York City are to or from Manhattan south of 59th Street. Corresponding figures are 84 percent for Boston; about 70 percent for Chicago, Toronto, Cleveland, and Philadelphia; and 64 percent for BART in San Francisco. Second, approximately 20-35 percent of all riders walk to and from stations on most existing systems (New York excluded). This range may reduce to 15-25 percent for proposed suburban-oriented systems.

Freeway Volumes

Freeway use depends upon the extent of the system and the size of the urban area. The proportion of vehicle-kilometers on freeways increases in general proportion to the relative amount of total capacity provided by freeways. For example (1 km = 0.62 mile):

Percent Freeway Capacity	Percent Vehicle-Kilometers on Freeways for Population	
	50 000-200 000	1 000 000+
20	7	12
40	20	30

From the ranges in observed maximum freeway volumes for urban areas of various sizes, it is clear that forecast daily volumes of over 200 000 should be carefully rechecked, since these are found only in a few very large cities today, on roadways with more than eight lanes.

OVERVIEW

This paper has set forth a summary of a broader research effort undertaken to provide meaningful guidelines and parameters for use in inputs to or verification of the demand forecasting process. The paper and the manual from which it has been extracted provide an important data resource for contemporary urban planning efforts.

Plans call for distribution of the manual to various transportation agencies for their use and review. This will be followed by progressive updating of the manual's 250 tables and charts as new data become available. In this way, the manual will be able to respond to ongoing needs and priorities.

ACKNOWLEDGMENTS

I wish to acknowledge Sam Zimmerman and Thomas Hillegass of the Urban Mass Transportation Administration for their guidance and many constructive suggestions.

REFERENCES

1. F. H. Wynn and H. S. Levinson. Some Considerations in Appraising Bus Transit Potentials. HRB, Highway Research Record 197, 1967, pp. 1-24.
2. Transportation and Parking for Tomorrow's Cities. Wilbur Smith and Associates. New Haven, CT, 1966.
3. Urban Transportation Concepts—Center City Transportation Project. Wilbur Smith and Associates. New Haven, CT, 1970.
4. B. Pushkarev and J. Zupan. Public Transportation and Land Use Policy. Indiana Univ. Press, Bloomington, 1977.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting and Committee on Traveler Behavior and Values.

Minimizing Error in Aggregate Predictions From Disaggregate Models

Fred A. Reid, University of California, Berkeley

This paper presents empirical tests of aggregated prediction error on a sample of work-trip mode choices for the San Francisco area and systematic criteria for choosing classification variables. It introduces a more efficient utility scale classification criterion for aggregated prediction. Aggregation error is found to be much larger than previous tests have indicated. The choice of classification variables that produces the smallest error is found to vary with the scale of the prediction aggregates. Level-of-service variables are more important for large aggregates, socioeconomic variables for smaller. Classification of the sample based on the scales of the total utility of the explanatory variables in each alternative is found to be much more efficient in error reduction than classification by individual variables.

A number of papers have developed the theory of making aggregate predictions from disaggregate (individual) demand models (1, 2, 3, 4). Koppelman (1) integrated these theories into general guidelines for aggregate prediction and tested the relative accuracy of different proposed approaches by Monte Carlo simulation and with empirical data for one situation—the demand for different travel modes by Washington, D.C., commuters. He developed the classification method as the most accurate approximate method for aggregate prediction and encouraged classification based on "choice-set availability" variables such as the number of cars per driver in a traveler's household. This contrasted strikingly with the traditional approach to data aggregation—classification on the basis of geographic groups—an approach shown to produce biased results.

Dunbar (5) and Koppelman and Ben-Akiva (6) have further promoted prediction based on cross-classification by the values of the most influential variables in choice.

There are two important problems with these results. The classification developments, although clearly emerging in principle to define relatively homogeneous choice segments, leave practitioners with no systematic method of choosing the type and number of classification variables, especially considering the difference in the relative variance of the variables at different geographic levels of aggregation. Koppelman's tests show that the number of cars per driver is the most important classifier for small geographic aggregates. Dunbar's and my

results show that level-of-service variables are more important classifiers for large aggregate forecasting. The empirical variability of these results and the newness of the procedures to planners suggest that systematic methods are needed.

A second problem is that the Koppelman study—the only empirical results isolating the error of the different approximate methods for aggregation—is based on only one choice model and setting and then for an optimistic case. His data were for work-trip mode choice restricted to one destination, the Washington, D.C., central business district (CBD). The data also happened to have nearly equal choice shares over the full sample (the zero aggregation bias condition). Thus aggregation error, at least for regional mode choice, should be higher than shown by Koppelman.

