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.

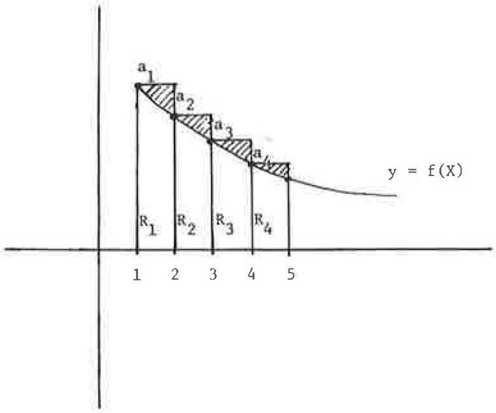
In 1970 Zahavi (1) noted that, when certain transportation-related quantities are ranked in decreasing order of value, a certain level of stability is attained in the relationship between rank and value that enables analyses that might not otherwise be possible. This relationship, well known in the sciences, 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

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

or a double exponential of the form:

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

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

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

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

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

an additive correctiion factor (A) would require

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

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

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

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.

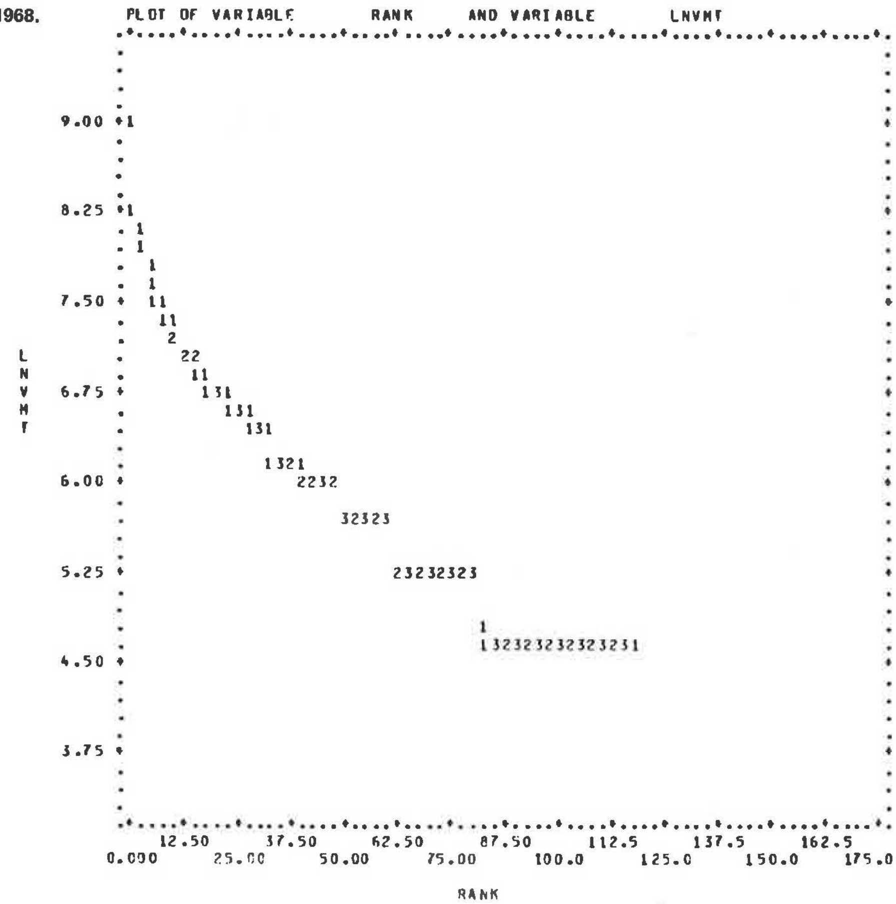


Figure 3. Log AADVMT by rank for 1974.

