (5).

As an example, consider the above traffic forecast made with assumptions of gasoline price as shown in Table 6. Suppose new trends show a revised steeper price profile (1978 dollars), and the analyst wishes to adjust the above forecasts. Table 6 shows that substantial underestimates of real price can lead to overprojections of travel, particularly in later years. The price error here is typical of many projections made in the early 1970s and leads to an overforecast of about 30 percent by 1990. This is about the same error as is made by failing to account for increased efficiency of cars; hence, the two tend to cancel out in many cases.

Table 6. Example adjustment for revised prices.

Year	Base Forecast Volume	Base Forecast Price ($)	New Price ($)	Change (%)	Elasticity	Adjustment Factor	New Volume
1980	1150	1.00	1.20	20	-0.15	0.970	1115
1985	1294	1.10	1.50	44	-0.35	0.846	1095
1990	1423	1.25	2.00	60	-0.50	0.700	996

Energy Supply Cutoffs and Embargoes

As shown in Figure 1, historical growth rates of New York State travel before 1974 and in the interim period 1975-1978 were 3-4 percent/year. These rates were temporarily curtailed by the 1973-1974 and 1979 energy shortages, during which time VMT fell 3.4 and 4.5 percent, respectively. However, in the same periods, gasoline use fell 5.1 and 5.3 percent, respectively, and prices rose 30 percent (6). These trends are summarized as follows:

Trend	Change (%) 1973-1974	1979
Maximum quarterly shortfall	-13	-11
Annual gasoline use	-5.1	-5.3
Annual travel	-3.4	-4.5
Annual nominal price	30	35

A number of separate effects result from such shortfalls:

1. Travel declines caused by the shortfall, and
2. If price rises are permitted, energy prices rise consistent with the shortfall.

When a short-term interruption in gasoline supply occurs, travel must fall, to the extent that the drop in supply cannot be taken up by increases in driving efficiency, or price rises do not reduce demand. During the 1973-1974 and 1979 shortages, supplies were down by 11-13 percent at the height of the shortfalls and averaged just over -5 percent for the year. However, the corresponding declines in VMT do not quite account for the entire reduction of the fuel supply. Some of the savings was achieved from increased purchases of fuel-efficient cars (especially in 1979), slower driving, and tune-ups. Approximately 70-80 percent of the reduction in supply showed up as reduced travel. Behavioral data from a panel of 1520 households interviewed in 1979 (7) also showed that 72 percent of the energy saved came from actions that entail a drop in travel. Based on this, we conclude that, in a crisis, the annual drop in travel will be approximately 75 percent of the reduction in supply. This will introduce a downward shift in the VMT growth curve, as in Figure 1. However, VMT will continue upward growth following the easing of the problem, thus leading to a saw-tooth picture of travel over time (8). The 75 percent factor would be applied to all future forecasts to account for the interim shortage (Table 7).

During a period of short-term shortage, the effects of price rises have been on the order of 2.5 to 3 times the immediate maximum shortfall and 5-7 times the annual shortfall. Although further evidence is scanty, we believe this past experience to be reasonably indicative of the immediate price impact of a future shortage, moderating somewhat at the high end. Considerable evidence exists to suggest that the price elasticity of gasoline use is about -0.15. This figure is higher in large cities such as New York City (-0.23) and lower in small cities and rural areas (-0.10). By using this and the ratios above, Table 8 shows factors to adjust prices upward in response to a shortfall.

We have purposely not shown adjustments for shortfalls above 20 percent because we do not believe that prices would remain decontrolled at that level of shortage. Several analysts, however, have calculated the equilibrium price at $2.15/gal for a 20 percent shortfall (5).

Inflation

Data are displayed in Table 9 (2) for the national consumer price index (CPI) and for total VMT in New York State. The data indicate that, in years of moderate growth in the CPI (3-6 percent annually), the associated annual growth in VMT is the highest. High growth in the CPI (6-9 percent annually) is associated with lower growth in VMT. Very high rates of inflation (above 9 percent annually) were associated with a decline in travel in 1974 and 1979 and stability in travel during 1975; both of these events were fueled by energy crises that triggered a rise in energy price and the subsequent inflation.

For the moderate inflation group, a change of +1 percent in the CPI was associated with a change of about +0.75 percent in VMT. A +1 percent change in the high inflation group was associated with a change of +0.6 percent in VMT. For very high inflation, a +1 percent change in CPI was associated with a decline in VMT of 0.25 percent. From the above information, VMT adjustment factors may be developed, as shown in the table below:

Table 7. Adjustment factors for travel reductions following energy-supply cutoffs.

Annual Supply Shortfall (%)	Annual Drop in Travel (%)	Adjustment Factor
2	1.50	0.9850
3	2.25	0.9775
4	3.00	0.9700
5[a]	3.75	0.9625
6	4.50	0.9550
7	5.25	0.9475
8	6.00	0.9400
9	6.75	0.9325
10	7.50	0.9250
15	11.25	0.8875
20	15.00	0.8500

[a]This figure is close to the experience of 1979.

Table 8. Adjustment factors for price effects of shortfalls.

Maximum Quarterly Shortfall (%)	Annual Shortfall	Annual Resulting Real Price Rise (%)
-5	-2	+13
-10	-4	+25
-12[a]	-5	+30
-15	-7	+42
-20	-8	+48

[a]-12% is close to the experience of 1979.

Table 9. Vehicle miles of travel for New York State and U.S. consumer price index.

Year	CPI[a]	Vehicle Miles of Travel (000 000 000s)	Annual Change (%) CPI[a]	Annual Change (%) Vehicle Miles of Travel
1968	83.2	62.3		
1969	87.6	64.6	5.2	3.7
1970	92.8	67.0	5.9	3.7
1971	96.8	69.5	4.3	3.7
1972	100.0	72.1	3.3	3.7
1973	106.2	74.7	6.2	3.6
1974	117.9	72.2	11.0	-3.4
1975	128.7	72.2	9.1	0.0
1976	136.1	75.0	5.7	3.8
1977	144.9	78.2	6.4	4.3
1978	155.9	81.5	7.6	4.1
1979	173.5	77.8	11.3	-4.5

[a]1972 base.

Annual Inflation Rate	Adjustment Factor
2	1.015
4	1.030
6	1.036
8	1.048
9	1.000
10	0.975
12	0.970
14	0.965

Employment

To determine the impact of employment levels, the labor force and resulting expected employment were projected for New York State for 1970-1995. By using the available population projections and the 1970 participation rates by cohort, projections of the labor force for 1980, 1990, and 1995 were made. The projected increase in population of 1.5 million is concentrated in the principal age groups of workers and, by using the 1970 rates, this yields an additional 1.23 million in the labor force: 947 400 men and 287 100 women. The overall participation of those age 16 or more rises from 57.1 percent in 1970 to 58.3 percent by 1995. The rate of unemployment in 1970 was 4.8 percent. This rose to 7.0 percent at the beginning of 1980. A long-run rate of 5 percent was estimated for 1990 and 1995. Under these conditions, New York State must provide jobs for an additional 1 134 500 workers (at constant cohort rates of labor force participation) if the population projections are to be realized.

The sensitivity of changes in the overall labor force participation rates and unemployment was also examined for the impact on travel. These results are shown in Table 10. Note that a 1 percent change in the unemployment rate results in a change of 147-171 million VMT during the 1970-1995 period; this is about 0.20 percent of the travel estimated in New York State (assuming changes in work trips only).

Table 10. Effect of a 1 percent change in labor force participation and unemployment rate in New York State.

Year	Unemployment Rate (%)	Labor Force (000s)	Employment (000s)	Change in Vehicle Miles of Travel: Millions	Change in Vehicle Miles of Travel: Percentage
1970	4.8	74.59	71.00	147.35	0.220
1980	7.0	81.34	75.65	156.97	0.247
1990	5.0	84.59	80.35	166.75	0.209
1995	5.0	86.93	82.58	171.36	0.207

Note: The addition (or change) in vehicle miles of travel for an additional worker is limited to the work trip and is calculated by 250 days of work x 10 miles round trip distance for work x 83 percent automobile mode split, or 2075 miles/worker per year.

Trip Length and Trip Rates

A comparison of changes in Buffalo and Rochester travel from the early 1960s to 1970s (9) indicates that, overall, both average trip length and average trip rates per household were relatively stable, and the gain in person miles of travel (PMT) generally resulted from increases in the number of households not changes in trip rates. This conclusion masks a number of individual factors that did change. Among these changes were the following.

1. Increased length of work trips, but on nearly the same travel-time budget, as new highway construction eased the move to suburbs and rural areas and permitted longer work trips with little change in travel time. A smaller share of trips for all purposes, including work, were made to central business district (CBD) destinations and destinations within the city; this indicates that trip origins and destinations are becoming more oriented to the suburbs.
2. A substantial decline in shopping trips to the CBD confirmed a trip reorientation to nearby suburban shopping centers. Shopping trips were slightly longer but constituted a smaller share of trips. Trips for personal business were similarly affected with respect to the CBD, but to a smaller degree.
3. The share of automobile driver trips rose and the share of automobile passenger trips and bus trips declined, which reflects greater affluence and increased automobile ownership. PMT per car declined substantially, and PMT per household rose slightly in Rochester and declined about 10 percent in Buffalo.
4. The share of social and recreational trips declined sharply and these trips were much shorter in length. Trips for the purpose of catching another mode of transportation (e.g., to bus stop or train station) declined by nearly 50 percent in length but the share was stable. Although these trip purposes and trip lengths showed a mixed pattern, the overall trip length and number of trips per household remained relatively stable.
5. Household size declined in nearly all automobile ownership classes; this was especially noticed in the zero-automobile households in Buffalo, where household size declined from 2.6 to 1.8 persons during the 1962-1973 period. This characteristic is also reflected in the growth of single-person households, which rose from 10 percent to 21.5 percent in Buffalo from 1962 to 1973. Nevertheless, trip rates and trip lengths for both oneand two-person households were either stable or rose slightly and overall stability in both trip rates and trip length was noted during the period.

Stability in household trip rates and trip length, as reflected in the Buffalo and Rochester comparisons in the early 1960s to early 1970s, although not updated for post-1974 events, suggests that constant travel per household is a reasonable expectation. Given the declining household size and the moderating gain in the number of automobiles (registrations), then a reasonable expectation for future travel trends is one where travel will increase by about the same magnitude as the growth in households. Under this expectation, the impact in New York State is as follows:

Year	Household Growth Index	VMT Adjustment: Rate	VMT Adjustment: Annual Percentage
1980	1.0	1.0	
1985	1.068	1.068	+1.31
1990	1.132	1.132	+1.25
1995	1.184	1.184	+1.00

Population growth in New York State is projected to be modest--about 8 percent from 1970 to 2000. However, the population declined slightly during the 1970-1980 portion of the period; population growth is projected to be evenly distributed after 1980 (Table 11).

Age Distribution

A gain of about 369 000 is expected in the elderly population category (65 and older) from 1970 to 2000. Estimates of baseline VMT have been projected by using a set of general trip rates for the general

Table 11. Households and population in New York State.

Year	Households (000 000s)	Population (000 000s)	Size of Average Household
1970	5.91	18.24	3.1
1975	6.13	18.08	2.9
1980	6.45	18.08	2.8
1985	6.89	18.34	2.7
1990	7.33	18.76	2.6
1995	7.77	19.23	2.5
2000	8.09	19.71	2.4

Table 12. Adjustment factors for elderly population.

Period	Increase in Elderly Population	Reduction in Travel from General Population (vehicle miles of travel 000 000s)	Reduction in Vehicle Miles of Travel, Final Year Estimate (%)	Vehicle Miles of Travel Adjustment Factor
1980-1985	66 101	271.4	0.391	0.996
1980-1990	164 695	731.6	0.919	0.991
1980-1995	229 765	1020.6	1.235	0.987

Table 13. Factors that will influence travel, 1980-2000.

Factor	Trend Direction	Likely Impact on Travel by 1995 Compared with 1975 (%)
Automobile efficiency	80 percent gain in efficiency, 1975-2000	+40 to +50
Gasoline price	Double real 1978 price by 1995	-40 to -50
Population	Growth in number	+8
Net baseline	Vehicle miles of travel	+45
Projection	Gasoline use	-10 to -20
Energy supply cutoffs	Periodic shortfalls	-10 to -20
Inflation	8-12 percent average over next 15 years	-10 to +10
Employment	Women working	+0.5
	Unemployment rates higher	-3
Households	Growth in number	+18
Urbanization	Increase ruralization	+9
Automobile ownership and use	Increase saturation and use	+5

population. Adjustments can be made to such assumptions if necessary. For instance, data from Albany show that the elderly trip rate there is about half that of the general population. If one wished to transfer the results of the diminished travel for the Capital District's elderly to the increase in the elderly population forecast for New York State, an estimate of travel by the elderly can be made. On the other hand, if the new group of elderly will retain travel patterns that are more consistent with those exhibited by the general population, rather than acquiring travel patterns associated with the existing group of elderly, then no reduction in estimated VMT for the elderly are necessary. The values for possible adjustments for reduced travel by the elderly are shown in Table 12.

Urbanization

The share of population living in the 10 standard metropolitan statistical areas (SMSAs) in New York State is projected to fall slightly from 1970 to 1990. This would suggest that the slightly higher population in the nonurban areas would tend to acquire the higher VMT of those areas and that a proportional gain in travel in these less urban areas would result in an increase of about 9 percent by 1995 (in addition to the change in households). Adjustment factors for urbanization trends are as follows:

Year	Population of SMSA: Urban (%)	Rural (%)	Rural Area Adjustment to VMT
1980	88.1	11.9	No adjustment
1985	87.6	12.4	1.04
1990	87.3	12.7	1.07
1995	87.0	13.0	1.09

Automobile Ownership and Use

The level of automobile ownership per household has increased steadily over time. By 1975 it had reached 1.25 cars/household on the national level. Some researchers (10) expect that the level of ownership will increase based on changing life-styles and household composition. Some (11) expect no further increase, and still others (12) expect an increase well into the latter part of this or the early part of the next century, to be followed by a stable plateau in the level of ownership per household.

The forces behind the changes in automobile ownership lead to changes in the use of automobiles as well. Because VMT per car is projected to level off earlier than the number of cars per household, we find that VMT per household reaches a stable plateau after VMT per car stabilizes but before the number of cars per household does so.

Existing forecasts of VMT can be adjusted in a very straightforward manner to account for changes in the level of automotive ownership or the use of the automobile. The national data described above, however, show too much growth for New York, which is a relatively densely developed urban state in which automobile saturation has already occurred. The table below presents a set of factors based on the growth rates inherent in these projections. These factors can be applied to a baseline forecast that assumes unchanged levels of ownership. Thus, they are most useful for separate studies of isolated links in which effects of automobile saturation have generally not been taken into account.

Year	Level of Ownership	Pattern of Use	Ownership and Use
1980	1.00	1.00	1.00
1985	1.01	1.005	1.02
1990	1.02	1.01	1.04
1995	1.03	1.02	1.05
2000	1.04	1.02	1.06

SUMMARY AND POLICY IMPLICATIONS

The above analysis suggests that broad forces at the national and state level are likely to have significant impact on the magnitude of travel in future years. Table 13 summarizes the likely magnitude of these impacts.

In particular are major changes in the efficiency of cars and the real price of gasoline. The effect of increased car efficiency is to permit more miles to be driven per unit of gasoline, as consumers trade off fuel efficiency and gasoline price. Price rises will reduce gasoline demand in the long run by accelerating fleet turnover further, but in the short run, low price elasticities prohibit major reductions in use through price alone. In the aggregate, these effects are likely to cancel out and result in growth in VMT of 40-50 percent higher than during the 1975-1995 period.

Other important demographic and economic factors could significantly change these projections. Trends toward greater automobile ownership per household but declining household size could add as much as 23 percent to travel projections over the same period; decreasing urbanization could add another 9 percent. But double-digit inflation, higher unemployment rates, and periodic energy supply shortages could more than cancel out these effects. The net direction of all of these factors is difficult to determine, but on the whole, our assessment is that, in spite of higher prices, the likelihood of a stagnant economy, and possible supply shortages, travel is likely to grow, albeit at a slower rate than in the 1960s and 1970s. Gasoline use (already down since 1979) is likely to continue to fall slowly.

U.S. energy policy, which so far has focused on new car efficiency and price decontrol, is generally correctly placed. Specific actions to reduce transportation demand (and hence energy use) through modal diversion or decreased travel have been historically cost-ineffective and will probably remain so. Such actions may be justifiable for other reasons, however.

In general, more attention needs to be placed on such factors than has been the case previously. Projections of travel made in the 1970s are not likely to include most, if any, of these concerns, particularly car efficiency, price rises, and inflation. Some of this work can be done with existing tools, but most of it cannot. New methods are needed that are sensitive to the joint interaction of these variables. Such methods need not be complicated: In fact, simplicity and ease of use are highly desirable attributes. We hope that this paper contributes to that effort.

ACKNOWLEDGMENT

I wish to express thanks and appreciation to those members of the New York State Department of Transportation who provided input to this document. By its nature this document is preliminary. Errors of interpretation, fact, or omissions remain with me. I also wish to recognize the assistance of Karen Rapp and Diana Davis in preparing this manuscript.

REFERENCES

1. Highway Travel Forecasts. FHWA, Nov. 1974.
2. N.S. Erlbaum, D.T. Hartgen, G.S. Cohen. NYS Gasoline Use: Impact of Supply Restrictions and Embargoes, New York State Department of Transportation, Albany, Prelim. Res. Rept. 142, Aug. 1978.
3. D.T. Hartgen and N.S. Erlbaum. Long-Range Forecasts of Transportation Energy Consumption in New York State (Rev. Ed.). New York State Department of Transportation, Albany, Prelim. Res. Rept. 158, April 1980.
4. Annual Automotive Fuel Economy Progress Report to Congress. U.S. Department of Transportation, April 1980.
5. Charles Rivers Associates, Inc. Methods for Analyzing Fuel Supply Limitations on Passenger Travel. NCHRP, Rept. 229, 1980, 132 pp.
6. D.T. Hartgen, A.J. Neveu, J.M. Brunso, J.S. Banas, and J. Miller. Changes in Travel in Response to the 1979 Energy Crisis. New York State Department of Transportation, Albany, Prelim. Res. Rept. 170, Dec. 1979.
7. D.T. Hartgen and A.J. Neveu. The 1979 Energy Crisis: Who Conserved How Much? In Considerations in Transportation Energy Contingency Planning, TRB, Special Rept. 191, 1980, pp. 157-165.
8. C.E. Meyers. NYS Transportation Energy Data. New York State Department of Transportation, Albany, Rept. No. 4, Nov. 1980.
9. G.S. Cohen and M.A. Kocis. Components of Change in Urban Travel. TRB, Transportation Research Record 775, 1980, pp. 42-47.
10. K.-W.P. Koeppel. Trends in Automobile Ownership Patterns, 1973-1977. TRB, Transportation Research Record 723, 1979, pp. 1-6.
11. G. Shunk and C. Loxley. An Analysis of the Automobile: Modeling the Long Run Determinants of the Demand for Automobiles. Wharton Economic Forecasting Associates, Inc., Philadelphia, Feb. 1977.
12. D.P. Sharp. Projections of Automobile Ownership and Use Based on Household Life-Style Factors. Univ. of Pittsburgh, Ph.D. Dissertation, 1978.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

The Shape of the 1980s: Demographic, Economic, and Travel Characteristics

FRANK SPIELBERG, EDWARD WEINER, AND ULRICH ERNST

Forecasts of economic and demographic conditions are the base for all forecasts of travel demand. During the 1970s many changes were observed in the demographics of the nation. This paper reviews the trends in pertinent demographic measures and projects the directions of these measures through the 1980s. The objective is to determine how transportation demand is likely to change.

Transportation analysis is based on the premise that demographic, social, and economic factors are major determinants of travel demand. For the past 25 years, metropolitan planning organizations throughout the world have conducted surveys of travel, performed analyses, and estimated models of travel demand, distribution, and mode choice. The projections of future conditions forecast by these models have been used to guide decisions on investments in new and improved transportation facilities.

Transportation planners have devoted extensive effort to ensuring that their models were statistically valid, replicated base-year conditions, and produced reasonable forecasts for future years. However, the validity of the forecasts of travel produced by any of these models can only be as good as the forecasts of the basic parameters on which the models depend (e.g., population and workers). Projections of these factors have most often not been the responsibility of the transportation professional; rather, they are most often developed by other staff members and used with only cursory review as inputs for the transportation forecasts.

Fawcett and Downes (1) showed that the projections of travel obtained by use of the carefully developed models are far more sensitive to changes in the values of the social and demographic parameters than to misspecification of the model or slight variation in the model coefficients.

During the past decade, significant changes have occurred in many of the social and demographic characteristics of the U.S. population. As we enter the 1980s, it is useful to identify these national trends in factors related to transportation and to examine some conjectures as to their likely effect.

APPROACH

To obtain a picture of trends in 1980s, a wide range of existing forecasts and analyses was reviewed. Evaluations were made of the most reasonable forecasts based on the methodology used, consistency with other information, historical trends, and forecasts of other factors.

Travel information was developed by tabulating data from the 1977 national personal transportation survey (NPTS). These data were stratified by household size and location. Average weekday trip rates by mode were calculated for each cell.

