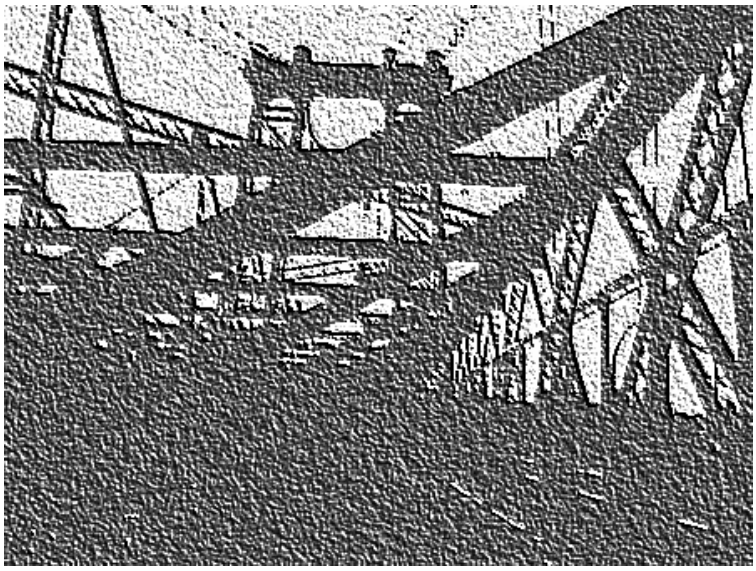
Household Travel Survey Research

Final Document



Sponsored by:
New York Metropolitan Transportation Council
Metropolitan Transportation Authority
University Transportation Research Center

Project #: Z-05-01

February 28, 2007

University at Albany
Dr. Catherine T. Lawson, PhD
Christine W. Fassman
Maria Y.C. Chau

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle	5. Report Date
Household Travel Survey Research	February 28, 2007
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Dr. Catherine Lawson, Associate Professor, Christine W. Fassman, and Maria Y.C. Chau	RF#49777-15-17

9. Performing Organization Name and Address	10. Work Unit No.
Geography & Planning, University at Albany, 1400 Washington, Albany, New York 12222 Region 2 – University Transportation Research Center, New York, NY 10031	11. Contract or Grant No.

12. Sponsoring Agency Name and Address		13. Type of Report and Period Covered
New York Metropolitan Transportation Council 199 Water Street New York, NY 10038	Federal Highway Administration U.S. Department of Transportation Washington, D.C.	Final Report Document 3/2006-3/2007
		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The reviews in this section are based on documents produced for the "1997-1998 Regional Transportation Household Interview Survey" (RT-HIS). This surveying effort was sponsored by the New York Metropolitan Transportation Council (NYMTC) and the North Jersey Transportation Planning Authority (NJTPA), two federally sanctioned Metropolitan Planning Organizations (MPOs) in the New York - New Jersey - Connecticut metropolitan area. The actual deployment took place between February 1997 and May 1998, under the direction of NuStats International of Austin, Texas, under subcontract to Parsons Brinckerhoff. In addition, NuStats employed Macro International, Ebony Marketing, and DBM Marketing during the data collection phase of the project.

The stated purpose of the data collection efforts was to provide primary data for use in a transportation planning model for the New York / New Jersey / Connecticut metropolitan area. These data were collected to consist of the essential elements in the "Transportation Models and Data Initiative Project". The RT-HIS used a household travel survey methodology that relied on the willingness of local residents to complete a travel diary over a 24-hour period. As with any surveying approach the documentation of the process in addition to the final output can be critical information for learning lessons and improving future surveying strategies. Where necessary, individuals or groups involved in the RT-HIS will be interviewed to add depth to the understanding of the identified issues and/or to determine if other previously undocumented issues need to be addressed.

17. Key Words	18. Distribution Statement
Travel survey, Household interview, Data collection	

19. Security Classif (of this report)	20. Security Classif. (of this page)	21. No of Pages	22. Price
Unclassified	Unclassified	43	

Form DOT F 1700.7 (8-69)

Table of Contents

Table of Contents

SECTION OVERVIEW

Part One: HOUSEHOLD TRAVEL SURVEY REVIEW relies on the various documents produced by the firms involved in the surveying effort to identify problems through the surveying process. Issues revealed in these documents are cited and discussed, with special attention paid to the possible consequences of the identified issues.

Part Two: COSTA RICA CONFERENCE (2004) DOCUMENT REVIEW summarizes the *Costa Rica Household Travel Survey Conference* and other research that examines or recommends alternative approaches or strategies dealing with those identified issues.

Part Three: HOUSEHOLD TRAVEL SURVEY SYNTHESIS matches the issues identified with potential strategies from the document review in Part Two. Matrix One developed in this section illustrates various combinations.

Part Four: DISCUSSION – REVEALED THEMES summarizes the major findings in the surveying process effort and recommends strategies which address these findings. Matrix Two developed in this section outlines the approach for each theme.

PART ONE

HOUSEHOLD TRAVEL SURVEY REVIEW

Introduction

The reviews in this section is based on documents produced for the "1997-1998 Regional Transportation Household Interview Survey" (RT-HIS). This surveying effort was sponsored by the *New York Metropolitan Transportation Council* (NYMTC) and the *North Jersey Transportation Planning Authority* (NJTPA), two federally sanctioned Metropolitan Planning Organizations (MPOs) in the New York - New Jersey - Connecticut metropolitan area. The actual deployment took place between February 1997 and May 1998, under the direction of *NuStats International* of Austin, Texas, under subcontract to *Parsons Brinckerhoff*. In addition, NuStats employed *Macro International, Ebony Marketing,* and *DBM Marketing* during the data collection phase of the project.

The stated purpose of the data collection efforts was to provide primary data for use in a transportation planning model for the New York / New Jersey / Connecticut metropolitan area. These data were collected to consist of the essential elements in the "Transportation Models and Data Initiative Project". The RT-HIS used a household travel survey methodology that relied on the willingness of local residents to complete a travel diary over a 24-hour period. As with any surveying approach the documentation of the process in addition to the final output can be critical information for learning lessons and improving future surveying strategies. Where necessary, individuals or groups involved in the RT-HIS will be interviewed to add depth to the understanding of the identified issues and/or to determine if other previously undocumented issues need to be addressed.

Identified Issues

ISSUE 1: Cooperation and coordination during data collection phase of all parties

Reference: *RT-HIS M & I: page 2*

Evidence: Special attention was given to the sampling design – including the explicit use of a methodology developed by Cambridge Systematics – "mode leadership densities". It intended to capture differences in mode utilization within the various residential density groups within counties. In other words, specific geographical characteristics (defined as density) formed the basis for the sampling frame. In the actual deployment, only 22 of the 28 counties used these density groupings. This occurred because six counties joined in the data collection effort after the sampling plan was completed.

The documentation does not include a reference on the research used in making the decision to use "density groupings". Nor is there a post-analysis to indicate what effect not using these groupings may have had on the outcome or quality of the data. It is not known is whether these six counties are more or less likely to have benefited from using the mode leadership approach. A key to avoiding future loss of survey strategies is making an extra effort to include all possible counties – or at least to evaluate the value lost by late comers as a result there was no indication of the possible loss to the data collection outcomes.

ISSUE 2: Weekend Surveys

Reference: *RT-HIS M & I: page 2*

Evidence: Weekend samples consisted of only 323 households – all located in New Jersey.

No explanation is indicated as to why the sampling frame did not include Saturdays and Sundays – with the exception of New Jersey households. This is particularly troublesome as many of the congestion concerns occur during the weekends.

Modeling efforts does not support weekend travel using this data. (Note: Use of NPTS to derive trips – see *RT-HIS Comparative Analysis Weekday and Weekend Travel with NPTS Integration*).

ISSUE 3: Adding participants after the sampling decisions are made

Reference: *RT-HIS M & I: page 3*

Evidence: In determining the effective sample size, the initial calculation assumed no increase in effective sample size – yet six counties were added to the surveying effort.

It is noted in the document that this is a conservative assumption, given that the effective sample size would "actually be a few hundred higher due to the contributions of the samples from these counties." This suggests the need to make an extra effort to include all possible participants as soon as possible.

It is not clear what impact including additional counties may have had. The *Recommended Survey Allocation Plan for Weekdays* in Table 1 displayed (1) the original proposed mode leadership-based plan, (2) additional requirement for county-level estimates and (3) the total weekday sample. (*8,104*, *2,846*, and *10,950*, respectively) At least for weekday travel activity it appears at least 160 households were sampled in each county.

ISSUE 4: Use of "Mode Leadership Methodology"

Reference: *RT-HIS M & I: page 4*

Evidence: Table 2 needs more explanation – why were the modes assembled in relation to densities in this way?

Is there research supporting the decision to place the modes in these densities? For all trip purposes, it is unclear why these modes were used. If this is only for the commute trip, should that have been factored into the decision?

Unable to determine if this caused problems at a later stage in the process.

ISSUE 5: Adding 30% excess sample

Reference: *RT-HIS M & I: page 5*

Evidence: The 30% excess sample is tied to which groupings? It is unclear what effect this decision may have on the representative-ness of the sample.

ISSUE 6: Final sample composition

Reference: *RT-HIS M & I: page 6*

Evidence: The sample objective is described with the hierarchy of the total number of households by:

- Specific county
- Specified mode leadership density

BUT not for each mode leadership density within a specified county – however, Kings County completed 157% of its original goal due to sample management error.

ISSUE 7: Order of operations on sampling

Reference: *RT-HIS M & I: page 6*

Evidence: *Mode leadership densities* were assigned based on the census tract variable "house address" and was assigned prior to fielding. When the sample was released and the household contacted, the household no longer lived in that particular type of place.

Since the order of operations resulted in an "after the fact" situation – the actual mode leadership density distribution did not follow the expected goals.

ISSUE 8: Impact on Hudson County

Reference: *RT-HIS M & I: page 7*

Evidence: *Mode leadership density 2* was only associated with Hudson County – and was thus an example of an extreme over sampling of the population. The goal was later reduced as it could not be met.

Determined to not adversely affect any modeling plans.

ISSUE 9: Overall implications of Table 6

Reference: *RT-HIS M & I: page 8*

Evidence: The relationship between the mode leadership density goals and the actual sample range from 47% to 320%.

Was there any expectation that household would actually use these modes?

It is unclear how mode leadership density's missed goals will impact the use of the data – and what the real relationship of the mode is to what the household actually uses to travel.

ISSUE 10: Using NPTS for weekend trips

Reference: *RT-HIS M & I: page 10*

Evidence: The information in Table 7 shows that Travel Day Distribution for Sunday was 1.3% and for Saturday was 1.3% (all households in New Jersey). The document states that the 1995 NPTS is to be used for that portion of the weekend travel analysis – however, the 1995 NPTS sampling frame was not based on the mode leadership method.

If the NPTS is sufficient for weekend travel which is more likely more complex than weekday travel, what does this mean for the use of the NPTS for the whole effort? If the mode leadership is not important to weekend travel, how is it justified for the surveying effort? (Note: Use of NPTS to derive trips – see *RT-HIS Comparative Analysis Weekday and Weekend Travel with NPTS Integration*).

It is unclear how this decision was made.

ISSUE 11: Weight Calculations for phones

Reference: *RT-HIS M & I: pages 11 and 12*

Evidence: The study area was not homogenous with respect to phone ownership turnover and rates of working phones within the home. For example, 14% of the sampled households had more than one working phone.

A weighting factor was created to account for multiple numbers per household. What is unclear is whether the number of phones in a household impacts travel behaviors – This question also needs to be answered for households that share one phone number. Is a simple factor an appropriate response? Episodic phone ownership where it was determined using the *Current Population Survey* (CPS) that 5.3% of the population were non-phone households. This was further adapted with anecdotal information, but no real research.

Factor application?

ISSUE 12: Another set of weights used 1990 Census to adj. household size, vehicles & income

Reference: *RT-HIS M & I: page 17 and 18*

Evidence: Tables 16, 17 and 18 show the magnitude of the adjustments to the 1990 Census. There are questions about the 1990 Census when compared to the 2000 Census – these characteristics should be checked for potential vulnerability to error as these populations are also hard to locate for the Census.