This paper gives aggregation tests by different methods on San Francisco region commuter data, presents a more systematic criterion for identifying the disaggregate data (variables) most important for low error, and shows that a classification method based on the total utility scales of the choice alternatives, rather than on the values of the individual variables, is much more accurate and efficient for aggregation.

AGGREGATION ERROR TESTS

I have used the models and data from the urban travel demand forecasting project at the University of California, Berkeley, to perform tests isolating error by different methods in different situations (7, 8, 9).

The tests of this study were done on a sample of 771 workers drawn from about half of the San Francisco Bay Area. Their individual mode-choice probabilities for commuting to work were described by the four-alternative logit model

$$p_j = \left(\exp \sum_K \beta_{jk} z_{jk}\right) \Big/ \left(\sum_J \exp \sum_K \beta_{jk} z_{jk}\right) \qquad (1)$$

where β_{jk} are model coefficients of the K variables z_{jk} for the jth alternative. These are shown in the table below, where 1 is auto alone (429 people), 2 is bus with walking access (134 people), 3 is bus with auto access (30 people), and 4 is carpooling (178 people). Each independent variable takes the described value in the alternatives listed in parentheses and zero in unlisted alternatives.

Independent Variable	Estimated Coefficient	t-Statistic
Cost divided by post-tax wage, in cents divided by cents per minute (1-4)	-0.019 06	(-2.646)
Auto on-vehicle time, in minutes (1, 3, 4)	-0.059 86	(-5.24)
Transit on-vehicle time, in minutes (2, 3)	-0.024 17	(-2.751)
Walking time, in minutes (2, 3)	-0.068 92	(-5.301)
Transfer waiting time, in minutes (2, 3)	-0.055 99	(-2.361)
Number of transfers (2, 3)	-0.060 39	(-0.442 9)
Headway of first bus, in minutes (2, 3)	-0.029 51	(-2.997)
Family income with ceiling of $7500, in dollars per year (1)	0.000 006 238	(0.070 16)
Family income minus $7500 with floor of $0 and ceiling of $3000, in dollars per year (1)	-0.000 066 25	(-0.492 0)
Family income minus $10 500 with a floor of $0 and a ceiling of $5000, in dollars per year (1)	-0.000 029 11	(-0.475 0)
Number of persons in household who drive (1)	0.989 4	(4.667)
Number of persons in household who drive (3)	0.975 6	(3.272)
Number of persons in household who drive (4)	0.859 3	(4.223)
Dummy if person is head of household (1)	0.716 4	(3.736)
Employment density at work location (1)	-0.002 888	(-3.945)
Home location in or near CBD (2 is in CBD, 1 is near CBD, 0 is otherwise (1)	-0.454 6	(-3.694)
Autos per driver with a ceiling of one (1)	4.992	(9.600)
Autos per driver with a ceiling of one (3)	2.333	(2.769)
Autos per driver with a ceiling of one (4)	2.385	(5.315)
Auto alone alternative dummy (1)	-5.210	(-5.868)
Bus with auto access dummy (3)	-5.584	(-5.476)
Carpooling alternative dummy (4)	-3.795	(-6.333)

The goodness-of-fit statistics are 0.4484 for the likelihood ratio index, -1069.0 for the log likelihood at zero, and -589.6 for the log likelihood at convergence. All cost and time variables were calculated for round trips. The dependent variable is the alternative chosen (value of one for chosen alternative, zero otherwise).

The overall exact mode shares were 55.6 percent auto alone, 17.4 percent bus with walking access, 3.9 percent bus with auto access, and 23.1 percent carpooling. The data were from household surveys and transportation network minimum-path simulations, modified to produce trip attributes temporally and spatially disaggregated to the commuters' schedules and trip ends.

Exact aggregate choice shares are obtained by summing the probabilities Equation 1 over the individual z_{jk} values for the prediction aggregate of interest (the enumeration method). Other (approximate) methods are motivated by the large data and computation requirements of this enumeration in practice. The extreme approximation—the "naive" method—assumes that Equation 1 represents aggregate shares when the z_{jk} are simply the average values for the prediction group.

The measure of aggregation error in the tests here is the percent root mean square (RMS) of choice shares:

$$\epsilon_{RMS} = \left(\sum_J \{[(\hat{P}_j - P_j)/P_j]^2\} P_j\right)^{1/2} \qquad (2)$$

where

J = set of choice alternatives,
$\hat{P}_j$ = aggregate share of alternative j estimated by the tested method, and
P_j = aggregate share by enumeration.

