Marginal Weighting of Transportation Survey Data

L. Taylor Ollmann,* John Hancock Mutual Life Insurance Company, Boston
Stephen M. Howe,* North Central Texas Council of Governments, Dallas
Kenneth W. Kloeber,* Calocerinos and Spina, Buffalo
Gerald S. Cohen, New York State Department of Transportation

Two procedures are discussed for calculating cell weights to be applied to sample data to ensure that the expanded sample jointly represents the population on several background distributions. Deming's procedures minimize the squared deviation of cell weights from unity, whereas Johnson's procedure iteratively produces weights that match the marginal distributions, are all positive, and have low variance from unity. The theory and calculation procedure of these two approaches are developed and are illustrated by a simple example. Advantages and disadvantages of the procedures are described, and applications to various transportation problems are highlighted. The paper concludes that, although the Johnson technique is less well grounded in theory, it produces reasonable weights, prevents negative estimates, and converges quickly. Its use is recommended.

A common problem in transportation planning is the selection and processing of a sample of observations so that they represent a larger population. Such activities are central in most analyses of transportation policies because planners are often concerned with the impacts of policies on the general population and its various subgroups. Methods of simple random sampling ensure representations within statistical limits, but may not be cost-effective or feasible. Other sampling procedures (stratification, clustering, multistage, or quota) can also be used, but each has its biases and limitations. These problems are further compounded by "micro" samples used in disaggregate modeling, and numerous sophisticated aggregation procedures have been evolved to ensure model application, even from highly nonrepresentative samples.

Many conventional analyses, however, still require that representativeness be demonstrable, if for no other reason than to increase the analyst's confidence in the validity of the data. In most surveys, representativeness is taken to mean that the distribution of sample data agrees (within statistical limits) with distribution of known demographic or geographic data for the population. The underlying assumption of such a test is, of course, that, if the demographics are representative, other data (such as trips and opinions) will also be representative of the population.

In many transportation surveys, however—particularly small-sample, telephone, and mail-out studies—it is difficult to achieve representations without very great cost. Hence, too many of one group may be sampled and not enough of another. To ensure that the data conform to census (or other population) distributions, one must generally weight the data by multiplying by appropriate weighting factors. For instance, 5 percent of a sample may consist of men aged 15 to 24 whereas census figures might show that 8 percent of the population is of that type. One would then weight the sample data for this group by a factor of 1.6.

MARGINALS AND CELLS

Suppose that a survey sample is categorized according to sex, age group, household size, and number of automobiles owned per household. The data may thus be tabulated into a four-dimensional contingency table, one cell of which might consist of males aged 15 to 24 from two-person households that own one automobile. A marginal group may be defined as an aggregate of all cells that share a single characteristic—for example, the marginal group of persons aged 15 to 24, the marginal group of persons from two-automobile households, or the marginal group of females. [In tabulation, the marginal groups would appear in the table margins as row or column sums (Table 1).]

It would be very useful if population data were available for each cell because then the sample data for each cell could be weighted to match the proportion of that cell in the entire population. Unfortunately, population (e.g., census) data are rarely cross tabulated for more than two categories simultaneously. Thus, the best we can do is to compute expansion factors for each of the cells so that the resulting weighted samples conform to census proportions in each marginal group. Marginal weighting is any procedure for artificially computing those cell expansion factors. This paper describes such a method recently programmed by the New York State Department of Transportation (NYSDOT) (1) and used extensively in small-sample transportation studies.

MARGINAL WEIGHTING

Several different procedures of marginal weighting are in use, and they generally produce different weights for the same data. An illustration is given below for two categories: age (in three marginal groupings with census proportions of 25, 50, and 25 percent) and sex (in two marginal groupings of 50 percent each). Table 1 gives sample proportions for each of the six cells and the five marginal-group target proportions mentioned above.

Table 2 gives four sets of weights that, when multiplied by the data in Table 1, yield proper marginals. For example, for the age 55 and over marginal group in Table 2, $0 \times 15 + 1.25 \times 20 = 25$ percent (procedure 1), $1 \times 15 + 0.5 \times 20 = 25$ percent (procedure 2), $0.82 \times 15 + 0.635 \times 20 = 25$ percent (procedure 3), and $0.79 \times 15 + 0.6575 \times 20 = 25$ percent (procedure 4). Since numerous sets of weights can be conducted, we must ask what criteria should be used to compare one possible set of weights with another and whether there is a best set of weights.