Economy

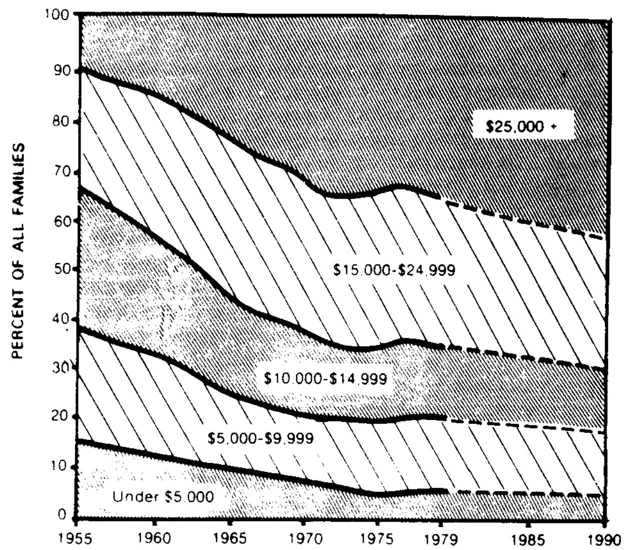
At the base of any projection of travel, explicit or implicit assumptions are made about the economic conditions of the area under study. Projections of economic growth and its characteristics typically are no more than attempts to explore the implications of a particular set of assumptions about various key factors. Changes in population, labor force, productivity, hours worked, tax policies, and exports are the forces that determine changes in the gross national product (GNP), personal income, and investments.

The period from 1950 through 1973 (during which time the data used for most U.S. travel models were collected) was a time of steadily increasing affluence combined with steadily decreasing costs for automobile travel. Continuous increases in per capita travel were observed in almost all groups in the U.S. population. The direction of economic trends has been less clear in the last half of the decade. The rate of increase in the GNP has been reduced, and for individual households, the primary issue has been how to increase or maintain real income in the face of significant inflation. GNP growth rates are expected to fluctuate about the average of 3.6 percent/year through the 1980s in real terms. The growth rates for personal income will mirror this pattern. Inflation, as measured by the consumer price index, is projected to continue at 8-9 percent/year throughout the decade.

Family Income

During the 1970s the relative growth of the upper-income brackets that occurred in the 1960s slowed down [Figure 1 (2)]. Since 1975, there has been an increase in the percentage shares of the two higher income categories. Changes in the lower-income brackets were almost negligible. Trends over the last two decades and current economic prospects suggest that the income distribution will change little during the 1980s. The percentage share of the highest income group is likely to increase slightly. Smaller upward shifts should affect the other categories. But, overall, the experience of the 1970s appears a better-suited model than the rapid gains of the 1960s. The income distribution forecast for the 1980s shows that close to 40 percent of U.S. families will have incomes below $15 000. These families are apt to be severely affected by real increases in fuel costs.

Figure 1. Family income distribution (in 1978 dollars).



Total Population

The total population of the United States will grow at a modest rate during the 1980s. In the second half of the 1970s, the average annual growth rate was 0.73 percent--a full percentage point below the rate for the late 1950s. Current census projections for 1990 imply average annual growth rates between 0.6 and 1.3 percent, depending on assumptions about fertility trends. Most experts expect that fertility will rise slightly from its current low to the replacement level (the level at which the population would exactly replace itself in the absence of net immigration). At that level the U.S. population is projected to grow at a rate of 0.9 percent/year, from 220.5 million in 1979 to 243.4 million in 1990.

Population Change

Of the three components of change in the total population (live births, deaths, and net immigration into the country), fertility trends have been the key to recent growth patterns [Figure 2 (3)]. Mortality rates have been relatively stable over the last few decades. There has been a recent downturn that has been broadly sustained since 1973. Death rates for persons 65 years old and over (who account for about 90 percent of all deaths) declined much more from 1973 to 1978 than from 1960 to 1973, primarily as a result of a reduction in mortality from major cardiovascular diseases. This change reflects improvements in health care as well as changes in lifestyle.

Figure 2. Components of population change.

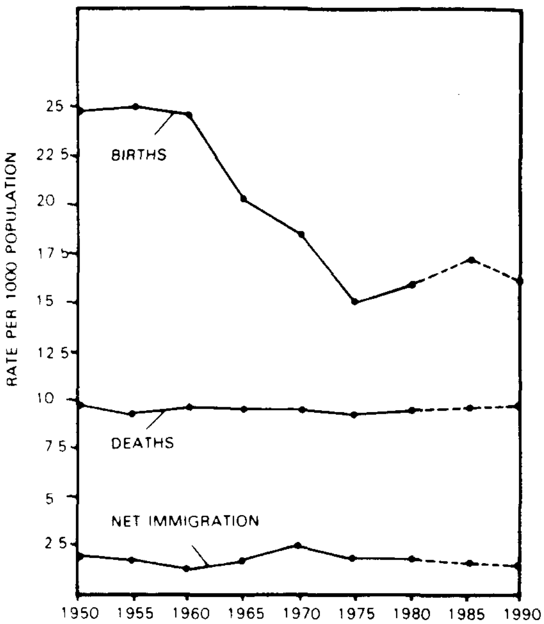
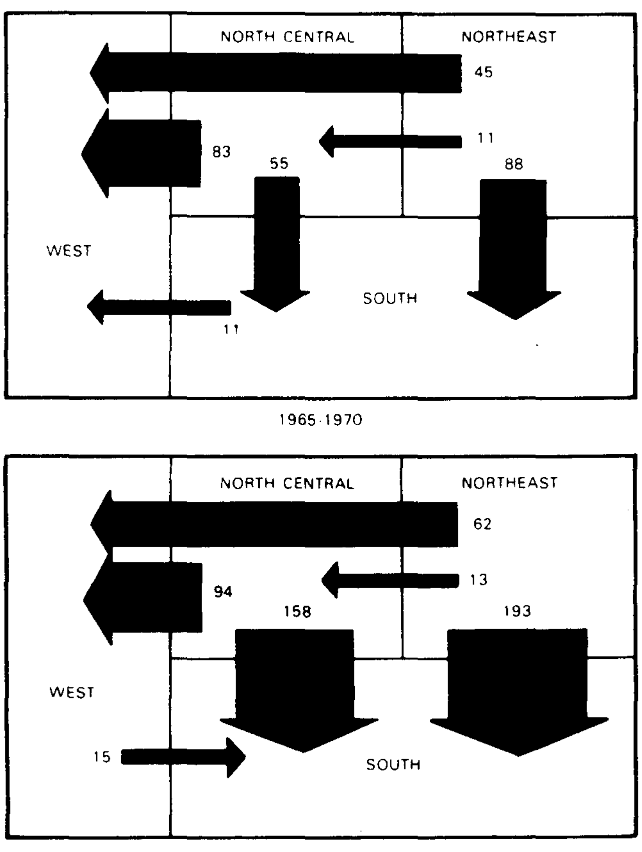


Figure 3. Net migration among census regions (annual averages, 000s).



The slowdown in population growth reflects a drop in birth rates. This drop is primarily the result of a decline in fertility rates for women across all age categories. The general consensus is that changing aspirations of women and couples have caused this decline in fertility rates. A popular assumption is that the decline in fertility rates in the younger age groups is the result of the postponement of childbearing and that the birth rate could rise again in the near future as many young women begin to make up desired births delayed from previous years. It is doubtful, though, whether the effect would lead to a substantial increase in the birth rate over the next few years. All indications are that fertility rates will increase slightly and approach the level at which the population replaces itself (in the absence of net immigration). This projection is consistent with most recent survey data of the birth expectations of young married women.

Regional Redistribution

The key factor in the regional redistribution of the U.S. population has been migration. The 1970s brought a radical shift in several long-term migration trends. First, the net out-migration from the South reversed. The South became the region with the largest in-migration. At the national level, however, total net annual regional change amounts to only 0.25 percent of the total population [Figure 3 (4,5)]. Second, the historical trend toward expanded urbanization reversed. Each past decennial census up to and including the 1970 count showed that a greater percentage of the population lived in urbanized areas. In the 1970s, for the first time, the number of migrants from urban to nonmetropolitan areas exceeded the number who moved in the opposite direction. Third, the rate of migration into the central cities from the suburbs increased in the second half of the 1970s. Migration from the suburbs to the central cities in the three years from 1975 to 1978 reached almost the level for the five-year period 1970-1975--3.6 versus 3.8 million. This relative acceleration may be an indication that the often-cited return to the city is more than a series of isolated phenomena.

As with regional migration, however, the net rate of migration is still quite small compared with the total population. In fact, in absolute terms, the annual rate of migration from central cities to suburbs was stable throughout the decade. The rate of turnover has increased, new groups that have smaller household size are moving to central cities, and larger households continue to seek the suburbs.

Population Growth by Type of Residential Area

Although the parameters of migration among types of residential areas have begun to change, the relative magnitude of migration flows implies a continuation of current growth trends through the 1980s. The full impact of these changes will be felt in the years that follow the next decade.

The projections shown here identify the suburbs as still the key growth area. Population in the suburbs (the noncentral portions of metropolitan areas) is projected to increase from about 63 million in 1975 to 86 million in 1990. This growth results in an increase in the share of the suburbs of the total population--from less than 30 percent in 1975 to more than 35 percent in 1990 [Figure 4 (6)]. Central cities are projected to show an absolute increase in the population--from 67 million in 1975 to 72 million in 1990. Their share in the total U.S. population, however, will decrease slightly--from 31 to 30 percent.

The population living in small-urban areas (cities that have a population between 2500 and 50 000) is expected to remain stable. Consequently, the percentage share of this category will decrease. The rural population is projected to decrease in ab-

Figure 4. Distribution of population by area.

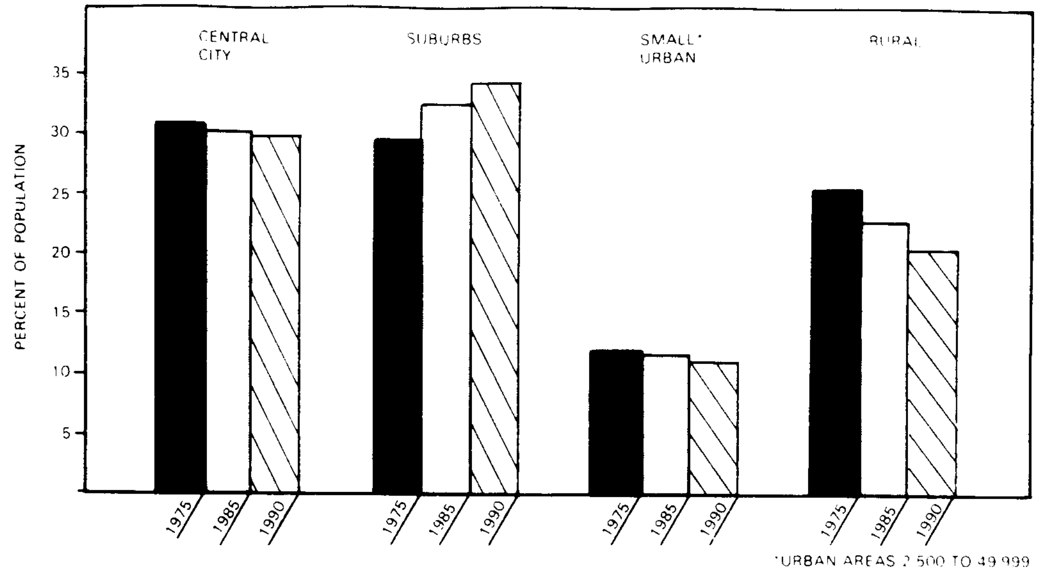
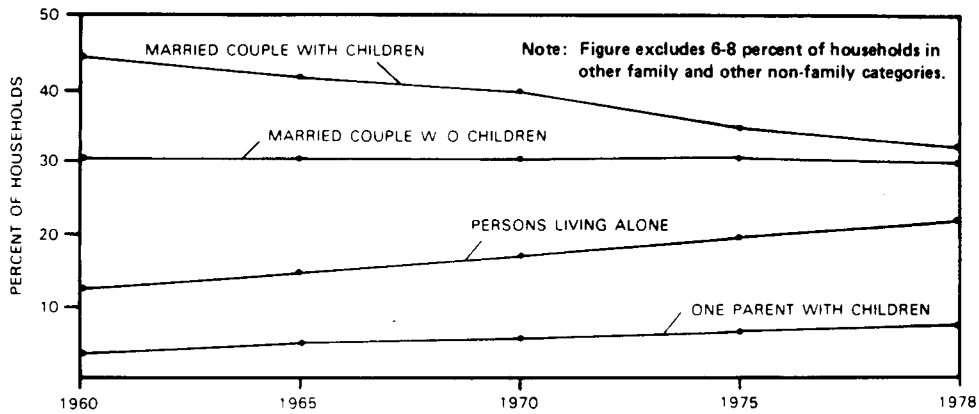


Figure 5. Trends in household composition.



solute terms. The percentage share of rural areas in the total population is projected to decline from more than 25 percent to under 22 percent. If recent shifts in migration patterns continue throughout the 1980s, this decline in the relative importance of the rural areas is likely to be milder.

Throughout the next decade the suburbs will continue to be the residence of the largest segment of the population.

Trends in Household Composition

The percentage of households that contain married couples who have children has declined from 44 percent in 1960 to 32 percent in 1978 [Figure 5 (7,8)]. Over the same period, the percentage of all households that consists of married couples who do not have children has remained constant at about 30 percent. The most significant gain has been in the category of single-person households. In 1960, persons who lived alone accounted for 13 percent of all households. By 1978 this percentage had increased to 22 percent. In the 1980s the growth in single-person households is likely to slow down only marginally, because the number and proportion of elderly in the population (who account for a large fraction of single-person households) will continue to increase.

Households that do not contain children have increased substantially since 1960 and are expected to represent an even larger proportion of all households by 1990. Such households place less emphasis on schools or the availability of play space in their decisions about residential location. Higher-density, central living will be more acceptable for this group.

This conjecture is supported, in part, by preliminary data from the 1980 census that show that, although central cities are experiencing a loss of population, they are, at the same time, showing growth in the number of households. Thus, we expect that in the 1980s the central cities will retain their relative share of households, but the suburbs will contain the larger proportion of population.

Labor Force

The total U.S. labor force is projected to increase from 102 million in 1980 to 114 million in 1990, which corresponds to an average annual growth rate of 1.1 percent/year. The major driving force behind this increase in the total labor force is the growth in the female labor force. From 1970 to 1980 the male labor force has grown at an average annual rate of 1.6 percent. Over the same time period, the female labor force has grown at a rate of 2.8 percent/year [Figure 6 (2)]. The growth in the labor force is expected to slow down for both groups during the 1980s--to a rate of 0.8 percent/year for the men and 1.5 percent/year for the women.

The fast overall growth in the past decade was largely attributable to the entry of the baby boom

Figure 6. Percentage of husband-wife families with a working wife.

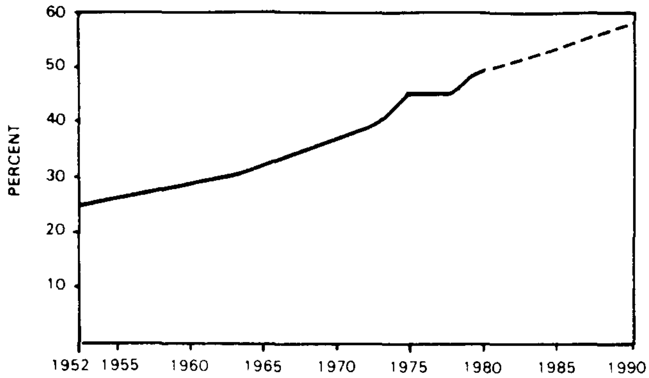


Table 1. Effect of female employment on household trip rates in 1977.

Household Size	Employed Persons Male	Employed Persons Female	Weekday Trips/ Household	Increase In Trip Rate
2	1	0	4.33	1.65
2	1	1	5.98	
3	1	0	5.73	1.64
3	1	1	7.37	

generation. This group no longer plays a role in the growth of the labor force. Significant growth in the 1980s is primarily due to the increased participation by women in the labor force.

Rate of Participation in the Labor Force

The long-term decline in the rate of participation in the labor force for men is expected to taper off in the 1980s. This rate is projected to stay virtually constant. In contrast, the rate of participation in the labor force for women is projected to continue its recent climb throughout the first half of the 1980s and to level off somewhat during the second half. By 1990, more than 51 percent of all women are expected to be in the labor force, which is up from 43 percent in 1970.

The growth of female participation in the labor force is the result of a series of complex factors. The implications for the social and economic fabric of our society and for changing demands on public services are only partially understood. The addition of another wage earner in existing households and the continued formation of small households as a result of greater economic independence of women require a reassessment of current notions about the travel behavior and transportation needs of households.

Working Wives

The proportion of married women who enter the labor force is projected to increase. This trend results not only from a change in attitudes regarding working wives but also from economic reasons, an expanding economy, equal-opportunity laws, increased female levels of education, and a corporate climate that provides women with improved wages and seniority.

Household Trip Rates

The effect of increased female employment on trip rates has been an open question during the 1970s. Some have suggested that the result of devoting many hours each day to employment would be to reduce the overall household trip rate. Data from the 1977 NPTS [Table 1 (9)] suggest that the results of increased participation in the labor force by married women is an absolute increase in travel. It appears that devoting time to employment does not suppress nonwork travel--rather, roughly 1.65 trips/weekday are added. This is almost exactly the number of work trips expected per employed person.

Migration of the Work Force

During the 1970s, heavily populated areas (the central counties of metropolitan areas that have a population of 2 million or more) experienced a net out-migration of jobs at a rate of almost 100 000/year. The situation was different for the smaller central counties. Those that have a population of 1 to 2 million people experienced a net in-migration of workers from other areas. Between 1970 and 1973, central counties that had a population between 500 000 and 1 million registered substantial job in-migration; by 1973-1976, this situation had changed to net out-migration.

The primary beneficiary of this change was the category of central counties that have a population under 500 000. The net in-migration of jobs almost tripled from 1970-1973 to 1973-1976, to a rate of about 50 000 workers/year. The remaining two categories, noncentral counties in metropolitan areas and nonmetropolitan counties, show a net in-migration in both periods. In both cases, total net in-migration increased. The relative increase is particularly significant for the nonmetropolitan counties.

These trends in geographical changes in job location are likely to continue into the 1980s. The dispersal of residential location has been followed by the dispersal of the location of the place of work. Although this change is gradual (less than 0.5 percent of the total labor force), it could begin to affect work travel patterns. The centers of metropolitan areas could continue to lose importance as a destination for the work trip.

Real Cost of Gasoline

In the absence of any new crises, such as those in Iran, total petroleum production during the 1980s is projected to increase from about 50 million barrels/day. The gap between world production and demand is projected to be small (about 200 000 barrels/day in 1980) but to rise steadily through the decade. Sporadic restriction of supply should also be expected during the 1980s.

Since demand will exceed supply, the price of a barrel of oil will continue to increase and reach almost $70/barrel by 1990. The price differential between domestic and imported crude will be eliminated by 1982. The increase in petroleum prices is expected to be accompanied by continued inflation. In the early fall of 1979 projections were that the 1990 barrel price (in 1979 dollars) would be on the order of $32. Subsequent events suggest that this projection might be low.

As shown in Figure 7 (10,11), the cost of gasoline relative to all items declined significantly between 1960 and 1973. During the 1973-1974 embargo the cost rose sharply but then once again declined slowly in real terms until the sharp price increases of 1979. Due to the continued increase in the real price of crude oil, the price of gasoline relative to all items is expected to continue its real increase through the 1980s, although the rate of increase will slow in the later half of the decade.

Figure 7. Price of gasoline relative to all items.

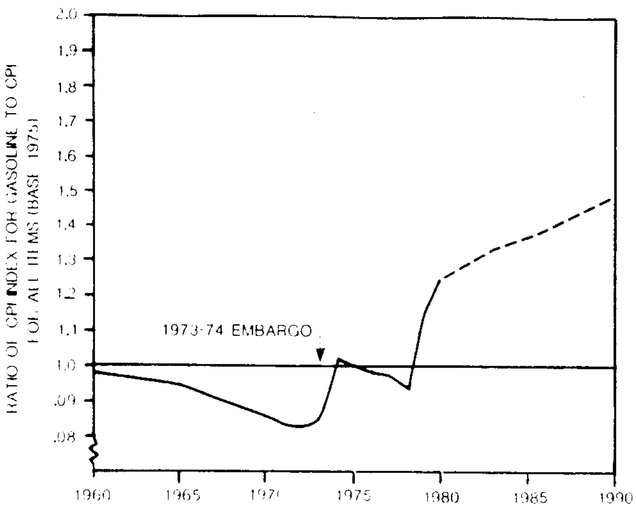


Figure 8. Automobile fleet average fuel economy.