The percentages of error from the naive aggregation method applied at three geographic levels on our sample are shown below. Predictions are all for the total region. The input data were the averages in the geographic aggregates shown. Thus, results represent either of two forecasting situations: geographic classification for full region predictions or the average absolute errors when making separate predictions for all the cells at a geographic (classification scale) level. These errors are approximately

Classification Scale	No. of Cells	Percent Error
Region	1	40.0
Cities	17	17.9
Traffic analysis zones	150	13.8

The errors are large. It is obvious that geographic classification alone is not an adequate aggregation method for this region and model. Errors decrease with geographic scale. Smaller area samples apparently do have less choice variability. Since classification-aggregation was done on the basis of residential district only, these results do not represent what would be expected from trip interchange forecasts. It is expected that average interdistrict aggregation errors would be about half those shown. Individual interdistrict aggregate errors would be worse.

Koppelman's results were lower in magnitude and did not vary with geographic scale. He showed an 8.5 percent error for a superdistrict scale similar to our cities, which show 17.9 percent. The regional errors were 10.4 and 40 percent, respectively. Several reasons account for this difference: he had a CBD-trip-oriented sample only; his average shares were nearly equal; the choice model he used was simpler; and the level-of-service data were less disaggregate (10).

Both tests aggregated (classified) the input data based only upon the origin of residence. Other sources of incomparability—the number of choice alternatives and the nonlinearity of the measure—are not large for these data. A much simpler model, equivalent to Koppelman's variable set, was found to produce 80 percent of the naive error on otherwise equivalent conditions for this data set (9, Chapter 6).

The error on our sample shown by five available methods of aggregation is

Method	No. of Cells	Percent Error
Naive	–	40.0
Taylor series	–	121.0
Classification by city	17	17.9
Classification by auto ownership	4	21.7
Classification by utility scale	4	3.1

Choice share prediction is again for the total (regional) sample. The naive method and the geographic classification method are the same as for the previous table. The Taylor series method is that of Talvitie (10). The by-variable value classification method is Koppelman's (1). The utility classification method is described later.

All of the methods except the Taylor series reduce error below that of the naive procedure. These results confirm earlier tests that the Taylor series approximation produces counterproductive results due to its poor convergence properties on large variance data. Unfortunately, it is in such data that the correction is important. Geographic classification, as shown above, is very inefficient in reducing aggregation error.

The predictions by classes of auto ownership also only reduce about half the naive error, giving unacceptable accuracy at the regional scale of forecasting. Koppelman showed auto availability classification to reduce a smaller regional naive error by two-thirds, with a result of 3 percent. Granting that the variable cars per driver will reduce two-thirds of the naive error, this classifier would still leave an unacceptable error in general regional aggregate travel predictions (for trips to all destinations and with unequal shares). Choice set availability classifiers used on predictions for subcity or city interchange level aggregates may yield more acceptable errors (below a third of 18 percent).

The method of utility classification gives a much greater reduction relative to naive error than the other methods. Only four class cells were used. This method is discussed below.

UTILITY SCALE CLASSIFICATION

Although classification by the values of one or two variables gives unacceptable error for large aggregate predictions, the error can be made small by cross-classification of more of the variables in a model. However, if more than a few variables contribute significantly to the variance of the utility of the choices, the number of cross-classifications must be large to achieve small error.

A more efficient method of classification is possible for the category of simply scalable models such as those of the logit form. Cross-classification between individual variables includes much information that does not matter to simple scales. The essential information needed to predict each individual's choice in these models is contained in the J utility scales of the attributes of the J alternatives of choice. Cross-classification between the utilities of the different alternatives picks up the full-scale variances and between-scale covariances, thus describing the full distribution of individual choice factors in an aggregate prediction sample. Regardless of scale complexity, this procedure bypasses those individual variable cross-classification trade-offs, which do not change the scale values.

Thus, the procedure requires fewer classes. Classification on the total utility includes the variances of the minor variables, not just the variance of the limited number of interactions feasible in classification by a subset of model variables. This further increases its efficiency.

To define relatively homogeneous classes of utility combinations across modes is to probe the essence of the classification approach: the grouping of individuals with uniform choice situations. Since the procedure operates on the utility scales, it is termed the utility scale classification method of aggregation.