The following criteria seem reasonable:

1. The marginal groupings of weighted cells should match closely (if not exactly) the corresponding population (census) proportions.
2. The individual weights should tend to be near one. That is, the individual cells (particularly those that contain a large proportion of the sample) should be changed as little as possible.
3. Weights should all be positive. Factors of zero fail to use all available information, and negative factors make even less sense.

Procedure 1 in Table 2 shows a particularly bad set of weights that satisfy the first criterion but violate the

Table 1. Sample data.

Age Group	Male (%)	Female (%)	Total (%)	Census Marginal (%)
16 to 24	10	10	20	25
25 to 54	20	25	45	50
55 and over	15	20	35	25
Total	45	55	100	
Census marginal	50	50		100

Table 2. Examples of marginal weights.

	Marginal Weighting					
	Age 16 to 24		Age 25 to 54		Age 55 and Over	
Procedure	Male	Female	Male	Female	Male	Female
1	5	-2.5	0	2	0	1.25
2	1	1.5	1.25	1	1	0.5
3 (Deming)	1.342	1.1575	1.214	1.029	0.820	0.635
4 (Johnson)	1.3644	1.1356	1.2253	1.0198	0.7900	0.6575

Table 3. Example cell weights for skewed samples.

Item			Census Marginals (%)	Deming Weights		Johnson Weights	
Cell percentages	20	30	30	0	1	0.2863	0.8091
	30	20	70	1	2	0.8091	2.2863
Census marginals, %	30	70					
Cell percentages	20	30	20	-0.5	1	0.1125	0.5917
	30	20	80	1	2.5	0.5917	3.1125
Census marginals, %	20	80					

next two. The weights in effect throw away all sample information about men over age 24 and reverse any sample input from young women! Procedure 2 gives a more reasonable set of weights, satisfying the first criterion and also doing a reasonable job on criteria 2 and 3. Generally, no set of weights is optimal with respect to all these criteria at once. Marginal weighting methods are known for optimizing some of the criteria but generally at the expense of others.

Deming Method

One procedure is designed to optimize the second criterion subject to the constraint of satisfying the first exactly (2). Deming measures the second criterion by the formula $\Sigma_{i_1} \Sigma_{i_2} \Sigma_{i_n} a_{i_1 i_2 \ldots i_n} (W_{i_1 \ldots i_n} - 1)^2$ where, in an n-dimensional table of sample data, $a_{i_1 i_2 \ldots i_n}$ is the proportion of data in cell $i_1, i_2, \ldots, i_n$ and $W_{i_1 \ldots i_n}$ is the computed weight applied to that cell. This term can then be minimized subject to the constraint of exactly matching the census target proportions. If there are m marginal groups in all, the computation of cell weights applies the method of Lagrange multipliers to obtain m equations in m unknowns, which are then solved to provide one coefficient for each marginal group. The cell weights are then computed by adding to one the coefficients of all marginal groupings to which the cell belongs. Hence, in the example from Table 1, five simultaneous linear equations are solved to yield two coefficients for sex ($c_1 = 0.1849$ and $c_2 = 0$) and three coefficients for age ($d_1 = 0.1576$, $d_2 = 0.0289$, and $d_3 = -0.3650$). The cell weights (given in procedure 4 in Table 2) are then computed by $W_{ij} = 1 + c_i + d_j$.

The advantages of the Deming approach are that the weights can be directly calculated, the number of computations remains manageable, and the results are "optimal." The major disadvantages are that some weights can be very far from one or can be even zero or negative, especially when the sample marginal proportions are sharply skewed. Deming's weights do seem reasonable when the sample is not so severely skewed (Table 3). In addition, the computation, which involves linear algebra, calculus, and the method of Lagrange multipliers, is harder to understand.

Johnson Method

An alternative weighting procedure proposed by Johnson (3) assigns a coefficient to each marginal grouping; the cell weights are then obtained by multiplying together the coefficients of all marginal groupings to which the cell belongs. Johnson describes a simple iterative procedure for producing such cell weights so as to match exactly the census marginal proportions (criterion 1).