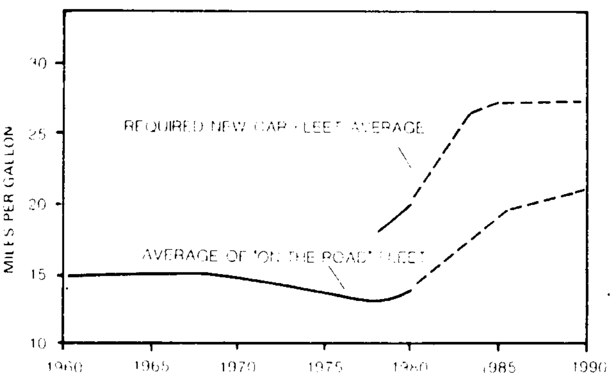
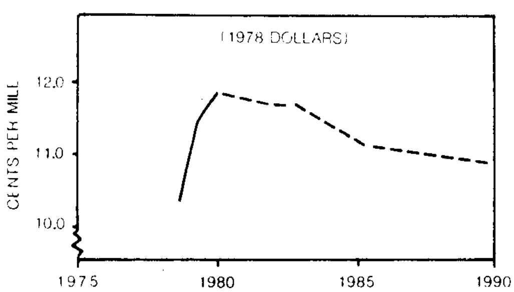


Figure 9. Automobile operating costs per mile.



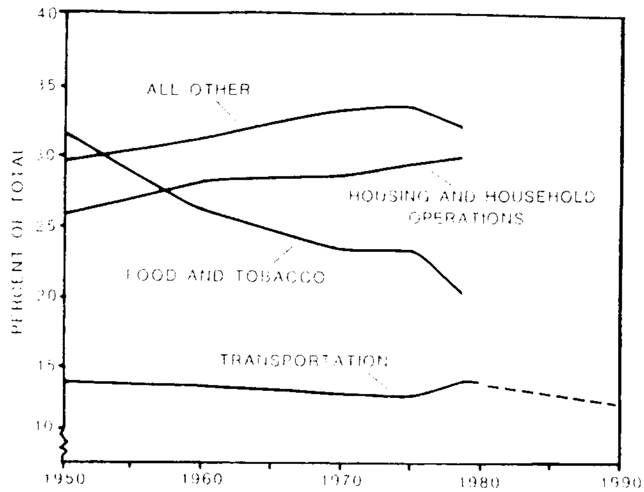
Note: Figure includes variable costs only; does not include fixed items such as cost of vehicle purchase, taxes, and licenses.

By 1990 the real cost of gasoline is projected to be approximately 1.5 times the 1979 cost. This forecast indicates that trends to smaller cars will continue in an effort to maintain existing life-style in the face of rising fuel costs.

Cost of Automobile Operation

The cost to a household of automobile operation depends on both the cost of gasoline and the fuel economy of the vehicles available. The fuel efficiency of the U.S. automobile fleet declined from

Figure 10. Transportation expenditures as a percentage of total personal consumption expenditures.



about 14.5 miles/gal in 1960 to a low of about 13.5 miles/gal in 1975 [Figure 8 (6)]. Mandated improvements in the fuel efficiency of new cars have begun to increase the average efficiency of the fleet. By 1980 the efficiencies of the early 1960s will be surpassed. The continued increase in new car efficiency coupled with the retirement of older vehicles will result in fleet efficiencies of 19 miles/gal by 1985 and almost 23 miles/gal by 1990.

The rate of increase in fleet efficiency is projected to be greater than the rate of increase in the real price of gasoline for the 1980s. Although the real cost of gasoline may rise more rapidly than shown in these projections, increases in fleet efficiency would compensate for price increases of 4 percent/year. The combined effect on automobile operating costs will be an increase through 1980, followed by stabilization in 1981-1982, and a decline in real terms for the remainder of the decade [Figure 9 (10)]. Real operating costs per mile in 1990 may be slightly less than in 1979.

Automobile operating costs in the 1990s will again increase. By 1985 new cars will have achieved fuel economy standards that can be easily implemented. Major technological innovation will be required to obtain additional efficiency. The gap between petroleum demand and production is also expected to widen more rapidly in the 1990s and lead to either higher real prices or supply restrictions.

Although the 1980s may be a period of relative stability, actions taken during the 1980s must recognize the likely problems of the next decade.

Transportation Budget

Between 1950 and 1970 the proportion of personal consumption expenditure devoted to transportation declined only slightly (from roughly 13 percent to 12 percent), even though the cost of travel in real terms fell substantially. A slight rise in the proportion of the budget devoted to transportation in 1974 was followed by a major increase from 1975 to 1977, when it reached a 20-year high of 14.3 percent [Figure 10 (12)].

How transportation expenditure will change in the 1980s is unclear. The conjecture shown in Figure 10 suggests that household travel expenditure will decline slightly during the 1980s, and households will seek to reestablish the historical values of 12-13 percent:

1. Studies have suggested that households over a wide range of urban areas and living conditions have a cost budget for transportation, and
2. Although the costs of gasoline will increase in real terms during the 1980s, the cost per mile of automobile travel will decline slightly as fleet efficiency improves.

The post-1974 increase in the transportation budget reflects the fact that households require some time to change dwelling locations and trip patterns. During the 1980s long-term decisions about residential location and activity patterns will be made that reflect travel costs. However, 80 percent of the housing units that will exist in 1990 are already in place. The availability and cost of certain types of transportation were implicit in the development of this housing. The residents of this housing will be forced to live with the economic and mobility consequences.

Nominal price increases or occasional supply restrictions during the 1980s may lead to some household adjustments. Households that do not change patterns in the 1980s should be able to maintain their transportation budgets within the historic range if household income keeps pace with inflation. Beyond 1990, due to real increases in operating cost, these households will need to either increase the proportion of income devoted to transportation or reduce travel. Within limits, the latter option is easier and more likely to be chosen. Some trips must be made, however, and in the 1990s a growing proportion of households will be seeking cheaper travel alternatives.

Gasoline Expenditures

The amount of money spent on gasoline is a direct function of travel. As would be expected, larger households make more trips and, therefore, have higher gasoline expenditures [Figure 11 (9)]. Similarly, higher-income households travel more and spend more on fuel, although the percentage of income devoted to gasoline expense shows a rapid decline with increasing income. Above the median income, gasoline expenditure increases more slowly than does income. Other factors, such as the amount of time available to travel, serve to constrain trip rates. These higher-income households could absorb significant real increases in gasoline prices without exceeding the average amount of the budget devoted to transportation.

The evidence, however, also suggests that lower-income households are already constrained by the cost of operations. The mobility of this group could be even more restricted as real costs increase.

CONCLUSIONS

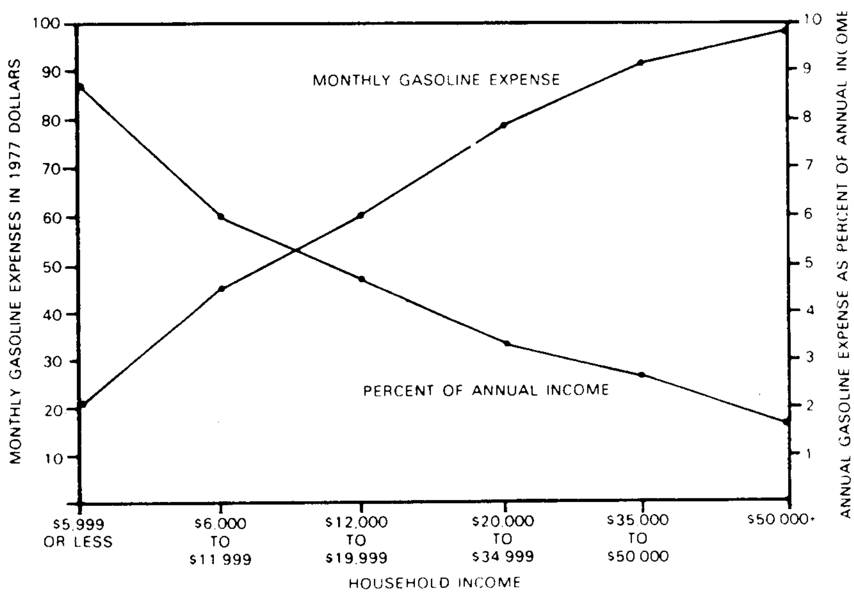
Based on analysis of trends observed during the 1970s and their projection to the next decade, the following findings are particularly significant:

1. Continued inflation will result in a halt in the rise of real income;
2. Growth in population will continue to slow;
3. Family size will decrease and the proportion of single-person households will increase;
4. Although some smaller households will seek dwelling units in higher-density areas (central cities and older suburbs), the majority of growth in both population and households will occur in areas characterized as suburban;
5. The proportion of women employed will continue to increase; and
6. Continued increases in the real cost of gasoline will be coupled with significant increases in vehicle fuel efficiency so that the per mile cost of automobile operation will be stable for the population as a whole.

Perhaps most significant, however, is that, in spite of definite trends in migration to the South, to nonmetropolitan areas, and, to a lesser degree, to central cities, the absolute magnitude of the net changes will be small compared with the total population. The vast bulk of the U.S. housing stock is in place. Substantial net change in residence locations over the decade is impractical.

Our view of the 1980s may be summarized as follows. Population in the South and West will continue to grow at a faster rate than will the nation as a whole, and most of this growth will occur in the lower-density developments that characterize new housing areas. The bulk of the population will live in suburban locations, dependent for travel on automobiles. Fuel costs will continue to increase in real terms so that households will take a variety of

Figure 11. Characteristics of household gasoline expenditure.



actions to maintain their transportation budgets within the historical range. The most popular action will be the purchase of more fuel-efficient vehicles. As a result, vehicle miles of travel in all areas will continue to increase, although the rate of growth will be slower than that observed in the past. Nonetheless, lower-income groups will face increased restrictions on their mobility.

Households that do not include children will increase as a proportion of the total. These households will find higher-density living more acceptable, and many will choose locations within the central city or older suburbs. These households, which frequently have two working adults, will find transit acceptable for many trips. As a result, transit ridership as a proportion of total travel will stabilize and grow in absolute numbers.

From a transportation viewpoint, these trends and forecasts indicate that the northeastern and midwestern regions will have to adapt to a low-growth future and concentrate on selective revitalization and rehabilitation of existing highway and transit facilities. The southern and western regions will need to encourage new development to occur at higher densities so that they can better serve travel with transit. Nationwide, extensive areas of low-density development will still exist, where paratransit options will be the only stable alternatives to single occupant use of the automobile. Ridesharing in carpools, vanpools, or taxicabs will be the most cost-effective transportation option. New institutional arrangements will be needed so that providers of such services can enter the market.

REFERENCES

1. F. Fawcett and J.D. Downes. The Sensitivity of Trip Generation Forecasts to Planning Inputs. Transportation and Road Research Laboratory, Crowthorne, Berkshire, England, TRRL Supplementary Rept. 389, 1978.
2. Current Population Reports. Bureau of the Census, U.S. Department of Commerce, Series P-60, No. 120, Nov. 1979.
3. Current Population Reports. Bureau of the Census, U.S. Department of Commerce, Series P-25, No. 802, May 1979.
4. B.J. Berry and D.C. Dahlmann. Population Redistribution in the United States in the 1970s. Population Development Review, Vol. 3, No. 4, Dec. 1977, p. 458.
5. Current Population Reports. Bureau of the Census, U.S. Department of Commerce, Series P-20, No. 331, Nov. 1978.
6. Office of Technology Assessment. Technology Assessment of Changes in the Future Use and Characteristics of the Automobile Transportation System, Vol. II: Technical Report. U.S. Congress, Feb. 1979.
7. Current Population Reports. Bureau of the Census, U.S. Department of Commerce, Series P-20, No. 340, July 1979.
8. Current Population Reports. Bureau of the Census, U.S. Department of Commerce, Series P-25, No. 805, May 1979.
9. Nationwide Personal Transportation Survey. FHWA, 1977.
10. Data Resource, Inc. Energy Review. Autumn 1979.
11. Transportation Energy Conservation Data Book (3rd Ed.). Oak Ridge National Laboratory, Oak Ridge, TN, Feb. 1979.
12. Survey of Current Business. Bureau of Economic Analysis, U.S. Department of Commerce, Nov. 1979.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

Estimating Vehicle Miles of Travel: An Application of the Rank-Size Rule

WAYNE R. UGOLIK

This paper suggests a simplified approach to the problem of estimating annual vehicle miles of travel without the need for extensive vehicle count data. By using detailed data on vehicle miles of travel per highway section of the New York State touring route system, it is shown that, when the highway sections are ranked in decreasing order of their section vehicle miles of travel, the vehicle miles of travel of the individual sections can be closely approximated by a function of section rank value. An approximation of the total sum of vehicle miles of travel on all highway sections is then obtained by integrating this function over all the rank values. The approach has potential for use as a forecasting tool. Tests on data for 1968, 1974, and 1976 show that the method can produce surprisingly accurate results.

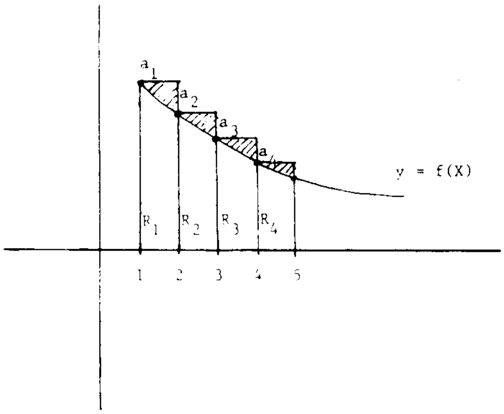
In 1970 Zahavi (1) noted that, when certain transportation-related quantities are ranked in decreasing order of value, a certain level of stability is attained in the relationship between rank and value that enables analyses that might not otherwise be possible. This relationship, well known in the sciences, is called the k-distribution. In this paper we explore a similar pattern of stability that is exhibited by a ranking (in decreasing order) of the annual vehicle miles of travel per section of the New York State touring route highway network. This stability enables us to develop a method for estimating annual vehicle miles of travel on the state route system by a simple scheme that has the potential to circumvent the costly and extensive vehicle-counting procedures that are currently in wide use. The method is based on what we term the rank-size rule. Its utility for producing accurate estimates is illustrated by an application to New York State vehicle miles of travel data for 1968, 1974, and 1976.

RANK-SIZE RULE

The rank-size rule can be described mathematically as follows:

1. Let a_1, a_2, a_3, ..., a_n be a listing of positive numbers with $a_1 \geq a_2 \geq a_3 \geq \ldots \geq a_n$; that

Figure 1. Graphical representation of the rank-size rule.



is, the numbers are ranked in decreasing order and

2. Suppose f(X) is a positive and decreasing function defined for $X \geq 1$, that has the property that $f(j) = a_j$ for $j = 1, \ldots, n$. Then,

$$\sum_{j=1}^{n} a_j \approx \int_{1}^{n+1} f(X)\,dX \qquad (1)$$

This can be easily understood from Figure 1. In the figure, the dot on the top left corner of each rectangle R_j is a_j units above the horizontal axis, and each rectangle has a width of one unit. Therefore, the area of each rectangle is a_j (square units). Hence, from Figure 1, it is clear that $\sum_{j=1}^{4} a_j$ equals the sum of the areas of the rectangles, and this area can be approximated by the area under the curve y = f(X) for X varying from 1 to 5. This area under the curve is precisely the value of $\int_1^5 f(X)dX$. The approximation will be quite good if the a_j's are large relative to measurement on the horizontal axis and if the rate of decrease of the a_j's is relatively low. This follows because the hatched area in each rectangle in Figure 1 would then be small relative to the value of a_j.

Application

Summaries were available of annual daily New York State highway system vehicle miles of travel, based on the 1968, 1974, and 1976 New York State highway sufficiency reports (2-4). The annual average daily vehicle miles of travel (AADVMT) is based on the annual traffic volume determined for each section of every touring route in the state network. A section is a particular length of highway, usually of uniform width; however, actual lengths of sections can vary greatly. The AADVMT is calculated for each section by multiplying section length by the traffic volume for that section. The sections (approximately 15 000 in number) were then arranged on the basis of descending section AADVMT, and successively aggregated into groups of 100 sections each in order to make the number of data points more manageable. The first group contains the first 100 top-ranked highway sections, the second group contains the second 100 top-ranked highway sections, and so on. A group usually contains sections from all over the state so that there is no general geographic pattern among the groups.

The AADVMT for each group (100 sections) was then calculated by summing over the sections in the group. The result for each of the years 1968, 1974, and 1976 was a ranking of (approximately) 150 groups based on the descending order of group AADVMT.

For each year, AADVMT per group rank was plotted vertically on a logarithmic scale against group rank. The resulting plots exhibited a negative exponential decay, as exhibited in Figures 2-4, which suggests that

$$\ln(\mathrm{AADVMT}_{\mathrm{group\ rank}}) \approx ae^{-b \cdot \mathrm{group\ rank}} \qquad (2)$$

or a double exponential of the form:

$$\mathrm{AADVMT}_{\mathrm{group\ rank}} \approx \exp(ae^{-b \cdot \mathrm{group\ rank}}) \qquad (3)$$

In this form, the a and b in the exponent are parameters that would vary from year to year, and e is the base of the natural logarithm.

If we assume that the approximation in Equation 3 is good, the rank-size rule would require evaluation of the integral

$$\int_{R_L}^{R_H+1} \exp(ae^{-br})\,dr$$

in order to estimate

$$\sum_{\mathrm{rank}=R_L}^{\mathrm{rank}=R_H} \mathrm{AADVMT}_{\mathrm{rank}} \qquad (4)$$

where R_L is the lowest group rank and R_H is the highest. This integral is not solvable in closed form by elementary functions, but its value can be approximated with a high level of accuracy by integrating a series approximation to the integrand after a change of variables or by using a numerical integration scheme such as Simpson's rule.

The estimation of AADVMT by such integrals would be facilitated if the parameters in the exponents did not vary from year to year but rather that the variations in AADVMT were accounted for by a correction factor, depending on the year. A multiplicative correction factor (A) would require

$$\mathrm{AADVMT}_{\mathrm{rank}} \approx A\exp(ae^{b \cdot \mathrm{rank}}) \qquad (5)$$

an additive correctiion factor (A) would require

$$\mathrm{AADVMT}_{\mathrm{rank}} \approx A + \exp(ae^{b \cdot \mathrm{rank}}) \qquad (6)$$

In either form, the constant A would generally vary from year to year, and the a and b would remain constant. A multiplicative factor alters the shape of the curves from year to year; an additive factor does not. In fact, given historical data (increasing vehicle miles of travel), a multiplicative factor would necessarily make the curves progressively steeper. Since this was not indicated by the data, the additive-factor approach is preferable, which indicates Equation 6, where A depends on the year, but a and b do not. In this form the rank-size rule becomes

$$\sum_{R_L}^{R_H} \mathrm{AADVMT}_r \approx A(R_H + 1 - R_L) + \int_{R_L}^{R_H+1} \exp(ae^{br})\,dr \qquad (7)$$

Assuming that the approximation in Equation 6 can be applied accurately, values of the integral on the right-hand side of Equation 7 for various ranges of the rank could be tabulated numerically and used for different years because a and b remain fixed. Thus, the rank-size rule in this form would produce an estimate of AADVMT based on a little simple arithmetic after the parameter A is estimated. In what follows, we indicate that this task may not be as difficult as it may seem.

Figure 2. Log AADVMT by rank for 1968.

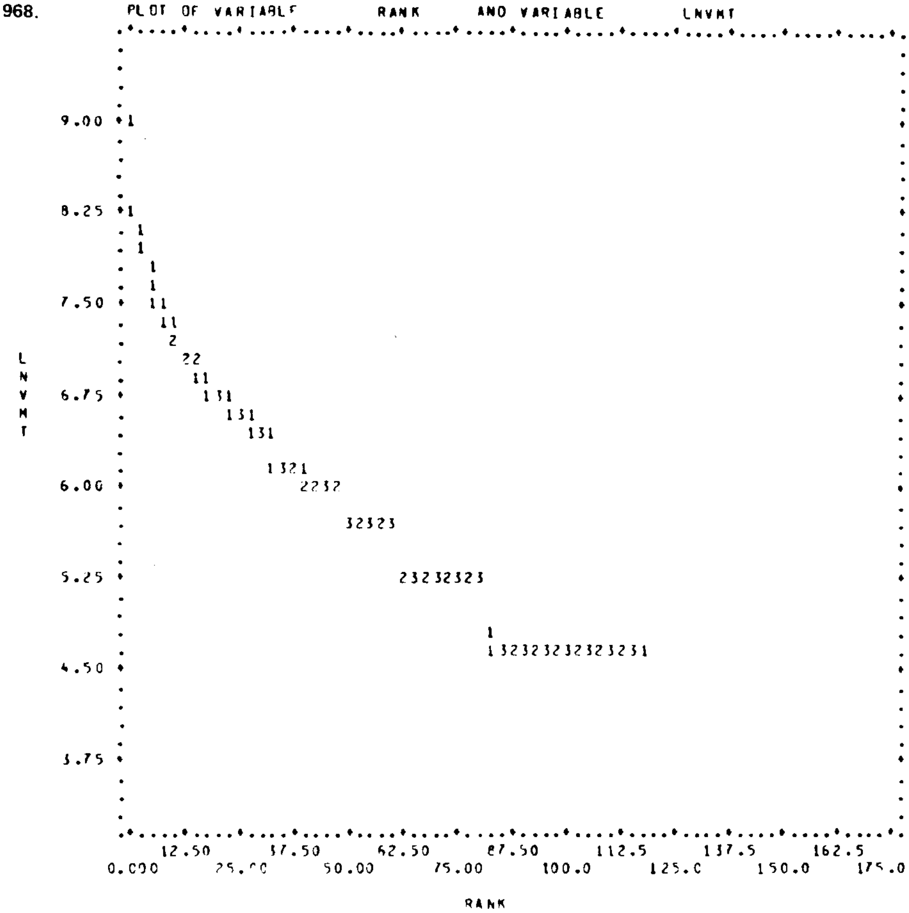


Figure 3. Log AADVMT by rank for 1974.