ISSUE 13: The final retrieval rate of only 26%

Reference: *RT-HIS M & I: pages 19 and 20*

Evidence: The total number of sample pieces was *52,390*. *9,454* were ineligible, *28,495* were not recruited but eligible samples – with *14,441* recruited samples. What was the spatial distribution of the unrecruited samples and was there a systematic bias or correlation with other important factors. *14,441/42,936 = 33.6%* as a recruitment rate. Of these *14,441*, *11,264* provided survey responses – 78.2%. The overall response rate was then calculated by multiplying the recruitment rate by the retrieved rate – yielding 26%. It is unclear how the distribution of these responses supports the underlying mode leadership densities.

ISSUE 14: Lessons learned from 1995 Pilot

Reference: *RT-HIS M & I: page 23*

Evidence: Early warnings regarding transit data collection: (1) the first boarding location for all transit trips should be highlighted, to ensure that all transit users provide this information. (2) Data elements with higher rates of non-response included ethnicity; cost to park at work; personal cost of transit use after transit subsidies; employer subsidies for taxis.

ISSUE 15: Change in Sample Goals

Reference: *RT-HIS M & I: page 25*

Evidence: The sample goals for individual mode leadership densities by county were determined by the Nustats staff to be too costly. However, the overall sampling goal remained at *11,199*.

ISSUE 16: Use of paper and pencil technology for retrieval interview

Reference: *RT-HIS M & I: page 26*

Evidence: The use of paper and pencil resulted in more time needed to integrate information into the system.

ISSUE 17: Human errors

Reference: *RT-HIS M & I: page 26*

Evidence: Human errors identified during the edit checking process included misreported times, missing trips, incorrect modes, geocoding errors, and other problems. It is not clear who was directly responsible for these errors, however, problem corrections were attempted.

ISSUE 18: Appropriateness of first question

Reference: *RT-HIS M & I: page 28*

Evidence: The assumption that respondents were sleeping at 3:00 AM should have been the default for the first activity as it seems inappropriate to treat this time frame in the same manner as the activities after 6:00 AM when normal activities begin.

A review of "*first questions*" should be conducted to look for a better approach to the initial start of survey sections.

ISSUE 19: Correction calls

Reference: *RT-HIS M & I: page 28*

Evidence: 1/3 of all households were called for corrections, with more need for these calls at the beginning of data collection stages.

ISSUE 20: Use of paper and sample management procedures

Reference: *RT-HIS M & I: page 28*

Evidence: The CATI system had a unique file format that was difficult to use. The use of paper and pencil for retrieval resulted in this information being excluded from the management system.

ISSUE 21: Geocoding problems

Reference: *RT-HIS M & I: page 32*

Evidence: A manual review was required for records associated with transit trips, out of area locations, and abnormal speeds.

ISSUE 22: Non-response for Income Data

Reference: *RT-HIS M & I: page 33 and 34*

Evidence: The key variable – household income – had a non-response rate of 23.6%.

Discussion: The suggested treatment of missing income data is to either eliminate the household or attempt to impute income. Table 21 lists Hensher's recommendations for imputing data.

ISSUE 23: Missing Secondary Work Location

Reference: *RT-HIS M & I: page 36*

Evidence: 18% of the secondary work locations were classified as non-response.

ISSUE 24: Audit Flags

Reference: *RT-HIS DATAUSER: page 8*

Evidence: Transit data was deeded to be misreported or misrecorded - requiring manual re-reporting of the data. This suggests a problem with the survey form itself with respect to transit sequencing. It is unclear from the description of the flagging whether the problems were systematic or random.

An examination of PASS1UPD and PASS2UPD – in comparison with the original retrieval information should reveal the nature of the problems –

ISSUE 25: Data Usage Flags

Reference: *RT-HIS DATAUSER: page 8*

Evidence: Flags were attached to the dataset where a speed violation occurred, corrections were made or the "responder-reported data was inconsistent with modeling expectations".

This suggests that the model is more reliable than the respondent. A determination of any systematic problems would be helpful. For example, the reference to speed violations due to rounding suggests a better way to get the information from the respondent.

ISSUE 26: Extent of problem with speed violations

Reference: *RT-HIS DATAUSER: page 11*

Evidence: According to Table 3, of the households with data user flags, 53.2% were associated with speed violations records, while 38.6% were associated with speed violations due to rounding – suggesting a large component with respect to households.

The extent of this problem may need to be revisited as it is unclear what the percentage of households really represents. The two types of problems with respect to households suggest 91.8% of the flagged households had some type of speed violation.

ISSUE 27: Missing zip codes

Reference: *RT-HIS DATAUSER: page 12*

Evidence: It was noted that respondents were not able to provide the zip code for many of their destinations.

This problem can be resolved by post-processing the GIS information and appending the zip code information electronically. It is unlikely that the respondent population typically know the zip codes of their destinations.

ISSUE 28: Inaccuracies in the location of places

Reference: *RT-HIS DATAUSER: page 12*

Evidence: Survey participants were unable to provide accurate location information.

New technologies, such as Google Earth or MapQuest, may aid in the retrieval process – as addresses could be entered in "real time" and verification could be confirmed by the survey respondent.

PART TWO

COSTA RICA CONFERENCE (2004) DOCUMENT REVIEW

Introduction

The principle theme of the Costa Rica conference was to review and further research recommendations on standards for travel research surveys and research needs in emerging travel survey issues. The following is a summary of the papers presented and the workshops attended at the conference. To date, the research team is unaware of any updates or reviews to these findings.

Handling Individual Specific Availability of Alternatives in State Choice Experiments

John M. Rose
David A. Hensher
Institute of Transport Studies
Faculty of Economics and Business
The University of Sydney
NSW 2006 Australia

Current travel surveys, (Stated Choice Experiments – "SC") allow for one or more alternatives provided in the choice sets that occur within the sampling population equal number of times. However this practice, while acceptable does not account for the non-availability of one or more labeled alternatives for each individual. This workshop/paper discusses a way to account for "observation specific non-availability of alternatives in the design of stated choice experiments and illustrates the method in the context of choice on main and access mode for commuting and non-commuting trips in the north-west sector of Sydney."

A limited number of alternatives in SC studies results in an analysis that results in an uneven distribution of alternatives across the given population. To rectify this and to create a more realistic picture of the circumstances, "the choice sets offered to the individual should reflect the specific context or situation faced or likely to be faced". This remedy however is difficult since most studies or surveys conducted are for a single respondent, despite their individual choices or alternatives available.

Other research that contributes to this argument include: *Bastell and Polking* (1985), *Raghovarao and Wiley* (1986), and *Anderson and Wiley* (1996). However research to date regarding the availability of alternative problems addresses "(1) all alternatives vary in their availability and (2) some alternatives vary whilst others are fixed in presence across all choice sets." It does not address the household constraints to the limitations of alternatives.

This research paper constructs SC experiments that have alternatives within the choice sets that are respondent specific and account for the individual differences likely to exist within real markets.

Findings of the Study

- "Strategies require in-depth probing of respondents prior to the SC experiment commencing."
- "Alternatives shown in the SC experiment are then tailored to the individual given the specific decision context and situation which exists or is likely to exist in the future."
- "Respondents are capable of completing complex choice tasks in a meaningful manner when the choice task presented are realistically framed, and hence, cognitively meaningful to the individual."
- "There is a need for further evaluation of task complexity"
- "Respondents are capable of meaningful participation independent of the design dimensionality… what counts is the believability of the choice task and the relevancy of the alternatives, attributes and attribute levels shown to the respondents." "Hence despite the complexity of the design, respondents are more capable of undertaking choice tasks if the tasks are realistically portrayed."
- In summary – tailoring the choice alternatives for various respondents that account for household limitations and real alternatives makes for a better and more accurate survey that respondents are capable and willing to complete despite preconceived notions that compensation and complexity decrease the response rate of surveys.

Quality Assessment

Resource Paper for Workshop A6
International Conference on Survey Methods in Transport
Costa Rica, August 2004

Peter Bonsall
Institute for Transport Studies, University of Leeds
Leeds, LS16 5NF, UK
00 44 113 3435335
pbonsall@it.leeds.ac.uk

At the previous transportation conference in South Africa in 2001, there was discussion that standards in survey design would improve the quality of travel surveys. However slow progress may be an indication that there is "that the imposition of inappropriate standards would bring additional costs and would an unwanted impediment to innovation."

The travel survey community should recognize that implementing and adhering to standards does not necessary indicate quality. The quality of a survey depends on the context of the situation and survey, a universal indicator of survey quality is not sought.

However this report suggests that: "efforts are needed to encourage survey practitioners and authors to provide more details of their data collection methods and that reviewers of academics papers should be encouraged to insist that authors provide full details of their data sources."

Conference organizers at both Costa Rica (2004) and South Africa agree that the following quality and standards issues should be analyzed:

- Developing a definition of data quality;
- Suggesting means by which the achievement of "quality" might be measured
- Determining the achievability, or otherwise, of universal measures of quality; and
- Assessing the potential role of benchmarking standards promoted by the International Standards Organization (ISO) or similar body

Processing, Analysis and Archiving of Travel Survey Data

Resource Paper for Workshop A5
7th International Conference on Travel Survey Methods
Costa Rica, August 2004

Gerd Sammer
Institute for Transport Studies
Department for Spatial, Landscape and Infrastructure Sciences
University of Natural Resources and for Applied Life Sciences
Peter Jordan Strasse 82, A1190 Vienna, Austria
gerd.sammer@boku.ac.at

This paper analyzes and identifies the travel survey activities at the state-of-the-art-level, to ensure the quality of data processing and data analysis. It also proposes quality standards in terms of guidelines for maintaining and archiving the results and the original data. The goal of this paper is to outline a standardized terminology of steps of a travel survey that serves as a basis for appropriate data documentation and archiving. It addresses: data base building, questionnaire editing, data entry, standardization for flexible coding, geo-coding, data editing, and cleaning and data weighting.

Suggested Main Seven Steps to Survey Design and Execution

1. Survey Design
2. Sample Design
3. Survey Instrument Design
4. Survey Implementation (Execution)
5. Data Processing
    a. Database building
    b. Questionnaire editing
    c. Coding
    d. Data editing/cleaning
    e. Weighting/correction
    f. Expansion
6. Data Analysis
    a. Validation
    b. Description and Explanatory Data Analysis
    c. Presentation of Results
7. Data Documentation and Archiving
    a. Documentation
    b. Archiving

The focus of this paper is on Data Processing and Data Documentation and Archiving

1. Data Processing

1.1 Data base building – includes the selection of the type of data structure, which is appropriate for the data collected and the software for the database system. It must determine the requirements for the database and the software in which the data will be analyzed and defined. (Helpful to consider requirements for archiving also)

Two types of data structures used for travels surveys (Richardson, et.al. 1995)

1. Flat File Database Structure
2. Related Database Structures

1.2 Questionnaire Editing – editing of completed questionnaires. Before the questionnaire goes to the coding procedure, the data needs to be checked for their completeness, consistency, and plausibility. This should consist of two stages (Richardson 1995): (1) Interviewer and (2) Supervisor editing.
If there is incomplete information the interviewer can contact the participant and obtain the complete information. The supervisor editing is important for quality control purposes.

1.3 Coding and Data Entry – fast development of methods can be observed – typically done through computer data entry

1.3.1 Coding Execution - coding should be completed as soon as possible so the information is fresh in the interviewers mind, rather than waiting a time period

1.3.2 Missing Values and Use of Zero it is important that missing data is flagged in the same way throughout the entire survey. At the 10th International Conference on Travel Behavior Research in Lucerne Switzerland, a paper was presented which discusses a proposal for some standardization to overcome those problems (Stopher et al 2003). They are as follows:

- "No blanks" standard – blanks are not a legit code and all data fields should contain an alphanumeric data
- "Missing data" standard – must be flagged by a specific code
- Correspondence between numeric values and the codes standard;
- Code standards for binary variables

1.3.3 Standardization for Flexible Coding of Complex Variables
It would be helpful to define an international standard for the values used which would enhance the comparability of surveys.

