AN INNOVATIVE SURVEY DESIGN TO UNDERSTAND SUSTAINABLE TRAVEL BEHAVIORS- POTENTIAL OF A ROLLING SAMPLE SURVEY TO REPLACE TRADITIONAL HOUSEHOLD TRAVEL SURVEYS

FINAL PROJECT REPORT

by

Cynthia Chen[1], Hejun Kang[2], Anne Moudon[1], and Qing Shen[1]
University of Washington[1], University of Idaho[2]

Sponsorship
PacTrans and WSDOT

for
Pacific Northwest Transportation Consortium (PacTrans)
USDOT University Transportation Center for Federal Region 10
University of Washington
More Hall 112, Box 352700
Seattle, WA 98195-2700

In cooperation with US Department of Transportation-Research and Innovative Technology Administration (RITA)

Disclaimer

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
2012-M-0003	01539930	

4. Title and Subtitle	5. Report Date
An innovative survey design to understand sustainable travel behaviors -potential of a rolling sample survey to replace traditional household travel surveys	October 7, 2014
	6. Performing Organization Code
7. Author(s)	**8. Performing Organization Report No.**
Cynthia Chen, Hejun Kang, Jin-hyun Hong, Anne Moudon, and Qing Shen	3-624881
9. Performing Organization Name and Address	**10. Work Unit No. (TRAIS)**
PacTrans Pacific Northwest Transportation Consortium University Transportation Center for Region 10 University of Washington More Hall 112 Seattle, WA 98195-2700 University of Washington and University of Idaho (List matching agency name and address) WSDOT	**11. Contract or Grant No.** DTRT12-UTC10
12. Sponsoring Organization Name and Address	**13. Type of Report and Period Covered**
United States of America Department of Transportation Research and Innovative Technology Administration	Research 6/16/2012-10/6/2014 **14. Sponsoring Agency Code**

15. Supplementary Notes

Report uploaded at www.pacTrans.org

16. Abstract

We propose an innovative survey with rolling samples to address a major fiscal challenge faced by many MPOs. Faced with a small, but continuous budget, MPOs are increasingly unable to continue the current survey practice: conducting a large survey every 10 years. A rolling sample design also has other benefits over the current practice. Yet, for its implementation in household travel surveys, many questions exist. Some are technical issues while others are cost and procedural-related. **The primary purpose of this project is to understand these issues and provide recommendations for a future household travel survey with rolling samples**.

It is also expected that a rolling sample design can help us understand travel behavior better for the purpose of VMT reduction. By sampling participants living in very different neighborhoods, it can help us devise better VMT reduction strategies. **The second purpose of this project is to assess the potential of a rolling sample design in addressing the potential of land use and infrastructure related strategies for VMT reduction**.

The research will help transportation planners and analysts to proactively reposition their service in light of the changing budgetary environment by developing a new approach to travel surveys based on small samples but continuous enrollment. This new approach is also more consistent with the recent changes in data collection methods used by the US Census Bureau. Our empirical results demonstrate the reliability of a rolling sample and points to the potential of using it to replacing the traditional household travel surveys. We also demonstrate an innovative parcel-based sampling procedure to allow an empirical examination of the effects of built environments on transportation outcomes based on data collected from continuous enrollment.

17. Key Words	18. Distribution Statement
sustainable travel behavior, rolling sample, parcel-based sampling	No restrictions.

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

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

We want to acknowledge the support of Pactrans to this project. Without the funding, this project would not be possible. Pactrans' funding supported four graduate students (two at the University of Washington and two at the University of Idaho).

We are also grateful to the support from PSRC. Throughout the project, we have had numerous discussions with Neil Kilgrin and Matt Kitchen, who provided us with ideas, advice, comments and suggestions.

The success of the two-wave rolling survey was largely attributed to the tireless work of the survey team throughout the survey period from June 2013 to October 2013. The survey team is led by Cynthia Chen (UW) and has four members: Jin-hyun Hong (UW), Xiangyang Guan (UW), Hang Zhou (UI), and Ensheng Dong (UI). Every day during the survey period, every member of the team processed a number of issues related to survey administration, updated statuses on various items, answered questions from the respondents, and held team conferences so that every one was kept abreast of the most recent status on survey administration. Hang Zhou and Ensheng Dong of UI also developed an elaborate process of procedures to ensure the successful completion of the two waves.