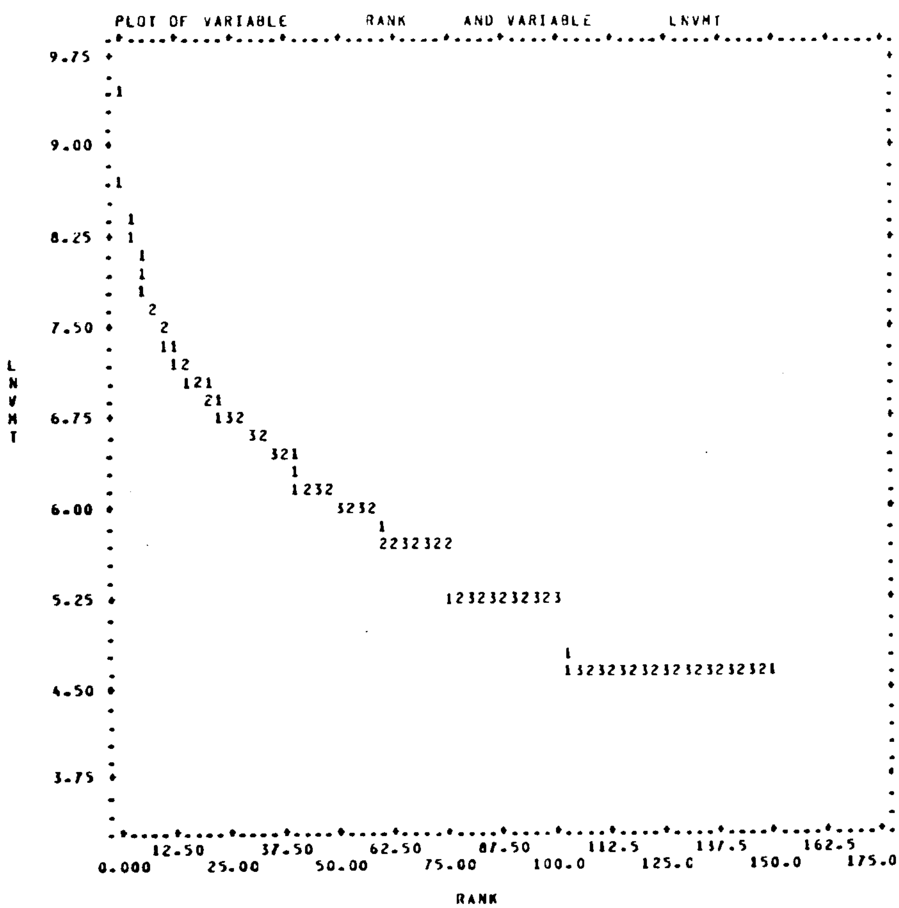
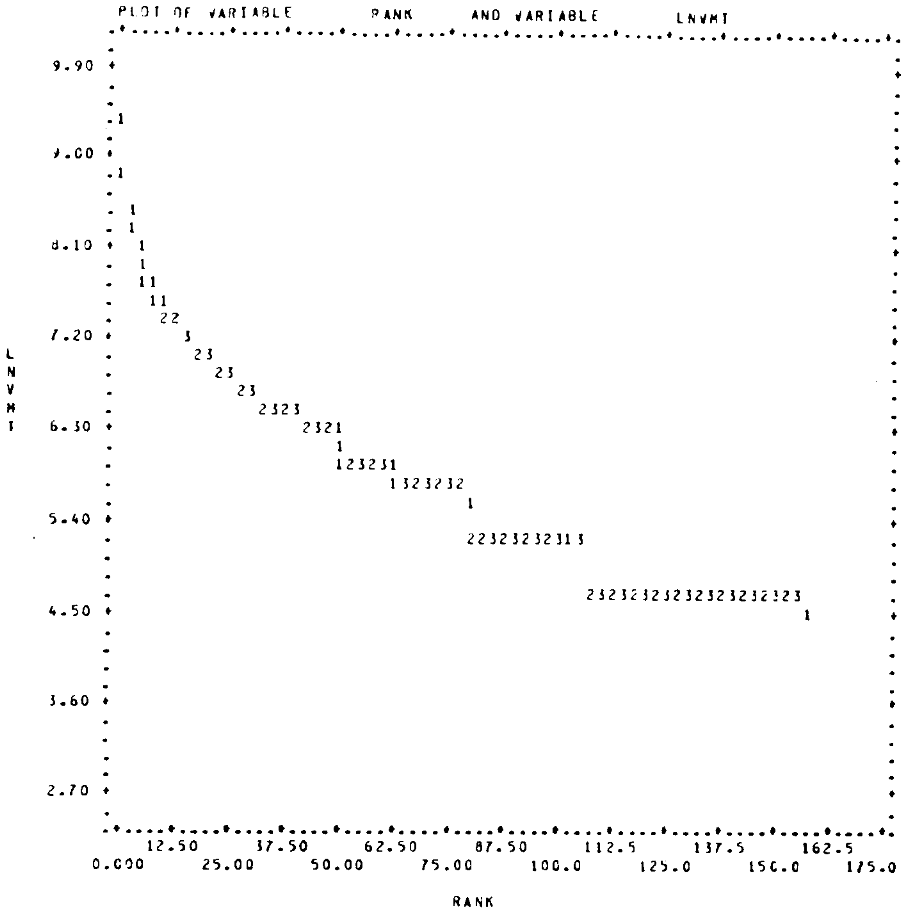


Figure 4. Log AADVMT by rank for 1976.



Calibration

Estimation of the constants a and b, and the shift parameters (A) was accomplished by nonlinear least squares by using a computer program available in the Biomedical Data Processing (BMDP) software system (5). The functional form (Equation 6) was calibrated on the 1968 data. The same a and b values obtained for the 1968 data were used for the 1974 and 1976 data, thus only the A's were calibrated differently for the years 1974 and 1976.

Initial tests and residual plots indicated that the ranked data would conform better to the functional form (Equation 6) if the data were partitioned into three segments:

Segment 1--group ranks 1-16 (first 1600 road sections),
Segment 2--group ranks 17-37 (sections 1601 to 3700), and
Segment 3--group ranks 38-150 (sections 3701 to 15 000).

This was basically due to the fact that the AADVMT over all ranks tended to level off more quickly than the functional form (Equation 6) would allow.

Final calibrations were performed separately on each of the segments for 1968, 1974, and 1976. As was originally done, the a and b for each segment were calibrated only on the 1968 data and were kept fixed for 1974 and 1976.

The results indicated that in each segment the ranked data conformed surprisingly well to the functional form (Equation 6), except that rank one (e.g., the first 100 ranked sections) in the first segment stood out as an outlier for each of the three years. This supports the hypothesis that the basic shape of the graphs of ranked AADVMT remains fixed, but that the graphs shift from year to year. In fact, the shift factors (A) appear to be extremely stable. That is, for each segment the three A's calibrated for 1968, 1974, and 1976 could be predicted from their least-squares line with an error of less than 0.01 percent [horizontal measurement equals the number of years from 1968 (e.g., 1974 is six years from 1968)]. Moreover, the percentage of total AADVMT contained in the first group rank (the outlier group) also remained stable and averaged about 12.0 percent of the total AADVMT for each of the years (actual percentages: 11.58 percent in 1968, 12.89 percent in 1974, 11.93 percent in 1976).

Integral Evaluation

By the change of variables $u = e^{-br}$, the integral

$$\int_A^B \exp(ae^{-bR})\, dR \tag{8}$$

is transformed to

$$-1/b\int_{e^{-bA}}^{e^{-bB}} (e^{au}/u)\, du = 1/b\int_{e^{-bB}}^{e^{-bA}} (e^{au}/u)\, du \tag{9}$$

By using the fact that

$$e^{au}/u = (1/u) + a + \ldots + [a^{n-1}u^{n-2}/(n-1)!] + (a^n u^{n-1}/n!) + \ldots \tag{10}$$

where the series on the right converges uniformly and quite rapidly on the interval of integration, we may approximate the integrand by a partial sum of

Table 1. Shift factors and integral values, 1968 and 1974.

Group Rank Segment	$A_{68,j}$	$S_{68,j}$	$A_{74,j}$	$S_{74,j}$
1	-268 641 000	4 056 179 461	-268 002 000	4 056 179 461
2	-270 474 000	5 694 482 805	-270 318 000	5 694 482 805
3	-271 270 000	21 989 340 000[a]	-271 176 000	30 668 225 250[a]

[a]There were only 118 full group ranks in 1968; there were 150 in 1974.

Table 2. AADVMT rank-size-rule estimates, 1968 and 1974.

Year	Group Rank	Actual (000 000s)[a]	Estimate (000 000s)	Error (%)
1968	2-16	27.806	26.564	-4.46
	17-37	14.789	14.529	-1.76
	38 and above	16.662	16.470	-1.15
	Total	59.257	57.563	-2.86
	1	7.759		
	Total	67.016		
1974	2-16	37.399	36.149	-3.34
	17-37	18.065	17.805	-1.44
	38 and above	25.510	25.337	-0.68
	Total	80.974	79.291	-2.08
	1	11.985		
	Total	92.959		

[a]Sum of section vehicle miles of travel from file.

the series on the right-hand side, and then integrate the resulting sum instead of the original integrand. Thus,

$$\int_A^B e^{ae^{-bR}} \approx 1/b \int_{e^{-bB}}^{e^{-bA}} (1/u) + a + \ldots + (a^n u^{n-1}/n!)du \qquad (11)$$

By evaluating the right-hand side, we obtain

$$\int_A^B e^{ae^{-bR}} dR \approx B - A + \Big\{[(ae^{-bA}/b) + (a^2 e^{-2bA}/4b) + \ldots + (a^n e^{-nbA}/n \cdot n!b)] - [(ae^{-bB}/b) + (a^2 e^{-2bB}/4b) + \ldots + (a^n e^{-nbB}/n \cdot n!b)]\Big\} \qquad (12)$$

The desired degree of accuracy will determine the size of the number of terms in the approximation (n). For our purposes, n = 10 was more than sufficient for each application.

Results

The values

$$S_j = \int_{R_L^j}^{R_H^j + 1} \exp\,[a_j \exp(b_j x)]\, dx \qquad (13)$$

were numerically calculated in the last section, where $R_L{}^j$ and $R_H{}^j$ represent the lowest and highest ranks in segment j (note that since rank 1 is an outlier, it is treated separately, and $R_L{}^1$ is taken to be 2). If we let A_{ij} represent the calibrated estimate of the shift factor for the jth segment of ranks for year i, the total AADVMT for year i is estimated according to the rank-size rule by the value

$$\text{AADVMT}_{\text{rank }1} + \sum_{j=1}^{3} [A_{ij}(R_H^j + 1 - R_L^j)] + S_j \qquad (14)$$

The resulting rank-size-rule estimates for 1968 and 1974 were within 3 percent of the actual values. The values of A_{ij} and S_j are presented in Table 1, the vehicle miles of travel estimates are presented in Table 2.

Table 3. Trend-line shift factors and integral values for 1976.

Group Segment	$A_{76,j}$	$S_{76,j}$
1	-267 789 000	4 056 179 461
2	-270 266 000	5 694 482 805
3	-271 144 000	30 668 225 250

Table 4. AADVMT predictions for 1976 based on the rank-size rule.

Group Rank	Prediction (000 000s)	Actual (000 000s)
2-16	39.344	38.508
17-37	18.897	19.082
38 and above	28.953	27.933
1	11.890[a]	11.587
Total	99.084	97.110

Note: Error is -2.03 percent.

[a]Based on the assumption that group rank 1 contains 12 percent of the total AADVMT.

In order to illustrate the predictive quality of this application of the rank-size rule, the stability of the A_{ij}'s and $\text{AADVMT}_{\text{rank }1}$ were exploited in order to make a forecast of 1976 AADVMT based on the 1968 and 1974 ranked data. Trend line estimates of the 1976 shift factors ($A_{76,1}$, $A_{76,2}$, and $A_{76,3}$) were constructed based only on 1968 and 1974 data, and an estimate of the $\text{AADVMT}_{\text{rank }1}$ for 1976 was based on the assumption that group rank 1 would contain 12.0 percent of the total AADVMT.

By using these trend line estimates, Equation 9 was calculated for 1976. The trend line estimates of the shift factors are shown in Table 3, the vehicle miles of travel predictions are shown in Table 4. Again, the exacting nature of the rank-size-rule procedure is exhibited.

CONCLUSION

The foregoing results indicate that four pieces of data (trend line estimates of the three shift factors and the AADVMT for rank 1) could be used in the rank-size rule to forecast AADVMT on the New York State touring route network. Although trend line approaches are crude by many measures, the apparent stability of ranked AADVMT and the fact that 1968, 1974, and 1976 span the 1973 energy crisis and the subsequent recovery, make the data sets studied here particularly suitable for such an approach. Nevertheless, other approaches that could take into account circumstances not detectable from historical trends are worthy of mention and future study.

In particular, the shift factors (A) could be estimated by actual vehicle counts on carefully selected sections of the New York State touring route network. This proposed approach would follow that of the bellwether polling districts, in that a very few statistically reliable sections would be taken to represent the whole. Such an approach might reduce considerably the extensive costs that are now incurred in monitoring and maintaining statewide vehicle counts. Another approach might entail estimation of the shift factors as functions of socioeconomic variables that take into account gasoline availability and price as well as other indicators of travel.

The initial application of the rank-size rule presented in this paper indicates that further study is warranted. The approach has the potential to greatly ease the very costly and burdensome task of estimating vehicle miles of travel, and its utility in forecasting vehicle miles of travel is yet to be fully explored.

Research is under way to determine effective bellwether sections to be used as a basis for estimating the necessary parameters for this application of the rank-size rule. In addition, the approach has been verified on vehicle miles of travel data for 1976-1979, and the development of a method for determining the shift parameters as functions of gasoline and diesel sales is currently under study.

ACKNOWLEDGMENT

I am grateful to David T. Hartgen for providing the basic idea and inspiration for this research. Special thanks are due to Dave Fifield of the Data Services Section of the New York State Department of Transportation for preparing the initial summaries of the data used in the work. Many thanks to Linda Unangst for her expert typing.

REFERENCES

1. Y. Zahavi. Introducing the Idea of the "K Distribution" to Transportation Patterns. HRB, Highway Research Record 322, 1970, pp. 162-170.
2. New York State 1968 Highway Sufficiency Ratings. New York State Department of Transportation, Albany, PR 020107, 1968.
3. New York State 1974 Highway Sufficiency Ratings. New York State Department of Transportation, Albany, 1974.
4. New York State 1976 Highway Sufficiency Rating. New York State Department of Transportation, Albany, 1976.
5. BMDP-79, Biomedical Computer Programs, P-Series. Univ. of California Press, Berkeley, 1979, pp. 484-498.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

Consideration of Nonresponse Effects in Large-Scale Mobility Surveys

WERNER BRÖG AND ARNIM H. MEYBURG

This paper continues the line of investigation of nonresponse problems previously presented. After a brief review of the context of the problem, namely the nonresponse effects on measured behavior in spite of demographic weighting, and the results of the previous research on this topic, the paper documents a broadening of the insights gained into the effects of nonresponse. These insights were applied to a large-scale nonresponse analysis of approximately 100 000 trips. The analysis included the nonresponse effects for the number of trips, trip purpose, travel mode, and seasons. Also, nonresponse effects are compared for written and interview surveys. Experience with the characteristics and impacts of nonresponse for intercity travel is presented. The insights gained could be used to clear up and correct past and present survey efforts and also to ensure that future data-collection efforts are conducted at lower costs, since corrections can also be made for smaller rates of return.

In principal, empirical surveys are not capable of providing an exact replication of measured reality: They only provide a picture that deviates more or less from this reality. The size and direction of these deviations are determined significantly by a variety of factors tied to the chosen survey design (see Meyburg and Brög in another paper in this Record).

Strict application of these basic facts shows the limits and possibilities of empirical research:

1. Precise determination of the distortions (biases) induced by the survey method will never be possible and
2. Systematic research into the biases caused by the survey method employed will lead to insights that will permit the estimation of the direction and order of magnitude of these deviations.

The corresponding measurement results will not be exactly correct, but they will be more correct (i.e., closer to reality). In order to reach results closer to reality, systematic research into survey methods is necessary.

Such methods research typically is very expensive. For that reason, these studies will have to be of an exemplary nature. This means that this fundamental research must be designed such that generalizable results (at least within reasonable limits) are obtained. These insights can be that

1. At least the direction of the bias in relation to the chosen survey method can be indicated;
2. Additional correction factors for the elimination of this bias can be provided, whose application would move the measured results closer to reality; and
3. An evaluation method is developed that would make it possible to estimate the relevant influences directly within the survey and to correct the survey data themselves.

The general level of knowledge about relevant factors of influence to survey methods in the determination of activities outside the home is rather limited to date. It has progressed only to a stage where we comprehend that a multitude of factors exists in the survey design that can be of significant influence on the measurement results. Furthermore, we begin to realize that, even in comparatively simple measurements of nonhome mobility, for example, regional and seasonal factors can generate

Table 1. Overall mobility by response increments.

Response Increment	Mobility per Increment	Cumulative Mobility	Index Cumulative Values[a]
First fifth	2.91	2.91	112.8
Second fifth	2.70	2.81	108.9
Third fifth	2.57	2.72	105.4
Fourth fifth[b]	2.41	2.64	102.3
Fifth fifth[c]	2.37	2.58	100.0

Note: KONTIV 1976 had approximately 105 000 person survey days.
[a]Estimated total value = 100.
[b]Partial nonresponse estimate.
[c]Complete nonresponse estimate.

specific survey situations that stand in the way of generalizing corresponding fundamental research.

For this reason, in the conduct of basic studies of survey-method-specific influence factors, special attention should be paid to the development of an evaluation method that allows the renewed examination of the results in a concrete case and that thereby expands the spectrum of corresponding insights.

THE NONRESPONSE PROBLEM

A significant bias in empirical surveys of nonhome mobility results from the fact that it is not possible to get all households or individuals to respond. This problem exists both in sampling and in total population surveys. In either case conclusions have to be drawn for a larger entity based on a smaller group of respondents. Unfortunately, the severity of the problem increases for the usual sampling situation. In practice, this problem is usually disguised by means of the indication of a significance level based on sampling theory (1). These statistical significance measures are valid only when information about each sampling element is available. This condition cannot be fulfilled in empirical surveys.

Therefore, it is important to deal with the nonresponse problem in a systematic fashion in order to be able to estimate how the observed results would change if corresponding information were available for each selected sample element. It is especially important that the information be relevant for the object of the investigation, in this case for nonhome mobility. Information that captures merely the sociodemographic characteristics does not fill this information gap and, therefore, is not sufficient for reliable estimation of the influence of nonresponse.

METHODOLOGICAL EXPERIMENTS

Since the nonresponse effect constitutes a significant bias of empirical results, we have investigated this problem area in several basic research efforts. One of these studies was presented in an earlier paper (2). It constituted the starting point for several additional analyses, including the research reported here.

For a survey of 984 households in West Berlin, the survey design corresponded exactly to the one employed in the national travel survey (KONTIV) and other large surveys in West Germany. The only objective of the West Berlin survey was to obtain as large a response rate as possible and to gather additional qualitative information about late or nonresponding households.

The survey used the mail-back technique with several follow-ups. Although in other large surveys in West Germany the number of reminders was usually limited to four (which normally results in a response rate of 65-75 percent), in this survey it was increased and the response rate increased correspondingly from 74 to 87 percent. Care was taken not to modify the survey design in order to truly measure the influence of the nonresponse effect and not that of a changed survey method.

In order to gain additional insights into the structure and motivation of the group of nonrespondents, interviews were conducted with late respondents and nonrespondents wherever possible. Otherwise, additional investigations were undertaken to obtain certain information about these groups of people.

The most-significant result of this methodological experiment was that the measured mobility (trips per person) decreased with the size of the response rate (see Table 1) and that this effect cannot be corrected sufficiently by means of simple sociodemographic weighting.

The follow-ups and additional investigations clearly indicated that people who had little nonhome mobility (i.e., few trips taken outside the home) did not feel sufficiently concerned and, therefore, did not participate in the survey. On the other hand, it could not be confirmed that late respondents, tired of the numerous follow-ups and reminders, simply report fewer trips than they actually perform.