1.3.4 Geocoding
This method is state of the art for coding addresses of origin and destination.
For local travel surveys the following methods of geocoding are generally used:

- Geocoding of full street address
- Geocoding of cross street address
- Landmarks
- Geocoding by sampling
- For Long Distance Travel Surveys
- Where spatial aggregation level needed is not as accurate

1.4 Data Editing Cleaning – (there is not a standardized method for this step)
The currently employed methods and terminology include: (1) data editing, (2) cleaning, (3) correction, (4) data correction (which is also used in the weighing procedure).

This section suggests that the following should be added to this step: (5) checking of the range error, (6) logical consistency and (7) missing data, re-contacting interviewees if data is missing.

1.4.1 Errors and their treatment
Five main types of bias can occur. (Richardson et. al. 1995) The errors are from various sources such as: sample drawing, respondent, interviewer, data coder/typist.

- "*Sample drawing biases* which affects a deviation in the principle of random sampling and results in a biased coverage of the population"
- "*Range Error* which is a result of typing and recording where the code value is outside of the permissible code for that response - double coding the questionnaire can help solver this problem"
- "*Logical consistency errors* occur if responses to different questions of one or several persons linked together give inconsistent results"
- "*Missing Data* can be caused by all party's"
- "*Coding error* which results in a false value, which is not outside the range of codes permissible for that response, can only be identified by double coding."

The editing and cleaning process must be documented in a qualitative and quantitative way.

1.4.2 Missing Data

"Unit non-response" is the failure of the respondent to answer the question.
"Item non-response" is the failure of the respondent to answer part of the question.
There are three methods suggested to deal with the missing information.

- Imputation by re-contacting the respondent
- Imputation of the missing information by logical consistency checks from other information
- Imputation of missing data based on associative rules or other procedures

1.5 Data Weighting – [is] "necessary when the sample is stratified or any disproportional sampling procedure is used. From a quality point of view the following questions are of interest:"

- "Under which circumstances is weighting of the data required?"
- "Which method of weighting is appropriate for the biased data?"

Weighting is related to key variables: household, person, number of journeys, distance traveled, trips traveled, travel time.

1.5.1 Criteria of Weighting

No weighting is required if statistical tests guarantee that no significant bias can be identified for the following proposed key variable:

- Distribution of household size and car ownership rate in relation to the number of cars
- Distribution of person age, gender and occupation, nationally etc.
- Regional distribution and distribution of the set dates of the survey
- Travel behavior between respondents and non respondents characterized by the number of journeys, trips, distance traveled, travel time budget etc.
- Item non-response

1.5.2 Weighting Methods

1.5.3 Efficiency of Weighting

2. Data Analysis

2.1 Validation Analysis - verifies the accuracy of the surveyed data and is one of the most important steps to the survey. A validation check is performed before the data is cleaned, weighted, and an expanded data set is ready for final use. This check can be performed by two different organizations.

- the organization that is responsible for the survey project
- and independent organization

The validation check must take into account the specific situation and framework of the travel survey.

2.2 Data Presentation

3. Data Documentation and Archiving

3.1 Status Quo Analysis

Currently there is no focus on the preservation of documentation of travel survey research and information for the scientific audience. Some of the reasons for this lack of documentation and awareness are as follows:

- the lack of awareness by public planning agencies
- the unwillingness of planning agencies, especially regional ones, to finance archiving surveys
- the absence of appropriate national organization in many countries
- the absence of well understood and adapted patterns

3.2 Objectives and Requirements on Documentation and Preservation

"The objectives and requirements have to be oriented toward the involved organization, persons or potential users."

3.3 Technology of Metadata and Data Archives on the Market

Data Documentation Initiatives is the most advanced, flexible, and user friendly metadata standard – "DDI pursues the goal of creating an internationally accepted methodology for the content, presentation, transfer, and preservation of metadata about datasets in the behavioral and social field of sciences. It uses the XML language (extendible Markup Language)…"

"NESSTAR *Networked Social Science Tools and Resources* is a web-based data analysis software which enables users and publishers of social data to exploit data and information via the web including survey data, online tabulations and relevant other information."

- "NESSTAR publisher is an advanced data managements suite that enables to convert enhance and manage data for publication to the NESSTAR Server"
- "NESSTAR server allows the data provider to set up an interactive data publishing and dissemination service."
- NESSTAR Explorer produces and integrated data discovery, browsing and retrieval platform, which runs on the users desktop and which interacts with the data on a server.
- "NESSTAR light allows the user to operate most of the services of the explorer but within the confinement of a standard web browser."

The DDI Codebook still needs a standardized codebook for transport surveys.

It has been available to the public since December 2003. (http://chi.fg.uni-mb/elmis/)

3.4 Future Needs in Data Archiving

There are both organizational needs and standardization needs in the archiving of travel research data.

3.4.1 Organizational Needs

The author suggests the formation of something called a "Charter for Transport Metadata" which would organize the metadata needs for transport surveys globally.

3.4.2 Standardization needs for Documentation and Archiving

The author also suggests that standardization should only define a minimum level. "This minimum level should be based on the identification of general essential elements and characteristics that are common to all surveys and which are accepted in the transport community.

- "Standardization of a typology of travel surveys."
- "Development of vocabularies for metadata documentation of transport data or a transport thesaurus."
- "Development of standardized flexible coding scheme"
- "Enhancement of metadata standard to accommodate the spatial dimension of data"

In conclusion, the author argues that creating international standards for *archiving, documentation,* and *metadata* would greatly increase the quality of the data as well as the utilization of data across all organizations and institutions.

Processing, Analysis and Archiving of Travel Survey Data

Prepared for the 5th International Conference on
Survey Methods to be held in Costa Rica 2004

David Levinson and Ewa Zofka
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455
Levin031@umn.edu
+01 612 625 6354

This paper discusses research standards for collecting travel survey data and describes efforts at organizing data and developing metadata. "Metadata standards used for documenting datasets using DDI (Data Documentation Initiative) for DTD (Document Type Definitions) is discussed." It also discusses recommendations for processing, analyzing and archiving data.

The paper discusses the importance of archiving data sets for future use and analysis. "To encourage archiving, the International Conference on Transport Survey Quality Workshop on Data Preservation and Data Presentation (2000) recommended: (1) Increase awareness of archiving issue (outreach) (2) Begin with survey process with the aim of archiving the data sets in mind and (3) Data protection regulations should be observed." While this is important for ongoing surveys and surveys in the future it does not address how to utilize or archive historical surveys.

According to the paper, the value in archiving survey data is as follows:

1. "The historical value and scientific understanding that can be obtained by analyzing this forgotten data:" – a new understanding of how current travel and activity patterns emerged
2. "Researchers today can apply new statistical modeling to old data sets to learn whether causal factors explained travel decisions"
    a. Travel budget hypothesis and commuting budget hypothesis need long-term data sets – these are necessary for creating and designing sound transportation policies.
    b. Modelers will have additional data to model and calibrate their data sets on.
3. "Inter-metropolitan comparison of travel behavior would be possible and would assist in adapting the next generation of travel models from one city to another."
4. "Allow the development of new performance measures that can actually be tracked over time by providing data in much more detail than the invaluable but geographically broad Nationwide Personal Household Transportation surveys of 1969, 1977, 1983, 1990, 1995, 2000; or the decennial Census Journey to work surveys that give great information on journey to work, but none on non-work travel."

Metadata as a way to archive travel survey data is a widely accepted method of archiving.
Metadata, which is information about information, forges the need for labeling, cataloging and descriptive information structured to permit data to be processed.

Two data formats have been created to facilitate this type of metadata archiving.

- "The World Wide Web Consortium (W3C) has created the Resource Description Framework (RDF) and the Platform for Internet Content Selection (PICS)"
- "PICS specify metadata in the form of "labels", which briefly describe data in machine readable format.
- "RDF treats metadata more generally, providing a standard way to use extended markup language (XML) to "represent metadata in the form of statements about properties and relationships of items."

"DDI is an effort to establish an international XLM based standard for the content presentation, transport and preservation of documentation of databases in the social and behavioral sciences." This process is timely and needs highly sophisticated technology.

There are several projects that deal with this issue and they include;

- DDI Initiative in the North America and Europe,
- NESSTAR (Networked Social Science Tools and Resources) a European social science project initiated by Norwegian Social Science Data Services,
- UK data Archive and the Danish Data Archive;
- FASTER (Flexible Access to Statistics Tables and Electronic Resources) project sponsored by the European Commission.
- ETHTDA (Eidgenossische Technische Hochschule Travel Data Archive) in Switzerland built upon the NESSTAR platform
- Space Time Research in Australia is sponsoring ComeIn promoting metadata standards – this will be accessible through CORBA, COM, and XLM.
- Europe at the Institute for Social and Economics Research affiliated with the University of Essex in UK (MTUS-Multinational Time Use Study) which is a similar project to the MTSA.
- The Survey Documentation and Analysis Software (SDA) developed by the University of California, Berkley is a set of programs developed and maintained by the Computer Assisted Survey Methods Program (CSM) to facilitate the documentation, analysis, and distribution of survey data on the World Wide Web.
- IPUMS (Integrated Public Use Microdata Series) at the University of Minnesota since 1997.
- Virtual Data Center at Harvard / MIT under development
- National Statistics are available through the Bureau of Transportation Statistics (BTS), and the US Department of Transportation, National Personal Transportation Survey (NPTS) and National Household Travel Survey (NHTS).

Scheduling Considerations in Household Travel Surveys

Resource Paper Prepared for Workshop A4
Transport Quality and Innovations Conference
Costa Rica, August 1-6, 2004

Stacey Bricka
Research Director and Principle NuStats
3006 Bee Caves Road, Ste A300
Austin, TX 78746
01 512 306 9065 ext 2240
sbricka@nustats.com

This paper outlines an eleven-stage household travel survey process in order to plan for, and execute a travel survey effectively and efficiently. "The paper is broken down into to two sections.

The first section discusses the issues to consider in scheduling the various travel survey activities, focusing primarily on telephone based survey administration." Topics discussed include; the planning and scheduling of events, training of project staff, quality control, preparing data files, and documenting/archiving the survey results.

The scheduling considerations considered include:

1. Training
2. Advance Notification
3. Recruitment
4. Placement of Materials
5. Reminder Call
6. Travel Period
7. Retrieval
8. Quality Control
9. Geocoding
10. Data File Construction
11. Documentation and Archival

This method of collecting and disseminating the data is considered to be "just in time" "both from the perspective of providing respondents with the information and survey materials at the appropriate times as well as performing quality control and geocoding on the data as retrieval takes place rather than waiting until data collection has been completed. It results in higher cooperation rates, and higher quality data collection focused through timely interviewer feedback.

"The second section addresses the questions of whether interviewers should focus on one contact with the respondents (i.e. Recruitment only) or whether they should serve as the main source of contact with the respondent through all interactions." This section focuses on the one administrative issue: that of "tailoring" respondent / interviewer contact by assigning a single interview to a respondent. The advantages and disadvantages are discussed and the potential impacts on establishing a rapport with the respondent, response rates, and types of error that could be introduced into the data set.

The paper concludes that the same interviewer would conduct the recruitment interview, reminder call, and retrieval interview for a given household. "The theory is that the single point of contact for the respondent will make the process easier and ensure higher cooperation rates." It is anticipated that the single interviewer will create a rapport with the respondent, which is an important aspect to the survey results and respondents willingness to participate completely.

The paper does not address GPS usage, administering follow up stated preference surveys customized from the revealed travel behavior in the travel logs, or conducting follow up surveys with non-responders.

