Estimating a Road-User Cost Function from Diversion Curve Data

HOWARD W. BEVIS, Alan M. Voorhees and Associates

A theory and methodology are presented for estimating a road-user cost function using diversion curve data. The theory of diversion curves is presented as being a probability density function with normal distribution. The point in the density function associated with a given proportion of freeway usage is related to a generalized cost difference equation describing the comparative costs of using the freeway as opposed to the best alternate. The parameters of the cost function are then derived using regression techniques.

The road-user cost equation derived from this analysis agrees quite well with theoretical predictions and with results obtained from other studies where relationships between the various components of road-user cost and speed of travel have been derived and priced out to obtain a user cost function. This general agreement of results implies that drivers tend to behave in a cost-conscious manner and that the cost function of the average driver approximates that obtained from the pricing studies.

•ESTIMATES of road-user costs are valuable to the highway planner in a number of ways. Apart from their use in benefit-cost analysis, other research (1) has indicated the desirability of user costs in predicting zonal interchange volumes and traffic volumes on the various links of an urban network. In fact, research in developing diversion curves for assigning a portion of the zonal interchange volume to an expressway as opposed to the best alternate route (2, 3, 4, 5, 6) indicates that the best results are most likely obtained by using a combination of both comparative distance and comparative time. The relative weighting of distance and time, or some alternative combination where distance and speed are used, may be interpreted as being equivalent to determining the appropriate weighting between out-of-pocket costs associated with distance and the value of personal time associated with travel time.

In developing a road-user cost function, two alternative approaches are available. The first requires itemizing the components of the cost function (e.g., gasoline and oil consumption, tire wear, maintenance and personal time), determining how consumption of these different components varies with respect to speed of travel, establishing prices or costs for the different components and hence component cost functions, and adding up these component cost functions to obtain an overall user cost function. The second approach describes theoretically the expected shape of the user cost function and then attempts to estimate the parameters of the function statistically by using data where the cost function should be operative in determining how drivers behave. The values of the parameters in the function do not necessarily represent the true prices or costs encountered by the driver but only the relative weights between the different components of the cost function. Comparison of these relative weights will, however, indicate the degree of agreement between the two approaches.

The research described in this paper follows the second approach using diversion curve data for the Shirley Highway, the Dallas Central Expressway, the Gulf Express-

Paper sponsored by Committee on Origin and Destination.

way in Houston and the Alvarado-Mission Valley Freeway in San Diego. In addition, an attempt is made to estimate changes in the value of personal time associated with changes in personal income, at least to the extent that these changes affect driver behavior.

The first section describes the theory behind using diversion curve data and why these diversion curves are S-shaped. The second section describes a generalized road-user cost function of mixed quadratic-hyperbolic form and why the function should take this form. In the third section, the statistical problems of estimating the parameters of the cost function are discussed along with the final form of the estimating equation. Empirical results are given here. Finally, the research conclusions are given along with a comparison of the results derived here and those obtained by the more usual method of pricing out the various components of the cost function.

THEORY OF DIVERSION CURVES

When faced with the problem of choosing among alternatives, a person generally adds up the gains and costs for the various alternatives according to this value system and picks the alternative that is "best" for him. Best in his sense may be that where gains are greatest if costs are zero, where costs are least if gains are zero, or where the gains-cost ratio is greatest implying the highest rate of return among the alternatives.

However, it is well recognized that different people, when faced with the same problems of choice, act differently. That is, they do not choose the same alternative, as has been adequately demonstrated by diversion curve data. These differences are usually ascribed to two factors: (a) imperfect information concerning what the true gains and costs are among the alternatives, and (b) underlying differences in the value systems used when determining the total gains and costs for each alternative.

In most instances, available data only describe variations in behavior and do not distinguish between the two sources of variation. This lack of differentiation is probably unimportant in predicting behavior in the steady state. However, it could become important in trying to estimate the impact that an improved information system would have on behavior, e.g., how many drivers would change their route of travel given advance warning of an impending bottleneck on their present route along with knowledge of free flow conditions on some alternate. This is the problem with diversion curve data; they only measure total variations in behavior.