SUBSEQUENT ANALYSES

The methodological experiment resulted in several important insights for subsequent work on the nonresponse problem:

1. Trend extrapolation on the basis of response speed proved to be a usable method of estimation and an acceptable method of evaluation;
2. Use of the term "mobility per person" (i.e., trips per person) proved to be too imprecise; variables such as "share of mobiles" (i.e., that share of the population surveyed who took a trip on the survey day) and "mobility per mobile" (how many trips the surveyed mobile person took on the survey day; i.e., trip rate per person) should be used instead; and
3. Stratification according to mode and trip purpose did not produce consistent results yet--probably due to relatively small sample size (this suggests further investigations on the basis of larger sample sizes).

In this experiment the specific influence of the survey area and the survey period could not be determined. The present level of knowledge suggested that the nonresponse investigation be repeated on the basis of this evaluation method for the KONTIV survey (3) [i.e., a sample representative of an entire region (in this case, West Germany) and distributed across all seasons].

In its basic version, as it is used in this paper, the KONTIV survey consisted of 105 000 person survey days, and it had a response rate of 72.4 percent. A stratification into five response segments of equal size was performed because the trend extrapolation can be performed most readily when only the last segment has to be estimated completely and the second to last has to be estimated partly. The results presented in the following sections permit a much-more-precise determination of the nonresponse effect. However they are only relevant for surveys of comparable methodological design (i.e., specifically for mail-back surveys).

SELECTED RESULTS

Measures of Nonhome Mobility

The average number of trips for all people surveyed shows the known effect that mobility decreases with increasing response rate. It is almost one-quarter higher for the first respondents than for the (estimated) population average.

The variable "mobility per person" is, however, a quasi-artificial average value that is composed of the "share of mobiles" and the "mobility per mobile." These two measures show modifications already. Although the value of "mobility per mobile" decreases with increasing response rates, the values in each response fifth (stratum) are relatively less exaggerated than they are for the measure "mobility per person" (Table 2). The "share of mobiles" reaches its highest value only in the second fifth, which indicates, among other things, that some people who have low mobility also answer very quickly (Table 2).

NONRESPONSE BY SEASON

The analysis of the nonresponse effect by season shows the importance of subdividing the average mobility into its two constituent components. If we look at the changes in "mobility per person," we also find a rather uniform decrease in values with an increasing response rate (Table 3).

This picture is largely reconfirmed in the analysis of the measure "mobility per mobile," but it is relatively different for the "share of mobiles." Two different tendencies become evident: Although the first two response fifths show the highest mobility in winter and spring, this effect only shows in the second and third fifth for summer and fall (i.e., it is delayed). The reason for this phenomenon lies in the fact that people who travel a lot and who belong to the group of fast respondents during winter and spring only answer with a delay during the summer and fall months when they are busy with nonhome activities or when they are more frequently on trips away from home (Table 4).

In connection with the seasonal variation development of the measure "mobility per mobile," we therefore observe quite different nonresponse effects dependent on the time of the survey. In a continuous year-long survey, it is therefore advisable to apply a nonresponse correction separately by time of year.

Nonresponse by Mode

Response or nonresponse behavior has a significant effect on the resulting frequency of modal use. Early respondents often use individual, often nonmotorized, travel modes, and a large number of public transit users apparently decide only relatively late (or not at all) to participate in a survey (Table 5). Also obvious here is that a response rate of, for example, 60 percent still contains tangible fluctuations in the modal split representation, in spite of relatively good representation of the total mobility.

Nonresponse by Trip Purpose

Inconsistencies due to nonresponse are even more pronounced in the analysis by trip purpose. A repeatedly indicated tendency of decrease with increasing response rate is evident for social-recreational and shopping trips, in spite of the fact that the initial values in the first fifths lie substantially above the estimated average value (Table 6). For mandatory trips, however, a substantial de-

Table 2. Share of mobiles and mobility of mobiles.

Response Increment	Share of Mobiles: Single	Share of Mobiles: Cumulative	Share of Mobiles: Index Cumulative Values[a]	Mobility of Mobiles: Single	Mobility of Mobiles: Cumulative	Mobility of Mobiles: Index Cumulative Values[a]
First fifth	75.7	75.7	102.3	3.84	3.84	110.0
Second fifth	78.4	77.1	104.2	3.45	3.65	104.6
Third fifth	76.8	76.9	103.9	3.34	3.54	101.4
Fourth fifth[b]	71.2	75.2	101.6	3.38	3.51	100.6
Fifth fifth[c]	69.3	74.0	100.0	3.42	3.49	100.0

Note: KONTIV 1976 had approximately 105 000 person survey days.
[a]Estimated total value = 100.
[b]Partial nonresponse estimate.
[c]Complete nonresponse estimate.

Table 3. Overall mobility by season.

Response Increment	Cumulative Value: Spring	Summer	Fall	Winter	Index Cumulative Value[a]: Spring	Summer	Fall	Winter
First fifth	3.00	2.79	2.90	2.98	113.2	106.1	112.8	120.7
Second fifth	2.82	2.73	2.85	2.83	106.4	103.8	110.9	114.6
Third fifth	2.69	2.64	2.80	2.73	101.5	100.4	109.0	110.5
Fourth fifth[b]	2.67	2.63	2.68	2.58	100.8	100.0	104.3	104.5
Fifth fifth[c]	2.65	2.63	2.57	2.47	100.0	100.0	100.0	100.0

Note: KONTIV 1976 had approximately 105 000 person survey days.
[a]Estimated total value = 100.
[b]Partial nonresponse estimate.
[c]Complete nonresponse estimate.

Table 4. Share of mobiles and mobility of mobiles by season.

Response Increment	Index of Cumulative Values[a] — Share of Mobiles: Spring	Summer	Fall	Winter	Mobility of Mobiles: Spring	Summer	Fall	Winter
First fifth	104.2	95.4	102.7	108.4	108.5	111.3	109.8	111.5
Second fifth	103.2	99.6	106.1	107.7	102.8	104.2	104.6	106.5
Third fifth	102.1	100.1	105.8	106.9	99.4	100.3	102.9	103.5
Fourth fifth[b]	101.1	100.0	102.7	103.0	99.7	100.0	101.4	101.5
Fifth fifth[c]	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Note: KONTIV 1976 had approximately 105 000 person survey days.
[a]Estimated total value = 100.
[b]Partial nonresponse estimate.
[c]Complete nonresponse estimate.

Table 5. Mobility of mobiles by principal mode of travel.

Response Increment	Total	Index of Cumulative Values[a] Nonmotorized Modes[b]	Individualized Travel Modes[c]	Public Transit
First fifth	110.0	115.2	107.2	100.0
Second fifth	104.6	107.2	104.8	94.7
Third fifth	101.4	103.6	101.2	97.4
Fourth fifth[d]	100.6	101.4	100.6	100.0
Fifth fifth[e]	100.0	100.0	100.0	100.0

Note: KONTIV 1976 had approximately 105 000 person survey days.
[a]Estimated total value = 100.
[b]Includes walking, bicycle, and motorized bicycle.
[c]Includes automobile driver, automobile passenger, moped, and motorbike.
[d]Partial nonresponse estimate.
[e]Complete nonresponse estimate.

Table 6. Mobility of mobiles by trip purpose.

Response Increment	Total	Index of Cumulative Values[a] Mandatory Trips[b]	Social-Recreational Trips	Shopping Trips[c]
First fifth	100.0	99.4	114.8	121.4
Second fifth	104.6	98.2	108.6	111.7
Third fifth	101.4	97.6	102.5	105.8
Fourth fifth[d]	100.6	98.8	100.0	102.9
Fifth fifth[e]	100.0	100.0	100.0	100.0

Note: KONTIV 1976 had approximately 105 000 person survey days.
[a]Estimated total value = 100.
[b]Includes work and school trips.
[c]Includes trips for shopping and personal business.
[d]Partial nonresponse estimate.
[e]Complete nonresponse estimate.

Table 7. Mobility of mobiles—mandatory travel by principal modal use.

Response Increment	Total	Index of Cumulative Values for Mandatory Travel[a,b] Nonmotorized Modes[c]	Individualized Travel Modes[d]	Public Transit
First fifth	99.4	104.1	97.6	100.0
Second fifth	98.2	100.0	97.6	92.3
Third fifth	97.6	100.0	97.6	96.2
Fourth fifth[e]	98.8	100.0	98.8	100.0
Fifth fifth[f]	100.0	100.0	100.0	100.0

Note: KONTIV 1976 had approximately 105 000 person survey days.
[a]Estimated total value = 100.
[b]Includes work and school trips.
[c]Includes walking, bicycle, and motorized bicycle.
[d]Includes automobile driver, automobile passenger, moped, and motorbike.
[e]Partial nonresponse estimate.
[f]Complete nonresponse estimate.

Table 8. Mobility of mobiles for social-recreational travel.

Response Increment	Total	Index of Cumulative Values for Social-Recreational Travel[a] Nonmotorized Modes[b]	Individualized Travel Modes[c]	Public Transit
First fifth	114.8	123.5	109.8	100.0
Second fifth	108.6	111.8	104.9	100.0
Third fifth	102.5	[illegible]	100.0	100.0
Fourth fifth[d]	100.0	[illegible]	100.0	100.0
Fifth fifth[e]	100.0	100.0	100.0	100.0

Note: KONTIV 1976 had approximately 105 000 person survey days.
[a]Estimated total value - 100.
[b]Includes walking, bicycle, and motorized bicycle.
[c]Includes automobile driver, automobile passenger, moped, and motorbike.
[d]Partial nonresponse estimate.
[e]Complete nonresponse estimate.

crease can be observed, especially in the middle fifths (Table 7). Evidently, people who have relatively simple activity patterns that do not go significantly beyond trips to and from work tend to belong to the groups of late or nonrespondents.

The indicated effects could be eliminated or, as is more common, enlarged due to the consideration of combined measures. The combination of the measures "trip purpose" and "predominantly used travel mode" illustrates even more clearly the problems of low response rates. For example, a rather good response rate of 60 percent generates a good result only for mandatory trips by nonmotorized modes. For all other modes the results are below average. The situation is exactly opposite for social-recreational travel. For the identical response rate, nonmotorized travel would be overrepresented, but the use of other modes would have been represented correctly (Table 8). For shopping trips only, public transit trips were captured correctly; however, they play a rather insignificant role for that particular trip purpose. The other much-more-important travel modes are substantially overrepresented for a 60 percent response rate (Table 9).

Sociodemographic Weighting

Prior to this nonresponse estimation, the results of the KONTIV survey were subjected to a detailed weighting process. First, an equal distribution of weekdays was performed. This equalization was followed by a reconstruction (replication) and correction of the selection procedures for the formations of the sample in the context of a free estimation of population values. Finally, the results were subjected to sociodemographic weighting on the basis of a cell plan with approximately 200 cells. These weighting efforts were relatively intensive and included all possibilities available on the basis of secondary statistical material.

When the results of the weighting process are compared with those obtained through the nonresponse estimation, it becomes evident that the weighting process does not lead uniformly in the same direction and that it results in substantial deviations from the estimates of the actual values in some cases.

Overall, the weighting procedure results in an overestimate of total mobility by 1.6 percent. This is a difference that looks relatively good compared with other nonresponse investigations (Table 10). A nonuniform picture arises for the individual seasons. In this instance it is particularly noteworthy that sociodemographic weighting is least precise for the winter, when the nonresponse effects are particularly strong. Substantial inaccuracies are also observed for the travel mode and trip purpose categories, where the unweighted results are not changed consistently in the proper direction. Therefore, a correction by means of sociodemographic characterstics does not ensure that the characteristics of the behavior under investigation are improved sufficiently accurately.

CONCLUSIONS ABOUT THE METHODOLOGICAL EXPERIMENTS

These results produce generalizable insights about the direction and the order of magnitude of biases that result from the survey method chosen. They also illustrate an evaluation method on the basis of which corresponding tests of the other empirical surveys can be undertaken. Recognize, however, that such a new nonresponse estimation requires sufficiently high response rates. In the application of the trend-extrapolation method, this rate should not lie below 70 percent if at all possible. Otherwise

Table 9. Mobility of mobiles for shopping and personal business travel.

	Index of Cumulative Values for Shopping and Personal Business[a]			
Response Increment	Total	Nonmotorized Modes[b]	Individualized Travel Modes[c]	Public Transit
First fifth	121.4	120.0	125.0	128.6
Second fifth	111.7	109.1	120.0	114.3
Third fifth	105.8	103.6	112.5	100.0
Fourth fifth[d]	102.9	101.8	107.5	100.0
Fifth fifth[e]	100.0	100.0	100.0	100.0

Note: KONTIV 1976 had approximately 105 000 person survey days.
[a]Estimated value = 100.
[b]Includes walking, bicycle, and motorized bicycle.
[c]Includes automobile driver, automobile passenger, moped, and motorbike.
[d]Partial nonresponse estimate.
[e]Complete nonresponse estimate.

Table 10. Influence of sociodemographic weighting on the measurement of nonhome mobility.

		Original Values		
Stratification	Characteristic	Sociodemographic Weighting	Nonresponse Estimation	Index
Total year	Mobility per person	2.62	2.58	101.6
	Share of mobiles	74.8	74.0	101.1
	Mobility per mobile	3.50	3.49	100.3
Spring	Share of mobiles	76.2	75.4	101.1
	Mobility per mobile	3.49	3.52	99.2
Summer	Share of mobiles	72.6	74.1	98.0
	Mobility per mobile	3.60	3.55	101.4
Fall	Share of mobiles	74.8	73.8	101.4
	Mobility per mobile	3.53	3.48	101.4
Winter	Share of mobiles	75.7	72.8	104.0
	Mobility per mobile	3.53	3.39	104.1
Travel mode	Nonmotorized	1.44	1.38	104.4
	Individualized	1.64	1.66	98.8
	Public transit	0.35	0.38	92.1
	Other	0.07	0.07	100.0
Trip purpose	Mandatory	1.54	1.65	93.3
	Social-recreational	0.85	0.81	104.9
	Shopping and personal business	1.11	1.03	107.8

Note: KONTIV 1976 had approximately 105 000 person survey days.

Table 11. Influence of response rate on the measurement of nonhome mobility.

		Original Values for a Response Rate of 33 Percent		
Stratification	Characteristic	Sociodemographic Weighting	Nonresponse Estimation	Index
Total year	Mobility per person	2.82	2.58	109.3
	Share of mobiles	76.7	74.0	103.6
	Mobility per mobile	3.68	3.49	105.4
Spring	Share of mobiles	77.9	75.4	103.3
	Mobility per mobile	3.66	3.52	104.0
Summer	Share of mobiles	72.7	74.1	98.1
	Mobility per mobile	3.75	3.55	105.6
Fall	Share of mobiles	77.2	73.8	104.6
	Mobility per mobile	3.68	3.48	105.7
Winter	Share of mobiles	78.4	72.8	107.7
	Mobility per mobile	3.64	3.39	107.4
Travel mode	Nonmotorized	1.50	1.38	108.7
	Individualized	1.74	1.66	104.8
	Public transit	0.37	0.38	97.4
	Other	0.07	0.07	100.0
Trip purpose	Mandatory	1.63	1.65	98.8
	Social-recreational	0.88	0.81	108.6
	Shopping and personal business	1.17	1.03	113.6

Note: KONTIV 1976 had approximately 105 000 person survey days.

a uniform trend might not be detected and, as a consequence, the final value will be estimated incorrectly.

In cases where a new nonresponse estimation cannot or should not be performed, corresponding values from the KONTIV survey are available. For a new empirical survey that has a comparable time frame, size of urban area, and response rate, corresponding correction factors (with or without sociodemographic weighting) can be computed and inserted. We have already tested such a procedure successfully.

Significance of Nonresponse Effects in Mail-Back Surveys

In the example presented earlier, the measurement error due to incomplete participation by the sample elements might appear comparatively small. To a large degree this is due to the high response rate of 72.4 percent achieved in the KONTIV survey.

Such high response rates will probably not be achievable in the future due to tightened data protection problems and due to the general public apathy (at least in Germany) toward the increasing number of poorly designed surveys. The measurement error will therefore increase substantially for lower response rates.

This problem can be illustrated by applying sociodemographic weighting only to the first third of the respondents and by comparing the results with the population estimates. The 33 percent response rate in Table 11 was selected because many travel surveys do not exceed that rate.

Such a response rate, although somewhat normal in general research practice, yet too low to produce reliable results, leads, for example, to an overestimation of the mobility of the average population of almost 10 percent (Table 11). The determining factors for this are the overrepresentation of the "share of mobiles" (by about 4 percent) and of the "mobility per mobile" (by about 5 percent). Accordingly, the mobility values by season, travel mode, and trip purpose are overrepresented with few exceptions. Low response rates in the absence of knowledge about effects induced by them constitute a substantial source of error in surveys of nonhome mobility.

On the other hand, more precise knowledge of the nonresponse effects in such surveys can also lead to substantial savings when such knowledge is implemented. A precise calculation of the funds required for such a survey indicates that the lowest survey cost per returned questionnaire is reached when two follow-up reminders are used. By applying the correction factors presented earlier, the bias introduced due to response losses can largely be compensated for and a substantially more advantageous cost-result (performance) ratio can be reached. The table below, from the West Berlin survey of approximately 45 000 persons for the transportation development plan, shows the relation of response rate and survey costs (4):

Survey Method	Reponse Rate (%)	Index of Cost/Usable Response
Mail-back questionnaire without follow-up action	30	100.0
Mail-back questionnaire with two reminder notices	60	88.5
Mail-back questionnaire with four reminders, including one additional questionnaire mailing	77	96.4

Significance of Nonresponse Effects in Interview Surveys

The trend-extrapolation method for estimating biases introduced due to nonrespondents by means of the response speed is based on the notion that, in a mail-back survey, a significant stimulus for participation (voluntarily, as a rule) lies in the object of the investigation. This insight has been documented through several research projects (5).

Interest in a survey on nonhome mobility is large when such nonhome mobility is practiced to a large degree, and it is small when such mobility is small or nonexistent. For this reason, it is only natural that, in a mail-back survey of nonhome mobility, many mobile people respond relatively faster and in larger numbers than do the immobile ones. This relationship holds only for this particular survey method, as we have stressed repeatedly.

The main reason for participation in an interview survey, on the other hand, is that the target person is reachable at home and can be convinced by the interviewer to participate. For relatively well-trained interviewers, the accessibility factor (meeting the interviewee) is the more important one here. Less-mobile people can be contacted more easily, and people who have a wide range of nonhome mobility are a definite problem group for interview surveys. They are hard to reach and often very busy--which means that they are potential interview refusers. For this reason, the nonresponse effect acts in precisely the opposite direction from that observed in mail-back surveys, where the respondents tend to provide too low a representation of their actual mobility.

These interrelationships were illustrated in a methodological experiment performed by Moolman (6). In the course of an interview survey about nonhome mobility the selected households were contacted until an interview actually was conducted. A response rate of 98.5 percent was attained by this method. When the respondents are stratified according to response speed into those who respond at the first, the second, or only at the third contact effort, it becomes evident that the mobility per person is substantially higher for the nonrespondents (6):

	Index for Number of Trips per Person	
Response Speed	Respondents	Nonrespondents
After one contact attempt	100.0	127.5
After two contact attempts	100.0	109.8
After three contact attempts	100.0	109.6

From that observation, we can conclude that, even with very good response rates (after three or four contact efforts), the observed mobility per person is 4-5 percent too low in interview surveys due to the nonresponse effect alone.

A further stratification of the unreported trips also shows that the largest underrepresentation occurs for the non-home-based trips (i.e., people who follow complicated trip chains are more likely to be nonrespondents). Expressed differently, we can state: Aside from those people who have little or no mobility, those people who have simple activity patterns are the primary respondents, and these are the groups that are relatively difficult to reach through mail-back surveys.

These substantial differences of nonresponse effect by survey method do not only belong to the absolutely necessary prerequisite basic knowledge in the area of nonresponse estimation, but they are also impressive proof of the fact that each survey method produces its own specific types of measurement errors. Therefore, discussion and comparison of empirical measurement results is practically not possible without knowledge of the survey method employed.

This also implies that identical numerical results that were achieved with different survey designs do not necessarily mean that the corresponding phenomena are represented identically. If, for example, in country X the mobility per inhabitant observed by means of an interview survey (with two contact efforts) is quantitatively identical to that obtained by means of a mail-back questionnaire (with a 30 percent response rate) in country Y, this implies that, on the basis of different nonresponse effects alone, without consideration of numerous other influencing factors, the mobility in country Y has to be set at 15 percent below that in country X. This insight could, for example, throw a different light on the recent discussion of international comparisons of non-home-activity time budgets.

APPLICATION TO INTERCITY PASSENGER TRAVEL

After an evaluation method, such as the one postulated in the first section of this paper, has been developed and tested successfully, it can be applied in general to other similar problem contexts. For the case of the problem of nonresponse, this means that in measurements of nonhome mobility we can typically expect that, on the basis of nonresponse effects, the mobility indices will be too high for mail-back questionnaires and too low for interview surveys. (Other influential factors exit, but they are not considered in this paper.) For an estimate of this nonresponse effect, the method of trend extrapolation on the basis of response speed presents itself. A distinction has to be made between mobiles and immobiles, and stratification according to trip purposé and travel mode are advisable.

Equipped with this knowledge and experience we can attempt to determine the biases in surveys of intercity travel generation caused by nonresponses. For this purpose, two continuous surveys are available that measure annual intercity vacation travel in the Federal Republic of Germany. One survey was performed by the Federal Statistics Office (Statistisches Bundesamt) in the context of the Microcensus. The second one was a privately conducted travel analysis (8). The two surveys are not exactly comparable with respect to the samples used, but both have been criticized for alleged underestimation of vacation travel volume.