Sample Design and Survey Error

Resource Paper for Sample Design Workshop
Seventh International Conference on Travel Survey Methods
Costa Rica, August 2004

Mira Paskota
The Faculty of Transport and Traffic Engineering
University of Belgrade
Vojvode Stepe 305, Belgrade, Serbia
M.Paskota@sf.bg.ac.yu
+381 11 3091 330

This article discusses the various solutions for both "typical" variances and selection of significance variables. The solutions presented and argued include: the application of principle component analysis, cluster analysis, and other multivariate statistical methods. The article also discusses the importance of the precision of data.

Sample Design and Survey Error

The sample is an important element of any survey especially with travel surveys because the quality of the data is directly associated with the quality of the sample. The article also suggests that, statisticians want to draw conclusions about the population based upon the information gathered in the sample. It is also important that with sample design, a statistician cannot be guided purely by theoretical reasons; they must take reality and the limitations it puts upon it into account.

Basic Concepts Related to Sample Design

Methods of interviewing are:

- Mail Survey
- Telephone Survey
- Face to Face Survey

According to the article the *first contact* and *respondent's motivation to participate* in the survey is closely tied.

Advantages and Disadvantages of Survey Modes

Mail

Advantages:

- Relatively cheap
- Requires a small labor force that is trained

Disadvantages:

- Possibility of non-coverage error and low response rates
- Address lists are difficult to obtain and unreliable
- Respondent's degree of interest
- No control about who fills out the questionnaire
- Cannot control if the form is filled out completely or correctly
- Illiteracy issues

Telephone Surveys

Dominant survey mode in countries where telephone penetration is quite high

Advantages:

- Faster compared to mail surveys and face to face surveys
- RDD *Random Digit Dialing* system enables design of a high quality sample for the given frame, as well as anonymity.
- Interviewers participation brings significant control into the interviewing process
- CATI systems enable simultaneous control of the course of survey, data entry, data control and immediate accessibility in desired format.

Disadvantages:

- It is not applicable in the case of every long questionnaire or diary surveys
- Within developed countries this method is a problem because of lack of telephone penetration
- In highly developed countries – a growing number of households, mainly young couples and individuals do not sue stable phone at home but mobile phones. Since these individuals are highly mobile this kind of non-coverage could cause serious problems when travel surveys are considered.
- "Experiences in some countries could have shown that surveys on mobile phones are not practically applicable, making this issue of high current interest, and still not solved in a satisfactory method."

Face-to-Face Surveys

This survey method is the most traditional, unsurpassed mode of survey, and is the only way to take a survey among non-listed population.

Advantages:

- "It enables completion of the longest and most complex questionnaires, possibly demanding some additional material to display to respondents (visual aids and similar)"
- "Interviewers presence ensures the correct filling in, and application of CAPI methodology and lap-top computers bring all the advantages of telephone CATI surveys.
- Convenient in cases of general population surveys and multistage samples

Disadvantages:

- High costs
- Demands that the interviewer strictly observes the rules of household selection and an individual in it, as well as rigorous control of field work.
- Non response differs drastically in urban and rural areas
- Smaller households are typically underestimated as a general rule
- "In highly developed countries, especially in urban and exclusive residential parts, it is very difficult to enter the household, unlike less developed countries in which this problem does not exist."

Internet Surveys

Difficult to administer travel surveys online – however it is possible

Advantages

- High quality internet surveys are possible only in case of very narrow and specific populations, the ones we have appropriate frame for, and for which we are confident about their having regular access to Internet, such as university population or population of employees in a company.

Disadvantages

- These surveys have no sample or control whatsoever and can not be regarded as valid research
- These surveys take much more planning, rigorous control and much more work than the other travel survey methods.
- Call for extreme caution

In order to decide what survey method is the best, depends largely on each individual survey and circumstance. Combining survey methods is sometimes the best way to achieve the results that are sought.

Example – can use advantages from one mode to compensate for the disadvantages of another mode. For example, "the issue with the telephone survey, the problem of non coverage can be resolved by Face to Face interviews of the respondents that have no telephone.

Sample Classifications

Since error is present with all types of samples, it is only possible to estimate error with probability samples. Probability samples are classified as follows:

- *Simple Random Sample* – elements are chosen completely randomly, independent from one another and without replacement, that is, an element that has been drawn once can not be chosen again.
- *Systematic Sample* – dividing the population by the sample size, we get the value called the step.
- *Complex Random Sample* – stratified sample, cluster sample, multistage sample.
- *Sub populations* or status are clearly defined in the population
- *Standard classification* to one-stage and multi-stage samples.
- *Biased sample* – this error is quite large and is defined

Before Starting a Sample Design it is Necessary to Know the Following

- The survey's aim should be the starting point for any sample design
- Clearly defined sample target population and the instrument to be applied
- Sample Size is an element that would like to be known, but is often a more complex issue than choosing a number. Sample Size is typically influenced by the following:
    - Maximum error allowed
    - Population size
    - Data variance
    - Identifying the smallest group and the exact size in the population for an estimate based on the sample
    - Intention of the data collected

Sample Frame

Sample frame is a group of elements (households or persons) that qualify for being chosen for the sample. The sample can be representative only if the frame equals the population, if there is no coverage error. (Ideal frames are seldom available).

A frame represents a list of elements. To make a good frame the following conditions have to be met:

- To cover a population as well as possible
- To exclude duplicates
- To exclude redundant elements (meaning that all the elements of the list are at the same time population elements)

The frame can be geographically defined as well. It is important to list everything that has been used as the frame when producing sample design for each individual sample, thus focusing attention to a possible cause of coverage error.

Survey Process and Possibility of Error
After survey goals are defined, one of the first steps in making survey design is determining the target population.

Four Types of Error which are most frequently brought up:

- *Coverage Error* – first opportunity for an error to occur
- *Sampling Error* – occurred due to the fact that instead of making conclusions about the population based on all its elements, we take only part of it.
- *Non Response Error* – the fact that the realized sample is always different from the one that has initially been projected
- *Adjustment Error* – the effort to remove errors from the process by weighting underrepresented and over represented population results in a sometimes unwanted effect.

Sample Design for Belgrade 2002 Survey
The survey goal was to collect a sufficient amount of data that would help solve the problems of public transport and city traffic in general.

Measuring device used - a questionnaire with questions addressing different target groups.
The following information was collected:

- Demographic data
- Habitation
- Motor and other vehicles ownership
- Parking problems
- Public transportation use
- All daily trips in the city
- Attitudes on traffic and public transport use

Collected data was analyzed on five levels:

1. Households
2. All household members 6+
3. Daily trips
4. Daily trips' segments
5. Chosen individuals 15+

And for each level a different sample was realized

For common basis, different samples were designed for each of the three levels:

- On the household level, two stage stratified household sample
- On the individual level, for questions on daily trips, three stage stratified cluster sample of persons
- On the individual level, for questions on attitudes in traffic and public transport, three stage stratified random sample of persons

Some definitions:
Target Population – Daily Trips Portion - urban population of Belgrade age 6+, Attitude on Traffic and Public Transport – urban population age 15+

Population Structure – was taken from the most previous census – data used was age and sex

Frame – the list of electoral units from the urban city territory was taken - each electoral unit in the city consisted of 200 households or 600 persons

Stratification – administrative municipal areas were used for stratification

Possible sources of error

- Out of date electoral lists
- Difference between voters and 6+ population

Field Work

- Two week period, Tuesday to Friday afternoons and evenings plus whole day on Saturday
- 105 interviews
- 10% of the sample was subject to random control
- Households with problematic or incomplete questionnaires were additionally controlled
- Data entry lasted 10 days, and 44 persons took part in it

Realized Sample Size

2,650 households 7,852 household members 6+, 16,181 daily trips, 28,272 trip legs, 2,650 persons 15+

Post stratification

Used weights with value from 0.8 to 1.11

Sample Error Estimates

Simple random sample error does not exceed 2% for households or 1.2% for all household members

Conclusions

Sample error is historically the first error from various survey errors to be identified. However, there are other errors in travels surveys, such as survey design which are typically not solved in a standard manner. Non response error is another common error found in travel surveys.
The error propagated through the process of modeling in traffic is the total survey error.

Important to note that non-sampling error behave conversely from sampling error, it grows with the increase of sample.

Survey Design, The Past, The Present, The Future

Henk van Evert
AVV Transport Research Centre
P.O. Box 3001
3000 BA Rotterdam
h.c.vevert@avv.rws.minvenw.nl
+31 10 2825960

Werner Brog
Erhard Erl
Socialdata Institut Fur Verkehrs-und Infrastructureforschung
Hans-Grassal-Weg 1
81375 Munchen, Germany
socialdata@socialdata.de
+49 89 716420

This paper discusses the history of the Dutch Travel Survey to describe the evolution of Survey Design in the past decades.

1978
The Dutch Statistical Office (CBS) started tracking the mobility of the Dutch Population in 1978. The Onderxoek Verplaasinggedrag (OVG) studied passenger transport. About 1500 households were visited each month for the survey. There were two visits per household. A travel diary was left for each householder over the age of 12. They had to fill in their personal data and the data on the trips they made over a two-day period.

1980
Mail back surveys were tested since they offered a lower cost alternative. The following conclusions were made based on this survey vehicle.

- The results are comparable with the standard face to face survey
- Signalized differences may partly be caused by the interview instruments
- The interview instruments employed can be refined
- Savings in financial terms amount to about 30%
- The response rate was, however only 50%

Two pilot studies were tested in June 1983 and June 1984. The objective of the pilot studies were to:

- To measure the response
- To determine the extent of zero travel
- To compare the trips thus measured with those from the method used at that time

The comparisons between the two surveys were completed and the following conclusions made:

- More trips and greater total distances were measured using the telephone/postal method
- The number of people with zero travel was more or less the same
- There is a relationship between the amount of travel and the number of telephone calls before initial contact was made; the more phone calls required the greater amount of travel
- Those with unlisted telephone numbers as well as those without a telephone deviate as far as their amount of travel is concerned. Those with unlisted numbers travel 8% more kilometers and those without a telephone 12% fewer kilometers than those with listed telephone numbers. It is assumed that the overall effect on the amount of travel is of no consequence;
- The response with regard to trip data is about the same
- The response at the household level is higher
- Non-response is higher than average among older people and those without their own means of transport
- Non-response households are on average smaller
- Car ownership is lower among the non-response householder.

Redesign No. One
1985 – Telephone interviewing (to reduce costs) – A *call* was preceded by a *letter* in the mail. The telephone interview portion included household information such as: the composition of the household, and vehicle ownership. A travel diary was sent to the household with specific days to track travel. In addition, personal and household information was asked, such as, income, education, and occupation. If the diary was not returned another one was sent, with the request to complete it one week later, if the second diary was not returned another diary was sent, with another week extension.

The benefits of the new survey method were:

- Decreased costs per unit of data gathered
- Information could be input immediately after the telephone interview and the information was checked for errors immediately

Sample was drawn from a stratified sample, and the stratification variables were Province and Urbanization. The telephone approach excluded those with unlisted numbers and those without a telephone.

This new design however limited the comparability between the 1978 surveys and 1985 surveys.

Redesign No. Two
Significant response rate decline between 1985 and 1998 for the OVG survey. There was an increasing unwillingness to participate, and an increasing number of unlisted telephone numbers. While the response rates were decreasing the yearning for this information was increasing. Therefore Statistics Netherlands in co-operations with AVV, looked at alternative design methods to increase response rates with enhanced research flexibility.

They contracted the German New KONTIVE Design, developed by the institute Social data in Munich. "This survey is a normal self-administered survey, with telephone motivation of respondents and subsequent follow up surveys for more detailed data per subgroup."

- Respondents are phoned when they receive the survey material and encouraged to fill in the questionnaire and diary
- The phone is an instrument to encourage participation, not as a tool to collect the information

The Basics of the NKD Survey Used

- Written questionnaire for the household
- Written questionnaire for each householder
- Respondents are asked to report their trips for a specific day
- The questionnaire is as simple as possible
- Respondents answer questions in their own words to make it as easy as possible
- Categories are given for mode and trip purpose and are clear and understandable
    - Ex. Work, education, work-related business, and shopping, return home
- If data is incomplete or require clarification then a follow up phone call is used
- Sometimes the NKD survey is followed up by satellite surveys for specific sub groups these are carried out through the telephone

The response rate for this survey was approximately 70%, a drastic increase from the previous surveys conducted.

