Travel Surveys: Moving from Tradition to Innovation  
Issues Paper on Sampling and Hard to Reach Populations

Table of Contents

I. Sampling ............................................................................................................................................... 1

A. How can sampling methodologies be refined to help ensure that all cells are within +/- 5%? .......... 1

B. What are the statistical significance and error tolerance that are acceptable or affordable for each step of the survey? ....................................................................................................................... 2

C. Should I plan for a geographically targeted sample based on mode accessibility (e.g. transit and walkability)? ................................................................................................................................ 2

D. What are the pros/cons of the various frames from which a random sample may be drawn? ........ 3

E. Given the wealth of research that suggests certain demographic groups respond poorly to HHTSs, how do DOTs/MPOs prioritize the wide range of demographic groups, so they go into the survey with a plan for how they will allocate resources to meet their goals................................................. 4

F. HHTs, like many other multi-dimensional surveys, have different data items that are not in the same level. For example, one may be able to get a good response on HH structure but a bad response for trips. Is it possible to design a survey to focus on more samples for essential data items and less samples for non-essential ones? Can a survey be designed that collects more samples on essential items and less non-essential items? ............................................................ 5

I. SAMPLING

A. How can sampling methodologies be refined to help ensure that all cells are within +/- 5%?

1. Chicago Regional Household Travel Inventory White Paper: Sampling Considerations (Englund, Vosha et al.)– this overview encompasses the following topics: Sample Quality; Stated Preference Survey; and Panel Survey. Recommendations are proposed for the following categories: Population; Sampling frame; Sampling Method; Sample Stratification; Sample Size; Sampling Quality; and Participation.

2. Sample Size Requirements for Multi-Day Travel Surveys: Some Findings (Stopher, Kockelman et al. 2008) - Using GPS-based PKT data collected from a random sample of households, this paper demonstrates how 7-day survey data may well diminish needed sample sizes by 65 percent (relative to a conventional, one-day diary
survey), and a 15-day duration (which may be optimal for a number of reasons) would reduce sample sizes by over 70 percent, in the case of a focus on PKT. Further, we argue that, while GPS surveys are still quite a bit more expensive to conduct on a per household basis than a conventional one-day survey using a diary method, the reduction in the required sample size results in significant cost savings if a GPS survey is used with the sample size that is necessary statistically, and that, even a substantial increase in sample size can be obtained without exceeding the costs of a conventional survey.

B. What are the statistical significance and error tolerance that are acceptable or affordable for each step of the survey?

1. Workshop on Quality Indicators (Kalifs, Meurs et al.)- The objective of this paper is to develop a number of indirect quality indicators for survey data, especially travel and activity diaries. The indicators are derived from a definition of quality as well as on the basis of results from studies on the factors that cause errors in transport survey data. This paper focuses on the less understood areas of nonresponse and measurement error. It will go on to discuss a definition of quality, nonresponse error, measurement error, meta-analysis and a set of proposed quality indicators for meta-analysis.

C. Should I plan for a geographically targeted sample based on mode accessibility (e.g. transit and walkability)?

1. Accessible Settlements Study for Greater Nottingham (Council, Council et al. 2010)- this government transport study in Great Britain establishes a consistent approach to accessibility measurement for all the settlements in the area. A form of weighting given to the facilities, for example, essential facilities such as doctors’ surgeries, Post Offices to be given more weight than access to leisure facilities, and more frequent travelling more than occasional visits. Also, the study establishes measures that represent a scale of accessibility for the facilities and services, and a travel time threshold that represents what people would consider to be a reasonable travel time by sustainable means. They consist of the accessibility (usually related to travelling time) to the nearest facility using walking, cycling or the public transport network (bus/heavy rail/light rail).

2. Household Travel Surveys in Focused Geographic Subareas of the Region (Griffiths 2012) - the project objective for this National Capital Region Transportation Planning Board work program was to “…collect additional household travel survey data in smaller geographic subareas of the region for the purpose of (1) permitting more intensive analysis of differences in day travel behavior in a wide variety of communities with varying densities, physical characteristics, demographic characteristics, and transportation options, (2) assisting local planners with current land use and transportation planning efforts and to provide information and illustrative examples that can also be used by all local jurisdictions in their future planning, (3) building a household travel survey database that can be used to measure...
changes over time in local community travel behavior resulting from major transportation improvements and/or significant shifts in the pattern of new development in these local communities.” At the conclusion of the Arlington County supplemental survey effort, planners in other TPB jurisdictions expressed similar needs for HHTS data for smaller geographic areas to support their local planning efforts. Further recommendations by the TPB Technical Committee called for the Unified Planning Work Program to include a Geographically-Focused Household Travel Surveys in three to seven areas of the region each fiscal year depending available funding.

D. What are the pros/cons of the various frames from which a random sample may be drawn?

1. **Assessing Sample Bias and Establishing Standardized Procedures for Weighting and Expansion of Travel Survey Data** (Nilufar 2003)- This review has focused specifically on the extent of bias observed in past survey data and highlighted the ambiguity that exists in bias assessment processes. In an effort to reduce bias from the sampled survey data this study has identified some specific design features of the past survey instruments that might be held responsible for low responses and biased sampled data. Finally, this study recommends weighting and expansion techniques to adjust bias in the sampled data as a final step in the data analysis process.