Utility class sizes and boundaries cannot be defined in the same way as can individual variable classifications. Utilities are not discrete, and intuition does not give the guide on thresholds of utility in choice that it does on a variable such as walking time to a bus in mode choice. However, utility values are clearly related to choice function thresholds. For the binary logit case, the optimum divisions would differentiate utilities near the maximum nonlinearity of the choice function (±1.6 on the utility difference scale). For multiple-choice logit model classification, the criteria should concentrate on pairs of alternatives with these same differences of utilities.

Cluster analysis techniques could be employed to achieve general isolation of classes (11, 12). The tests here and in supporting research show that ad hoc class divisions are adequate and more appropriate, considering the above nonlinearities and the press of obtaining predictions (9).

One such procedure is the successive division of an aggregate sample about the median (or mean) values of the differences of the utilities of pairs of alternatives, starting with the pair with the largest variance in utility difference. The procedure cycles through all pairs of utilities, further subdividing the first pairs if the variance of utility differences in the resulting classes are still large. The variance criteria for desired accuracy can be estimated with the covariance analysis procedure discussed in the next section.

The utility ranges gave a much smaller error than any of the individual variable or geographic classifications considered in the tests in the tables with only two utility classes on a four-alternative model. Clearly this method is more efficient in error reduction than the others. This is to be expected at the regional level, where the total utility variance of all of the variables is present. It would also be true for predictions for small aggregates unless only a minor subset of model variables dominated the variability of these subsamples. Such is not the case for within-city aggregates, as will be seen in the following percentages of aggregation error by cell count.

Class Cell Count	Percent Error
1 (naive method)	38.4
2	6.4
4	2.3
6	2.6
8	0.5

This shows the variation in aggregation error with the number of class cells used in the utility classification method. These results are shown for a three-alternative subset of the four-mode choice model used in the previous tests (auto alone, bus with walking access, and carpooling). The cell definition criteria were those above. The errors seem acceptable with only four cells. With eight cells error becomes negligible. The instability in error reduction with cell count is due partly to the ad hoc cell definition criteria used but also to the nonlinearity of the RMS measure. An error measure composed of a linear sum of absolute errors in each alternative decreased monotonically with cell count (9).

Since utility scale classification does not differentiate the joint distributions of all variables in a model, it does not apply to general models that are not simply scalable. Probit models and others that allow estimation of different coefficients for the same variable in different alternatives are examples (13, 14). Aggregate predictions with these, in general, require individual variable cross-classification or covariance procedures (see the paper by Bouthelier and Daganzo in this Record). However, even non-simply-scalable models are not likely to estimate cross-alternative coefficients for many variables.

In view of the laborious nature of cross-classification and the efficiency of utility classes on the simple scales, the latter may be the most effective even in this case. Alternatively, it may be efficient to cross-classify utility classes with an important cross-alternative variable.

The focus of this aggregation method on the total utility value need not preclude the retention and association of the values of specific explanatory variables with the classes. This makes it possible to intuitively interpret the cells, analyze the effects of policy changes on choices, and predict subsegment choices. These can all be accommodated by characterizing each class cell by the average value of any desired model variable or socioeconomic descriptor, in addition to the average value of the J total utilities of its members. Together with the values of the underlying model scale coefficients, the mean utilities may be adjusted for policy changes on specific variables to simply analyze their effects on aggregate choices. Population subgroup shares may be predicted by summing the products of the proportions of these groups and the shares in each cell. A complete description and examples of the use of utility classification method in subsegmented regional travel prediction and policy analysis are given in a related report (15).

This, like any accurate classification method, requires that individual choice-maker data observations be available to compute the cell mean values. Hence, the principal advantage of the method is in the efficiency of the predictive calculations, given the data. This is no small advantage, since predictions can require exponentiation over many choices for each individual in large samples or they may be desired for numerous policy (input) changes. Once the disaggregate data information is reduced to cell means and parameters, predictions for different policies or output segments or both are simple enough to make hand calculation feasible (15).

It is tempting to believe that by-variable-value class means are simply available from aggregated planning or census data. However, data cannot be directly collected in autos per driver classes or other cross-classes necessary for accurate aggregation. These require individual observations. This method may also indirectly aid the data-collection process. By identifying the minimum amount of information necessary for prediction, it focuses on these requirements. What is necessary for logit models are only the utility scale variances, not the unique within-alternative distributions of variable values. It may be possible to directly extract the utility scale distribution information in a sample directly from census cross-tabulations (9, 16) or from minor subsets of the model variables. The latter approach is discussed in the next section.