The various chapters in this report are contributed by the following people:
1. Chapter 1: Cynthia Chen, Jin-hyun Hong, Hejun Kang, Qing Shen
2. Chapter 2: Cynthia Chen
3. Chapter 3: Cynthia Chen, Jin-hyun Hong, Xiangyang Guan
4. Chapter 4: Cynthia Chen and Anne Moudon
5. Appendix 1: Cynthia Chen, Jin-hyun Hong, Hejun Kang, Qing Shen

ABSTRACT

We propose an innovative survey with rolling samples to address a major fiscal challenge faced by many MPOs. Faced with a small, but continuous budget, MPOs are increasingly unable to continue the current survey practice: conducting a large survey every 10 years. A rolling sample design also has other benefits over the current practice. Yet, for its implementation in household travel surveys, many questions exist. Some are technical issues while others are cost and procedural-related. **The primary purpose of this project is to understand these issues and provide recommendations for a future household travel survey with rolling samples**.

It is also expected that a rolling sample design can help us understand travel behavior better for the purpose of VMT reduction. By sampling participants living in very different neighborhoods, it can help us devise better VMT reduction strategies. **The second purpose of this project is to assess the potential of a rolling sample design in addressing the potential of land use and infrastructure related strategies for VMT reduction**.

The research will help transportation planners and analysts to proactively reposition their service in light of the changing budgetary environment by developing a new approach to travel surveys based on small samples but continuous enrollment. This new approach is also more consistent with the recent changes in data collection methods used by the US Census Bureau. Our empirical results demonstrate the reliability of a rolling sample and points to the potential of using it to replacing the traditional household travel surveys. We also demonstrate an innovative parcel-based sampling procedure to allow an empirical examination of the effects of built environments on transportation outcomes based on data collected from continuous enrollment.

EXECUTIVE SUMMARY

Metropolitan Planning Organizations (MPOs) in the nation are responsible for providing updated information on the nation's travel patterns. The current practice is to conduct a travel survey every 10 years. This practice is not sustainable because it is increasingly difficult for MPOs to identify a large budget at one time. We propose an innovative survey design that will enroll survey participants continuously over time to address this issue. **The primary purpose of this project is to understand the issues related to a future implementation of a travel survey with continuous enrollment.**

It is also expected that this new survey design can help us devise better strategies to reduce VMT (vehicle miles traveled) by generating a sample of survey participants living in very different neighborhoods. **The second purpose of this project is to assess the potential of this new survey design in helping us identify strategies for VMT reduction**.

We conduct an extensive review of the relevant issues, a pilot data collection effort with the proposed survey design, and a detailed analysis of the survey process as well as the data collected.

The research will help MPOs to proactively reposition their service in light of the changing budgetary environment by developing a new approach to travel surveys based on small samples but continuous enrollment. It will also enable researchers to gain a much better understanding of the potential of designing a new methodology for empirical examinations of the effect of the built environment on people's travel behaviors, which will lead to better strategies toward a sustainable future society.

CHAPTER 1. ROLLING SAMPLE SURVEYS

1 INTRODUCTION

For over half a century, travel surveys have played an essential role in obtaining data required for transportation planning and policymaking. However, travel surveys are facing increasing challenges (Griffiths et al. 2000; Stopher and Greaves 2007; Ortuzar et al. 2011). In particular, the existing approach of conducting one large cross-sectional survey roughly every decade raises questions about the timeliness and reliability of the data, the dependability of government funding, the continuity of experienced and skilled staff, and the effectiveness of communication with survey respondents. To address these issues, an innovative survey design with rolling samples is proposed as an alternative to the existing scheme (Griffiths et al. 2000; Stopher and Greaves 2007; Ortuzar et al. 2011).

A rolling-sample design offers both promises and challenges. The primary purpose of this project is to understand these promises and challenges, and then provide recommendations for implementing future household travel surveys with rolling samples. The second purpose of this research is to assess the potential of a rolling sample design in helping us understand travel behavior better and thus improve our analytical tools for evaluating urban planning and transportation policy strategies for coping with congestion, air pollution, and climate change. An initial step toward our research objectives is to review the relevant literature.

