Title: **HOUSEHOLD TRAVEL SURVEYS: USING DESIGN EFFECTS TO COMPARE ALTERNATIVE SAMPLE DESIGNS**

Authors: James Leiman, Ph.D.
Morpace Inc.
31700 Middlebelt Rd, Suite 200
Farmington Hills. MI 48334
248-737-3216 - phone
248-737-5326 – fax
jleiman@morpace.com

Karen Faussett
Michigan Department of Transportation
P.O. Box 30050
Lansing, MI 48909
517-353-2956 - phone
517-373-9255 - fax
faussettk@michigan.gov

Submission Date: August 1, 2005

Modified Date: November 8, 2005

Word Count: 3,116 plus 4 tables = 4,116

ABSTRACT
A common sample design used in household travel surveys is a geographically stratified RDD sample. Quotas are usually established within each geographic stratum using the number of persons in a household and/or the number of vehicles in the household to define cell quotas. Sampling continues within a stratum until the quotas are met. Due to differential response rates associated with the quota variables, this strategy can significantly increase overall survey costs. How does one measure the benefits of this commonly used approach? Two methods are examined. The first consists of calculating the correlations between key survey variables and quota variables. The second consists of calculating design effects treating the quota variables as post-stratification variables. Design effects index gains or loses in precision for a particular design compared to the use of a simple random sample. Implications for evaluating alternative sample designs are presented. The data come from the 2004 Michigan Travel Counts travel inventory.

Stratified random sampling is commonly used in household travel surveys. This approach usually consists of the identification of a number of geographic strata coupled with RDD (random-digit-dial) sampling within each stratum. Allocation across the geographic strata may be proportionate or disproportionate to the respective household populations. Quotas are usually established within each stratum using variables with known correlations with key survey measures. Typical quota variables are the number of persons and the number of vehicles in a household. Some researchers have also used household income. The quotas are usually based on a population model using U.S. Census data. Again, the allocation of quotas across cells may be proportionate or disproportionate, depending on the study objectives and, in some cases, variance calculations from previous research.

[A good discussion of various sampling approaches used in household travel surveys can be found in the *Travel Survey Manual*, developed under contract to the Federal Highway Administration (FHWA) as part of the Travel Model Improvement Program (TMIP). This manual is available online at http://ntl.bts.gov/index.html.]

TABLE 1 Typical Travel Sample Design For A Given Geographic Stratum

Household Size	0 Vehicle Households	1 Vehicle Households	2 Vehicle Households	3+ Vehicle Households
1 person	quotas	…	…	…
2 persons	…	…	…	…
3 persons	…	…	…	…
4+ persons	…	…	…	…

If it is decided to maximize the efficiency of the sample design a strict proportional allocation based on a population model across the cells of the design will be made, i.e., the desired sample size within a given cell will correspond to its relative proportion of the population across all the cells. However, if variance estimates for key parameters in the survey are available, the allocation may be weighted by these estimates as a further strategy for reducing overall sampling error.

The primary motivation behind this process is to insure that the sample is as representative as possible for some key demographics that have been shown to be related to some important travel survey variables (e.g., the number of trips generated by a household) and are important predictors in travel demand forecasting models.

This quota sampling process is based on some of the same motivations behind the use of stratified sample designs. However, in the typical travel survey, beyond the initial geographic stratification, the process does not constitute true stratification because the frame elements (e.g., households) cannot be divided a priori into separate groups based on characteristics – like household size or the number of vehicles – which are then sampled independently. Instead, one simply uses a single RDD frame within each geographic stratum.

Due to differential response rates associated with the household characteristics used to establish the quotas, successful implementation of the approach – meeting the pre-defined quotas – generally requires significant effort and can increase overall survey costs. This raises the question of how to measure the benefits of this approach.

Stratification is recommended in survey sample design because it almost always leads to increases in precision. (See a statistics text on sampling for a discussion of stratification, e.g., Kalton (1983) or Groves et. al. (2004).) One indicator of possible benefits is simply the degree of association between key survey variables on the one hand, and the frame or strata variables on the other. A more precise indicator is the calculation of design effects.