COVARIANCE MATRIX ANALYSIS OF THE EXPLANATORY VARIABLES

Whatever the method of aggregation, the requirement for disaggregate data against a past experience of aggregate data collection suggests that data collection may have to focus on the major sources of error. Thus, it is necessary to know which variables in a model contribute the greatest part of the covariances of the utilities within any prediction aggregate. Precedent and intuition have shown the importance of some variables such as socioeconomic descriptors and transit access variables in mode choice. The problem is that the important variables for classification vary with the geographic scale of the aggregate predictions and, to the extent that there is no universal model of behavior, with the model being used. The limited precedents available may be for the wrong cases. It is unreasonable to expect expert judgment or research-level optimization of aggregation accuracy in practical planning situations.

For effective use of cross-classification methods, where necessary, one must have knowledge of the variables contributing the most to utility covariances at the level of aggregation.

A systematic picture of priorities of individual variable data can be gained by looking at the covariance matrixes of the within-aggregate variation of the explanatory variables at the desired level of aggregate. Binary choice aggregation theory has established the direct relationship between this matrix and aggregate shares for probit models (7). I have shown that this simplified analytic form of the aggregation method for probit models and normally distributed explanatory models can be extended from probit to an approximate relation for logit models with an error less than 1.2 percent of total shares (17).

$$S_j/N = 1 + \exp\left((\overline{V_i - V_j})/\{[1 + \mathrm{Var}(V_i - V_j)/2.79]^{1/2}\}\right) \quad (3)$$

where

S_j/N = choice share of the alternative j,
V_j and V_i = utilities of the two choices, and
$\overline{V_i - V_j}$ and $\mathrm{Var}(V_i - V_j)$ = mean and variance of their difference in the aggregate of N travelers.

The assumption of normality need apply only to the scale utility difference, not to each explanatory variable. Tests have shown this assumption to contribute 4.0 percent error for regional aggregate predictions that had 43 percent error by the naive method, suggesting that normality is a reasonable assumption (9). The method is still limited to binary choice.

Although only a more complicated relationship of this type exists in the multiple choice case, the covariance matrixes are still better indicators of which variables are appropriate classifiers than a priori knowledge, especially if these vary with aggregate level.

Table 1 shows a normalized covariance matrix of the intracity aggregate values of the major explanatory variables in two alternatives of a prediction model. The model used was shown at the beginning of the paper. The covariance elements are normalized by dividing the largest of their values (in this case it is the variance of the

Table 1. Normalized covariance matrix of explanatory variables for intracity portions of individual observations.

Variable[a]	Cars/ Driver	Auto On-Vehicle Time	No. of Drivers	Employment Density	Cost-Auto	Bus With Walking Access	Bus On-Vehicle Time	Headway	Transfer Waiting Time
Cars/driver	1.0	0.14	<[b]	<	<	<	<	<	
Auto on-vehicle time	0.14	0.59	<	0.10	0.10	-0.09	0.27	<	0.06
No. of drivers	<	<	0.25	<		<	<		
Employment density	<	0.10	<	0.11		-0.05			
Cost-auto	<	0.10			0.04	<			
Bus with walking access	<	-0.09	<	-0.05		0.98	0.38	0.05	
Bus on-vehicle time	<	0.27	<	<	<	0.38	0.25	<	0.07
Headway	<	<	<			0.05	<	0.07	<
Transfer waiting time	<	0.06				<	0.07	<	0.15

[a]Covariance values are normalized by multiplying the variances of the variables by their respective model coefficients and dividing by the maximum such utility component ($\beta_0 S_{00}^2 \beta_0$) in this case the utility variance of cars per driver having the value 1.92. Individual observations are minus the mean variable values for all travelers in sample with work trips between the same cities of origin and destination.

[b]< indicates the element is less than 0.04 (0.08 unnormalized); blank indicates the value was not computed but judged to be less than 0.04.

utility of cars/drivers in the traveler's household) in order to give a picture of the ranking of the contributions of each variable to the overall utility variance.

The underlying model is a logit function, linear in its explanatory variables. Hence, each element in the matrix is the product of the covariances of the individual variables and their corresponding model coefficient is

$$\sum_{kl}^{2} = \beta_k S_{kl}^2 \beta_l / \beta_0 S_{00}^2 \beta_0 \qquad (4)$$

where

K and l = matrix indexes of the K variables in the model ($l \le k$, $l \le K$),

$\beta_k \beta_l$ = linear utility coefficients for the kth and lth variables, and

$\beta_0 S_{00}^2 \beta_0$ = corresponding utility variance for the largest matrix element.