Assume that sample proportions for n marginal distributions are given (r_k marginal groups in the kth category). There are $(r_1)(r_2)\ldots(r_n)$ cell proportions $a_{i_1 \ldots i_n}$ and $r_1 + r_2 + \ldots + r_n$ marginal target proportions M_{k,i_k} ($k = 1, n$, $i_k = 1, r_k$). First, sum the cell sample proportions for each marginal group in the first marginal:

$$S_j^1 = \Sigma_{i_2} \ldots \Sigma_{i_n} a_{ji_2 \ldots i_n} \qquad j = 1, r_1 \tag{1}$$

and then rescale the cells in each of these marginal groups to match target marginal sums in the first marginal:

$$a^1_{ji_2 \ldots i_n} = a_{ji_2 \ldots i_n} \cdot M_{1,j}/S_j^1 \qquad j = 1, r_1 \tag{2}$$

The rescaled cell values a^1 then match the first marginal target sums exactly but probably do not match targets in other marginals. One then rescales a second time to match the second marginal targets:

$$S_j^2 = \Sigma_{i_1} \Sigma_{i_3} \ldots \Sigma_{i_n} a^1_{i_1 j, i_3 \ldots i_n} \qquad j = 1, r_2 \tag{3}$$

$$a^2_{i_1 j, i_3 \ldots i_n} = a^1_{i_1 j, i_3 \ldots i_n} \cdot M_{2,j}/S_j^2 \tag{4}$$

and again for each marginal until the last:

$$S_j^n = \Sigma_{i_1} \ldots \Sigma_{i_{n-1}} a^{n-1}_{i_1 \ldots i_{n-1} j} \qquad j = 1, r_n$$

$$a^n_{i_1 \ldots i_{n-1} j} = a^{n-1}_{i_1 \ldots i_{n-1}} \cdot M_{n,j}/S_j^n \tag{5}$$

This completes the first iteration.

For the data in Table 1, these calculations are as follows:

$S_1^1 = 0.1 + 0.2 + 0.15 = 0.45$

$S_2^1 = 0.1 + 0.25 + 0.2 = 0.55$

$a^1_{1,1} = 0.1(0.5/0.45) = 0.1111$

$a^1_{1,2} = 0.2(0.5/0.45) = 0.2222$

$a^1_{1,3} = 0.15(0.5/0.45) = 0.1667$

$a^1_{2,1} = 0.1(0.5/0.55) = 0.0909$

$a^1_{2,2} = 0.25(0.5/0.55) = 0.2273$

$a^1_{2,3} = 0.2(0.5/0.55) = 0.1818$

$S_1^2 = 0.1111 + 0.0909 = 0.2020$

$S_2^2 = 0.2222 + 0.2273 = 0.4495$

$S_3^2 = 0.1667 + 0.1818 = 0.3485$

$a^2_{1,1} = 0.1111(0.25/0.2020) = 0.1375$

$a^2_{1,2} = 0.2222(0.5/0.4495) = 0.2472$

$a^2_{1,3} = 0.1667(0.25/0.3485) = 0.1196$

$a^2_{2,1} = 0.0909(0.25/0.2020) = 0.1125$

$a^2_{2,2} = 0.2273(0.5/0.4495) = 0.2528$

$a^2_{2,3} = 0.1818(0.25/0.3485) = 0.1304$

At the end of the first iteration, the rescaled cell values $a^n_{i_1 \ldots i_n}$ will not match exactly any target marginal sums except those of the last marginal. However, they will tend to match other marginal targets more closely than did the original cell values $a_{i_1 \ldots i_n}$. A second iteration can then be performed that repeats the above process for each category in turn and finally obtains cell values $a^{2n}_{i_1 \ldots i_n}$, which will tend to match each marginal target more closely than did the values $a^{n}_{i_1 \ldots i_n}$. One continues to iterate until all marginals are matched simultaneously to a desired degree of accuracy. In our example, three full iterations suffice to obtain values that match all five marginal targets with a tolerance of 10^{-7}. Weights for each of the cells are then obtained by dividing the final scaled cell values by the original values, as follows:

$$W_{i_1 \ldots i_n} = a^N_{i_1 \ldots i_n} / a_{i_1 \ldots i_n} \qquad \text{if } a_{i_1 \ldots i_n} \neq 0 \tag{6}$$

For our example, those weights appear in procedure 4 in Table 2.