2 PROBLEMS IN EXISTING TRAVEL SURVEYS

In most countries, conventional cross-sectional travel surveys are "one-off" exercises conducted once every ten years, and typically travel information for only one weekday is collected from each respondent (Ortuzar et al. 2011). This approach, however, has major shortcomings. First, it is not suited for providing timely information to help policy makers address current issues related to transportation. The low frequency of data collection is incompatible with the new environment within which travel surveys are conducted, which is characterized by policy makers' increasing demands for more frequent and reliable information (Griffiths et al. 2000). For example, a travel survey conducted five or ten years ago is unlikely to provide adequate answers to questions of the time, such as, what is the level of acceptability toward tolling in the metropolitan region?

Second, one-off travel surveys can produce unreliable data because they are especially vulnerable to the impacts of unpredictable events, and because they tend to mix long-term trends and short-term fluctuations (Ortuzar et al. 2011). Familiar events that unpredictable include strikes and extreme weather conditions, whereas common short-term fluctuations include economic booms or recessions.

Third, the existing approach often runs into major financial barriers. Conducting a large travel survey once every decade requires a large amount of funding reserved for a one-time undertaking, but securing a massive irregular budget allocation tends to be difficult (Stopher and

Greaves 2007). Indeed, the survey could be postponed or even cancelled if the political climate is not favorable for funding it (Ortuzar et al. 2011). As demands for high- quality data continue to increase, and rates of survey response continue to fall, the financial burden that comes with such a travel survey is getting heavier and government funding is becoming less dependable.

Fourth, large but infrequent travel surveys are a discontinuous operation associated with the inevitable loss of experienced staff in the time period between surveys (Griffiths et al. 2000). This discontinuity undermines the capacity to develop and preserve technical and managerial skills in the conduct of travel surveys (Griffiths et al. 2000). It is costly to assemble and cultivate a team of skilled workers every time a survey is conducted.

Finally, infrequent travel surveys are especially ineffective as a means of communication between transportation planners and survey respondents. They often fail to obtain sufficient information about people's opinions about novel transportation solutions such as ITS and alternative-fuel vehicles, because many respondents may have little or no experience with these innovations when a survey is conducted (Griffiths et al. 2000), and because these respondents are not given another opportunity to answer the questionnaire once they have become knowledgeable about the new transportation technologies and services. The existing approach also reduces the opportunities for transportation planners to use the survey to help the public explore more sustainable travel options by informing them of innovative transportation technologies, services, and policies.

3 INTRODUCTION TO ROLLING SAMPLE SURVEYS

The rolling sample design was championed by Kish (Kish and Verma 1983; Kish 1990) for the census as well as the current population survey (CPS), a monthly survey on the nation's labor force. It was primarily motivated by the need to obtain information within the 10-year interval of the decennial censuses. In Kish's definition, a standard rolling sample design is one that selects k non-overlapping probability samples, each of which represents 1/F fraction of the population. Each sample is interviewed in a single time period until all samples are interviewed after k periods. Depending on the precision requirement as well as the area size, one or more samples together may provide estimates of a population. For greater precision or for smaller domains (e.g. smaller geographical areas), accumulations of larger numbers of consecutive samples can be used, up to k/F of the population; a rolling sample design with k=F is called a "rolling census" (Alexander 2002).

This innovative survey design has many potential advantages in comparison to the traditional approach. It holds the promise to make the data more current and incorporate topics of the time into updated questionnaire (Kish 1986; Griffiths et al. 2000). It can facilitate the adjustment of seasonal effects on certain types of travel, such as bicycling, walking, and making long distance trips for holidays) (Ortuzar et al. 2011). It can also rely on an experienced staff to keep the operation continuous and efficient, while directly addressing the budgetary concerns over the need to obtain an enormous amount of funding for a one-time large survey. There are other potential benefits too, including the capacities to provide better estimates for small communities (Kish and Verma 1983), and reach rare or hard-to-reach populations (Rust 2010).

Challenges also abound. While the concept of a rolling sample design was developed decades ago, it has only been recently implemented at a large scale. The most notable example of application in the U.S. is the American Community Survey (ACS), which has replaced the long-form component of the decennial census. For the implementation of a rolling sample design in household travel surveys, many questions and concerns exist. Some of the potential issues are technical in nature, such as questionnaire design, sampling, and inference; others are operational, such as survey administration and cost.

3.1 Pros and Cons

The most fundamental advantage of a rolling sample design is that it enables more frequent data collections and publications. Conventional surveys or censuses that are conducted at a 10-year interval typically provide data that are 2 to 12 years out-of-date (MacDonald 2006) and, therefore, do not provide timely information. In contrast, rolling-sample surveys are conducted continuously and, therefore, can provide data that are much more up-to-date. Accessing timely information is especially crucial for understanding the effects of projects and policies.