Present Survey Use
2004 a new survey was adopted. This survey is called the *Mobiliteitsonderzoek Nederland* (MON). "It is an opportunity to introduce new subjects into the investigation of the mobility of the Dutch population." The new design applied to the MON survey follows nine quality principles:

1. Respondents are customers
2. Continuous improvement
3. Fit for the foreseeable future
4. High response rates
5. Immanent Validation
6. Total Households
7. Linked and unlinked trips
8. Attitudes and behaviors
9. Human being desperately needed

"The MON has three principles which will be of general importance in the future"

- Flexibility – survey design that is able to react to demands of the users.
- Quality Standards
- Respondent - Orientation

"The MON uses a quality scheme that is specific to it, but is also based on principles that constitute useful general guidelines." The elements of the scheme are:

- For each day of the year households will be surveyed
- The households are selected randomly
- Only members of the households living at the sample address are interviewed
- All members of the household living at the sample address are interviewed;
- The net response rate must be 60% for each month. The net response is defined as quotient from:
  - Number of usable returned household forms
  - Number of sample of selected households
- The share of households responding by phone must be a maximum of 20% of all responding households per month;
- Validation surveys are part of the basic survey for; public transport trips (collecting of stages), problems of completing the forms, non-response and validation of non-reported trips;
- The response rate for the public transport follow up is at least 80% of the respective sample per month (usable returns)
- The respondents of the survey do not get any incentives

In conclusion this section of the paper found that "[t]he investment costs for the survey organization are quite high, the validity of the data is good, but the range of the data is very limited."

Panel Surveys

Workshop A8
Paper presented for Costa Rica Conference 2004

Jimmy Armoogum
Department of Economics and Sociology of Transport (DEST)
National Institute of Research on Transports and Safety

Bastian Chlond
University of Karlsruhe (TH), Institute for Transport Studies

Jean-Loup Madre
Department of Economics and Sociology of Transports (DEST)
National Institute of Research on Transport and Safety

Dirk Zumkeller
University of Karlsruhe (TH), Institute for Transport Studies

Historically travel surveys and data collection revolved around traffic modeling and peak hours for an average weekday analysis, when the traffic volume is maximal. However, this article argues that travel research should work at obtaining a better understanding of individual behavior and look at ways to influence behavior that better utilizes the infrastructure in place. This article argues that travels surveys need to capture the same people in different situations in order to capture "the flexibility of a person to react and the identification of constraints and regimens". The article also argues that that cross sectional and snap shot oriented surveys of one day give poor descriptions of ongoing changes and do not distinguish between real changes in behavior from external factors.

The article identifies ways to measure change in behavior.

- "By asking persons or firms how they would change their behavior in the future. The prospective methods are interactive interviews, stated preference, and stated response."
- "By observing these changes in the past on a long enough period; these longitudinal methods are based on panel surveys, repeated surveys or mobility biographies, but also on administrative data"
- "It is not enough to do one cross section to understand the conditions of change, but it is necessary to repeat the observations or "moving pictures of trajectories" (generic panel surveys)"
    - Problems with this approach is that it is quite costly, time consuming, and analysis is not straight forward

Definitions of Survey Designs

Repeated Cross Sections

- Repeated surveys within a population at different points in time with different individuals are not panels – this is simply a series of cross section surveys. This is not a satisfactory approach since this method assumes that people of the same age groups are assumed to have similar lifestyles travel experiences/choices.

The Longitudinal Survey

- Measurements of behavior over one week, or even longer.
- "Short Term Panel".
- Multi-day data of one week.

The Panel Survey

- "Observation at different discrete points in time of the same items"
- "It distinguishes from repeated cross section surveys following the same design but with independent samples."

Panel Surveys and Panel Data

- Panel data is helpful as it allows for a historical account of travel patterns, however this information can be biased based upon biases by errors of the respondents, (selectivity phenomenon, transfiguration of the past, and simply forgetting) – this has been excluded from this study.

According to the article, there are many survey weaknesses with longitudinal character when they are compared against cross section or snapshot surveys. Some of these weaknesses include:

- "The regularity of respondents' activities is unknown, since the frequency of trips is not known."
- "An underestimation of long distance trips since they are not typical."
- "Cannot distinguish between interpersonal and intra-personal variability. Therefore the analyzer can not determine the behavior changes on an individual level."
- "Only net changes between two cross sections are identifiable and quantifiable in the sense of the margin values or changes in the margin distributions."
- "Assume complete reversibility of effects (e.g. Persons who will buy a car will behave like a person who never owned a car and vice versa)."

These problems and methodological weaknesses of snap shot/cross section surveys are avoidable by using panel surveys in their closer definition.

- Panels measure the effects of any changes in external factors for individuals and households
- They allow for causal analysis, as the temporal sequence causes and effects are known
- It is possible to capture gross changes and their potential compensation(s) with panel surveys
- Once a panel is installed it allows for quick identification of mobility demand
- General panels can be used as comparison against special panels (control group)

The study suggests that longitudinal surveys should concentrate on the length of a week.

Works Cited

Armoogum, J., Chlond, B., Madre, J. L., Zumkeller, D. (2004). *Panel Surveys;* Workshop A8 Paper presented for Costa Rica Conference 2004

Bonsall, P. (2004) *Quality Assessment*; Resource Paper for Workshop A6 at the International Conference on Survey Methods in Transport.

Bricka, S. (2004). *Scheduling Considerations In Household Travel Surveys*; Resource Paper Prepared for Workshop A4.

Brog, W., Ehard, E. & Van Evert, H. (2004). *Survey Design, The Past, The Present, The Future*

Griffith, R., Lee-Gosselin, M., & Richardson, A.J. *Travel Survey*. Committee on Travel Survey Methods. Federal Highway Administration.

Hensher A. D., & Rose, J. M. (2004) *Handling Individual Specific Availability of Alternatives in Stated Choice Experiments*.

Levinson, D. & Zofka, E. (2004). *Processing, Analysis and Archiving of Travel Survey Data*.

Paskota, M. (2004). *Sample Design and Survey Error*; Resource Paper for Sample Design Workshop.

Sammer, G. (2004). *Processing, Analysis and Archiving of Travel Survey Data*; Resource Paper for Workshop A5.

National Cooperative Highway Research Program. Research Result Digest. (2002).*The Case for Standardizing Household Travel Surveys*.

PART THREE

HOUSEHOLD TRAVEL SURVEY SYNTHESIS

Introduction

Using the identified issues and the review of available knowledge on household travel surveying techniques, this section begins the process of matching issues with potential strategies. The layout for Matrix One lists: *First Author*; *Date*; *Article Title*; *Topic Areas*; and *Corresponding Identified Issue.* This matrix is intended to illustrate various combinations of issues and resources.

MATRIX ONE - Article-Issue Index

First Author	*Year*	*Title*	*Topic Areas*	*Corresponding Identified Issue*
Armoogum	2004	Panel Surveys	Individual behavior Utilization of existing infrastructure Measuring a change in behavior Cross sections Longitudinal Surveys Panel Surveys	8,14,18
Bonsall	2004	Quality Assessment	Archiving Documentation Quality Standards - data quality, quality measurement Universal Indicator of travel survey quality	1, 3, 7, 8, 13 thru 18
Bricka	2004	Scheduling Considerations in Travel Household Surveys	Planning and execution of travel surveys Travel survey schedules Training project staff Quality control Preparing data files Documenting and archiving survey results	1, 3, 7, 8, 9, 14 thru 19, 22, 23, 24, 27, 28
Brog	2004	Survey Design: The Past, The Present, The Future	Dutch Travel Survey History	8,14
Griffith	N/A	Travel Survey		
Hensher	2004	Handling Individual Specific Availability of Alternatives in Stated Choice Experiments	Stated choice experiments Non-availability of alternatives Respondent's capability Tailoring choice alternatives to account for household limitations	4,14,18

MATRIX ONE - Article-Issue Index (Continued)

First Author	*Year*	*Title*	*Topic Areas*	*Corresponding Identified Issue*
Levinson	2004	Processing, Analyzing and Archiving of Travel Survey Data	Data Organization Developing Metadata Data Documentation Initiative (DDI) to DTD Document Type Definition Archiving Data	1, 3, 6, 7, 9, 10, 12, 15, 16, 17, 20, 21, 25
Paskota	2004	Sample Design and Survey Error	Sample Design and Survey Error Typical Variance Mail, Telephone, Face to Face Surveys Sample Classifications Survey Design Non-Response Error	1, 3, 4, 5, 6, 7, 8, 9, 10, 11 12 thru 17, 20, 23, 25, 27, 28
Sammer	2004	Processing, Analysis and Archiving of Travel Survey Data	Database building, questionnaire editing, data entry, standardization for flexible coding, geo-coding, data editing, and cleaning and data weighting.	1, 3, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 19, 20, 21, 23, 25, 27, 28
National Cooperative Highway Research Program	2002	The Case for Standardizing Travel Surveys		

PART FOUR

DISCUSSION – REVEALED THEMES

Introduction

After reviewing the available documents, summarizing the international materials and holding discussions with persons involved in the original deployment, several themes have emerged. These themes are discussed in the following segments: *Need for Custody/Control Management System – Using IT as a Foundation, Rethinking the Methodologies for Capturing Transit Data: Using GPS, Sampling Frame Concerns.* Matrix Two lists the revealed themes, with suggestions for first steps and follow-on strategies to address the issues and reduce the risk of reoccurrence in the up-coming household travel surveying effort.

Need for Custody/Control Management System – Using IT as a Foundation

A common problem described by internal staff and evidenced in the documentation of the flow of work was the high turn-over of personnel both inside the NYMTC and with the various consulting firms. Staff changes are to be expected in any large agency; however, when a long-term project such as the household travel survey and subsequent model building, testing, and implementation is underway, such changes can cause problems.

The problems identified below relate to the changes in staff and consulting firms, or problems that arose as a result of inconsistent oversight/management. :

- Issues of cooperation and coordination during data collection phase of all parties
- Whether weekends were to be included or not in surveys
- Desire to add participants after the sampling decisions are made
- Concerns over the final sample composition
- Acknowledgment of lessons learned from 1995 Pilot
- Changes in sample goals by consultants
- CATI system failure requiring the use of paper and pencil technology for retrieval interviews

Since no policy or program can guarantee permanency in personnel, a custody/control management system can act as an oversight tool for the length of the data program. New ways of using Information Technologies (IT) can aid in the design and use of such a system. For example, if the addresses of the sampling frame households are "marked" in real time on a map, then it is possible to monitor and/or track the coverage of any number of factors regarding the progress of the surveying effort. The system would provide the transparency for the entire operation, reducing the risk of error, unrecognized mistakes, or unintended over-sampling of particular locations, days of the week, etc.

The ability to track and trace the data from its creation to its final uses offers security advantages as well. A custody/control management system may already be part of some work practices. However, any subsequent or smaller subsystem would need to be accessible by internal staff and not considered property of various consulting team. This would defeat the purpose of agency oversight and increase the risks of future failure.

The first step in the design of a custody/control management system is to trace out the various players, products, and protections needed for a travel surveying effort. A careful description of the functionality (but not necessarily in absolute terms) required, will make it possible to include such a system in future contracts and requests for proposals. A pilot test of such a system is also essential in determining the completeness and effectiveness of the system.

Rethinking the Methodologies for Capturing Transit Data: Using GPS

A review of the documentation from the Regional Household Travel Survey in 1997 revealed issues with the accuracy of transit data. In most American cities; even those with advanced transit systems, transit usage is relatively minimal. However, in the New York region, transit plays a large role in household trip-making activities.