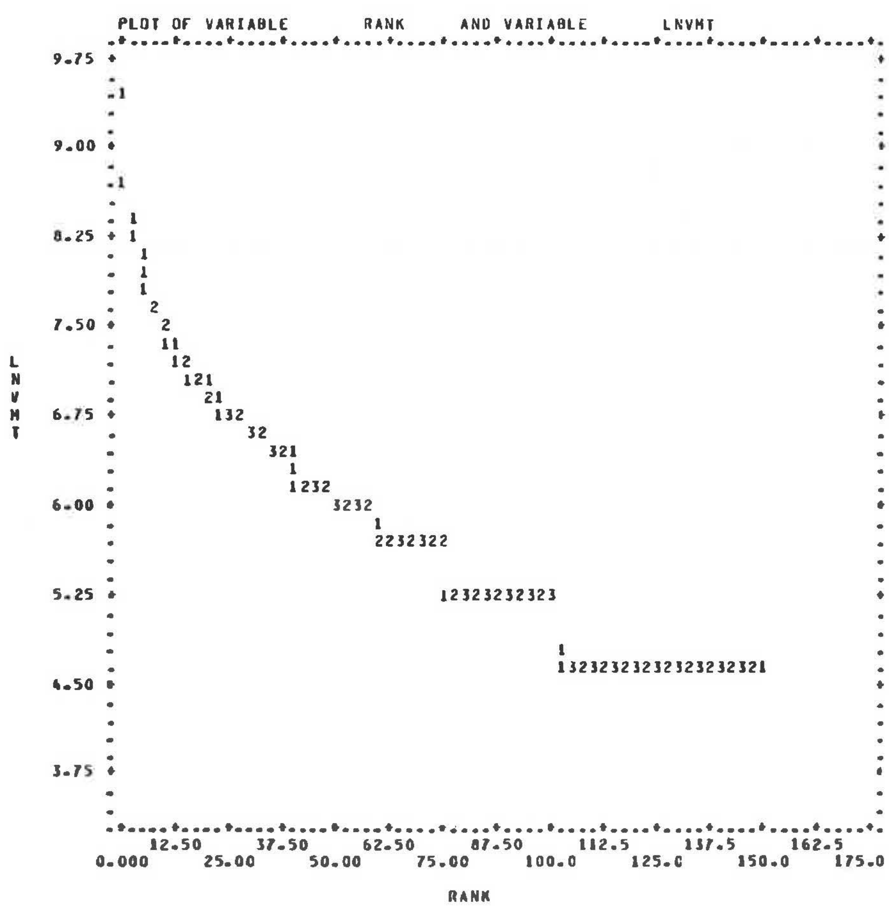
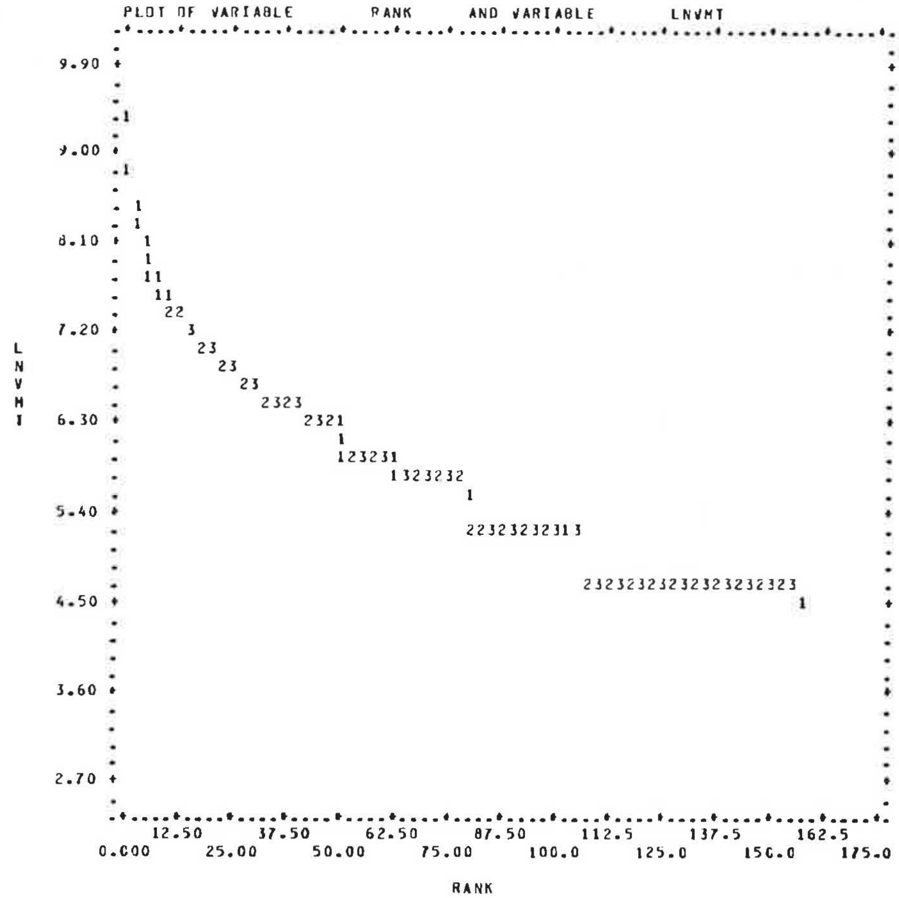


Figure 4. Log AADVMT by rank for 1976.



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,

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

By evaluating the right-hand side, we obtain

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

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

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

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

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

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