These elements are called "utility component covariances," since each expresses the part of the total sample variance of the linear utility for one mode contributed by a single variable (diagonal variance elements) or pairs of variables (covariance elements). The sum of the matrix elements for any mode equals the total variance of the utility for that mode.

For binary logit choice models the relationship between these covariances and aggregation error is understood under the assumption of a normal distribution of the utility scale of the explanatory variables (Equation 3). Thus Table 1 provides a convenient way to rank the individual variables, and their combinations, by their importance for reducing aggregation error for binary choice.

The larger the variance of the difference of utilities between the two modes, the larger the aggregation error.

Since most of the variables in this model are unique to a mode, variances of intermodal utility differences nearly equal the individual modal elements. Where the same variables appear in the utility expression for two modes, the differences of utilities can be obtained from the table by this expression for the variance of such a difference: Var (X - Y) = Var X + Var Y - 2 Cov (X, Y).

The variables that have the individual values most important for reducing aggregation are the ones with the largest variance values in Table 1, such as bus with walking access, cars/driver, and auto on-vehicle time. The intermodal covariances show the relative importance of pairs of variables in aggregation. However, since intermodal elements contribute to the variance of binary utility difference with negative sign, it is the negative values such as that between cars/driver and bus with walking access that increase the aggregation correction. Positive values indicate a compensating reduction in aggregation error.

The amount of aggregation error that will result in a forecast using the individual values of only selected variables can be estimated by comparing Equation 3 with and without the unnormalized values of the sum of the matrix elements for the variables considered. [Only the individual values of some variables or their variance may be available in a forecasting situation. In some cases, variances are possible to obtain or estimate where covariances are not (10).]

Inspection of Table 1 shows that a large part of the variance of the utilities, and hence aggregation correction, can be recovered by considering only a few of the variables in the model. For example, three of the eleven variables in the model used here—cars/driver, bus with walking access, and bus on-vehicle time—constitute 64 percent of the utility variance. When only their individual values are used for binary prediction, they correct for 80 percent of the aggregation error produced by the naive method. Auto on-vehicle time, in contrast, contributes nothing to the correction of aggregation error by its inclusion with this set, since its covariance with bus with walking access cancels out its variance contribution alone.

The covariance matrix also contains the information for guiding the classification method in its implied correlations between the classifier variable(s) and the others. The additional power of a classifier from this effect is given by the sum of the product of covariances and correlation coefficients of all variables correlated to the classifier. With this procedure, one can minimize aggregation error, given limits on the number of classification variables, or minimize the number of classifiers, given tolerable limits on error. This is systematically outlined elsewhere (11).

This analysis is only simply applicable to binary choice. For multiple choice it requires a relationship analagous to Equation 3 between covariances and choice. The only example of this is the approximate formulas of Bouthelier and Daganzo for probit models.

In principle, it is possible to use any of the above effective subsets of disaggregate variable data for aggregation accuracy using these formulas and covariances between all pairs of alternatives. In practice, or if the model in question is not multinomial probit, simpler approximate guides to the important aggregation variables can be gained by observing the covariance matrixes of major pairs of the alternatives. These should be alternatives that have aggregates different from 1/J and considerably different from one another.

The procedure also does not account for skewed utility distributions. Skewedness will weaken the analysis. This does not appear to be a problem in the tests on this data set (9).

Covariance analysis on a small population sample for

which policy-forecasting data are received can guide the subsequent cost-effective emphasis on which variables to focus collection of individual observations in the larger forecasting sample.

CONCLUSIONS

On the basis of the above exposition, I have drawn the following seven conclusions.

1. The tests show that aggregation error by all methods is larger than previously revealed. Naive method error can be larger than other errors in prediction, such as model specification or data measurement.
2. The relative ranking of the accuracy of existing approximate aggregation methods by Koppelman is confirmed. However, error is found to grow substantially with aggregate size. Previous studies have underestimated this and general error level because their samples had limited trip types, choice share ranges, and data disaggregation.
3. None of the existing simple approximation methods gives even marginally acceptable (≤13 percent) error at the urban regional scale of mode-choice aggregation using multinomial logit models.
4. The method of classifying on the differences of utility scales of major choice alternatives gives at least a five times lower error to class cell-count ratio than other classification criteria. It can reduce naive error by a factor of 77 (to 0.5 percent) with only eight cells.
5. The choice homogeneity of utility classification cells eliminates the need to collect data and make separate aggregations for each sector of a population for which share predictions are desired. It replaces such repetitive choice function evaluation with a smaller number of class segment proportion calculations.
6. Analysis of the covariance matrix of the utility components of the individual model variables answers the question of which variables should be stressed in disaggregate data collection. Doing this on a small sample for the prediction population of interest can guide the subsequent data collection to the subset of variables that gives the least accuracy return. This analysis can also give systematic guidance to the selection of by-variable classification methods for predictions with non-simply scalable models.
7. Predictions, even for large aggregates, can be made accurately with disaggregate data on only a minor subset of critical variance variables. Aggregation can be accomplished with the minimum of choice function computations using utility classification on the full data sample rather than by-variable classification on every prediction segment.