2. **Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples** (Yeager, Krosnick et al. 2011)- This study assessed the accuracy of telephone and Internet surveys of probability samples and Internet surveys of non-probability samples of American adults by comparing aggregate survey results against benchmarks. The probability sample surveys were consistently more accurate than the non-probability sample surveys, even after post-stratification with demographics. The non-probability sample survey measurements were much more variable in their accuracy, both across measures within a single survey and across surveys with a single measure. Post-stratification improved the overall accuracy of some of the nonprobability sample surveys but decreased the overall accuracy of others.

3. **Improving Accuracy in Household and External Travel Surveys** (Pearson, Hard et al. 2010)The Texas Department of Transportation has a comprehensive on-going travel survey program, this research examines areas within two select travel surveys concerning quality control issues involved in data collection and sampling error in the data caused by various assumptions, survey methods, and issues such as non-response. Quality control issues, sampling errors, and non-response in external and household travel surveys conducted in Texas are identified, examined, and evaluated. The state-of-the-practice in these types of surveys relative to quality control during and after the surveys are conducted and how sampling errors and non-response are treated (or corrected) in the survey analysis are reviewed and documented. The results are assessed to formulate a set of recommendations for incorporating into survey designs for the travel survey program.

4. **Cell Phone-Only Households and Problems of Differential Nonresponse Using an Addressed-Based Sampling Design** (Link and Lai 2011)- Identifying and contacting cell-phone-only (CPO) households in an efficient and cost-effective manner is critical
for many studies to ensure the representativeness of final survey estimates. With the advent of address-based sampling (ABS)—that is, the sampling of addresses from a database with near universal coverage of residential households—researchers have an alternative to cell-phone frame sampling for identifying and surveying these households. Using a large ABS mixed-mode study, we explore three key questions. The study shows that, while CPO homes can be reached and interviewed using an ABS design, it is important to assess and correct for the potential biasing effects of differential nonresponse.

E. Given the wealth of research that suggests certain demographic groups respond poorly to HHTSs, how do DOTs/MPOs prioritize the wide range of demographic groups, so they go into the survey with a plan for how they will allocate resources to meet their goals.

1. *Surveying Hard to Reach Groups* (Riandey and Quaglia 2008)- Surveying hard to reach groups is difficult but necessary for preventing a selection effect and biased sampling. The topic may be analyzed as five challenges concerning people missing from the sampling frames; people out of home during hours or days of data collection; people refusing contact; people refusing to answer the questionnaire; people who cannot answer to the interviewer.

2. *Respondent driven sampling – where we are and where should we be going?* (White, Lansky, et al. 2012) - In this editorial, the authors highlight that ‘RDS’ includes both data collection and statistical inference methods, discuss the limitations of current RDS inference methods for generating representative estimates, highlight other applications of RDS for which it may be more reliable, propose and request feedback on a draft RDS reporting check-list, and finally suggest priority areas for RDS research. The authors call for a clearer distinction between the methods used for RDS sampling and the methods used for statistical inference, as well as a systematic reporting of RDS studies.

Additional papers related to RDS include (and can be accessed here): Assessing respondent-driven sampling (Goel and Salganik 2010) [http://5harad.com/papers/assessing-rds.pdf](http://5harad.com/papers/assessing-rds.pdf) and supporting information ([http://5harad.com/papers/assessing-rds-si.pdf](http://5harad.com/papers/assessing-rds-si.pdf))

F. HHTs, like many other multi-dimensional surveys, have different data items that are not in the same level. For example, one may be able to get a good response on HH structure but a bad response for trips. Is it possible to design a survey to focus on more samples for essential data items and less samples for non-essential ones? Can a survey be designed that collects more samples on essential items and less non-essential items?

1. Multilevel Modeling of Refusal and Non-Contact in Household Surveys: Evidence from Six UK Government Surveys (Durrant and Steele 2009) – The authors analyzed household unit non-response in six major UK Government surveys by using a multilevel multinomial modeling approach. The models are guided by current conceptual frameworks and theories of survey participation. One key feature of the analysis is the investigation of the extent to which effects of household characteristics are survey specific. The analysis is based on the 2001 UK Census Link Study, which is a unique source of data containing an unusually rich set of auxiliary variables. The study contains the response outcome of six surveys, linked to census data and interviewer observations for both respondents and non-respondents.

Endnotes


Griffiths, Robert E. (2012). Household Travel Surveys in Focused Geographic Subareas of the Region. Memorandum to Transportation Planning Board, National Capital Region Transportation Planning Board. May 9, 2012


White, R.G., A. Lansky, et al. (2012) Respondent driven sampling - where we are and where should we be going? Sex Transm Infect, Oct. 2012, vol. 88, no.6 Published by group.bmj.com


Preferred Citation: Issues Paper on Sampling and Hard to Reach Populations. Household Travel Survey Symposium: From Tradition to Innovation (Dallas: Texas A&M Transportation Institute, 2012)