Quick Policy Evaluation With Behavioral Demand Models

Frederick C. Dunbar, Charles River Associates, Cambridge, Massachusetts

Many of the important policy issues confronting present day urban planners involve regionwide transportation system changes that will have many effects. Conventional urban transportation planning models do not capture the full range of travel impacts, and are cumbersome and resource consuming for evaluation of these policy options. In response to this, new behavioral travel demand models have been developed; these are policy sensitive and can be generalized among urban areas. However, there are several unresolved questions about these disaggregate demand models that prevent their widespread application. These problems are:

1. Models estimated in one urban area have not been validated on other urban areas to test their generality.
2. Models estimated on small data sets have not been applied to other small data sets to predict regionwide travel behavior.
3. Disaggregate logit models give biased forecasts when applied to sketch plan or district sized zones.
4. Disaggregate demand models estimated on automobile drive alone and transit mode choices will not predict the full range of choices available to trip makers, which may include car pooling, chauffeuring, and walking, in response to a change in system performance.

This paper presents methods that apply disaggregate, probability choice demand models to a sample of sketch plan zones to evaluate various automotive pollution control strategies in the Los Angeles region. Each of the above problems is considered.

THE AGGREGATION PROBLEM

In the logit specification of probability choice models, the probability of an individual choosing any given mode has the following functional form:

$$P(A) = 1 \left[1 + \sum_{\substack{b=1 \\ b \neq a}}^{n} \exp(-Y_{ab}) \right] \qquad (1)$$

where a is one among n alternatives and Y_{ab} is the relative costs and attributes between alternatives a and b. Each Y_{ab} represents the log of the ratio of the probability of a to the probability of b. The prediction of travel behavior in a zone of T individuals requires estimates of individual probability choice of a:

$$N_a = \sum_{t=1}^{T} P_t(a) \qquad (2)$$

Typically, the only information available about the arguments of the Ys is their means for a zonal interchange and, possibly, the variances and covariances of the terms in Y. There is no analytical form to translate this information into an estimate of N_a.

A Taylor's series approximation of equation 1 evaluated about the zonal means of the data has been suggested to adjust for this problem (1). The expected value of the resulting expression, truncated after the third term, gives the following equation:

$$E[P(a|Y)] = P(a|\bar{Y}) \left\{ 1 + \sum_{\substack{b=1 \\ b \neq a}}^{n} \mathrm{var}[Y_{ab}] * P(b|\bar{Y}) - ½ \right\} \qquad (3)$$

where

E [ ] = expected value operator and
var [ ] = variance.

The variables with bars over them are means. Equation 3 is somewhat different from those derived by Talvitie. The operational difference is that stochastic independence between the attributes of alternative a and other alternatives is not assumed here but is by Talvitie (1).

The expected value of choices other than $\bar{a}$ is

$$E[P(c|Y)] = P(c|\bar{Y})\left\{1 + \sum_{\substack{b=1\\b\neq a}}^{n} \mathrm{var}[Y_{ab}]\,[P(b|\bar{Y}) - \delta]\,[P(b|\bar{Y})] - ½\right\} \qquad (4)$$

where $\delta = 1$ if $b = c$ and 0 if $b \neq c$.

Consider the comparison function, Y_{ab}. For the mode choice, the functional form for Y (2) is

$$Y_{ab} = \alpha_0 + \alpha_c(C_a - C_b) + \alpha_T(T_a - T_b) + \alpha_s(S_a - S_b) + \beta 0 \qquad (5)$$

where

- C = operating cost of the trip,
- T = waiting and line-haul time of the trip,
- S = walking time for the trip,
- 0 = availability of an automobile, and
- a, β = estimated constants.

There are 28 possible variance-covariance terms for this equation. About half of these can be presumed to be zero because of stochastic independence or constancy over a zone. Of the other half, there is a presumption that most are proportional to, or simple functions of, the variance of the distance traveled in a zone interchange.

With the information available, one cannot be very precise in measuring the variance in distance in zonal interchanges. The approach here is to assume that distances (or origin and destination points) are distributed over the area of a zone pair according to a well-defined probability density function. This approach, ultimately, allows estimates of the variance of distance as a function of the areas of the two zones in a zonal interchange.