Conceptually, as long as an accumulation of samples provides a reliable estimate of a population characteristic, meaningful data are obtained (Alexander 2002). Higher levels of data currency can be achieved by increasing the frequency of surveys and/or the size of rolling samples.
A second, and related, advantage of a rolling sample survey is that it creates the flexibility to add topics of the time to the questionnaire and get answers quickly (Kish 1986). It can serve as a convenient and powerful tool for obtaining pertinent information for addressing current policy questions. For example, when tolling is being considered for implementation in a region, policy makers naturally are interested in understanding the public's opinions about tolling. With a rolling sample survey, relevant information can be obtained in a timely way by promptly updating the questionnaire for the next survey. For different policymaking contexts, different data accumulation schemes can be designed in corresponding to the most relevant spatial and temporal scales.[1]

It is important to understand that the flow of timely information is not one directional (Griffiths et al. 2000). Frequent surveys with current topics can also help inform the public about new ideas and options, such as policies for VMT (vehicle miles traveled) reduction and congestion relief, facilities for biking, and schedules of weekend transit services. Moreover, a rolling sample survey can be designed differently according to its purposes. For example, if one is interested in understanding people's travel behavioral change in response to tolling, panels[2] are needed. On the other hand, if the main question is only about the change of average commuting time in a region, we will need multiple samples over time, but the samples do not need to be panels. For surveys with multiple purposes, Kish (1986) proposed the Split Panel Design (SPD). SPD adds non-overlapping samples (e.g., *a-b-c-d*) to a panel *p*, which is thus symbolized by *pa-pb-pc-pd*.

---

[1] For example, one can use 1-year estimates to understand the level of acceptability toward tolling in the state of
[2] In panels, individuals are surveyed more than once.

Through this novel design, individual changes can be captured through the panel p; meanwhile, larger samples can be obtained by cumulating the non-overlapping samples, *a-b-c-d*, over time.[3]

A third strength of a rolling sample design is that it facilitates adjustments of seasonal effects. Rolling samples are evenly distributed into years, quarters, months, even weeks, and days. While transportation data collected on a continuous basis has tremendous values for planners and policy makers, there is concern that seasonal variability in variables could be continuously collected in a rolling sample survey. Using two transportation variables ("journey time to work" and "transportation mode selected for commuting") in ACS data (2000-2002), Parkany and Murakami (2004) compare monthly aggregations, quarterly aggregations, and the released 3-year estimates and concludes that there are no seasonal trends observed in the aggregated data products. This is supported by Ortúzar et al. (2011) who suggest that seasonal effects could be controlled in aggregated data provided by a continuous survey.

A fourth advantage of this sampling approach is that it lowers the budgetary barrier by substituting small continuous survey expenditures for a large one-time cost. The traditional decennial survey in the U.S. is overburdened with rising cost because of decreases in the mail response rate, increases in population and housing units, and increases in highly labor-intensive efforts (e.g., face-to-face interviews to reduce undercounts) (Edmonston 2001).[4] The enormous and rising one-time cost puts great pressure on budgeting. Different from the one-off survey, the cost of a rolling sample survey is much smaller at a given time and remains largely the same from year to year. In an era of continuing public finance austerity, it is politically easier to obtain a relatively small continuing budget allocation than a massive allocation at various points in time (Griffiths et al. 2000; Stopher and Greaves 2007; Ortuzar et al. 2011).

Closely related to the attribute of continuous budget allocation, a fifth advantage of a rolling sample design is that it can depend on a permanent, better-trained staff. Setting up a competent survey team requires substantial investment in recruiting, training, and development of the staff. Unlike the traditional one-off survey, where loss of experienced staff and institutional knowledge in the inter-survey period is unavoidable and must be re-developed for the next survey, a rolling sample survey continuously improves the technical and managerial skills of the staff (Griffiths et al. 2000). For this reason, it can gain significant economies of scale from undertaking the survey on a continuous basis (Ortuzar et al. 2011).