The advantages of the Johnson technique are many. The iterative procedure is easy to understand and to program: One simply multiplies each marginal in one category (age, for example) by a factor that makes the weighted cell total match the census target for that marginal. One then repeats this process for each of the other categories and then again for all categories and so on until all marginals match their census proportions simultaneously. Further, the procedure converges rather quickly: Weights have been obtained to an accuracy of six decimal places in 5 to 15 iterations on small-sample data that have three categories of three, four, and six marginal groupings (72 cells).

Criteria 2 and 3 are also well satisfied: All cell weights are positive and tend to be near 1.0. Values of Deming's criterion tend to be low although not the minimum obtainable by Deming's method. The weights obtained look reasonable and seem to be uniquely determined by the requirement that a fixed weight be assigned to each marginal grouping and each cell weight be simply the product of the weights of those marginal groupings to which the cell belongs. Reasonable weights are obtained even in matrixes where 50 to 60 percent of the cells contain no samples as long as each marginal grouping is reasonably well sampled. The method is robust: The inputted sample can initially be highly skewed from desired marginals as in the case of daytime telephone samples.

DISCUSSION OF RESULTS

One disadvantage of both the Deming and Johnson procedures is that the act of subdividing a marginal group into several smaller groups changes the weights assigned not only to cells of that group but also to cells of other groups that were left unchanged. For example, if the two lower age groupings in Table 1 are combined, the Deming and Johnson weights are as given below:

Age Group	Male (%)	Female (%)	Total (%)	Census (%)
16 to 54	30	35	65	75
55 and over	15	20	35	25
Total	45	55	100	
Census	50	50		100

Deming Weights		Johnson Weights	
1.2556	1.0667	1.2709	1.0535
0.8222	0.6333	0.7916	0.6563

Notice that the weights assigned to the cells for the group aged 55 and over are not the same as the corresponding ones in procedures 3 and 4 in Table 2 even though the cell descriptions are the same. The differences are only a few percent, but the examples show that the numbers produced by either method have no natural validity and artificially depend on the way one chooses to define marginal groupings.

We attempted to overcome this difficulty by applying the Johnson technique to each cell individually. To weight one cell, we aggregated all marginals in each category except for the marginals that contained the particular cell in question. The result was a $2 \times 2 \times \ldots \times 2$ table of sample data (categorized only by the marginals that defined the cell and the complementary marginal groups), one cell of which was identical to the original cell. The Johnson technique was then applied to obtain a weight for that cell. Cell weights for each cell in the original data table were obtained one at a time in this way and are given below:

Age Group	Male	Female
16 to 24	1.3675	1.1325
25 to 54	1.2372	1.0102
55 and over	0.7916	0.6563

The resulting weights depended only on the description of the particular cell, the sample proportions of that cell's own marginal groups, and the census targets for those groups. Hence, changing the definitions of some marginal groups would not change the weight of an unaltered cell. This approach and the pure Johnson procedure were then applied to small survey samples from four small upstate New York cities: Plattsburgh, Watertown, Glens Falls, and Elmira (4). The sample sizes ranged from 209 to 245 and were tabulated in three categories divided into three, four, and five marginal groups. Between 22 and 32 of the 72 cells in each survey were empty, and many of the remainder contained only one or two samples. The weights so produced tended to be closer to 1.0 than those produced by the pure Johnson algorithm; Deming criterion values were about 20 to 50 percent less. Unfortunately, this individual treatment of cells was unable to reproduce the target marginal totals (criterion 1). The marginal totals after weighting differed from the desired values by 1 to 10 percent and in a few cases by as much as 30 percent! This is unacceptably high. It may well be that such losses of marginal accuracy are necessary in any weighting procedure whose cell weight function depends only on the cell and directly related marginals and does not depend on the classification of all other data in the sample.

Returning to the two original weighting procedures that do yield correct marginals, we favor the Johnson approach for its previously mentioned advantages. Johnson points out that the procedure can fail to converge if a large proportion of the cells are empty. Our results on the small-city surveys showed that this proportion can be as much as 60 percent and the algorithm will still converge in less than 15 iterations.

The chief distinction of the Deming approach is the criterion it optimizes. We conjecture that the Johnson method also minimizes a function of the following form:

$$\Sigma_{i_1} \ldots \Sigma_{i_n} a_{i_1 \ldots i_n} \cdot W_{i_1 \ldots i_n} \cdot \log (W_{i_1 \ldots i_n}) \tag{7}$$

This is not unreasonable as a criterion. But we can give no "elegant" reason why this is an ideal function to minimize.