In deriving the appropriate density functions, it is presumed that, for a given zone pair, trips are distributed over a range that reflects both the distance between the zone centroids (geographic centers) and the sizes of the zones. In symbols this is

$$D \in [D' - (Y_i + Y_j)/2,\ D' + (Y_i + Y_j)/2] \qquad (6)$$

where

- D = a stochastic variable that represents the distance between zones i and j for person trips,
- D' = the distance between the geographic centers of zones i and j, and
- Y_i, Y_j = some measure of the size of zones i and j.

If another stochastic variable, X, which can have values in the range from 0 to $(Y_i + Y_j)$, is now introduced, the distance for any trip can be represented by the sum of two variables:

$$D = \{D' - [(Y_i + Y_j)/2]\} + X \qquad (7)$$

The above relation indicates that trips must travel a nonstochastic minimum distance (the term in [ ]) and that the rest of the distance varies randomly between 0 and $Y_i + Y_j$. The variance of D is

$$\mathrm{var}[D] = E[X^2] - (E[X])^2 \qquad (8)$$

The distribution function for X is assumed to be

$$f(X) = \{3/[2(Y_i + Y_j)]\} - [X/(Y_i + Y_j)^2] \quad \text{for } 0 < X < Y_i + Y_j \qquad (9)$$

The premise of the density function is that the distribution of trips can be approximated by a linear declining function over the range bounded by $Y_i + Y_j$.

The first two moments about zero of the distribution are

$$E(X) = [5(Y_i + Y_j)]/12 \qquad (10)$$

$$E(X^2) = (Y_i + Y_j)^2/4 \qquad (11)$$

The variance of distance can then be calculated from equation 9 and the above moments:

$$\mathrm{var}[D] = [11(Y_i + Y_j)]/144 \qquad (12)$$

The measures of zone size over which trips are distributed should reflect the length of the zone. Use of an intuitive measure of length, the square root of area, leads to the following:

$$\mathrm{var}[D] = [11(\sqrt{A_i} - \sqrt{A_j})^2]/144 = [11(A_i + 2\sqrt{A_i A_j} + A_j)]/144 \quad \text{for } i \neq j \qquad (13)$$

For intrazonal trips, the above equation must be modified to account for the fact that the stochastic part of the range is only half, on an average, of that of interzonal trips:

$$\mathrm{var}[D] = 11A_j/144 \quad \text{for intrazonal trips} \qquad (14)$$

The remaining terms in the Taylor expansion tend toward zero. However, the truncation of the series after the third term opens the possibility that, for values of Y that diverge rather far from $\bar{Y}$, equations 2 and 3 will not provide a measure of probability that increases monotonically with $P(a|Y)$. In symbols, $E[P(a|y)]$ must satisfy the following three conditions:

$$\sum_{a=1}^{n} E[P(a|Y)] = 1 \qquad (15)$$

$$0 < E[P(a|Y)] \quad \text{for all a} \qquad (16)$$

$$\partial E[P(a|Y)]/[\partial P(a|\bar{Y})] > 0 \quad \text{for all } P(a|\bar{Y}) \epsilon [0, 1] \qquad (17)$$

Conditions 15 through 17 ensure that $E[P(a|Y)]$ is a probability measure. Stronger conditions are required if $E[P(a|Y)]$ is to have plausible properties in terms of individual choice behavior. One of these is that the elasticity be greatest at $E[P(a|Y)] = 0.5$, i.e.,

$$\partial^2 E[P(a|Y)]/\partial Y^2 = 0 \quad \text{at } E[P(a|Y)] = 0.5 \qquad (18)$$

$$\partial^3 E[P(a|Y)]/\partial Y^3 < 0 \quad \text{at } E[P(a|Y)] = 0.5 \qquad (19)$$

Of the above five conditions, 15 and 18 hold for all values of the variances. The following constraints on the variances are sufficient to ensure that the other conditions are met:

$$\sum_{\substack{b=1\\b\neq a}}^{n} \mathrm{var}[Y_{ab}] < 16 \qquad (20)$$

$$\mathrm{var}[Y_{ab}] < 1 \quad \text{for all } b \neq a \qquad (21)$$

The variance and covariance of terms in equations 3 and 4 are computed as proportions of 13 and 14 subject to the above constraints.

Table 1. Estimated versus actual work trips.