The hypothesis used in explaining the variations in behavior evidenced by diversion curves is simply that the variations in individual value functions caused by the two factors mentioned previously are normally distributed. In other words, individual estimates of the cost of travel on a given route, which differ due to these two factors, are normally distributed.

Given a normal distribution of estimates of travel costs on a given route, the distribution of estimates of comparative costs between two alternative routes, i.e., a freeway and best alternate, will also be normal or approximately normal (7), depending on how the cost comparison is made. For an algebraic difference, the comparative cost distribution will be normal, and for a cost ratio, the comparative cost distribution is approximately normal.

Support for this hypothesis is given in Figures 1 and 2 where the percentage of people using the freeway is plotted against distance difference and time difference, respectively. Both distance difference and time difference represent a major portion of a user cost function. Both plot linearly on probability paper, indicating that the comparative cost function is indeed normal.

It is to be expected that this distribution of route preferences with respect to comparative costs would, in addition to being normally distributed, have a zero mean. Even though the distribution may not have a zero mean for distance difference or time difference when considered separately, it should when considering the total cost function. This is because one would expect half the people to use the freeway and half to use the alternate when the comparative costs are equal, always remembering that costs are defined by user value systems.

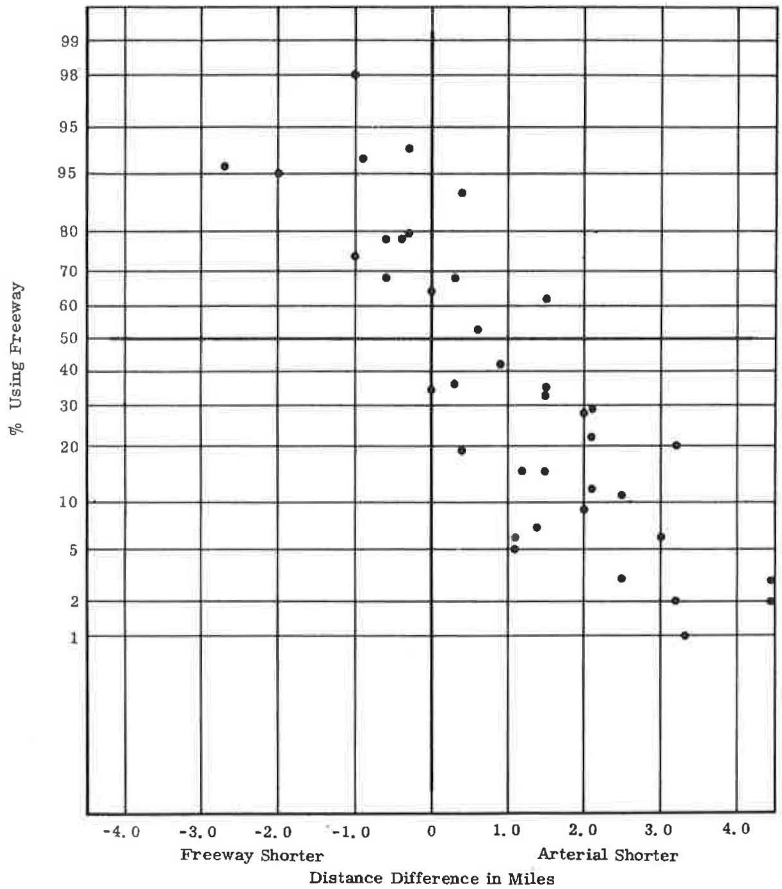

Figure 1. Relationship between percent using freeway and distance difference.