The mere knowledge of the direction of valid nonresponse effects already indicates one possible cause for deviations of the estimates from reality. However, in this case at least two other significant influences have to be recognized in addition to the nonresponse effect that must result in the underestimation of travel, given that the interview survey technique was applied. Respondents are asked to report vacation trips performed during the preceding 12 months. Memory gaps are known to show in this retroactive technique. These gaps have proved to be greatest for interview surveys, largely due to the required instant recall. The mail-back method gives the respondent more time to recall past travel. In addition, the interview method typically requires that a household member also report the behavior of other members, which again will lead to underreporting of trips.

These three influences suggest that it is advisable to survey intercity travel behavior by means of the mail-back questionnaire technique. We need to be aware that, initially, the travel volume will be overrepresented. It will have to be scaled down

Table 12. Vacation travel behavior.

	Index of Cumulatives Values		
Response Increment	Average No. of Vacation Trips per Person	Share of People with Vacation Trips	Average No. of Vacation Trips per Vacation Traveler
First fifth	132.0	118.9	111.5
Second fifth	116.5	110.2	105.5
Third fifth	108.2	104.9	103.0
Fourth fifth	103.1	102.0	101.2
Fifth fifth	100.0	100.0	100.0

Note: Index of estimated total travel = 100.

Table 13. Modal choice for vacation travel.

	Index of Cumulative Values for Principally Used Travel Mode			
Response Increment	Automobile	Train	Airplane	Other
First fifth	112.7	116.7	105.0	100.0
Second fifth	105.5	112.5	100.0	100.0
Third fifth	102.7	108.3	100.0	100.0
Fourth fifth	100.9	104.2	100.0	100.0
Fifth fifth	100.0	100.0	100.0	100.0

Note: Index of estimated total travel = 100.

Table 14. Intercity travel behavior.

	Index of Cumulative Values for Share of People			
Response Increment	Who Did Not Make Intercity Trips	Who Had Vacation Travel Only	Who Made Other Private Intercity Trips Only	Who Made Vacation and Other Private Intercity Trips
First fifth	64.4	100.0	108.6	150.9
Second fifth	80.1	99.7	107.4	128.0
Third fifth	89.7	99.2	106.2	114.7
Fourth fifth	95.8	99.7	102.5	106.0
Fifth fifth	100.0	100.0	100.0	100.0
Index[a]	94.9	97.3	109.9	108.3

Note: Estimated total value of index of sociodemographically weighted values = 100.
[a]Response rate was 67 percent.

(corrected) in the course of a nonresponse analysis. This method was applied in a large survey of intercity travel behavior in the Federal Republic of Germany (7).

When we look, for example, at the number of vacation trips per person by response segments (fifths), the familiar nonresponse influence is very pronounced, especially among early respondents to a 1979 intercity travel survey of approximately 60 000 people (Table 12). This effect is mainly due to the fact that people who have not made vacation trips apparently can only be enticed very late, if at all, to respond to the survey. Among the people who have made vacation trips, those people who made more than one vacation trip are more likely to respond.

A stratification, from the same survey, of vacation trips by primary mode of travel shows that the nonresponse effect varies by mode. Particularly overreported are train trips in the case of low response rates. Automobile trips develop largely analogous to the total distribution, due to their dominant share of all vacation trips (70 percent) (Table 13).

The nonresponse effect is even more significant in the stratification according to trip purposes. In this case, observe how late respondents answer a mail-back questionnaire if they did not undertake any intercity trips (Table 14). Equally obvious is that people who have a particularly high intensity of intercity travel respond to this type of survey. The share of people who engage in private intercity trips other than for vacation purposes develops in a relatively moderate fashion. People whose intercity travel consists of their annual vacation trip are represented equally (i.e., largely correctly) in all response increments.

Therefore, the special problem groups with respect to the nonresponse influence in measuring vacation travel are again composed of those people who do not, or who rarely, undertake the activity under investigation. In light of the present level of understanding, these groups are overestimated in mail-back surveys and underestimated in interview surveys, which results in too high a volume for mail-back responses and too low a volume for interview surveys. Again, sociodemographic correction can only provide limited compensation for these deviations.

Although in this mail-back survey this effect was compensated for by appropriate nonresponse factors, such corrections were absent in the other two intercity travel surveys mentioned earlier. As a consequence of this uneven treatment, the results for the three surveys are substantially different. A comparison of our results with those of the travel analysis conducted during the same year (8) shows that the share of people who made vacation trips (travel intensity) was underestimated by 4-5 percent and the number of trips per vacation traveler (trip frequency) by more than 20 percent. This underestimation is not only due to the nonresponse factor but is also a consequence of the distortions that result from memory gaps and from reporting by a household member about activities of other members. The latter sources of errors could have been controlled in a systematic analysis of response behavior. In any event, the documents currently in use concerning intercity vacation travel in the Federal Republic of Germany understate that travel category by about 25 percent. A substantial part of this underestimation can be tied directly to the fact that nonresponse effects were not taken into account.

SUMMARY

This paper reports on an ongoing investigation into the effects of nonresponse on the accuracy of empirical survey results. A number of examples were presented to show the types and magnitude of distortions that result from ignoring nonresponse effects. The different impacts of stratification by season, mode, and trip purposes were demonstrated. Also, the varying results generated by the mail-back versus interview method were analyzed.

Given the limited experience with nonresponse effects, a useful procedure was developed and presented by means of which the relevant influences can be reestimated and corrected for each survey. Aside from these methodological experiments, applications of their results to the study of intercity passenger travel behavior were shown.

The paper concludes that fundamental methodological survey research must be designed so that generalizable results can be obtained. Through systematic research into the distortions caused by nonresponse, at least the direction of the bias in relation to the chosen survey method can be determined, and additional correction factors for the elimination of this bias can be generated. Sociodemographic weighting is shown not to be a satisfactory remedy for the effects of nonresponse.

REFERENCES

1. M. Wermuth. Effects of Survey Methods and Measurement Techniques on the Accuracy of Household Travel-Behavior Survey. In New Horizons in Travel-Behavior Research (P.R. Stopher, A.H. Meyburg, and W. Brög, eds.), Lexington Books, D.C. Heath and Co., Lexington, MA, 1981.
2. W. Brög and A.H. Meyburg. Nonresponse Problem in Travel Surveys: An Empirical Investigation. TRB, Transportation Research Record 775, 1980, pp. 34-38.
3. Sozialforschung Brög. Kontinuierliche Erhebung zum Verkehrsverhalten, KONTIV 76 [Continuous Survey of Travel Behavior, KONTIV 76]. Ministry of Transport, Bonn, West Germany, 1976.
4. SOCIALDATA GmbH. Verkehrsuntersuchung Spandau--Auswirkungen des U-Bahn Baues auf das Verkehrsverhalten [Traffic Study Spandau--Effects of Subway Construction on Travel Behavior]. Senator for Construction and Housing, West Berlin, West Germany, 1979.
5. R. Höttler. Ausschöpfung von Stichproben und die Non-Response Problematik in Theorie und Praxis [Response Rates and Problems of Non-Reponse in Theory and Practice]. Paper presented at the 5th Workshop of the German Transport Science Society, Grainau-Eibsee, West Germany, Nov. 1979.
6. B.A. Moolman. A Fresh Look at Non-Sampling Errors in Survey Research. Paper presented at the 5th Workshop of the German Transport Science Society, Grainau-Eibsee, West Germany, Nov. 1979.
7. SOCIALDATA GmbH. Marktanalyse des Verhaltens im Personenfernverkehr [Market Analysis of Intercity Travel Behavior]. Ministry of Transport, Bonn, West Germany, 1979.
8. Studienkreis für Tourismus e.V. Urlaubsreisen 1979, Ergenbnisse der Reiseanalyse [Vacation Travel 1979, Results of a Travel Analysis]. Starnberg, West Germany, 1980.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

Validity Problems in Empirical Analyses of Non-Home-Activity Patterns

ARNIM H. MEYBURG AND WERNER BRÖG

Validity problems of empirical data have been neglected to a large extent in the transportation planning field. This paper illustrates the impact that choice of survey method has on the validity of the data. It shows that the recorded data depend directly on the method selected for obtaining them. An uncritical application of survey methods is not justifiable and, in fact, can lead to incorrect survey results. Basic research in the area of empirical survey methods is long overdue. An international exchange of experiences in this regard is considered most beneficial, as illustrated by this paper. The exchange of information and insights is often hampered because the survey methods used for specific investigations tend to be inadequately documented. This deficiency makes subsequent assessment of data validity very difficult, if not impossible. Furthermore, the use of such data without consideration of the underlying survey method is dangerous. The paper cites examples where the results of analyses can be manipulated by means of different survey methods. Greater efforts should be made to integrate data collection with the research effort performed on the basis of these data.

Many transportation planners, engineers, and modelers have, for all too long, ignored the quality of the basic input to their research efforts, namely the data. Since much of this research is of an empirical nature, the data are obtained through empirical surveys. This paper is intended to add to the efforts concerned with survey methodology for empirical analyses of travel activities. It presents a number of examples that show how the survey method and design can influence the results of an investigation.

We recognize that only a limited set of examples can be shown in the context of a paper of this nature. Further, the basis of comparison is yet another survey, albeit one that is generally recognized as representative of the state of the art in survey methodology [e.g., the national travel survey (KONTIV) (1)]. We need to accept the quality of that survey and the validity of its results in order to believe in the results of the research presented in this paper.

Even if the argument is made that the different survey results obtained through two surveys do not prove the correctness (or lack thereof) of one or the other survey results, the disturbing fact remains that different results about mobility were obtained when the study objectives were identical and only the survey method differed. This result alone is worth keeping in mind. The transportation planner or modeler is well advised to pay careful attention to the procedure used to generate the inputs that are used in any modeling effort or in the generation of simple mobility statistics. The validity (correctness) of the data will also determine the validity of any model or statistical results, which in turn might be used as the basis for policy and investment decisions.

VALIDITY VERSUS REPRESENTATIVENESS

The paper concentrates on the validity aspects of survey results because problems of validity of empirical data have been neglected to a large extent in the transportation planning field. The issue of validity of survey data transcends that of representativeness. Representativeness addresses the question of whether we have enough data points for each of our strata or cells. The concept of validity is aimed at questions of whether the data obtained are valid (correct or relevant) or whether they are an accident ascribable to a particular method of data collection. We attempt to show in this paper that empirical results based on survey data typically contain substantial errors that result in severe misrepresentation of reality.

The error sources addressed in this research lie exclusively in the survey method and design employed to generate these data. A fair assessment of present survey and research practice in transportation probably is that these distortions of reality go largely undetected. This is in part due to the researcher's unawareness of any problems of bias in the data and to a great extent to the preoccupation with the modeling and analysis phase of the typical research undertaking. Since data collection is such a tedious and costly element of any empirical research effort, it is somewhat understandable, yet not justifiable, that some researchers are unwilling to look back when a data base has been generated that can be analyzed. Any additional effort that questions the validity of the information seems to detract from the real research work and progress.

VALIDITY PROBLEMS OF EMPIRICAL SURVEY RESULTS

The following sections contain a number of examples of the impact of survey technique and survey instrument on the validity of the results of studies of travel behavior, irrespective of the quality and structure of any models used in the analysis phase. The items illustrated here do not represent an exhaustive set of influence factors, but they are intended to provide an array sufficiently broad enough to show the severity of the problem. The issue of nonresponse, which is certainly an element of validity problems, is not dealt with in detail in this paper because it is discussed extensively elsewhere (2) and in our other paper in this Record.

Use of Perceived Versus Actual Values

Investigation of travel distances on the basis of information from survey respondents is very difficult, due to the difference between actual and perceived values. A Dutch study (3) shows an overall overestimate of approximately 10 percent based on perceived (reported) distances [Table 1 (3)]. Of course, these results also vary by mode (ranging from 0.8 percent for transit to 15.7 percent for moped) and with the actual distance traveled. Obviously, the use of perceived distances as input to other investigations, such as determination of travel speed, produces an error at the input stage. The problem becomes even more pronounced if airline distances are used.

In a separate investigation in West Germany that dealt with the effects of using reported rather than actual (measured) values [a study performed to determine price elasticities for travel by transit (4)], similar discrepancies were detected. For reported distances by the mode actually used, a close match between the Dutch and the German results can be observed for automobile (+10.7 percent versus +10.2 percent); however, the difference is more significant for public transit (+0.8 percent versus +4.5 percent). Of course, keep in mind that in the latter study only estimates by users of the mode under consideration were employed [Table 2 (4)].

A comparison of reported and actual travel times by automobile and transit shows significant overestimates [Table 3 (4)]. This is particularly true for automobile drivers in the judging of travel times by transit--61 percent overestimated transit travel time by more than 20 percent. Even transit riders overestimated their travel times on the average by 10.4 percent (Table 3).

The picture is much less dramatic for estimates of travel time by automobile, where drivers overestimate their times by 8.4 percent and transit riders judge automobile travel time on the average to be 4.0 percent more than the actual values (Table 3). When evaluating this information, keep the stratification of misestimates by percentage (provided in the table) in mind because the averages tend to veil a number of interesting details. For example, the largest percentage misestimate (28 percent) of automobile travel time lies in the 21 percent-plus stratum of transit riders.

For travel cost by automobile and transit [Table 4 (4)] we observed the well-known phenomenon that automobile travel costs tend to be underestimated by both automobile users and nonusers. The absolute misestimation is most pronounced for automobile drivers (62.4 percent) and the largest underestimation is in the 21 percent-plus stratum (57 percent). Naturally, transit riders report their transit fare accurately because the out-of-pocket costs for that mode are obvious and easy to recall.

Influence of Elapsed Time on Reported Trip Volume

Another serious influence on the results of mobility

Table 1. Difference between reported and actual travel distance.

Travel Mode	No. of Trips	Total Distance (km) Reported	Total Distance (km) Measured	Difference (%)
Car	1342	12 352	11 162	10.7
Truck or delivery van	40	860	814	5.6
Moped	100	464	401	15.7
Bicycle	843	1 583	1 467	7.9
Walk	1004	626	594	5.3
Public transit	104	735	729	0.8

Table 2. Reported travel distance by mode used for all trip purposes.

Estimate	Automobile Driver (n = 930) (%)	Transit Rider (n = 2327) (%)
Overestimation		
By 21 percent and more	22	17
By 11-20 percent	8	13
By 1-10 percent	3	6
Total	33	36
Correct	42	28
Underestimation		
By 1-10 percent	7	7
By 11-20 percent	12	18
By 21 percent and more	6	11
Total	25	36
Average misestimation	+10.2	+4.5
Absolute misestimation	|18.5|	|18.2|

Table 3. Reported travel time for all trip purposes.

Estimate	Perceived Travel Time by Transit (%) Automobile Driver (n = 400)	Perceived Travel Time by Transit (%) Transit Rider (n = 2380)	Perceived Travel Time by Automobile (%) Automobile Driver (n = 891)	Perceived Travel Time by Automobile (%) Transit Rider (n = 1306)
Overestimation				
By 21 percent and more	61	27	28	20
By 11-20 percent	10	16	11	11
By 1-10 percent	5	16	7	2
Total	76	59	46	33
Correct	9	9	7	14
Underestimation				
By 1-10 percent	3	14	10	9
By 11-20 percent	4	9	22	16
By 21 percent and more	8	9	15	28
Total	15	32	47	53
Average misestimation	+28.9	+10.4	+8.4	+4.0
Absolute misestimation	|36.5|	|20.6|	|25.8|	|31.2|

Table 4. Reported travel cost for all trip purposes.

	Perceived Travel Cost by Automobile (%)		Perceived Travel Cost by Transit (%)	
Estimate	Automobile Driver (n = 870)	Transit Rider (n = 841)	Automobile Driver (n = 331)	Transit Rider (n = 2397)
Overestimation				
By 21 percent and more	21	7	21	1
By 11-20 percent	4	5	15	1
By 1-10 percent	1	2	1	0
Total	26	14	37	2
Correct	5	22	38	93
Underestimation				
By 1-10 percent	5	4	2	2
By 11-20 percent	7	4	18	0
By 21 percent and more	57	56	5	3
Total	69	64	25	5
Average misestimation	-3.1	-25.4	+12.4	+0.2
Absolute misestimation	|62.4|	|39.7|	|22.2|	|2.4|

Table 5. Influence of elapsed time on reported trip volume.

	Travel Mode				
Elapsed Time Before Survey	Total	Automobile	Train	Air	Other
Intercity Vacation Trips During 1 Year					
First quarter, travel directly before survey	100.0	100.0	100.0	100.0	100.0
Second quarter, 3-6 months	100.0	96.8	100.0	100.0	100.0
Third quarter, 6-9 months	95.7	90.3	100.0	100.0	100.0
Fourth quarter, 9-12 months	87.0	83.9	85.7	100.0	100.0
Other Personal Intercity Trips During 3 Months					
1 month	100.0	100.0	100.0	100.0	100.0
2 months	92.1	91.2	100.0	100.0	100.0
3 months	80.3	79.4	100.0	50.0	100.0

studies can be traced to memory lapses that obviously increase with the length of time for which travel activities are to be reported by the respondent. An example from the investigation of intercity vacation and personal travel is selected to illustrate this point [Table 5 (5)]. For intercity vacation trips the recollection of trips decreases by up to 14.3 percent for train travel and 18.1 percent for automobile travel undertaken more than nine months prior to the reporting date. The average underestimate of travel by all modes is 4.3 percent after a six-month and 13.0 percent after a nine-month time lapse. Also note that the more significant underreporting occurs for the more common modes, namely automobile and train, and air and other constitute more memorable (less frequent and costlier) events and result in accurate reports of vacation trip making.

In general, the reporting of other personal trips is even less reliable than that of vacation trips. After only three months, 19.7 percent of all trips are lost due to memory lapses, with the automobile and air modes being the main factors. One implication seems to be that vacation trips are more memorable and, therefore, more easily recollected.

Panel Effects on Reported Mobility

Data obtained through the use of a survey panel generally are considered to be a reliable source of information input for research studies. Aside from the fact that time-series data can be obtained by this method, it also represents a certain level of efficiency in sampling. The standard virtue of this survey technique is that a more or less consistent set of sample elements is available.

Unfortunately, a number of disadvantages are also associated with panel surveys, aside from the problem of setting up a willing set of respondents. As is illustrated in Table 6 (6), which represents the results of travel activity and trip frequency reports for a three-year panel survey performed in Munich, West Germany, substantial decreases in reported mobility can be observed over the three-year reporting period. This apparent decrease in overall activity and trip frequency represents a special hidden form of nonresponse influence. The respondents who are basically willing to participate in the panel survey, more and more frequently return their second- and third-stage questionnaire with the remark that they did not partake in any out-of-home activities during the survey day. Consequently, the share of immobiles (i.e., respondents who claim not to have performed any trips at all) increased by 6 percent from year to year, or more generally from phase to phase, and the average trip frequency of the mobiles (i.e., those who report out-of-home activities) remained relatively stable. Due to the fact that survey panels tend to measure artifacts of the methods rather than results, survey researchers are becoming more hesitant to use panels.

Comparison of Oral Versus Written Responses

Recent methodological research into survey methods has established that mail-back and interview surveys will produce different results for the identical reported phenomenon (7, and Brog and Meyburg in this Record). Underreporting and poor reporting tend to be the rule for oral responses. Table 7 illustrates the substantially different level of accuracy for length of time of travel generated by the two survey techniques. The average deviation from the correct (objective) travel time is -11 percent and +36 percent for automobile and transit, respectively, when an oral survey is used; however, deviations of only +3 percent and +12 percent are registered for written surveys.

How fundamentally oral interviews influence the result of the survey can also be seen in the parameter of "number of activities listed." For this purpose, diaries were kept for a week in two random samples. The first group, after receiving adequate instructions, filled out each day's activities by themselves. In the second group, the persons were orally questioned on the first day, and then filled out days two through seven by themselves. The surveys differed from each other only in this procedural method on the first day (i.e., case 1 = written survey, case 2 = oral and written survey). If the number of listed activities in the first case is set equal to 100, we get the results shown in column 2 of the table below (7).

	Index of Activities	
Day	Case 1 (n = 1162)	Case 2 (n = 882)
One	100	85
Two	100	99
Three-seven	100	100

Although the answers on days two through seven are comparable, oral interviewing lowers the result noticeably. Responsible are the following factors:

1. Unconscious mistakes on the part of the interviewees, who are forced to completely remember something within a short period of time;

Table 6. Panel effects on reported mobility.