A common practice used by survey statisticians to evaluate a design is to compare sampling variances from a survey with the sampling variances had the survey been conducted as a simple random sample (SRS). This ratio is known as the design effect (*deff*). The design effect for a survey is simply the sampling variance calculated from the actual survey design divided by the sampling variance using an SRS design. For example, in the case of a stratified design, the design effect for the mean of a particular estimator *y* would be the following:

$$deff = v_{st}(\bar{y}) / v_{srs}(\bar{y})$$

Stratification can lead to design effects less than one – lower sampling variances compared to an SRS survey – when the survey variable is correlated with the variables defining the strata. (See the discussion in Groves et. al. (2004).) Perhaps this can provide us with a process for evaluating the sample designs commonly used in household travel surveys.

In the balance of this paper we take some actual survey results and first compute the correlations between key survey variables and the quota variables as though the sample frame had been stratified by these quota variables. Evidence of significant correlations would suggest that there is some merit to the process of establishing and reaching certain quotas. Following this, we will compute the associated design effects treating the quota variables as post-stratification variables. This should provide a way to index the potential benefits of some alternative quota designs.

In both exercises, comparisons will be made across qualitatively different types of geography. If substantively different outcomes are obtained for different types of geography, it may suggest that different quota strategies should be considered across different strata in a design. For this demonstration we will use data from the 2004 Michigan Travel Counts household travel inventory.

BACKGROUND ON THE 2004 MICHIGAN COUNTS TRAVEL INVENTORY

The 2004 Michigan Travel Counts travel inventory, sponsored by MDOT (Michigan Department of Transportation) consisted of 14,280 household travel diaries. A 48-hour (Monday through Thursday) location-based activity hybrid diary format was used. Data was collected for seven geographic strata. The original design called for a minimum of 2,040 households per stratum. Within each geographic stratum, quotas were monitored for household size, the number of household vehicles, and the number of workers in the household. The quotas reflected a proportionate allocation based on households for the cross of these three variables. U.S. Census PUMS data was used to establish the quotas.

CORRELATIONS BETWEEN SURVEY VARIABLES AND FRAME VARIABLES

Let's first look at some correlations between household size, the number of vehicles, and the number of workers on the one hand, and three survey variables on the other – the number of trips per household, the number of work trips per household, and the number of transit trips. The correlations computed for two strata – the SEMCOG (Southeast Michigan Council of Governments) region consisting of the City of Detroit and seven suburban counties, and the rural counties in the Upper Lower Peninsula region of the state – are presented in Table 2. These areas were selected because they represent two distinct types of geography – urban/suburban and rural. All of the correlations presented in Table 2 are significant at p<.0001 unless noted with an "*" where p<.001.

TABLE 2 Correlations Between Survey Variables and Post-Stratification Variables

Potential Post-Stratification Variables	Survey Measures		
	Number of Trips	Number of Work Trips	Number of Transit Trips
Urban/Suburban			
Persons per Household	0.724	0.310	n.s.
Vehicles per Household	0.392	0.444	-0.246
Workers per Household	0.467	0.692	-0.068 *
Rural			
Persons per Household	0.683	0.350	n.s.
Vehicles per Household	0.151	0.174	n.s.
Workers per Household	0.504	0.669	n.s.

Source: 2004 Michigan Travel Counts

As one would expect, two of the largest correlations occur between the number of trips and the number of work trips and the frame variables persons per household and workers per household, respectively. Moreover, this is consistent across the two types of geography.

As one would also expect, the number of work trips is most highly correlated with the number of workers per household across both types of geography. In the urban/suburban region, it is more highly correlated with the number of vehicles per household than it is with the number of persons per household. In the rural region, however, the reverse is true – the number of work trips is more highly correlated with the number of persons per household than it is with the number of vehicles per household.