A sixth advantage of this sampling design is that it can improve data quality by reducing non-sampling error. One major reason is that sampling sample surveys typically employ permanent professional staff to follow up with non-respondents, which results in higher rates of survey completion (MacDonald 2006). Compared to lightly-trained temporary workers hired for one-time survey, skilled professionals can effectively use survey techniques to greatly increase the chance of obtaining information from unresponsive households. This helps to alleviate the

[3] SPD has some other merits. With the cumulated samples, it allows for checking and correcting biases which a panel survey may suffer if a respondent is unable or unwilling to keep accurate recording over time. Additionally, it can help recruit replacements for panel mortality and for panel renewals, by using the rolling samples.
[4] In the case of the U.S. Census, in 2000 dollars, the cost per household increases from $13 in 1970 to $56 in 2000, for a total of $7 billion. By 2010, the cost per household was projected to increase to $72 if the long-form census continued (Edmonston 2001).

common problems of higher nonresponse rates from minority groups and under-coverage of children in large households. Another reason for the reduction of sampling error is that a rolling sample survey requires the master address file to be regularly updated. Inaccurate address files have been a significant source of non-sampling error in one-time surveys. In particular, many hard-to-reach households (especially in rural areas) and new or converted units may not be added to the sampling frame in time and therefore tended to under counted (MacDonald 2006). This kind of bias is minimized by a current and more complete address file, which is expected for a rolling sample survey.[5]

Finally, a rolling sample survey has the additional benefits to provide better estimates for small communities and reach rare or hard-to-reach populations. Small communities traditionally rely on the decennial census for rather limited and infrequent descriptions of their demographic and socioeconomic characteristics. However, census data is insufficient for understanding many important aspects of small communities, such as residents' travel patterns and trends. One-time regional travel surveys often serve as another data source, but typically do not have a large enough sample to provide reliable data for small communities. In comparison, a rolling sample survey allows data to be cumulated over time and, thus, obtains more accurate estimates for small domains (Kish and Verma 1986). Of course, there may be errors in calculating an estimate based on multi-year samples, but still it is likely to be better than using estimates from census and one-time surveys.

Similarly, the new sampling design can facilitate data collection from small population groups (e.g., low-income minorities). Evidence from the U.S. indicates that responses are lower for tracts with high proportions of African American or Hispanic populations (Alexander 2003). To provide reliable estimates for small population groups, a rolling sample survey can employ an oversampling strategy, which assigns higher sampling rates to areas with lower response rates. In addition, a rolling sample survey with a continuously updated sampling frame can provide the basis for any follow-up surveys of relatively small or hard-to-reach populations (Citro and Kalton 2007).

A rolling sample design also has its known limitations and potential issues. Obviously, the resulting timeliness of data is achieved by increasing survey frequency and reducing sample size at any given time. This necessarily means larger sampling error. Depending on the sample size and the time period for which samples are cumulated, the resulting data may or may not be reliable. For example, some data may not be suitable for transportation modeling and planning.[6] As discussed previously, rolling sample surveys can reduce non-sampling error. Therefore, adopting the rolling sample approach will likely involve a tradeoff between sampling error and non-sampling error (MacDonald 2006). In general, the advantage of providing timely data is weakened for small geographical areas, because they must accumulate samples over a longer

[5] In the case of the ACS, for example, interviewers are equipped with handheld GPS units and can update addresses regularly and simply (MacDonald 2006).

[6] In the case of the ACS, for example, the annual sample of 3 million housing units leads to a five-year cumulative sample of roughly 12.5% of all addresses, compared to about 17% for the Census 2000 long form. According to Stopher & Greaves (2007), errors in the annual ACS data for 2000–2005 are quite large, which make the data not useful for transportation planning, and ACS data is significantly inferior to the Census long form data as far as transportation modeling is concerned.

time period to produce reliable data. Users from small geographic areas may have to wait for several years to obtain reliable data products, although samples are collected continuously on an annual, monthly, weekly, or daily basis.

Even the flexibility of adding topics of the time to the questionnaire of a rolling sample survey is constrained by the need for data accumulation. In particular, to obtain reliable estimates for small geographic areas, questions in the questionnaire should remain stable for samples to be accumulated. On the other hand, topics of interests invariantly change over time; and in today's fast-changing world, policy makers have very high demands for timely data for helping addressing a wide range of new questions (Griffiths et al. 2000). Apparently, some of these demands for current information, especially information concerning small geographic areas, cannot be adequately met.