Classification methods of prediction, especially those based on selecting homogeneous utility or choice groups, realize the full potential for efficiency of disaggregate models that was first seen when it was found that they could be calibrated on a fraction of the observations necessary for aggregate models. Now prediction can also be done on samples large enough to identify only a limited number of utility classes, that is, homogeneous choice markets.

Predictions for subaggregates require larger samples but are simply a process of determining the relative numbers of the population in each of these classes. The prediction problem has been reduced from one of computing choice shares for great numbers of aggregate outputs to locating the proportions of these aggregate cases in a small number of behavioral market segments. Selected disaggregate data are necessary to gain the accuracy of behavioral models, but predictive computations are greatly reduced. Reid (9) and Reid and Harvey (15) discuss these aggregation methods and their application to policy analysis in greater depth.

ACKNOWLEDGMENT

This research was supported in part by a National Science Foundation grant through the Research Applied to National Needs Program to the University of California, Berkeley. I am indebted to Daniel McFadden and Ibrahim Hasan for discussion and comment on these subjects. Mr. Hasan, Chris Murano, and Rafi Melnick performed valuable supporting computations. I am solely responsible for the results.

REFERENCES

1. F. S. Koppelman. Travel Predictions With Models of Individual Choice Behavior. Center for Transportation Studies, Massachusetts Institute of Technology, Cambridge, Rept. 75-7, 1975.
2. D. McFadden and F. A. Reid. Aggregate Travel Demand Forecasting From Disaggregated Behavioral Models. TRB, Transportation Research Record 534, 1975, pp. 24-37.
3. A. Talvitie. Aggregate Travel Demand Analysis With Disaggregate or Aggregate Travel Demand Models. Proc., Transportation Research Forum, 1973.
4. R. B. Westin. Predictions From Binary Choice Models. Journal of Econometrics, 1974, pp. 1-16.
5. F. C. Dunbar. Policy-Contingent Travel Forecasting With Market Segmentation. TRB, Transportation Research Record 637, 1977, pp. 27-32.
6. F. S. Koppelman and M. E. Ben-Akiva. Aggregate Forecasting With Disaggregate Travel Demand Models Using Normally Available Data. Paper presented at the World Conference on Transport Research, Rotterdam, April 26-28, 1977.
7. D. McFadden and others. Demand Model Estimation and Validation. Urban Travel Demand Forecasting Project, Institute of Transportation Studies, Univ. of California, Berkeley, Final Rept. Series, Vol. 5, June 1977.
8. F. A. Reid. Disaggregated Supply Data and Computation Procedures. Urban Travel Demand Forecasting Project, Institute of Transportation Studies, Univ. of California, Berkeley, Final Rept. Series, Vol. 3, 1977.
9. F. A. Reid. Aggregation Methods and Tests. Urban Travel Demand Forecasting Project, Institute of Transportation Studies, Univ. of California, Berkeley, Final Rept. Series, Vol. 7, 1978.
10. A. Talvitie and Y. Dehghani. Supply Model for Transit Access and Line Haul. Urban Travel Demand Forecasting Project, Institute of Transportation Studies, Univ. of California, Berkeley, Working Paper No. 7614, July 1976.
11. P. E. Green and Y. Wind. Multi-Attribute Decision Marketing: A Measurement Approach. Holt, Rinehart and Winston, New York, 1973.
12. H. P. Friedman and J. Rubin. On Some Invariant Criteria for Grouping Data. American Statistical Association Journal, Dec. 1967.
13. J. A. Hausman and D. A. Wise. A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences. Department of Economics, Massachusetts Institute of Technology, Cambridge, Working Paper No. 173, 1976.
14. D. McFadden. A Closed Form of Multinomial Choice Model Without the Independence from Ir-