Item	Spring 1977 Total (n = 1938)	Spring 1977 City (n = 1152)	Spring 1977 Region (n = 786)	Spring 1978 Total (n = 1938)	Spring 1978 City (n = 1152)	Spring 1978 Region (n = 786)	Spring 1979 Total (n = 1938)	Spring 1979 City (n = 1152)	Spring 1979 Region (n = 786)
Travel activity									
Did not leave home on survey day (%)	12	10	14	18	15	21	25	21	31
Left home on survey day (%)	88	99	86	82	85	79	75	79	69
Trip frequency									
No trips (%)	12	10	14	18	15	21	25	21	31
One trip (%)	0	0	1	1	1	1	1	2	1
Two trips (%)	42	40	47	39	38	40	34	36	30
Three trips (%)	8	10	5	7	8	5	7	8	5
Four trips (%)	20	22	18	20	21	19	18	19	16
Five trips (%)	7	8	5	6	7	5	5	5	5
Six trips (%)	8	8	8	7	7	6	6	6	7
Seven or more trips (%)	3	3	3	2	2	3	4	3	4
Average mobility									
Average trips per person for all survey days	2.91	3.02	2.74	2.70	2.78	2.58	2.50	2.57	2.38
Average trips per person for survey days that had trips	3.29	3.36	3.18	3.27	3.28	3.26	3.35	3.27	3.47

Note: Percentages may not total 100 due to rounding.

Table 7. Deviations of oral and written responses with respect to length of travel.

Estimate	Length of Travel: Oral Survey, Automobile (n = 800)	Oral Survey, Public Transportation (n = 520)	Written Survey, Automobile (n = 1100)	Written Survey, Public Transportation (n = 538)
Correct, within the permissible limit[a] (%)	72	79	93	88
Wrong, not within the permissible limit[a] (%)	28	21	7	12
Average deviation[b]	89	136	103	112

[a]Permissible limit is ±25 percent. [b]Correct (objective) time = 100.

2. Conscious mistakes on the part of the interviewees, who are unwilling to give a stranger certain information; or

3. Influence of the interviewer, who attempts to complete the interview as quickly as possible.

The diary filled out in the presence of an interviewer measures 15 percent fewer trips than the one completed with more leisure by the interviewee alone. We may conclude that interview surveys will produce substantially reduced mobility levels due to survey characteristics alone. To that element other factors have to be added, such as nonresponse influences (see our other paper in this Record).

Influence of Survey Instrument Layout on Reported Mobility

The KONTIV travel survey (1) conducted in the Federal Republic of Germany is generally considered an excellent example of the application of state-of-the-art survey design, implementation, and instrument layout. Therefore, a number of comparisons have been performed to test other results against those generated by KONTIV. The comparison of the KONTIV results with another survey of the travel behavior of senior citizens in Germany showed some interesting differences with respect to the mobility characteristics of that segment of the population. These differences could be traced to the somewhat inferior design of the survey instrument of the second survey.

That survey instrument has the following characteristics: The beginning of the survey form contains a filled-in example that leaves the respondent unsure as to whether return trips are to be reported. Also, the trip sequence in the example does not fit chronologically. Another problem is that the form has to be flipped over after the third trip. As a result of these problems, the survey shows many more survey days with an uneven number of trips, as can be seen in the table below (1,8).

No. of Trips Reported	KONTIV Survey Days (%) (n = 13 710)	EMNID Survey Days (%) (n = 6411)
One	2	17
Two	52	47
Three	8	16
Four	23	14
Five	5	3
Six	6	2
Seven and more	4	1

In the EMNID survey the number of days when four or more trips were taken is 18 percent smaller than in the KONTIV results (largely due to turning the page). The total mobility is reported to be 23 percent lower (2.52 versus 3.10 trips/person on the survey day) and this may be due largely to poor design of the survey instrument.

Substantial research on the effects of different survey instrument layouts and survey administration has been performed at the Socialdata Institute for

Empirical Social Science Research in Munich. A voluminous paper would be required to detail all research findings on these topics. The pitfalls a survey designer might encounter range from the obvious and trivial to the very subtle. The simple example of column versus row arrangement of the survey instrument illustrates how lack of methodological insight can lead to incorrect results. The KONTIV survey (1) and a comparative experimental study by Socialdata showed that a mobility difference per household of 9 percent was observed when a column arrangement was used in conjunction with simple check-off response possibilities. Also, the response rate proved to be 12 percent larger with the column layout.

CONCLUSIONS

We have presented selected results of past and ongoing research in the area of empirical survey methods. It is hoped that this paper can make a contribution to an increased level of awareness and knowledge of the dangers of uncritical use of data in travel behavior research.

This paper indicates that methodological experiments are necessary to improve the generated base information. Use of perceived versus actual values, influence of elapsed time, panel effects, nonresponse, oral versus written responses, and survey instrument layout have been shown to severely affect the validity of the survey data. These error sources can severely undermine the relevance and validity of any models and modeling results based on invalid data. One should keep in mind, though, that these factors constitute only a subset of factors that can influence survey and, therefore, modeling results.

It is quite common in our research community to use data collected for other than our own purposes or collected by some other organization that provides insufficient or no knowledge as to the survey method, design, administration, or questionnaire layout characteristics that were employed in the generation of the data set. In short, frequently the analyst or modeler is completely removed from the data source.

Clearly, given the various influencing factors on the validity of survey data, this fact produces some doubts as to the validity of many modeling results. In the course of preparing this paper, we encountered substantial evidence of a certain amount of obliviousness by some researchers as to how the data they used were generated. This problem becomes particularly serious when data of unknown or questionable validity are used in modeling efforts that result in claims of providing new insights into travel behavior. In our opinion, there is substantial reason to question the validity of these claims. It is quite evident that research results can be subject to manipulation by means of choice and execution of a survey procedure. For example, the results of ignoring the influence of nonresponse were illustrated by us for a number of cases (2, and in our other paper in this Record). We might also point out that sometimes one encounters a certain level of puzzlement and bemusement when one tries to get the researcher to detail the methods used to obtain the data that form the fundamental source of claims to new research insights. It is more than careless to conduct surveys without proper documentation of all details of the data collection method used or to use data without knowledge about the source and the survey method.

ACKNOWLEDGMENT

We gratefully acknowledge the useful inputs by Alfred Nickesen, Stephen Ritchie, and Orlando Strambi of Cornell University, who reviewed an earlier draft of this paper.

REFERENCES

1. Sozialforschung Brög. Kontinuierliche Erhebung zum Verkehrsverhalten, KONTIV 75,76,77 [Continuous Survey of Travel Behavior, 1975, 1976, 1977]. Ministry of Transport, Bonn, West Germany, 1977.
2. W. Brög and A.H. Meyburg. Nonresponse Problem in Travel Surveys: An Empirical Investigation. TRB, Transportation Research Record 775, 1980, pp. 34-38.
3. The Extent to Which People Can Estimate Their Trip Distances. Institute for Road Safety Research, Voorburg, Netherlands, Feb. 1978.
4. W. Brög and N. Meyer. From Price Elasticity to Price Susceptability--A Reality-Oriented Model for Influencing Transportation Fees According to Demand. Socialdata GmbH, Munich, West Germany, 1978.
5. Socialdata GmbH. Marktanalyse des Verhaltens im Personenfernverkehr [Market Analysis of Intercity Travel Behavior]. Ministry of Transport, Bonn, West Germany, 1979.
6. KONTIMUC: Survey of Transit Use in Metropolitan Munich. Socialdata, GmbH, Munich, West Germany, 1977.
7. W. Brög and K.H. Neumann. The Interviewee as a Human Being--A Critical Contribution to the Discussion of Methods in Empirical Social Research. Paper presented at the 44th European Society for Opinion Surveys and Market Research Seminar on Ways and New Ways of Data Collection, Jouy-en-Josas, France, Nov. 1977.
8. EMNID Institute. Type, Duration, Frequency as Well as Purpose and Destination of Travel Activities by People of Age 60 and Older in Selected Regions of the Federal Republic of Germany (in German). Bielefeld, West Germany, Sept. 1976.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

Transit Trip Distribution Model for Multimodal Subarea Focusing

STEPHEN M. HOWE, YEHUDA GUR, AND DAVID L. KURTH

This paper describes the development of a model for distribution of transit person trips, separately from automobile trips, as an integral component of a multimodal subarea focusing methodology. The extension of subarea focusing to multimodal transportation planning is reviewed. The components of the multimodal transportation analysis process, particularly the interaction between disaggregate mode choice estimation and mode-specific trip distribution, are then described. Transit trip distribution model theory and design are then presented, featuring the simultaneous distribution of distinct transit trip classes. An analysis of observed travel patterns supports hypothesized differences between segments of the travel market. Calibration results show a high degree of accuracy in estimated trip patterns and subsequent transit assignment and, therefore, point to greater precision in assessing the effects of transit-oriented actions.

In the past five years the North Central Texas Council of Governments (NCTCOG) has invested heavily in the development of a refined travel forecasting methodology, designed specifically for detailed subregional planning. The methodology is designed to answer many of the planning needs that arise from increased emphasis on (a) transportation system management strategies for more efficient short-term use of existing facilities, (b) transportation control measures for improvement of air quality, and (c) analysis of alternative major transit investments at the subregional or corridor level. To integrate the various evaluation and decision-making processes into one consistent, efficient work effort, NCTCOG relies heavily on macrosimulation models analogous to the Urban Transportation Planning System (UTPS).

The NCTCOG methodology was first implemented in the form of the thoroughfare analysis process (TAP), which was designed for highway planning at the subregional level (1-4). The extension of subarea focusing to multimodal analysis was undertaken in response to Urban Mass Transportation Administration (UMTA) guidelines of 1976, which require rigorous analysis of alternatives in support of proposed major transit investments. A fully operational feature of the new Multimodal Transportation Analysis Process (MTAP) is the focusing of transit networks. The design of MTAP and the calibration and validation of the models for the evaluation of proposed transitway technologies and alignment along Dallas' North Central Expressway are documented in a forthcoming report (5).

MTAP

The transit trip-distribution model described in this paper could be presented as a stand-alone development effort. In reality, however, this work was an integral part of the overall MTAP development effort. This section provides a brief overview, with particular attention to the challenges posed by subarea focusing and the rationale behind the model structure ultimately adopted.

MTAP features that pertain specifically to transit analysis include the following:

1. Incorporation of the UTPS program INET to use TAP's detailed highway networks in construction of compatible transit networks;
2. Computerized construction of transit approach links;
3. Specification of transit access and egress impedances in the form of frequency distributions, rather than zonal averages, to permit explicit treatment of intrazonal variation;
4. Preparation of stripped transit skim trees, to reflect line-haul impedances only (i.e., those impedance components that have little or no intrazonal variation);
5. Specification of trip maker characteristics in the form of frequency distributions, again to permit sampling of intrazonal variation;
6. Disaggregate mode choice estimation, driven by Monte Carlo sampling and weighting of individual trips; and
7. Distribution of transit person trips, separately from automobile vehicle trips, and explicit treatment of distinct transit trip classes.

Design

The overall MTAP design is schematically presented in Figure 1. The zonal and network master files are structured hierarchically. The zonal hierarchy ranges from 40 jurisdictions at the coarsest level to 5000 traffic analysis zones at the finest level. The hierarchical structuring of the highway network data is analogous to that for zones--level 1 links provide connectivity for jurisdictions, the addition of the level 2 network provides connectivity for the second coarsest zone level, and so on. The structuring of the transit line data base is somewhat simpler and serves mainly to identify those lines likely to be accessed by automobile.

Subarea focusing is then performed to extract subfiles tailored specifically to the analysis. Full detail is preserved within the area of interest, and progressively less detail is given away from the area of interest. Approach links are automatically constructed after zonal and network subfiles have been extracted from the master file.

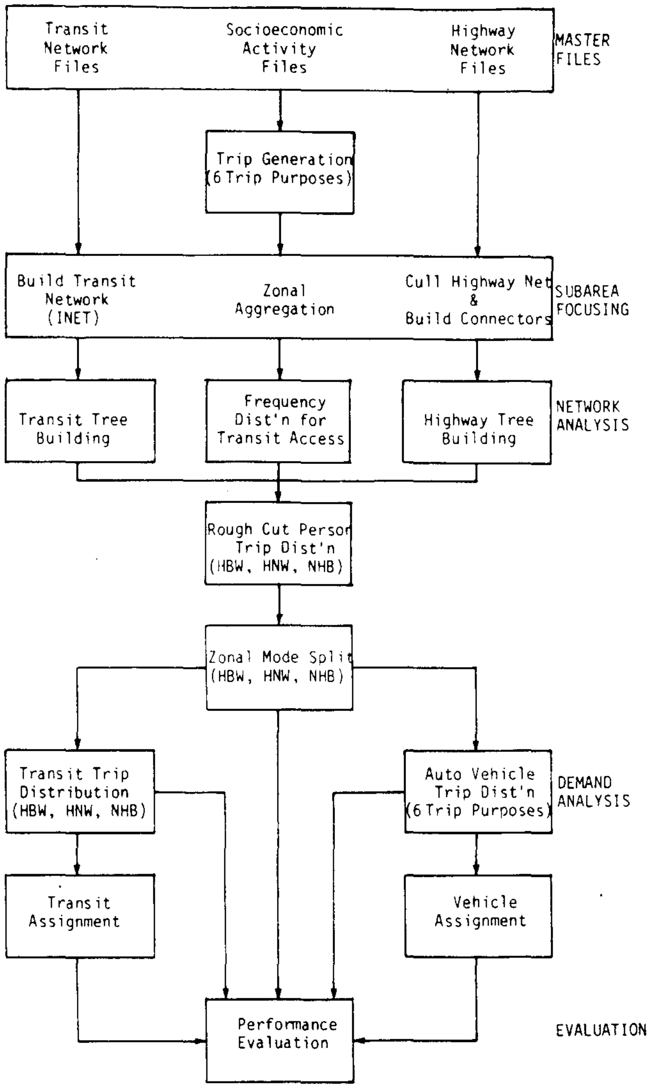
The network analysis phase follows, in which the following functions are performed:

1. Highway and transit network tree building and skimming,
2. Preparation of stripped (line-haul) transit skim trees, and
3. Calculation of zonal frequency distributions for transit access-egress impedances.

Certain criteria are used in the creation of the stripped skim trees to remove network impedances that are expected to have large intrazonal variation (e.g., access-egress by walk or by short feeder-bus legs). The intrazonal variability is explicitly dealt with later, by Monte Carlo sampling from the access-egress frequency distributions.

In the demand analysis phase, MTAP departs from the conventional modeling sequence in that mode split precedes trip distribution and is performed at the zonal (trip end) rather than the zone-pair (trip-interchange) level. The mode split thus provides trip ends stratified by mode. Automobile vehicle trips and transit person trips are subsequently distributed separately, each based on their respective networks. Finally, vehicle and transit trip tables are assigned to their respective net-

Figure 1. MTAP design.



works and evaluation of the results takes place.

Principles

MTAP aids in the analysis of problems when estimates of highway link volumes and transit line boardings are needed for evaluation. Mode-specific assignments and, hence, mode-specific trip tables are necessary to address such problems. Further, MTAP is designed for detailed but cost-effective analyses at the subregional or corridor level. Thus, the challenge has been to find a model structure for producing mode-specific trip tables that is robust enough to withstand drastic variation in zone structure without necessitating extensive calibration or adjustment for each subarea. Finally, a time element was introduced by the need to meet the schedule for upcoming alternatives analyses. Thus proven methodologies were used as much as possible.

The conventional modeling sequence (i.e., multimodal person trip distribution followed by interchange-level mode split) was considered with skepticism. Aggregate models for multimodal trip distribution abound but are more complex than mode-specific distribution (e.g., accounting for transit captivity effects, or measuring multimodal impedances) and, in light of our experience with automobile vehicle trip distribution under focusing (4), did not appear promising for short-term implementation. Disaggregate models for destination choice would, in theory, be more promising under variable zone structures, but experience with such models in producing trip tables suitable for accurate assignment remains somewhat limited.

On the other hand, disaggregate sampling for mode split seemed suitable for use either at the trip-end or trip-interchange level. Disaggregate mode-choice models, applied in a disaggregate sampling and weighting framework, permit accurate treatment of intrazonal variability in socioeconomic characteristics and access-egress impedances and immediately resolve problems incurred in the application of disaggregate models at aggregate levels (6). Performance of mode split at the trip-end level permits the sampling of intrazonal variability more cost effectively than would be possible at the trip-interchange level and also places less demand for accuracy on the mode split and the preceding calculation of person trip weights.

With these considerations in mind, we opted for trip-end mode split followed by separate trip distributions for automobile vehicles and transit persons. Separate distribution of transit person trips allows the unique characteristics of transit to be addressed free from the overwhelming effects of automobile travel. We also capitalized on proven methodologies for mode split and automobile vehicle distribution, and the distribution of transit person trips remained a relatively low-risk research and development project. A relatively simple model for multimodal person trip distribution was devised for constructing matrices of person trip weights.

Mode Split

The inputs required by the mode split program include the highway and transit network skim trees and, for each zone, frequency distributions for transit access-egress impedances and trip-maker (socioeconomic) characteristics. To describe transit access-egress for each zone, the primary frequency distribution required is that for travel distance. Other terminal impedance variables (e.g., travel time and travel cost) are calculated as functions of distance. Ultimately, as many as four competing transit access-egress submodes are represented for each zone, depending on mode availability: walk, feeder-bus, park-and-ride, and kiss-and-ride. For representation of trip-maker characteristics, a multivariate frequency distribution for income, household size, and automobile ownership is derived for each zone from input zonal averages by a process similar to marginal weighting (7). The theory and process by which the access-egress impedances and the trip-maker characteristics are represented are somewhat involved and a complete description lies beyond the scope of this paper. The procedures are well-defined and quite efficient, however, and have been successfully applied in mode split models for the Chicago Area Transportation Study (8) and the Northeast Ohio Areawide Coordinating Agency.

An additional input required by the mode split program is a matrix of person trip weights. The weighting matrix is derived by a preliminary distribution of person trips, based on a relatively crude multimodal impedance formulation that is sensitive to both highway and transit level of service. For a given origin zone, individual trips are sampled and weighted in proportion to the weighting matrix vector that corresponds to this zone. Since this matrix is used solely as a destination weighting matrix, the demand for accuracy is far less than is required in a trip interchange mode split model. The weighting matrix can only affect the final, mode-specific tables through the trip-end mode splits. The latter are not affected significantly by inaccuracies in the matrix of person trip weights as long as the multimodal trip table is relatively correct [for example, in the distribution of origins for trips to the central business district (CBD) and in the proportion of central city versus suburban destinations for trips that originate in various parts of the region].

For a given origin zone, the mode split program ultimately derives aggregate mode splits as follows. First, potential destinations from this zone are sampled in proportion to the corresponding weighting matrix vector. A sample of 50-200 individual trips results, along with the weighting factors necessary to expand the sampled trips to the total trips for the zone. For each individual trip, highway impedances and transit line-haul impedances are immediately available from the skim trees. Transit access and egress travel distances are sampled from the frequency distributions for the origin and destination zone. Other impedance values for the available transit access-egress modes are then calculated as functions of access-egress distance. Trip-maker (socioeconomic) characteristics are also determined by sampling from zonal frequency distributions. When all necessary values are thus determined, disaggregate choice probabilities are calculated for each individual trip by using a nested multinomial logit model. Finally, the sample results are weighted and summed to obtain the aggregate (zonal) mode splits.

For each zone, the following mode splits are obtained for trip productions: (a) automobile versus transit and (b) transit access mode (walk versus feeder bus versus park-and-ride versus kiss-and-ride). For input to transit trip distribution, the transit trip productions are also broken down by class:

1. Trips to the CBD,
2. Corridor trips (non-CBD trips that do not include transfers), and
3. Other trips (non-CBD trips that include one or more transfers).

The classification is obtained by a straightforward tabulation during the sampling and weighting of individual trips: Each trip is classified based on type of destination (CBD versus non-CBD) and number of transfers. The mode splits calculated for trip attractions for each zone are somewhat simpler: (a) automobile versus transit and (b) transit egress mode (walk versus feeder bus). The trip attraction mode splits are obtained by accumulation of destination-end effects during origin-zone processing. The user also has the option of processing each zone as a destination, analogous to the processing of origins; but, simple accumulation is obviously cheaper and has proven effective.

TRANSIT TRIP DISTRIBUTION

Theory

The classification of transit trips was deemed to be an important prerequisite for enhanced simulation of transit travel patterns because of the dual nature of transit travel. Transit patrons include transit captives and noncaptives (those who are free to choose between automobile and transit). If we assume that noncaptives choose the best mode available, then transit trips by noncaptives are likely to be concentrated where transit service is competitive with automobiles. Transit captives, on the other hand, are limited in the selection of their destinations to areas that are served by transit. Transit trips by captives are largely dictated by activity pattern rather than by choice of mode, hence are likely to be less concentrated than trips by noncaptives.

The classification of trip productions performed in mode split accounts for captivity effects without the difficult task of definition, measurement, and prediction of captivity. The three transit trip classes are defined below, in order of decreasing transit travel intensity, transit level of service, and, most likely, percentage of transit patrons who are noncaptive.

The CBD trip class includes all trips for which the attraction end lies in the CBD. Since the CBD trip class generally receives the most-intensive transit service, this class would be expected to have the highest proportion of noncaptives.

Among transit trips that are not attracted to the CBD, those made without transfer generally enjoy better service than those that incur transfers. Thus, the corridor trip class is defined to include non-CBD trips that do not include transfer, although not all such trips lie within clearly delineated radial corridors. Only line-haul transfers are used to identify corridor trips. For example, a non-CBD trip that uses short feeder access to priority service would be considered corridor unless a subsequent transfer to a different line-haul mode was made. For this class, the proportion of noncaptives is perhaps less than that for CBD trips but is expected to be greater than for trip interchanges with transfers.