The problem of estimating the proportion who will use which route then becomes one of determining where the measure of comparative costs falls under the distribution and determining the amount of area under the curve that falls on either side of the comparative cost measure. This is shown in Figure 3. The curve is the normal distribution curve with mean zero corresponding to the case where comparative costs are equal for the freeway and the alternate. Two illustrative cases are given, ΔC_1 and ΔC_2, where ΔC refers to alternate costs minus expressway costs. In each case, the area under the curve to the left of the line corresponding to a given ΔC represents the proportion of trips taking the freeway and the area to the right represents the proportion taking the alternate. Thus, for ΔC_1, where freeway costs are lower than arterial costs, the area to the left of the vertical line indicates that a larger proportion of persons will use the freeway than the arterial. The converse holds for the case of ΔC_2.

Estimating the amount of area under the curve on either side of the vertical line (ΔC) may be accomplished by integrating the equation for the normal curve to some standard score (x) that is equivalent in position to ΔC. Thus, using the equation for the normal curve with mean zero and unit variance, the probability (or proportion of trips) of using the freeway P(F) or using the alternate P(A) are

$$P(F) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2}} dx \tag{1a}$$

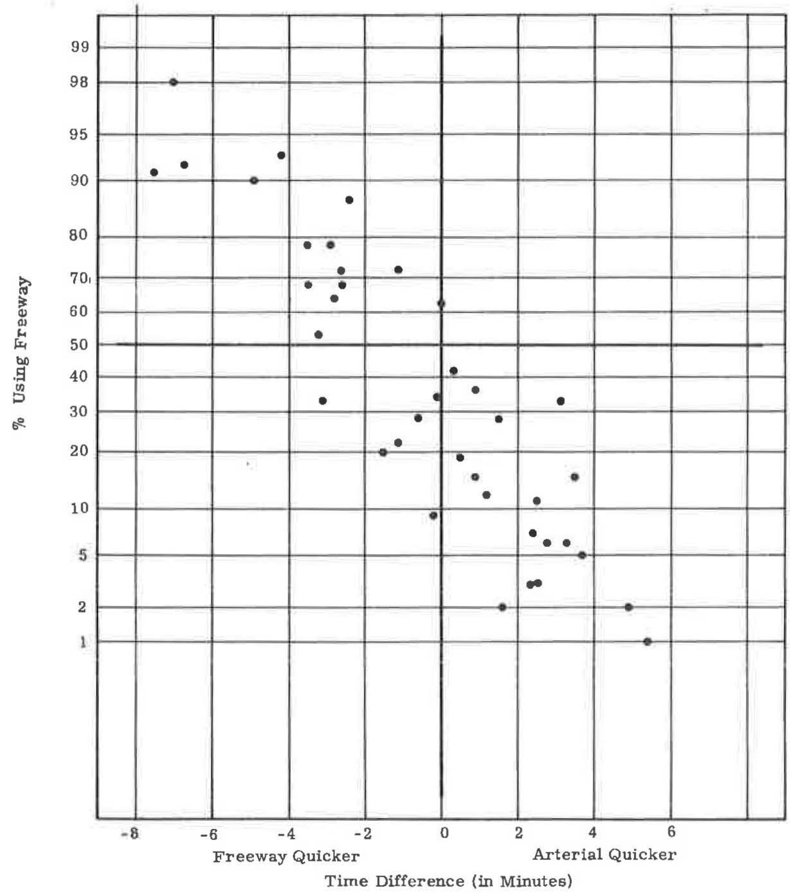

Figure 2. Relationship between percent using freeway and time difference.

$$P(A) = \int_x^\infty \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2}} dx \qquad (1b)$$

It is clear that P(F) plus P(A) equal one. It is also clear, though not presented here, that similar reasoning could be extended to the multivariate case where more than two alternate routes are considered.