According to the recently published *Commuting in America III* (Pisarski 2006), the national mode share for transit to work in the New York metropolitan area has risen over the decade of 1990 to 2000, from 37% to 38%. (89) Pisarski notes that the use of transit in New York City itself represents 30% of the national share. (90) As a share for Central City destinations, transit in New York is reported at 45%, while the transit share of Downtown commuters is 76.5% (93-94). The area, only about one square mile, is defined as New York Central Business District (CBD). The area referred to as "midtown" reports 78%, and the Wall Street area of Lower Manhattan is less at 74%. Although the shares of transit usage are impressive, consider the impact on household activities for the estimated actual number of riders: 2,065,120 persons taking transit to the Central City and 290,390 taking transit to Downtown each day!

The sheer magnitude of this level of rider-ship means travel surveying efforts need to be successful in understanding the transportation services provided by all types of transit in the New York region. Good information will inform modeling, policy, and transit operations. No information or worse, bad information, will not only corrupt modeling efforts, but could result in poor policy choices and overall reduction in service satisfaction through poor operating decisions.

A review of the available documentation on the household travel surveying efforts revealed early warnings regarding the collection of transit data during the 1995 pilot tests. For example, it was noted that the first boarding location for all transit trips should be highlighted to ensure all transit users provide this information. It was also noted that data elements with higher rates of non-response including personal cost of transit use after transit subsidies. (RT-HIS M & I: page 23).

During the actual data collection effort, problems in the Computer-Aided Data I (CADI) system resulted in the use of paper and pencil rather than the intended computerized data collection process (RT-HIS M & I: page 26). It is possible that this alone could have corrupted the data. In any event, during the data quality review, it was noted that a manual review of the geocoding was required for records associated with transit trips (RT-HIS M & I: page 32). And finally, the transit data was deeded to be misreported or miscoded – required manual re-reporting of the data. The major problem was the sequencing and the data was flagged for future users to identify the problem cases (RT-HIS DATAUSER: page 8).

A set of analyses need to be conducted to examine the red-flagged transit trip data to better understand the impact of using proxy reports, trying to recall any transit trips in particular parts of the NYMTC region, and an overall examination for systematic problems in the 1997 dataset. Since there are limited strategies available to more accurately capture transit trips using a traditional household travel survey processes, it is recommended that new technologies are considered.

One approach is to scope and pilot test the use of on-body GPS for household members taking transit. This would provide a visualized record of the travel patterns of each person as the movement can be mapped using GIS. An enhanced understanding of actual movement would include, dwell time, walk distances, and in-vehicle times. It is not clear what types of equipment provide the best results in difficult environments like Manhattan. Pilot tests for the equipment need to be conducted using a managed set of participants (not general public) as patterns will need to be replicated and instructions for using the equipment followed without variation. Development of on-body data capture will require inputs from the general public, after the technology has proven its ability to produce satisfactory output. The GPS technology still needs to be used by the population expected to participate in the surveying process.

Sampling Frame Concerns

In the 1997 deployment, a concept referred to as Mode Leadership was suppose to ensure an appropriate sampling frame for the household travel survey effort. It appears to have been attempted without prior experience and although laudable, it does not appear to have accomplished the stated goals. Part of the problem was revealed in the order of operations for developing the sample. After all the characteristics of the desired household were assigned, it was discovered that the particular household had moved.

Consideration could be given to "mapping" a surface of households - using a synthetic population base model. The original synthetic populations used a methodology based on the Public Use Microdata (PUMS), using the Census long-form data. The long-form has since been replaced with the continuous census (the American Community Survey) methodology. A test bed could be developed for assembling an ACS synthetic population or a hybrid population, such as the newly tested Longitudinal Employer-Household Dynamics (LEHD)[1] methodology for determining journey to work. Advanced GIS tools can provide a number of measures for access to various modes, including network features.

Another approach, for at least some of the participants in the next household travel survey could be similar to a consumer panel model. According to Elaine Murakami, consumer panels offer a potential for panel survey work – and could provide a set of household behaviors with a rich set of travel pattern information and intrahousehold travel patterns. Using these inputs and other "transferable" travel behaviors from other sources, may change the very nature of household travel survey recruitment.

Reference:

Pisarski, A. (2006). *Commuting in America III: The Third National Report on Commuting Patterns and Trends*. Washington, D.C.: Transportation Research Board.

[1] LEHD, U.S. Census http://lehd.dsd.census.gov/led/about-us/FAQ.html#lehd, Page dated: 02/28/07

MATRIX TWO – Revealed Themes

Revealed Theme	*Evidence*	*First Step*	*Next Steps*	*Expectations*
Need for Custody/Control Management System *Cradle to Grave* Strategy	Changes in internal staff, management team and consulting firms contributed to lack of control over survey design, deployment and outputs	Draw up chain of custody and describe integrity, rigor, reliability, and transparency requirements	Automated system elements for tracking and monitoring all aspects	Team bidding will need to demonstrate ability, system, and pilot test results of all system elements, including IT features. Complete process will be understood by internal staff – reducing risk of failed data collection due to sampling issues (missed populations), deployment issues (equipment/system failures), and unrecognized losses during process
Transit Trip Innovations	Systems failed during surveying process, reliance on proxy reporting, and other location-based problems reduced the value of the transit trip data.	Investigate the feasibility of using new technologies, such as on-body GPS to capture transit trips in Manhattan area	Combine information on technology, and survey participant needs to scope usefulness.	The next household travel surveying effort will include a range of survey instruments and technologies, customized for different market segments to ensure the best capture rates, and highest quality of information
Sampling Frame	Mode leadership concept was not successful because households moved and relationships with mode useable were not well understood	Determine feasibility of establishing geographic base/synthetic population	"Grow" synthetic populations for testing	Use of ACS and other administrative databases will provide foundation for surveying effort – and allow for enhancements

Appendix I:

Recent Travel Survey Research with an International Perspective

In the paper, "National Transport Surveys: What can we learn from international comparison?" Bonell, Armoogum and Madre analyze National Personal Travel Surveys conducted all over the world to identify the best practices of Travel Surveys used throughout the world to assist France with their 2007 survey.

The findings of this survey concluded the following:

Objectives and Definitions
The objective, which is important to define early in the process, typically dictates the methodology. The two major types of surveys include the trip based approach and the activity based approach. Trip based approach asks questions regarding the actual trips the respondents take, while the activity based approach analyzes the places and activities people go to outside the home. The United States, Norway, Finland, Great Britain, The Netherlands, France, Belgium, and the Czech Republic all use trip based approaches, while Demark, Germany, Austria, and Italy use activity based surveys. Sweden Switzerland and New Zealand all use stage based travel surveys.

The two main objectives for conducting a travel survey in most counties are:

- *describing travel behavior and analyzing its main explanatory variables,*
- *providing inputs for developing a national travel demand forecasting model, important objective in the case of Austria, Great Britain, Norway, Sweden, The Netherlands and Spain.*
- *other objectives are inter-related: the analysis of long-term trends is necessarily linked to modeling just as modeling is frequently used as a planning tool.*

While the objectives vary, so do the organizations involved in the survey process. Typically the 'National Transportation Authority' is involved in the process, while the local, regional, research organizations, and the national statistics bureau participation varies. It is generally accepted however that the funding source has an impact on the survey design and objective.

Various definitions are used for both local and long distance travel and the information gathered in travel surveys. Most counties do not have a trip distance limit, however Great Britain, Switzerland, Spain and Italy each have limitations on distance ranging from 10 meters to 300 meters.

Purposes and Modes
In most counties the purpose of each trip is identified. "In Germany purposes are ranked on the basis of those which are judged to be the most constrained. The main mode is defined either according to a modal hierarchy (e.g. public, transport, car, motorized two-wheeler, bicycle, walking) or to the mode used for the longest distance in the trip. The number of alternative items varies, for example from 6 modes (Spain, Czech Republic) to 31 (Denmark) and from 7 purposes (Spain) to 54 (The Netherlands). Thus, the detail of the data differs between countries."

Survey Modes include face-to-face interview, telephone survey, postal survey or a combination of these). In addition to these classic modes of gathering data almost all methods use computer-assisted technology as well. They identified that whether or not a respondent was notified before the survey was administered affected the response rate of the survey. The following details the survey modes found among the countries surveyed.

- *the group of the telephone surveys is the largest one (Denmark, Finland, Norway, Sweden and Switzerland). In the case of the four Nordic countries, the sample is selected from the national register of individuals, then the telephone number must be found, while in Switzerland the phone is used exclusively;*
- *three other countries (Germany, Italy and the USA) use the combination of telephone and postal mode. In Germany, the sampling base is of the same type as in the Scandinavian countries, but the telephone number can only be obtained for 60% of households (in order to avoid excluding 40% of the sample, the questionnaire is sent by post, entirely self-administered, with no assistance for respondents). In the case of the telephone survey, the process used is similar in Germany, in Italy and in the USA. The initial contact is established by telephone (socio-economic data on the household and its members). A travel diary is then sent by mail to be completed for a clearly defined period; after this period, a telephone contact is made a second time in order to collect the data from the travel diary;*
- *the third group includes three countries (Austria, Great Britain and France). The survey uses both administered questionnaires, which are completed during face-to-face interviews, and self- administered travel diaries. During the first visit, socio-economic data and some travel data are collected. After this initial contact, the interviewer leaves the travel diary to be completed by respondents; the interviewer comes back one week later to collect the travel diary. Depending on the country, during the second visit the interviewer may either merely check the travel diary and add additional information when necessary, or collect the data, in which case the travel diary is used as memory jogger;*
- *in the fourth group we find Belgium and The Netherlands which run a self-administered postal survey. When the phone number is available, it is used for reminders as well as for a validation call in the case of The Netherlands. In Spain, household characteristics and daily mobility were collected for a first survey during the last trimester of year 2000. A second survey has been conducted on long distance by phone a few months later (March 2001 - February 2002).*

TABLE 5: Survey Mode

Country	Mode of contact	Computer-aided interview	Proxy
USA	1st telephone contact travel diary sent by post 2nd telephone contact: to collect diary data	CATI	< 14 years old: yes 14-15: possible 16-17: not in general >= 18 years old no
Norway	Telephone survey	CATI	No
Sweden	Telephone survey	CATI	children < 15 years old
Denmark	Telephone survey	CATI	No, but children may be helped by parents
Finland	Telephone survey	CATI	No, but children may be helped by parents
Great Britain	Face-to-face + diary self administered for local travel	CAPI	Allowed for long distance and possible for diary
The Netherlands	Self administered + phone call reminder (when telephone number is available)	No	No except specific case
Germany	Telephone survey when phone number is available (60%); if not postal survey	CATI for telephone survey	< 9 years old: yes 9 to 14: possible
Austria	Face-to-face contact + self administered questionnaire + 2nd contact to collect and check questionnaire	No	In principle not, but impossible to verify
Switzerland	Telephone survey	CATI	Strictly limited to specific cases
France	2 face-to-face visits + self administered long-distance diary	No	No
Belgium	Self administered + phone call reminder	CATI for phone call	In principle not
Czech Republic	Face-to-face	No	
Spain	Daily mobility and household characteristics: face to face survey Long distance: telephone survey	Daily mobility: CAPI Long distance: CATI	No
Italy	Telephone recruitment + travel diary self administered + CATI: diary collection with memory prompting	CATI	No proxies in principle
New Zealand	2 face-to-face visits with memory jogger given during the 1st visit	CAPI	No proxies in principle

Sources: survey by LET and INRETS-DEST + Kunert et al., 2001

The duration of which the travel survey is conducted varies by country. The following table summarizes these findings according to the article.