Marginal weighting techniques have many possible applications. The New York State Department of Transportation has found them useful in many of the following contexts:

Analytical Problems	Studies
Nonrandom samples	Daytime telephone surveys
Surveys with missing data	Small area surveys
Limited budget or small samples	Transit on-board studies
Need for quick turnaround	Disaggregate data for mode choice
Limited population data	Community attitude studies
	Travel update studies
	Corridor studies

NYSDOT has applied these techniques in many contexts, including a 1000-household survey, a 300-household survey on community attitudes toward transit, a survey of 1200 transit riders in Albany, and 200- to 300-household surveys of community attitudes in six towns. In all cases, the method has proved to be useful and time saving and to provide reasonable results. The procedure develops reasonable weights with manageable effort for real-world data even if those data are highly skewed with many empty cells. Examples may exist that cause the method to fail, but they must be concocted for the purpose, would probably not occur in practice, and would probably cause other methods of marginal weighting to fail also. We feel that both methods are valuable additions to the repertoire of survey analysis techniques used by transportation analysts, and we particularly recommend the Johnson procedure because of its ease of understanding and programming.

ACKNOWLEDGMENTS

We wish to recognize the important contribution of Richard Johnson of Market Facts in constructing and describing his procedure. The assistance of Wilma C. Marhafer in preparing this manuscript is gratefully appreciated.

REFERENCES

1. K. W. Kloeber and S. M. Howe. Marginal Weighting Procedures for Expanding Small Sample Surveys. Planning Research Unit, New York State Department of Transportation, PRR 97, Nov. 1975.
2. W. E. Deming. The Statistical Adjustment of Data. Dover, New York, 1964.
3. R. M. Johnson. Marginal Weighting. Market Facts, Chicago, Oct. 1972.
4. W. C. Holthoff, S. M. Howe, and E. P. Donnelly. Documentation of Expansion and Cross Tab Set Up for Small Urban Area Study, 1974-75. Planning Research Unit, New York State Department of Transportation, memorandum, Dec. 1975.

Publication of this paper sponsored by Committee on Transportation Information Systems and Data Requirements.

**Messrs. Ollmann, Howe, and Kloeber were with the New York State Department of Transportation when this research was performed.*

Demonstration of a Simplified Traffic Model for Small Urban Areas

Donald Rhodes, New Hampshire Department of Public Works and Highways
Thomas Hillegass, Urban Mass Transportation Administration, U.S. Department of Transportation

A simple and straightforward method for forecasting highway traffic that was applied in Plaistow, New Hampshire, is described. The method requires an external origin-destination survey, demographic forecasts, and good ground counts. No home interview survey or demand model calibration is required. The approach combines a Fratar expansion of external trip tables with a factoring of internal travel on a link-by-link basis to produce a forecast of total automobile travel by link. A unique feature is that the link-by-link factoring of internal travel is directly related to the growth of population and employment in internal zones. This is done by means of an artificial trip-generation and gravity model that requires no calibration. The method produced logical results for Plaistow and appears to have promise for other applications.

There is widespread recognition today that traditional transportation planning techniques are somewhat more complex and costly than warranted by the types of problems to be addressed in smaller urban areas. This recognition and the desire to direct more of limited planning resources toward the solution of short-range problems have led to a search for simplified approaches to long-range travel forecasting that are tailored to the needs of smaller urban areas.

Planners for smaller urban areas are urged to perform an analysis of problems, growth, and related factors and to design study techniques that best suit these problems. Yet, lacking proven step-by-step alternatives to current forecasting methods, planners for these areas are justifiably reluctant to risk new approaches. The profession badly needs demonstrations of new methods, or new ways of using old methods, that are relatively straightforward and are suited to typical problems of small urban areas. Both the method and the outcome of these demonstrations as well as the types of situations to which they are adaptable need to be well documented. This paper attempts such a documentation of an approach taken by the New Hampshire Department of Public Works and Highways to forecasting automobile traffic for the town of Plaistow, New Hampshire.

PLAISTOW

Plaistow is a 27-km^2 (10.5-mile^2) community with an estimated 1975 population of 5330 persons. Located on