Since P(F) and P(A) are precisely determined for a given value of x, the problem is one of estimating the value of x for corresponding values of ΔC. This is accomplished by determining the value of x, i. e., the standard score, corresponding to the percent of trips using the freeway and relating this to ΔC using multiple regression:

$$x = f(\Delta C) \qquad (2)$$

GENERALIZED COST FUNCTION

The two major components of user costs are out-of-pocket costs and the value of personal time. Out-of-pocket costs include such items as gasoline and oil consump-

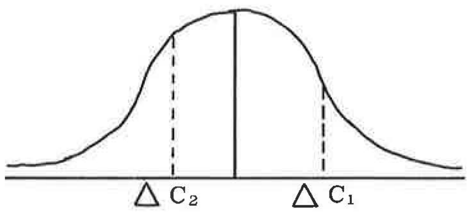

Figure 3.

tion, tire wear, and maintenance. Previous research has indicated that these costs are parabolic with respect to speed (8, 9, 10). This is to be expected because energy requirements for moving a vehicle are proportional to the square of the velocity. The constant and linear terms found in empirical estimating equations result from the loss of operating efficiency at low speeds. This loss is due to stop-and-go travel with the consequent consumption penalties due to idling, acceleration and deceleration and to vehicular design itself. Most automobiles are designed to operate most efficiently in the middle speed range. The value of personal time is usually assumed to be linear so that, for a given distance, its value is hyperbolic with respect to speed. The same assumption is made here.

A generalized cost function can, therefore, be described as follows:

$$C = D[a_0 + a_1 S + a_2 S^2 + a_3(1/S)] \tag{3}$$

where

C = cost,
D = distance,
S = speed, and
a_i = parameters to be estimated.

In this equation $D(a_0 + a_1 S + a_2 S^2)$ represents out-of-pocket costs and $D[a_3(1/S)]$ represents the value of personal time. In addition, the estimates of the parameters indicate the relative importance of the different parts of the cost function. These can be compared to the estimates obtained by the usual pricing procedures for consistency.

Empirical Results

The theory has been developed explaining choice between alternative routes of travel in terms of a probability density function whereby the position of the comparative cost of travel on the alternate routes under the normal distribution curve determined the probability of choice of a given route. It was also shown how this probability measure could be linearized by using a standard score corresponding to the percent using the expressway.

This standard score could then be estimated using the usual single equation multiple regression techniques with a set of independent variables representing a comparative cost function. For this analysis, a cost difference equation was used, i. e., cost via alternate minus the cost via the expressway. The cost difference is preferred to the cost ratio because: (a) we see no theoretical basis for choosing one over the other, particularly in urban travel; and (b) use of the cost difference is more tractable mathematically. Thus, we can write a generalized cost function, as in the previous section, for use of the expressway and another function for using an alternate, and take the difference between the two. This yields the following equation used in the multiple regression analysis:

$$P(F) = k + a_0(D_a - D_e) + a_1(D_a S_a - D_e S_e) + a_2(D_a S_a^2 - D_e S_e^2) + a_3(T_a - T_3) \tag{4}$$

where

$P(F)$ = probability of using expressway;
D, S, and T = distance, speed, and time for the expressway and alternate as denoted by the subscripts a and e;
a_i = parameters of equation to be estimated; and
k = constant term in regression equation which, according to our hypothesis, should equal zero.

TABLE 1

COMPARISON OF PARAMETER ESTIMATES OF COST FUNCTION FOR DIFFERENT REGIONS OF THE COUNTRY

Parameter	California	Texas	Washington, D. C.	All
a_0	0.608	0.806	0.656	0.726
a_3 (T)	0.562	0.687	0.385	0.737
a_1 (S)	-0.0174	-0.0252	-0.0189	-0.0158
a_2 (S^2)	0.000362	0.000648	0.000375	0.000298
r	0.820	0.867	0.946	0.813

The data used in this analysis were from the Shirley Highway (2), the Gulf Freeway in Houston (3), the Central Expressway in Dallas (4), and the Alvarado-Mission Valley Freeway in San Diego (5). Certain zonal interchange volumes were not used because it was felt that the sample size represented by them was too small to be reliable, regardless of whether or not they fitted the equation well. This provided a total of 197 sample cases.