relevant Alternatives Restrictions. Urban Travel Demand Forecasting Project, Institute of Transportation Studies, Univ. of California, Berkeley, Working Paper No. 7703, 1977.
15. F. A. Reid and G. Harvey. Regional Policy Analysis Case Study. Urban Travel Demand Forecasting Project, Institute of Transportation Studies, Univ. of California, Berkeley, Final Rept. Series, Vol. 9, 1978.
16. D. McFadden and others. Demographic Data for Policy Analysis. Urban Travel Demand Forecasting Project, Institute of Transportation Studies, Univ. of California, Berkeley, Final Rept. Series, Vol. 8, June 1977.
17. F. A. Reid. A Set of Models for Optimizing the Benefits of a Transportation Plan. Univ. of California, Berkeley, M.S. thesis, 1974.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting and Committee on Traveler Behavior and Values.

Disaggregate Demand Model for Nonwork Travel

Joel Horowitz, U.S. Environmental Protection Agency

Daily nonwork travel by urban households frequently involves visits to several destinations during a single roundtrip from home or several round trips. This paper describes a disaggregate approach to modeling the demand for nonwork travel, including multidestination travel. The approach is presented in two parts. First, a theoretical framework for modeling the demand for nonwork travel is developed that uses tours (i.e., round trips from home) and sojourns (i.e., visits to nonwork destinations) as the basic units of travel, and relates tour frequency, sojourn frequency, and destination choice to household characteristics, destination characteristics, and transportation level of service. An empirical model of nonwork travel demand that is based upon the theoretical framework is then presented. The empirical model enables tour and sojourn frequencies, destination choice, and nonwork vehicle-kilometers traveled to be computed as functions of household characteristics, destination characteristics, and transportation level of service. Several tests of the validity of the empirical model are described. The model was found to perform well.

One of the problems that must be addressed in modeling the demand for urban nonwork travel is that of incorporating multidestination travel into the modeling framework. Current operational models of nonwork travel demand use one of two methods to treat multidestination travel. In the first, only two nonwork travel options are explicitly modeled: the option of doing no nonwork travel during a day and the option of taking a single round trip between home and nonwork destination during a day. Multidestination travel is ignored or its effects are approximated through relatively crude factoring procedures. This method is used frequently in connection with disaggregate models of nonwork travel demand (1, 2, 3).

The other method of treating multidestination travel uses individual trip links as the basic units of travel and models the demand for these links. Effects of multidestination travel are represented by changes in the demand for home-based and non-home-based trip links. This method is used in conventional aggregate demand models (4) and in at least one disaggregate demand model (see the paper by Ben-Akiva, Sherman, and Kullman in this Record). It also is used in Markovian activity sequencing models (5, 6, 7). The method implicitly assumes that travelers' decisions are based only on the characteristics of individual trip links without regard to the characteristics of groups of links or travel patterns. Distortions that this assumption causes in travel demand models are described elsewhere (8, 9).

Neither of the foregoing methods provides a satisfactory treatment of multidestination travel: the first virtually ignores multidestination travel, and the second ignores relationships between the trip links that make up multidestination travel patterns.

Two other approaches to treating multidestination travel have been suggested as means of overcoming these difficulties. In one of these approaches complete travel patterns of households during 24-h periods are used as the basic units of travel, and the demand for individual travel patterns is modeled (8). This approach offers a theoretical framework in which travel decisions depend on relationships between trip links as well as on characteristics of individual links. However, it provides no means of distinguishing between the set of travel patterns actually considered by households in the process of making travel decisions and the virtually infinite set of travel patterns that are, in principle, available for consideration.

The second approach uses the tour and the sojourn as the basic units of travel (10). A sojourn is a visit to a place other than home or work. A tour is a movement that begins and ends at home or work and includes one or more sojourns. The approach consists of modeling the demand for tours and sojourns. The detailed sequence of sojourns within tours is not modeled, thus avoiding the need to consider a potentially infinite variety of tour structures. On the other hand, because the tour is a multidestination unit of travel, it is not necessary to assume that travel decisions are made without regard to collective characteristics of trip links.

This paper describes a disaggregate model of the demand for nonwork tours and sojourns. The model is presented in two parts. First, using the household as the basic decision-making unit, a theoretical framework is developed that relates tour frequency, sojourn frequency, and destination choice to household characteristics, destination characteristics, and transportation level of service. An empirical model of nonwork travel demand that is based upon the theoretical framework is then presented. The empirical model, which was estimated using travel data from the Washington, D.C., area, enables tour and sojourn frequencies, destination choice, and nonwork vehicle-kilometers traveled to be computed as functions of household characteristics, destination characteristics, and transportation level of service.