The other trip class includes all non-CBD trips that incur one or more (line-haul) transfers. The proportion of noncaptives is expected to be the smallest for this class.

In the modeling process, the class distinction permits control over the range of attractions from which a given transit patron may select a destination. CBD patrons are forced to find an attraction in the CBD, corridor patrons must choose from a relatively tight range of non-CBD attractions, and all remaining attractions are available to other patrons. If a zone has high automobile ownership, hence fewer transit captives, then in mode split, transit will capture a significant share only when transit is competitive with the automobile. The CBD and corridor classes will tend to dominate among the transit trips produced by such a zone. Later, when transit trips from this zone are distributed, a large percentage will be restricted to CBD or corridor interchanges where transit service is competitive. Conversely, for a zone that has low automobile ownership (more captives), the portion of other trips will be larger and a generally broader range of attractions will be available.

Design

In MTAP, transit person trips are distributed by the program TTDGRAV, a gravity-model formulation adapted from the access and land development model originally developed by Schneider (9). The basic gravity formulation may be expressed as

$$T_{ij} = P_i [G(F_{ij})A_j / \sum_r G(F_{ir})A_r] \quad (1)$$

where

T_{ij} = number of trips produced by zone i and attracted to zone j;
P_i = total number of trips produced by zone i;

G(F) = travel (decay) function, which represents the decline in attractiveness as a function of increasing travel impedance;
F_{ij} = impedance of travel from zone i to zone j;
A_j = attractiveness (number of trip attractions) for zone j; and
r = one through total number of zones.

In the TTDGRAV model, the basic formulation is extended to handle simultaneous distribution of distinct trip classes,

$$T_{ij} = P_i^{c_{ij}} [G(F_{ij})A_j / \sum_{r \epsilon c_{ij}} G(F_{ir})A_r \quad (2)$$

where

c_{ij} = class of the ijth interchange,
$P_i^{c_{ij}}$ = number of trip productions of class c from zone i, and
A_j = number of trips attracted by zone j (all classes).

As in typical gravity applications, the model is applied iteratively to balance trips received to trip attractions (i.e., $\sum_i T_{ij} = A_j$ for each zone j).

The specific inputs required by TTDGRAV are the following:

1. Transit trip productions and attractions, with productions stratified by trip class;
2. Stripped transit skim trees; and
3. Standard (UTPS) transit skim trees.

The trip ends, stratified as indicated, are provided by the mode split program. The stripped skim trees provide the number of line-haul transfers used in the classification of trip interchanges and a priority mode indicator that is used to identify nontransit interchanges. The UTPS skim trees provide total weighted impedance, which is the basic zonal separation measure used in transit distribution. The total impedance measure is a weighted sum of access times, run times, wait times, and transfer times.

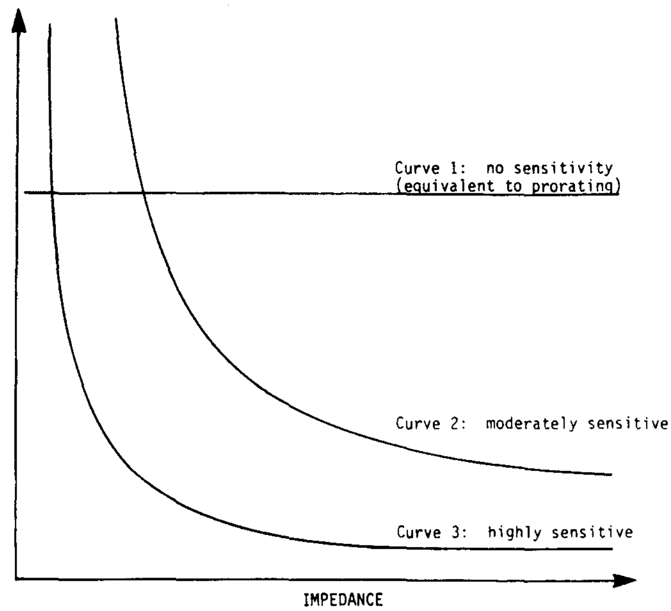
TTDGRAV CALIBRATION AND SENSITIVITY ANALYSIS

The calibration of TTDGRAV was carried out within the context of a major transit investment alternatives analysis for the North Central Expressway corridor in Dallas, Texas. The subarea focusing methodology was used for a highly detailed presentation of this corridor within the NCTCOG study region.

Extensive on-board survey data (10,11) provided an exceptionally rich data base for observed travel patterns on the Dallas Transit System (DTS). Approximately 30 000 boardings were surveyed and weighted to represent approximately 87 000 total trips and 110 000 total boardings, including transfers. From these data, observed transit trip ends and trip tables [home-based work (HBW), home-based nonwork (HNW), and nonhome based (NHB)] were constructed for calibration of distribution of transit trips. For the sake of brevity, only results for HBW will be discussed in detail.

To assess goodness of fit, the 504 zones were aggregated to 61 districts and 10 superdistricts. The districts and superdistricts conformed to logical boundaries in the region (e.g., major thoroughfares and the CBD). In addition, districts and superdistricts were also focused on the corridor to reduce the possibilities of improving fit simply through aggregation. The calibration procedures and criteria were structured as follows:

Figure 2. Alternative decay functions.



1. Comparison of observed versus estimated trip tables: R^2 calculated for district-level interchanges, average trip length by trip class within superdistrict, percentage of trips received by trip class for each superdistrict, and possible trends or biases in average trip lengths for individual zones; and
2. Comparison of ULOAD assignments of observed versus estimated trip tables: percentage root mean squared error (RMSE) of boardings by line and maximum loading points and load volumes on individual lines.

The parameters to be calibrated included those that define the shape of the decay function [G(F) in Equation 1]. The decay function is a concave, monotonically decreasing function that reflects traveler sensitivity to impedance in selecting among potential destinations. Examples are shown in Figure 2. The steeper the curve, the greater the (simulated) sensitivity.

Other parameters to be considered and calibrated, if necessary, were two types of fixed penalty (i.e., surcharges) added to the basic zonal separation measure in certain instances. The two penalty types are (a) a transfer penalty, possibly stratified by trip class, and (b) production-end and attraction-end penalties for certain groupings of zones. The penalties permit adjustment to account for phenomena not otherwise represented in the basic impedance measure.

Decay Function Options

The spatial allocation of trip attractions around the production end of a single trip production is referred to here as the opportunity surface. The opportunity surface governs the probabilities for selecting among competing trip attractions and consequently affects average trip length and other characteristics of trip distribution from the given production zone. If travel impedance did not have a deterrent effect, then all attractions in the opportunity surface would have the same probability of

Table 1. Opportunity surface analysis for HBW trips.

	CBD			Other			Corridor		
Zone	Avg Trip Length	Avg Opportunity Length	Ratio	Avg Trip Length	Avg Opportunity Length	Ratio	Avg Trip Length	Avg Opportunity Length	Ratio
18	57.1	57.8	0.99	120.6	115.5	1.04	74.8	85.3	0.88
77	53.5	54.8	0.98	100.8	109.5	0.92	74.4	93.8	0.79
90	76.1	75.9	1.00	120.0	133.0	0.90	71.4	80.0	0.89
132	122.8	122.7	1.00	163.1	177.4	0.92	117.5	112.1	1.05
223	79.6	78.2	1.02	127.3	129.6	0.98	80.9	73.0	1.11
228	61.6	61.3	1.00	124.0	133.7	0.93	92.1	101.7	0.91
246	57.8	57.2	1.01	142.0	126.2	1.13	82.0	101.7	0.81
294	58.8	56.0	1.05	107.4	111.4	0.96	68.1	87.0	0.78
335		20.6		84.7	73.9	1.15	78.7	69.9	1.13
374	75.8	79.9	0.95	120.6	134.7	0.90	77.6	98.7	0.79
421	97.6	98.5	0.99	166.9	168.7	0.99	139.8	142.7	0.98
433	74.7	75.5	0.99	121.3	131.2	0.92	78.8	76.3	1.03
473	87.4	87.5	1.00	142.8	141.0	1.01	84.1	85.0	0.99

Note: Length is total weighted transit impedance in minutes.

serving as destinations (i.e., the average trip length would be equal to the average distance of opportunities from the production end). The amount by which the average opportunity distance exceeds average trip length indicates the extent of traveler sensitivity to impedance.

Table 1 lists the average HBW trip impedance, the average HBW opportunity impedance, and the ratio of trip impedance to opportunity impedance for a small sample of zones. The zones were selected judgmentally to represent different parts of the study region and also to ensure at least five trips in each trip class. The impedance ratio is relatively high in all trip classes and ranges from 0.79 to greater than 1.0. This phenomenon marks a clear distinction of transit travel from highway or person-travel. The sensitivity to impedance is much less than observed elsewhere (2) for automobile vehicle trips.

For the CBD trip class, the ratios in Table 1 are all close to 1.0, which indicates virtually no sensitivity to impedance. Simple prorating is thus considered for distribution of CBD trips, separate from the decay function for corridor and other trip classes. In the other trip class, the ratios range from 0.89 to 1.0, which indicates a small degree of sensitivity and suggests use of a very shallow travel function. The corridor trips display more sensitivity to impedance; this does not necessarily imply that distinct decay-function parameters are required, however. Corridor trips are generally shorter than other trips, and the slope of a typical decay function (Figure 2) decreases with increasing impedance, which means that the portion of the travel function applicable to corridor trips will generally be steeper than that for other trips.

Sensitivity Analysis

The sensitivity issues to be explored include the following:

1. Sensitivity to parameters that govern the steepness of the decay function,
2. Selection of travel functions for different trip classes, and
3. Sensitivity to fixed penalties.

For the initial attempt to simulate HBW transit trip patterns, it was decided to use a relatively shallow decay curve for all three trip classes. No fixed penalties were applied in this base run. The goodness of fit attained by this initial run is evidenced by an R^2 of 0.98 on all nonzero district interchanges. In the superdistrict summary of trips and trip lengths, by class, nearly all items were accurate to 5 percent or better. The following discrepancies, however, were noted:

1. Trips from the CBD were too short;
2. Average length for corridor trips was understated by 3 percent, and
3. Average trip length was consistently short for zones served by the main crosstown route.

Interestingly, problems seem to occur where the distinction between corridor and other trips is unclear (trips from the CBD, trips served by the crosstown).

To assess sensitivity to the shape of the decay curve, two additional test runs were made in which decay-curve steepness was first increased and then decreased. The goodness of fit was excellent in all cases and there was surprisingly little difference in the test results despite order-of-magnitude variation in decay function parameters. These results suggest that much of the robustness of the model is attributable to the trip classification scheme.

Next, simple prorating (constant decay function) was considered for the CBD trip class. The base run parameters slightly underestimated average impedance for the CBD class, and prorating overestimated the average impedance. Some sensitivity to impedance, however slight, appears to be required in the travel function. To correct for underestimation of average impedance by the base run parameters, a transfer penalty is imposed on the CBD class.

Remaining discrepancies noted in the base run were addressed by introduction of additional fixed penalties. A transfer penalty was imposed on the corridor trip class, which reduced the error in the (underestimated) average impedance by 46 percent. A production-end penalty was assessed for trips produced by the CBD, which reduced the underestimation of average impedance by 64 percent. To increase trip lengths on the crosstown service, error was reduced by 76 percent via imposition of a production-end penalty on zones thus served.

Final Calibration Results

The final comparison of observed versus estimated trip tables is shown in Table 2. The overall R^2 exceeded 0.98 for nonzero interchanges. The superdistrict comparisons show better than 10 percent accuracy for nearly all entries. Average trip length for total trips by class is estimated within 2 percent accuracy in all cases. The final results generally show a high degree of accuracy in replicating

Table 2. Observed versus estimated trip lengths.

Superdistrict	CBD Trips Sent	CBD Trips Received	CBD Avg Length Sent	CBD Avg Length Received	Other Trips Sent	Other Trips Received	Other Avg Length Sent	Other Avg Length Received	Corridor Trips Sent	Corridor Trips Received	Corridor Avg Length Sent	Corridor Avg Length Received	Total Trips Sent	Total Trips Received	Total Avg Length Sent	Total Avg Length Received
Observed HBW Trips																
1	25	34 662	31.6	73.3	52	0	93.8	0.0	857	0	78.3	0.0	934	34 662	78.0	73.3
2	2 455	0	55.2	0.0	603	1 357	112.1	112.0	235	506	76.9	86.8	3 293	1 863	67.1	105.1
3	1 834	0	74.2	0.0	317	1 827	125.4	129.9	185	240	84.5	79.9	2 336	2 067	82.0	124.1
4	1 142	0	74.7	0.0	52	430	140.8	147.5	48	64	100.0	110.1	1 242	474	78.5	142.7
5	3 366	0	84.7	0.0	1 443	4 275	138.5	124.7	750	1174	105.1	105.8	5 559	5 449	101.4	120.6
6	4 229	0	71.4	0.0	3 166	1 187	120.2	114.0	442	784	78.2	75.5	7 837	1 971	91.5	98.7
7	6 515	0	66.9	0.0	1 664	2 320	116.5	116.2	509	741	84.7	86.4	8 688	3 061	77.4	109.0
8	3 889	0	84.4	0.0	213	334	144.7	165.0	246	71	119.9	96.8	4 348	405	89.3	153.1
9	5 993	0	78.8	0.0	1 580	868	124.3	120.6	547	484	98.4	83.9	8 120	1 352	81.8	107.4
10	5 214	0	69.3	0.0	4 240	732	122.8	124.7	834	589	80.1	80.1	10 288	1 321	92.2	104.8
Total	34 662	34 662	73.3	73.3	13 330	13 330	123.2	123.2	4653	4653	88.6	88.6	52 645	52 645	87.3	87 3
Estimated HBW Trips																
1	25	34 662	30.6[a]	73.3	52	0	89.0[b]	0.0	867	0	75.6	0.0	944	34 662	75.3	73.3
2	2 455	0	56.2	0.0	603	1 412	110.6	111.1	235	464[b]	80.4	83.2	3 293	1 876	67.9	104.2
3	1 834	0	74.4	0.0	317	1 798	128.4	131.0	185	283[a]	84.2	83.0	2 336	2 081	82.5	124.4
4	1 142	0	74.4	0.0	52	438	135.0	150.0	48	54[b]	108.1[b]	105.0	1 242	492	78.2	145.5
5	3 366	0	84.6	0.0	1 443	4 053[b]	139.0	124.6	750	1321[a]	101.8	102.2	5 559	5 374	101.1	119.1
6	4 229	0	71.0	0.0	3 166	1 221	118.9	115.2	442	766	79.9	75.5	7 837	1 987	90.8	99.9
7	6 515	0	66.9	0.0	1 664	2 300	117.4	116.8	509	790[b]	88.7[b]	84.5	8 688	3 090	77.9	108.5
8	3 889	0	84.8	0.0	213	370[a]	151.7	164.5	246	74	121.2	112.0[a]	4 348	444[b]	90.1	155.8
9	5 993	0	78.6	0.0	1 581	920[b]	127.3	121.4	547	411[a]	102.5	90.5[b]	8 121	1 331	89.7[b]	111.9
10	5 214	0	68.9	0.0	4 240	819[b]	122.1	119.2	834	500[a]	76.8	78.3	10 288	1 319	91.4	103.7
Total	34 662	34 662	73.3	73.3	13 331	13 331	123.2	123.2	4663	4663	88.3	88.3	52 656	52 656	87.3	87.3

Note. Length is total weighted transit impedance in minutes.
[a]Error more than 10 percent. [b]Error more than 5 percent but less than 10 percent.

Table 3. Assignment of observed and estimated trip tables.

Sector	No. of Routes	HBW Observed	HBW Estimated	Percentage Difference	Percentage RMSE
Southwest[a]	7	22 072	22 088	0.7	0.7
Southeast[a]	7	9 978	10 014	0.4	1.7
Northeast[a]	4	9 485	9 517	0.3	1.3
North Central-East[b,c]	6	6 123	6 171	0.8	1.8
North Central-West[b,c]	7	6 619	6 719	1.5	2.7
Northwest[a]	6	11 972	11 929	-0.4	1.6
Crosstown[c]	2	956	913	-4.5	9.6
Total	39	67 205	67 351	0.2	1.4

[a]Unit of comparison = cluster of Dallas Transit System (DTS) routes.
[b]Within North Central Expressway corridor.
[c]Unit of comparison = individual DTS routes.

observed travel patterns.

In an additional calibration test, observed and estimated trip tables were assigned to the transit network by the UTPS transit assignment program ULOAD. Results were compared to examine the propagation of transit trip table errors through assignment. Summary results are shown in Table 3. Estimated boardings were compared with observed boardings (actually, boardings that resulted from assignment of the observed trip table) for each route. The comparison shown in Table 3 generally indicates that the effect of errors in trip table assignment is minimal. The overall RMSE is remarkably low--1.4 percent.

In more detailed examination of the assignment results (not shown), the maximal load point and the maximal load were compared for each route. In nearly all cases, the maximal load point was the same in both the estimated and observed assignments, and the maximal load was estimated within 5 percent accuracy. Thus, even at a highly disaggregate level, evaluation of the accuracy of TTDGRAV, in terms of the effect on assignment results, seems quite favorable.

CONCLUSIONS

In response to a growing need for detailed planning of multimodal transportation facilities at the subregional level, NCTCOG has undertaken the development of a multimodal subarea focusing methodology, MTAP. For greater stability under focusing and validity in simulating travel patterns for distinct segments of the travel market, MTAP employs disaggregate mode choice estimation in conjunction with mode-specific trip distributions.

The separate distribution of transit person trips, with explicit treatment of distinct transit trip classes, affords several advantages. First, travel patterns of transit users are clearly distinct from those of automobile users: The spatial allocation of opportunities is dictated by the presence of transit service; further, observed sensitivity of transit travelers to impedance is markedly less than for automobile users. Second, additional distinctions may be drawn between captive and noncaptive users of transit, since the latter group is more likely to use transit in CBD and intracorridor interchanges where transit is more competitive with the automobile. The MTAP classification scheme permits the capturing of differences in travel patterns.

Specific conclusions from the calibration of the transit trip distribution model include the following:

1. The model replicates observed travel patterns with a high degree of accuracy;
2. Model performance is relatively insensitive to decay-function parameters, which suggests that much of the apparent robustness is attributable to the classification scheme;
3. Despite observed differences between trip classes, use of a common travel function together with the fixed penalties for selected categories appears feasible;
4. More precise evaluation of transit-oriented policies is possible with such a model.

The success in formulating a reliable transit trip distribution model validates a major MTAP design decision for attaining robustness under subarea focusing: It is worthwhile to forgo the estimation of interchange specific mode splits and to concentrate on improved estimation of transit trip ends. With the improved estimation of trip ends, the transit trip distribution model can provide excellent trip tables for analysis purposes.

The results in this paper indicate that a workable travel demand model structure can be obtained by integrating the strengths of disaggregate (primarily mode split) and aggregate (e.g., trip distribution) models. For modeling within a subarea focusing framework, in particular, this approach shows much promise for obtaining credible results for practical applications.

ACKNOWLEDGMENT

The preparation of this report was financed in part through a grant for technical studies from the Federal Highway Administration, the Urban Mass Transportation Administration, and the U.S. Environmental Protection Agency.

REFERENCES

1. John Hamburg and Associates, Inc. and R.H. Pratt Associates, Inc. Functional Specifications for the North Central Texas Thoroughfare Planning System. North Central Texas Council of Governments, Arlington, Nov. 1976.
2. S. Howe and Y. Gur. Trip Distribution in Subregional Analysis. TRB, Transportation Research Record 673, 1978, pp. 165-171.
3. S. Howe, T. Ryden, and D. Penny. Subarea Diagnostic and Evaluation Procedures for Programming Short-Range Transportation Improvement. TRB, Transportation Research Record 698, 1979, pp. 9-15.
4. D.L. Kurth, S.M. Howe, and Y. Gur. Stability Tests and Enhancements in Trip Distribution for Subarea and Corridor Level Planning. TRB, Transportation Research Record 771, 1980, pp. 6-12.
5. S. Howe. Dallas North Central Subarea Transportation Study--Travel Model Development Report. North Central Texas Council of Governments, Arlington (forthcoming).
6. F.S. Koppelman. Methodology for Analyzing Errors in Prediction with Disaggregate Choice Models. TRB, Transportation Research Record 592, 1976, pp. 17-23.
7. L.T. Ollmann, S.M. Howe, K.W. Kloebar, and G.S. Cohen. Marginal Weighting of Transportation Survey Data. TRB, Transportation Research Record 677, 1978, pp. 73-76.
8. Y. Gur, D. Kopec, E. Ryan, A. Vyas, and E. Lowe. Estimation of Demand for Public Transportation. TRB, Transportation Research Record 728, 1979, pp. 76-79.
9. M. Schneider. Access and Land Development. In Urban Development Models, HRB, Special Rept. 97, 1968, pp. 164-177.
10. Wilbur Smith and Associates. Dallas Bus Operational Study. Dallas Transit System, Dallas, 1971.
11. Wilbur Smith and Associates. Transit Operational Study. Dallas Transit System, Dallas, 1978.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.