TABLE 6: Design for Local Travel

Country	Type of questionnaire	Surveyed period	Choice of the day/period	Week day/weekend
USA	Memory jogger + CATI data collection	1 day	Randomly predefined day	All days of the year
Norway	Memory collection with memory jogger sent in advance	1 day	Randomly predefined day	All days of the year
Sweden	Memory collection with memory jogger sent in advance	1 day	Randomly predefined day	All days of the year
Denmark	Activity/trip of the day before	1 day	Uniform distribution over week	All days of the year
Finland	Memory collection with memory jogger sent in advance	1 day	Randomly predefined day	
Great Britain	Self administered + control by interviewer during the 2nd visit	1 week	Randomly predefined week	All days of the year
The Netherlands	Diary for a pre-defined day	1 day	Day of the week pre-defined randomly	All days of the year including weekend and bank holiday day
Germany	Telephone: memory jogger postal: standard trip log diary	1 day	Randomly predefined day	All days of the year
Austria	Self administered diary with control by the interview during the second visit	1 fixed day	Randomly predefined day	Week
Switzerland	By memory	1 day	Randomly predefined day, but could be postponed	
France	Trips of the day before and the last weekend by memory	1 day + 2 weekend day	No, but distribution of days checked	All days of the year
Belgium	Diary for a pre-defined day	1 day	Randomly predefined day	All days of the year
Czech Republic	Average number of trips	Trip habits	No	Trip habits
Spain	Week day before + 1 weekend day by memory	1 week day + 1 weekend day	Randomly predefined day	Both week day and weekend
Italy	Self-administered diary with control during the CATI	1 day	Day of the week pre-defined randomly	All days of the year except Saturdays and bank holidays
New Zealand	Memory jogger + collection during the 2nd visit	2 days	Randomly predefined days	All days of the year

Sources: survey by LET and INRETS-DEST + Kunert et al., 2001

Long Distance Travel

Long distance travel is typically recorded by memory over a designated period of time. Of the countries surveyed the time period at which long distance travel is recorded varies from 1 day to three months. In addition, the detail of the long distance travel data also varies. Some countries gather the specific local trips on long distance travel, while others only assess the egress and ingress to the area.

Periodicity and seasonality

Typically most countries gather travel data over an entire year, encompassing all four seasons. Austria and Spain are the only countries that do not gather this information throughout the year. "Only Denmark, Sweden, Great Britain, The Netherlands, the Czech Republic and New Zealand have set up continuous nationwide transport surveys (**Table 8**), although Denmark and Sweden

have stopped their data collection (Denmark is about to start a new continuous one (Madre, 2005)). In all the other countries, the survey is carried out periodically, except in Belgium (first and only survey in 1999) and Spain (2000, but the next one is scheduled for 2006)."
The methodology of the surveys change over time among all countries, however limiting the number of changes to the survey each year make it easier to compare results among the surveys.

Geocoding
Geocoding also varies among the countries surveyed. Many countries use specific x and y coordinates for both long distance travel and local travel, while others use information provided by the respondent. According to the article "So far, no country has ever used real time monitoring systems with the geolocation of trips by means of GPS or cell phones. However, this is under consideration for the next survey in the United States and France."

Sampling Issues

Sample Size
"The scope of the survey consists of the entire population of permanent residents in the country, irrespective of nationality. All the surveys aim to be exhaustive, even though some groups are excluded because of the content of the sampling base or particular difficulties in conducting the survey." The table below summarizes the sample sizes used in each country.

4.1 Target Population
TABLE 10: Scope

Country	Sampling data base	Statistical unit	All household members?	Individuals excluded?
USA	Random digit dialling	≈ Household but more precisely phone number	All household members	People in institution (except students)
Norway	National register of individuals	Individual over 13 years old	Only selected individual	< 13 years; people in institution
Sweden	National register of individuals	Individual	Only selected individual	<6 years and > 84 years
Denmark	National register	Individual	Only selected individual	< 10 years and > 84 years
Finland	National register of individuals	Individual	Only selected individual	< 6 years
Great Britain	Royal Mail address database	Household	All household members	People in institution
The Netherlands	Address database	Household	All household members	People in institution
Germany	Community register of individuals	Individual of 14 years old and more	All household members	No, but people in institution are often not in the register
Austria	Census + national register	Individual	All household members	< 6 years
Switzerland	Swiss telephone register (including unlisted)	≈ Household but more precisely phone number	1 or 2 individuals	< 6 years people in institution
France	Census + new dwelling	Household	1 or 2 individuals	< 6 years people in institution
Belgium	National register of individuals	Head of household	All household members	< 6 years, collective household
Czech Republic	Census	Individual	Only selected individual	< 15 years
Spain	Statistical National Institute community register	Household	All household members until four persons	No age limit
Italy	National register of households	Household	1 person drawn randomly	People in institution; < 18 years
New Zealand	Census	Household	All household members	People in institution

Sources: survey by LET and INRETS-DEST + Kunert et al., 2001

Sampling Techniques

In most countries, the size of the sample correlates with how often the surveys are performed. In countries where the survey is ongoing the sample size is smaller. In countries where the survey is not performed as frequently the sample size is greater. The table below summarizes the sample sizes used in each country.

TABLE 11: Sampling

Country	Sampling method	Sample size
USA	Random sampling with geographical stratification (telephone sector)	26,038 households + 44,779 households (over-sampling in some States)
Norway	Random sampling without stratification	2005: 17,000 individuals
Sweden	Random sampling with geographical stratification at region level	11,000 individuals (2003 communication survey) 41,000 individuals (2005 transport survey)
Denmark	Random sampling	25,000 individuals before 2002 14,400 individuals in 2002 13,200 individuals in 2003
Finland	Random sampling	18,250 individuals (1999)
Great Britain	Random sampling with geographical and car ownership stratification Sampling applied every month	2003: 8,300 households (19,500 individuals) 1989 to 2001: 3,500 households
The Netherlands	Random sampling without stratification	2004: 43,202 households
Germany	Random sampling with geographical stratification	26,000 households 62,000 individuals (2002)
Austria	Random sampling with geographical stratification	32,211 individuals (1995)
Switzerland	Random sampling with geographical and socio-demographic stratification	29,000 individuals (belonging to 28,000 households): in 2000
France	Random sampling with car ownership stratification	14,200 households
Belgium	Random sampling with geographical, household structure stratification	3,064 households, 7,037 individuals (1999)
Czech Republic	Random, with quotas for gender, age, education and income based on census	1,000 individuals
Spain	Random sampling with geographical, household structure stratification	2000: 23,635 households (62,473 individuals)
Italy	Random sampling with geographical stratification	21,779 individuals (2004/5)
New Zealand	Census meshblocks random selection. Over a 7/8 years cycle every household in the selected meshblocks will be surveyed	2,200 households per year

Sources: survey by LET and INRETS-DEST + Kunert et al., 2001

Non-response

"A considerable number of studies (Bonnel, 2003; Ampt, 1997; Sammer, 1995) have shown that travel behaviour is generally correlated with a failure to respond or with the number of attempts that are required in order to reach a respondent. Even worldwide, little research has been undertaken to investigate non-response in the context of transport surveys (Richardson and Ampt, 1993; Richardson, 2000; Ampt, 1997)."

The recommendations outlined for the 2007 French survey by Bonell et al. based on the comparison study of travel surveys conducted around the world is summarized below.

- The survey needs to address the access of mass transit, and personal vehicles.
- Changes in the survey mechanism must be gradual over time to ensure that they are comparable from year to year.
- The data collected will be synthesized using CAPI systems.
- "The travel survey will be conducted over 12 months to account for seasonality."
- A survey day will be allocated to each household rather than a random day picked by the interviewer.

The 2007 French NPTS will involve 4 survey instruments:

- *For the household or all its members: a questionnaire administered by the interviewer using a CAPI system which records the socio-demographic characteristics of the household; commuting habits (home-to-work, home-to-study or to nursery); biography for the oldest member of the household; driving license holding and driving practices, traffic accidents; season tickets and price reductions for public transportation; a description of the vehicles available to the household; a description of the residential environment;*
- *A vehicle diary (for a car, motorcycle or bicycle), kept by the driver(s) of one of the household vehicles for 7 days (it could be selected with unequal probabilities: e.g. secondary vehicles were over-represented in 1993 because of an interest in the potential market for electric vehicles);*
- *During the second visit, an individual "Kish" over 6 years old (selected at random; if he/she is under 12 years, an adult will help) will be asked to describe his/her trips on the previous day and during the previous Saturday or Sunday using a CAPI system, and,*
- *During the second visit as well, the same individual Kish, will be asked to describe his/her long distance trips in the last three months (from memory) with a CAPI system. The interviewer will make two visits in order to complete the questionnaire (about 2 hours in total). The "vehicle diary" is handed out during the first session and checked during the second. The two visits are separated by at least seven days. We are planning a follow-up of trips for a sub-sample of about 1,500 volunteers using a GPS receiver. The interviewer will hand over the GPS receiver during the first visit and collect it during the second one. Between these two dates, the GPS receiver will record the individual's trips.*

Madrigal and Monzon (2007) present the benefits and advantages of using an activity based diary rather than a trip based diary for personal travel surveys conducted for local and national travels in Madrid Spain. “The aim of this comparison is to check if the trip based approach of the present travel diaries in Spain is related with the low trip rate reported.” Based on the pilot study conducted in May 2006, the study concludes that higher response rates were evident among the activity based journals compared to the trip based journals.

The research methodology employed for this research project included an activity based travel journal that was based on a trip based travel historically used. This journal was then administered within the same sample population. The survey was conducted among two samples, one located in the densely population city center, and the other located in the suburbs in a sparsely populated area. For the purpose of this pilot study conducted the survey “define a trip as any travel movement along public routes for a particular purpose”.

Diary Design

The design of the diary consisted of two parts. Both diaries included a socioeconomic section gathering information about all members in the household. “Information such as age, sex, driving license, cars per household, employment status and educational level were asked for.” The second part of the diary consisted of information about mobility that was gathered by all household members over the age of four.

The trip based diary used was the previous diary implemented during the last Madrid travel survey. “This diary was presented in an A3 landscape format, with data running horizontally. It was presented in the form of a table taking up almost the whole of the printed sheet, and in which each row records information relative to one trip. This diary went with an instructions sheet showing how to fill it in and with an introductory letter about the survey and explaining its aims.”

The activity based diary was designed to look similar to the trip based diary; therefore it was printed on an A3 landscape format, with similar tables taking up an entire sheet. It is noted in the study that this format it not necessarily the best choice for a person to carry around with them, however pervious studies revealed that most people did not carry the diary with them despite the small size of the journal.

The diary layout came in a three-block design:

• *What was I doing? endeavoring to reconstruct all the activities carried out outside the home by the respondent throughout the day*

• *Where did I go? seeking to reconstruct the places where the activities were carried out and the time each activity started and finished*

• *How did I get there? taking up the thread of the trip questionnaire: stages, modes, parking, etc.*

The data in the first two blocks are written down vertically whereas in the last block, recording information specific to travel movements, not only is the horizontal run of the TB diary data regained but also the same structure for data reporting on the trip stages.

The Study Areas

In the pilot survey two zones with very different characteristics were picked. The first zone consisted of high density development close to the city center. The second zone was located in the suburbs in a less populated area. In addition to the geographic differenced of these two areas analyzed, the first zone (Zone C) has a more mature population compared to the second zone (Zone P). Additionally, zero households in zone C have more than one automobile per household, whereas it is typical that households in zone p have more than one vehicle per house.

Data Collection Procedure
The following stages were involved:

• *Telephone contact 1, designed for introducing the study and explaining what the collaboration sought consisted of. The idea of this was to recruit households that would be prepared to take part in the survey while taking advantage of the contact to make an initial data collection (household and members).*
• *Household visit 1 for handing out the documentation, explaining the concepts, questions and options contained in the travel diary, filling in any data on the questionnaires relative to household and member questions that could not be collected during the telephone recruitment contact and setting the day of reference for which the mobility data would have to be reported for all individual members of the household over four years old.*
• *Telephone contact 2 for collecting the mobility data.*
• *Household visit 2 (where necessary) for putting right any shortcomings contained in the information provided by respondents.*

The survey fieldwork was carried out over two weeks from May 19-28, 2006. The data reported corresponded to an average working day, so respondents were asked to report the travel information related to one day from Monday to Thursday marked as day of reference in the household visit 1.

Results
The two formats were evaluated on their trip-reporting rates, share of so-called immobile, trip purpose, modal share and rounding of time for the two selected transport zones in the urban area of Madrid. The data obtained of the samples have been expanded to the overall population of each zone.