Mode	Mode Shares for 172 Zonal Interchanges		Shares for LARTS Region (%)	
	Actual	Estimated	Actual	Estimated*
Automobile driver	1040	960	84	79
Transit	47	54	1	4
Automobile passenger	123	196	12	16
Driver serve-passenger	–	40	–	–
Walk	–	10	–	–

*Excludes driver serve-passenger and walk

THE NEW MODE PROBLEM

If mode choices are constrainted to be automobile drive alone and bus transit, the model will not predict the range of responses caused by a policy that significantly alters system performance. The consideration of new modes requires a heuristic approach that results in the construction of new comparison functions, Y_{ab}.

The new comparison functions are formed by attributing to each new mode a variable cost per mile, a time spent in-vehicle and waiting, and a walking time for person round trips between each i, j zone pair. Each of these trip system performance variables can then be substituted for their transit counterparts in the estimated mode split equations to derive the odds between automobile choice and the given new mode choice, and from equation 1, the probability of choosing any alternative among all modes—automobile alone, transit, car pool, serve passenger, and walk—can be derived.

The values for level of service for the new modes are largely the results of assumptions about extra time penalties involved with car pooling and chauffeuring. Such assumptions are required because of the paucity of data about these alternatives. In particular, for individuals who currently drive alone, virtually nothing is known about the availability and attributes of potential car pools.

TRAVEL DEMAND MODEL WITH LOS ANGELES REGIONAL TRANSPORTATION STUDY DATA

This section compares actual Los Angeles Regional Transportation Study (LARTS) data for 1967 against predictions from the demand model. The data given are the number of person round trips between zone pairs by travelers surveyed in the 1967 household survey (a 1 in 100 sample). The level of aggregation is sketch plan zones defined by LARTS in 1970; there are about 12 traffic analysis zones to each sketch plan zone and 69 sketch plan zones for the analysis area (Los Angeles and Orange Counties—the Los Angeles Air Quality Control Region).

The tests described below attempt to determine whether (a) application of the disaggregate demand model estimated on Pittsburgh data and adjusted for zonal variations and new modes can be generalized to Los Angeles, and (b) a small but representative sample of zonal interchanges can be used to predict regionwide effects.

The approach to applying the model is summarized in the following steps:

1. Odds functions for automobile versus other modes are estimated for each zonal interchange using the zonal averages for system performance;
2. Probabilities of each mode choice for each zonal interchange are calculated from application of equation 1;
3. The mode shares for each zonal interchange are calculated using the probabilities, the calculated variance-covariance terms from the formulas in the previous section, and equations 3 or 4, to adjust for aggregation; and
4. The estimated mode shares are multiplied by total trips in the zonal interchange to derive predicted trips by mode.

A random sample of 172 zonal interchanges was chosen for testing and applying the approach. Because the policies in the study were evaluated by their effects on vehicle-miles traveled (VMT) the model also was tested by placing the most emphasis on actual versus predicted VMT.

Table 1 compares the mode shares predicted from the sample to the work-trip mode shares for the entire LARTS region. The model gives reasonable predictions of mode split. Although the total vehicle trips were underpredicted by 7.69 percent, the VMT (estimated, 15 302; actual, 15 211) were predicted with virtually no error.

CONCLUSION

The results presented in this paper indicate that disaggregate demand models hold promise for quick evaluation of transportation-related policies. Although actual policies are not discussed (3, 4), the models were used to simulate the effects of various pollution control strategies by projecting 1974 base case trip behavior and the changes that would have been caused by gasoline taxes, emissions taxes, parking surcharges, and bus system improvements. The resulting predictions were comparable with other research efforts; for example, the implied elasticity of gasoline in Los Angeles in 1974 was between 0.19 and 0.24, which corresponds with many econometric estimates of short-run gasoline demand.

ACKNOWLEDGMENTS

The research reported in this paper was part of a project performed for the Washington Environmental Research Center, Environmental Protection Agency on Policies for Controlling Automotive Pollution in Los Angeles. The author wishes to thank Eleanor Maxwell for performing the calculations in this paper.

REFERENCES

1. A. Talvitie. Aggregate Travel Demand Analysis With Disaggregate Demand Models. Proc., Transportation Research Forum, Vol. XIV, No. 1, Oct. 1973, pp. 583-603.
2. A Disaggregated Behavior Model of Urban Travel Demand. Charles River Associates, 1972.
3. F. C. Dunbar. Evaluation of the Effectiveness of Pollution Control Strategies on Travel: An Application of Disaggregated Demand Models. Proc., Transportation Research Forum, Vol. XVI, No. 1, 1975.
4. F. C. Dunbar. Policies for Controlling Automotive Air Pollution in Los Angeles. Charles River Associates, in press.