In testing this hypothesis, both theory and previous experience indicate the range of values the parameter estimates should take. Thus, k, the constant term in the equation, should approximate zero because, if the comparative costs of the two routes are equal, 50 percent of the people should take each route. The standard score associated with 50 percent usage is zero. The three parameters, a_0, a_1, and a_2, are the three parameters in the parabolic equation relating per mile out-of-pocket cost to speed of travel. Thus, a_0 should be positive. The value of a_1 should be negative and smaller in magnitude than a_0. The value of a_2 should be positive and smaller in magnitude than a_1. Both the sign and relative size of these parameters has been indicated in prior research (8). Finally, the value of a_3, which corresponds to an estimate of the value of personal time, should be positive.

In estimating the parameters of the equation a departure from the usual type of multiple regression analysis was done because of the high degree of intercorrelation among the independent variables. Briefly, the procedure followed was to obtain orthogonal estimates of each parameter, assuming independence among the independent variables, and then to multiply these parameter estimates by the variance-covariance matrix for the parameter estimates obtained through the usual multiple regression analysis. The variance-covariance matrix was scalar multiplied by an iterative procedure to constrain the constant term (k) in the equation to be equal to zero.

The final parameter estimates obtained by this procedure are presented in Table 1. For these purposes Dallas and Houston were combined under the assumption that there would be no regional differences in behavior between the two cities and to increase the sample size for that region. It will be noted that there is general agreement among the parameter estimates for the different areas. They are, in fact, statistically not significantly different from each other.

The comparative cost equation was further modified to incorporate median income of the home zone to determine if the value of personal time was related to income. This was accomplished by adding an additional term to the right-hand side of Eq. 3 of the form a_4YT where Y corresponds to median income, and incorporating it in Eq. 4 as $a_4Y(T_a - T_e)$. The results are not presented here because the parameter estimate of a_4 was not found to be significantly different from zero; in the three different regional samples and the combined sample, the parameter estimate was in the order of 10^{-10} or greater.

CONCLUSIONS

The theory and methodology have been explained for deriving a road-user cost function using diversion curve data. The correlation coefficients, which all exceed 0.80, indicate that the data fit the theory very well. Indeed, when the sampling error inherent

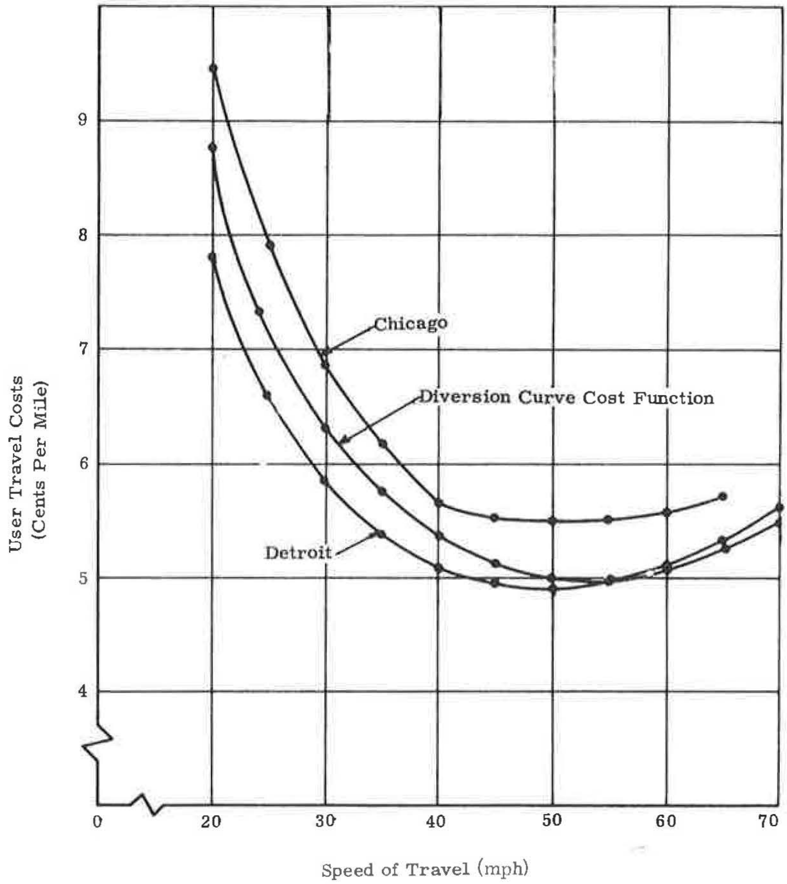

Figure 4. User cost curves for Chicago, Detroit and this study.

in estimating the zonal interchange volumes and, hence, the proportion using the expressway is considered, these correlation coefficients are probably close to the upper limit that could be achieved from the data.