In the urban/suburban stratum, the number of transit trips is most strongly related to the number of vehicles in a household, and only weakly correlated with the number of workers in a household. There appears to be no relationship with the number of people in a household. In the rural stratum, the number of transit trips is not related to any of our post-stratification variables. The results for transit trips probably reflect the greater number of transit options available in the urban/suburban stratum compared to the rural stratum.

DESIGN EFFECTS FOR SOME ALTERNATIVE DESIGNS

So, conventional wisdom is confirmed, all three of these frame variables are related to some key survey travel variables. This, of course, is a requirement for good stratification. However, just how much precision is gained by insuring that certain quotas are met with regards to these variables? If we consider our quota variables in terms of a post-stratification, we can estimate the gains in precision due to the stratification by calculating the design effects (*deff*) for various stratification options. This should provide even more insight into the efficacy of our overall survey process.

Design effects are presented in Table 3 on the next page for five hypothetical stratified designs – three designs where the stratification is based on each of the three frame variables individually, and two designs based on a pairing of two of the frame variables. Again, comparisons are made across the two types of geography.

The design effects in Table 3 were calculated as follows. The sampling variance calculated for each hypothetical stratified survey design was divided by the sampling variance assuming an SRS design. The sampling variance for an SRS design is simply the sample element variance divided by the sample size, multiplied by the finite population correction (*fpc*). (Because the *fpc* is the proportion of frame elements <u>not</u> sampled it can be ignored in the RDD portion of most household travel surveys.) The sampling variances for the stratified designs were calculated as the sum of the sampling variances for each cell in the design treated as an SRS weighted by the square of the population proportion (W_h) of each stratum.

$$deff = v_{st}(\bar{y}) / v_{srs}(\bar{y})$$

where

$$v_{srs}(\bar{y}) = s^2 / n$$

and

$$v_{st}(\bar{y}) = W_1^2 (s_1^2 / n_1) + W_2^2 (s_2^2 / n_2) + \ldots + W_h^2 (s_h^2 / n_h)$$

The calculation of the *deff* for the number of work trips in the first design (Urban/Suburban – Persons per Household) in Table 3 is outlined below.

Persons per Household	Sample Size	Number of Households	Population Proportion	W^2	Element Variances	Sampling Variances
1	682	450,324	24.4%	0.0595	2.375	0.0002
2	757	559,649	30.3%	0.0920	6.364	0.0008
3	459	330,348	17.9%	0.0320	6.508	0.0005
4+	636	505,091	27.4%	0.0749	6.740	0.0008
	2,534	1,845,412	100%			0.0022

$v_{srs}(\bar{y}) = 6.146 / 2{,}534 = 0.0024$ and

$deff = 0.0022 / 0.0024 = 0.919$

TABLE 3 Design Effects for Some Hypothetical Stratified Designs

Alternative Post-Stratified Designs	Survey Measures		
	Number of Trips	Number of Work Trips	Number of Transit Trips
Urban/Suburban			
Persons per Household 4 cells (1, 2, 3, 4+)	0.527	0.919	1.025
Vehicles per Household 4 cells (0, 1, 2, 3+)	0.769	0.630	0.982
Workers per Household 4 cells (0, 1, 2, 3+)	0.786	0.550	0.945
Persons by Vehicles 16 cells	0.548	0.851	1.021
Workers by Vehicles 16 cells	1.009	0.555	1.502
Rural			
Persons per Household 4 cells (1, 2, 3, 4+)	0.587	0.906	1.055
Vehicles per Household 4 cells (0, 1, 2, 3+)	0.830	0.679	1.410
Workers per Household 4 cells (0, 1, 2, 3+)	0.696	0.472	1.068
Persons by Vehicles 16 cells	0.660	0.809	1.161
Workers by Vehicles 16 cells	0.789	0.526	1.891

Source:2004 Michigan Travel Counts

There are significant design effects for the number of household trips due to stratification for nearly all of the alternative designs. The stratified design based on persons per household has a sampling variance about 53% of what it would be for a simple random sample of the same size in the urban/suburban region, and 59% in the rural region. These correspond to reductions in standard error of 27% and 23%, respectively.