The average mobility rate for those with the Activity Based travel survey was higher than the average obtained by the Trip based travel survey across both samples.

"The trip rate obtained per TB diary led the researchers to believe that mobility was virtually the same in zone C and zone P (2,29 against 2,21). The opposite occurred when changing of diary. A pronounced difference appeared between the two zones (3,59 in zone C against 2,83 in zone P). In other words, TB diary did not disregard a uniform mobility percentage for the entire environment; in the denser populated central areas highly representative percentages of the total real mobility could be being omitted."

Trip Purpose and Home Based Trips
Non essential trips are not being captured with the trip based diary as much as the activity based diary. Therefore non recurring trips, which are not routine everyday, are not being recorded.

Modal Share
In both zones a higher rate in non-motorized trips were recorded with the use of the "activity based" diary, however this mode was not the most neglected mode of the "trip based" diary. In the suburban zone the least reported mode was the public transport, unlike the center city zone which reported this mode the most. In the suburban zone the most reported mode was the use of the private vehicle. Additionally as expected this zone also reported the least number of home trips between trips. This portion of the study also highlighted that the activity based diary was less likely to miss some short, not recurring trips made since respondents were more likely to remember stopping by to see someone rather than the journey to get there, especially in the suburban zone since most trips are not home based and in a private vehicle, not taking up much time from there day. In the city center, it was more likely for respondents with the trip based diary to "forget" certain trips especially since in the city center public transport is very efficient and

wait times are quite short. It is more difficult to forget trips with the activity-based diary since respondents have already committed to the activity.

Conclusions

- Small differences in the design of diaries may have substantial impact on response rates and data quality. An increase of 46.40 % in the trip rate recorded was experimented with the activity travel diary, 56.77% in central zone and 28.5% in the outlying one. As was to be expected, the trips tending not to be reported with the TB diary correspond to non-essential mobility and to trips corresponding to run-on travel movements.
- Trips typically not reported with the trip-based diary correspond to non-essential mobility and to trips corresponding to run-on travel movements.
- Public transport was not reported in the trip-based diary in the city center zone, while private vehicle mode was not reported as much in the suburban zone.
- *An increase occurred in the daily mobility reporting as a result of the change in diary type used. This fulfilled the aim of the research, which was to show that the TB diary was one of the reasons affecting trip-rate underreporting.*

References

Bonnel, Patrick, Madre, Jean-Loop, and Jimmy Armoogum. (2007). National transport surveys: What can we learn from international comparisons. Presented at the 86th Annual Transportation Research Board Meeting in Washington, D. C., January 21-25, 2007.

Madrigal, Esther and Andres Monzon. (2007). Applying an Activity-based Travel Diary Compared to a Trip-based Travel Diary in both a Central and an Outlying Zone in Madrid. Presented at the 86th Annual Transportation Research Board Meeting in Washington, D. C., January 21-25, 2007.

Appendix II - Household Travel Survey: Conversation Summary with Select Agencies Listed in the Minnesota Survey Archive

This information was produced as part of a study to examine if the NYMTC region is within the nation-wide trend in conducting the Household Travel Survey (HTS). Comparisons of various groups listed in the Minnesota [HT] Survey Archive[1] identify seven agencies who have conducted multiple HTS within a 10 -12 year span. The MPOs for Washington DC and Chicago were also included since they are currently in the process of conducting their regional HTS. Each of these agencies were contacted and asked (i) the reasons multiple HTS were conducted, (ii) what were some findings, i.e. behavior shift, and (iii) whether the data was used to estimate new models. The agency, contact information, year the Household Travel Surveys were conducted along with a summary of the conversation is included in the table below.

Agency	Year HTS Conducted	Household Sample Size	Conversation Summary
Atlanta Regional Commission Contact: Guy Rousseau Phone: (404) 463-3274 E-mail: grousseau@AtlantaRegional.com	1991, 2001	**Household Population 2000 Census:** 3,053,952 **Sample Size:** 8,069 (.26%)	Atlanta Regional Commission (MPO) conducts HTS approximately one year after the U.S. decentennial Census in order to use the information as a source of validation for their HTS efforts. Recent surveys were conducted in 1991 and 2001. With the tremendous growth of the Atlanta region, the survey was conducted to capture the most current reality of the locale, including infrastructure improvements, such as new highways, rail, housing developments, additional households, income, trip chaining, telecommuting, as well as other typical demographic characteristics. Data from the 2001 survey was used to update the travel forecast model as an activity base model. The next HTS will be conducted in 2011.
California State DOT Contact: Greg Miyata Phone: (916) 654-5089 E-mail: greg.miyata@dot.ca.gov	1991, 2001	**Household Population 2000 Census:** 11,704,742 **Sample Size:** 17,040 (.15%)	California State DOT conducts HTS every ten years, approximately a year after the U.S. decentennial Census, in order to use the Census information as a source of validation for their HTS efforts. This is a state-wide HTS which serves all 58 counties in California and is conducted in part to aid smaller MPOs. This survey is also used to validate and update existing models. The travel forecast model re-haul in 2000 modified their existing trip-based model to an activity based model, an effort which the Southern California Association of Governments developed and then was promoted statewide after the 2001 HTS was conducted.
Chicago Metropolitan Agency for Planning Contact: Kermit Weis Phone: (312) 386-8820 E-mail: kwies@chicagoareaplanning.org	1990, 2007	**Household Population 2000 Census:** 3,500,000 **Target Sample Size:** 13,000 (.37%)	Chicago Metropolitan Agency for Planning (MPO) conducted their HTS in 1990 and is currently in the process of conducting an updated HTS for 2007. The information from this survey will be used to update the dataset in their travel forecast model. Industry standards, constituent expectations, MPO structural changes and the U.S. decentennial Census was also cited as reason to conduct this survey. The four step model is still employed at the MPO for this region, however the HTS format is activity based which will be compatible for activity based models to perform retroactive analysis. Concerns in the current survey effort include a low response rate especially in "hard to reach" segments of the populations, using 1990 survey methods which are less effective, and in-vehicle GSP units with dairy responses may not be in agreement.

[1] www.surveyarchive.org

Appendix II - Household Travel Survey: Conversation Summary with Select Agencies Listed in the Minnesota Survey Archive

Agency	Year HTS Conducted	Household Sample Size	Conversation Summary
Minneapolis & St. Paul Twin Cities Metropolitan Council Contact: Mark Filipi Phone: (651) 602-1725 E-mail: mark.filipi@metc.state.mn.us	1982, 1990, 2001	**Household Population 2000 Census:** 1,226,229 **Sample Size:** 6,386 (.52%)	Twin Cities MTC (MPO), conducted their survey to correspond with the U.S. decentennial Census. Surveys were conducted in 1970, 1982, 1990, and 2001. These surveys were used to calibrate and update their travel forecast models. Between 1970 and 1990 tremendous growth in the region estimated a 66% increase in auto travel by trip per person and household[2]. Current trends reveal 'auto availability' has surpassed a 1:1 ratio per person even though the trips per person, per household has flat lined. The four step model is still employed at the MPO for this region, however the HTS format is activity based which will be compatible for activity based models to perform retroactive analysis.
Sacramento Area Council of Governments Contact: Bruce Griesenbeck Phone: (916) 321-9000 E-mail: bgriesenbeck@sacog.org	1991, 2000	**Household Population 1998 Estimate:** 740,000 **Sample Size:** 3,942 (.53%)	Sacramento Area Council of Governments (MPO) conduct HTS every ten years to correspond with the U.S. decentennial Census. The 1991 HTS was conducted in conjunction with the California State DOT HTS, however in 2000 SACOG conducted their own regionally focused HTS. Currently, SACOG employs both the four steps and an activity/zone based model. In 1994 they recalibrated their four step model and in 2006 they adopted the activity based travel forecasting model. Concerns in conducting the next HTS include bias sampling and response rates. New HTS s in the future may entertain the use of focus groups, panel surveys, and other alternate methods along with the HTS to counterbalance the inefficiencies of HTS efforts.
Saint Louis, East-West Coordinating Council of Governments Contacts: Marty Altman, Lubna Shoaib Phone: (314) 421-4220 E-mails: marty.altman@ewgateway.org, lubna.shoaib@ewgateway.org	1990, 2002	**Household Population 2000 Census:** 968,533 **Sample Size:** 5,094 (.52%)	Saint Louis, East-West Coordinating Council of Governments (MPO) conducts HTS every ten years, approximately a year or two after the U.S. decentennial Census, in order to use the information as a source of validation for their HTS efforts. Recent surveys were conducted in 1990 and 2002. The 1990 survey was used to recalibrate the four steps travel forecast model. The 2002 survey was used to update the travel forecast model as an activity base model and additional collar areas were also included as part of the survey coverage. On-board surveys were also conducted in part to align with preferences for FTA funding. The data from the HTS and On-board survey was used to develop equations for their new model. Changes in infrastructure improvements such as new rail lines, highways, housing developments, sprawling communities and longer commutes were considered in the initiation of new HTS. The next HTS efforts will be revisited for 2012.

[2] 1990 Twin Cities HIS Report, Introduction, p.1

Appendix II - Household Travel Survey: Conversation Summary with Select Agencies Listed in the Minnesota Survey Archive

Agency	Year HTS Conducted	Household Sample Size	Conversation Summary
San Francisco Metropolitan Transportation Commission Contact: Chuck Purvis Phone: (510) 817-5700 E-mail: cpurvis@mtc.ca.gov	1990, 2000	**Household Population 2000 Census:** 2,466,000 **Sample Size:** 15,064 (.61%)	San Francisco, MTC (MPO) conducted their survey to correspond with the U.S. decentennial Census. Surveys were conducted in 1965, 1981, 1991, and 2000. Both the data from the 1990 and 2000 survey was used to modify their travel forecast models. The 1990 survey was used to recalibrate their four step model. The 2000 survey was used to re-haul the travel forecast model to an activity base model.
Tucson, PIMA Association of Governments Contact: Tom Cooney Phone: (520) 792-1093 E-mail: tcooney@pagnet.org	1993, 2000	**Household Population 2000 Census:** 821,712 **Sample Size:** 4,883 (.59%)	PIMA Association of Governments (MPO) conducted their HTS in 1993 and 2000. In 1993 the survey was conducted to gain a better understanding of the region's growth. In 2000 the survey was conducted to correspond with the U.S. decentennial Census and to update their travel forecast model. The new survey included new mode choices in compliance with FTA standards which differ from the 1993 survey. Their next survey effort in 2008 will be a purchase of the NHTS add-ons.
Metropolitan Washington Council of Governments Contact: Robert E. Griffiths Phone: (202) 962-3280 E-mail: reg@mwcog.org	1988, 1994, 2007	**Household Population 2000 Census:** 2,200,000 **Target Sample Size:** 10,000 (.45%)	Metropolitan Washington Council of Governments (MPO) conducted their HTS in 1988, 1994 and is currently in the process of conducting an updated HTS for 2007. This effort will be used to calibrate their existing four-step trip based model, however an activity based survey is being deployed in consideration of adopting a new activity based travel forecast model. The coverage of this deployment will extend beyond their region to account for extraneous traffic from other regions. MWCG has also partnered with Baltimore Metropolitan Council (MPO) to include an additional 4,700 households from their region. A longitudinal survey of 2,000 HH was conducted from 1988 to 2003. Results from this panel survey indicated no significant behavioral changes and reaffirmed findings from their HTS in 1994. Lessons learned from other MPOs implemented in this current HTS account for coverage measurement of populations who are "hard to reach". Strategies employed include using an *address based list* rather than the traditional *random digit dialing*, a 2-day *in vehicle* GPS add-on, and an over sampling of "hard to reach" populations. These populations include cell phone only households who tended to be younger and travel significantly, non-traditional housing units, mix-use, multi-unit HH, lower income, and Hispanic HH. If a phone number and address match could not be found then households would receive postal correspondences. This mailing includes 3 postcard reminders and a $50 incentive for completing the survey. If they received no response or a soft refusal from the HH then an in-person follow up would be dispatched. To date using these methods have yielded a closer to census estimate representation of households.