The implications of this research are three-fold. First is the obvious one in using the results obtained for assigning traffic to a proposed new facility where it is desired to use diversion curves. Second, and probably more important, is that the user cost equation developed may be used in predicting zonal interchange volumes or traffic assignment with the minimum path programs currently in use, or both. Finally, user trip costs, which can be estimated using the equation, represent another measure of trip length rather than either distance or time alone and can, therefore, be combined with other data in estimating trends in trip length.

The derivation of the user cost equation from the results is straightforward. Since the hypothesis presented here states that the proportion of trips using the expressway is a function of comparative costs on the two routes and the comparative cost function is well defined, the costs on either route can be arbitrarily set equal to zero and the total cost for the remaining route can be estimated. It should be pointed out that the user costs estimated by this equation are parametric estimates; i. e., the parameters indicate the relative importance of the different components in the user cost equation. The final measurement is not in dollars and cents but in some arbitrary unit—maybe gilders. However, a direct comparison may be made between the user cost equation derived here and that obtained by the more usual methods, such as employed in Detroit (8) and Chicago (10), by multiplying each of the parameter estimates by a constant to place them in the same range as the Detroit and Chicago equations and then comparing the shape of the curves. This is shown in Figure 4, where each of the parameter estimates was multiplied by 3.2 to place the resulting curve in the same range as the other two. As can be seen, all three curves have a similar shape.

The multiplication by the constant of 3.2 provides a basis for estimating the value of personal time, at least in its relationship to the cost of other portions of the cost function. The precise value of personal time chosen is, of course, dependent on the value of the constant multiplier chosen, but this analysis indicates its value is on the order of $1.25 to $1.40 per hour. The greater steepness of our curve, as well as the fact that it crosses the Detroit curve, indicates that the value of personal time is greater than the $1.20 per hour used in pricing out the latter.

In conclusion, we have developed a road-user cost equation from data reflecting driver behavior where the driver's choice of route is determined from comparative costs on the two routes. There is a high degree of agreement between the cost equation developed from this approach and those obtained in the usual pricing studies. This indicates both that drivers tend to behave in a cost-conscious way and that the cost function of individual drivers approximates that obtained from pricing studies.

REFERENCES

1. Hansen, Walter G. Recent Applications of Gravity Model Procedures. Paper presented to O-D Committee of Highway Research Board, Jan. 1964.
2. Trueblood, Daniel L. Effect of Travel Time and Distance on Freeway Usage. Highway Research Board Bull. 61, pp. 18-37, 1952.
3. Traffic Assignment to the Gulf Freeway. Texas Highway Dept., Dec. 15, 1954.
4. Texas Highway Department, photostatic copies of tabulations for Central Expressway in Dallas, May 14, 1952.
5. Moskowitz, Karl. California Method of Assigning Diverted Traffic to Proposed Freeways. Highway Research Board Bull. 130, pp. 1-26, 1956.
6. Campbell, E. Wilson. A Mechanical Method for Assigning Traffic to Expressways. Highway Research Board Bull. 130, pp. 27-46, 1956.
7. Duncan, Acheson J. Quality Control and Industrial Statistics. Richard D. Irwin, Inc., 1959.
8. Bevis, Howard W. The Application of Benefit-Cost Ratios to an Expressway System. Highway Research Board Proc., Vol. 35, pp. 63-75, 1956.
9. Road User Benefit Analyses for Highway Improvements. AASHO, 1960.
10. Chicago Area Transportation Study, Final Report. Vol. 3, 1962.