Reduction in S.E. = 100 x (1 – SQRT(deff))

There are also significant design effects for the number of work trips due to stratification for all of the alternative designs. The reduction in standard errors is greatest for the workers per household stratified designs in both the urban/suburban region and the rural region. There, the reductions are 26% and 31%, respectively. Notice that the designs with more stratification – the workers by vehicles designs – do not show any gains in precision. In fact, the design effects for these designs suggest that they are somewhat less precise. This can be accounted for by the fact that design effects depend in part on the sample size in each stratum. As the number of strata increases, the sample sizes in some, if not all, strata will obviously decrease.

The design effects for the number of transit trips due to the various stratifications demonstrate that stratification need not lead to increases in precision for all survey variables. Only in the urban/suburban region would we expect to see increases in precision due to stratification since transit is not a significant mode option in the rural region. The increases we do see in the urban/suburban region are modest – a 3% reduction in standard error for the workers per household stratification and a 1% reduction for the vehicles per household stratification. With respect to some of the designs we actually see a reduction in precision compared to a simple random sample. Again, this is largely due to the small sample sizes associated with some of the strata in the proportionate allocation, and the lack of a significant correlation between the survey variable – transit trips – and the stratification variables.

The only design where the design effect suggests there is no gain due to stratification in estimating the number of trips is the number of workers by number of vehicles design in the urban/suburban region (*deff* = 1.009).

SUMMARY DISCUSSION

The method of treating the quota variables commonly used in household travel surveys as post-stratification variables and then calculating the design effects clearly demonstrates that some quota variables with modest correlations with important survey variables can lead to significant decreases in standard errors. For example, in the rural region the correlation between the number of household trips and the number of vehicles per household was only 0.151 (see Table 2). The corresponding design effect, however, was 0.83, which translates into a 9% reduction in standard error.

The same quota design does not need to be used in all geographic strata in a household travel survey. Moreover, the quota design used within a stratum should reflect the objectives of the survey. A design using the number of persons per household by the number of vehicles per household seems to work well for surveys focused on trip generation. However, a design using the number of workers per household by the number of vehicles per household might be preferable for surveys focused on work trip generation.

The size of a design effect depends in part on how large of a sample is drawn from each stratum. In a quota design, this suggests that one needs to be careful about the number of strata defined by the quota variables. If the sample allocation for some of the strata becomes too small, the designs may actually lead to increases in standard errors. This analysis suggests that all of the gains in precision due to the stratification could have been accomplished by simply establishing independent quotas for the frame variables. Establishing quotas based on the cross of the variables did not enhance all of the designs. Strategies for collapsing across strata reducing the overall number of cells in a design need further examination.

One area that clearly requires some additional analysis is the benefits to be gained by the use of other quota design variables such as household income. Another area would be to apply this type of analysis to a region where transit plays a larger role in overall travel patterns and planning than it does in the SEMCOG region reported here. While the SEMCOG region is served by an extensive bus system, there are no subways or commuter rail. Applying this analysis to a region with a more clearly defined central business district might also provide other insights.

Finally, an analysis incorporating survey costs is called for. In the typical household travel survey costs tend to increase at the end of data collection. This is largely due to the need to fill certain difficult quotas. What is needed is a method for associating measures of survey effort (e.g., survey completions per interviewer hour) with the design effects for various stratification strategies so that the economics of realizing potential reductions in standard errors due to stratification can be quantified.

Of course, at the end of the day, transportation planners and modelers use the travel inventory databases developed using survey methodology for a variety of tasks. The considerations presented here have focused on improving the survey quality of these databases. Planners and modelers may have other needs which impact sample design.

REFERENCES

Groves, R. M., F. J. Fowler, M. P. Couper, J. M. Lepkowski, E. Singer, and R. Tourangeau. *Survey Methodology*. John Wiley & Sons, Hoboken, 2004.

Kalton, G. *Introduction to Survey Sampling*. Sage Publications, Beverly Hills, 1983.