Similarly, the benefits of having a frequently updated address file are not easily attainable in practice. While it is true that a current sampling frame is essential for conducting rolling sample surveys, there is no single convenient way to obtain timely and complete updates of addresses. Rather, it takes real efforts by data collection agencies and often requires collaborations with various organizations in both public and private sectors (Ampt and Ortúzar 2004; MacDonald 2006).

Another shortcoming of the rolling sample approach is that the survey becomes more complex and presents many technical challenges. The sampling usually requires a complicated design with multiple stages to draw samples appropriately and avoid repeating the same address within a cycle of surveys. Data accumulation across geographic units or time periods involves sophisticated weighting and adjustment schemes. The complexities will obviously increase if special efforts are made to provide better estimates for small communities and hard-to-reach populations, or to incorporate a panel component.

Cost and funding remains a major concern. Although budgetary consideration is a major reason for moving from a one-time survey to a continuous survey, budget constraints still apply. For instance, in the case of the ACS, a monthly sample size of 500,000 addresses was initially suggested to produce multi-year accumulations for small areas at the same level of data reliability as the long-form census sample (Alexander 1993), but the budget only allows 250,000 addresses monthly. In fact, one may argue that rolling sample surveys, such as the ACS, may cumulatively cost more than a traditional one-time survey (Edmonston 2001; Kincannon 2003). In addition, funding could be withdrawn at any time during a continuous survey (Ortuzar et al. 2011). Alternative avenues to seek funding, including user charges, therefore, are worth exploring (Ampt et al. 2009).

Also worth exploring is the role of technologies in rolling sample surveys. Can the Internet, GPS, and other technological tools help reduce the respondents' self-reporting burdens and improve data quality? Early studies suggested that GPS helps collect more accurate travel data than traditional data collection methods. It is especially effective in collecting data on travel time, distance, speed, and route (Battelle 1997a; Battelle 1997b; Hato and Asakura 2001), and in picking up short distance walk and bicycle trips (Stopher et al. 2008b). However, it appears to negatively affect survey response rate (Rofique et al. 2011). Much research work is needed to gain a better understanding of the new possibilities created by technological advances.

Finally, one should keep in mind that rolling sample surveys are not the only possible alternative to the traditional one-time survey. There are possibilities for taking different approaches to meet different demands for data and information, or combining some approaches. For example, in the context of transportation modeling, Stopher & Greaves (2007) suggests the use of national panels, supplemented with a continuous cross-sectional survey and a simulation exercise, to augment samples for transportation analysis in local areas. This is another exciting area for future research relating to rolling sample surveys.

3.3 Past Applications

A number of rolling sample surveys have been conducted around the world. They serve different purposes, and have different sampling designs. Four of these past applications, the American Community Survey (ACS), the Great Britain National Travel Survey (NTS), the New Zealand Household Travel Survey (NZHTS), and the Sydney Household Travel Survey (HTZ), are briefly reviewed here. The purposes and basic design features of these surveys are discussed. Because much more information, especially information about the technical details, is available for the American Community Survey than for the others, we place the detailed information about ACS in Appendix 1.

American Community Survey (ACS)

Purpose. The primary purpose of the ACS is to measure the changing social and economic characteristics of the U.S. population. Formally implemented in 2005, the ACS has replaced the long-form component of the decennial census. 1-, 3-, and 5-year estimates are produced from the continuously collected data. ACS data are similar to census long-form data in terms of subjects and geographic levels (with block group as the smallest unit for 5-year estimates), but more current than census long-form data.

Sample size and frame. The ACS sample consists of 250,000 addresses monthly, or 3 million per year. The sampling frame used for the ACS is an extract from the national Master Address File (MAF), which contains mailing and location addresses, geocodes, and other attribute information for all known living quarters in the United States. Addresses in the MAF are maintained dynamically through addition, deletion, or revision, based on information collected by the U.S. Postal Service's Delivery Sequence File (DSF), the Community Address Updating System (CAUS), and the Census Bureau's other household surveys (U.S. Census Bureau 2009).

Sampling method. The ACS sampling has two phases (U.S. Census Bureau 2009). First-phase sampling starts with assigning each census block, and its constituent addresses, to one of five sampling strata, each with a unique sampling rate. This is followed by a two-stages sampling process. In the first stage, the addresses on the sampling frame are assigned systematically to five distinctive subframes, each containing roughly 20 percent of the frame. Addresses from only one of these five subframes are eligible to be in the ACS sample for the current year and each subframe is used every fifth year. This ensures that no housing unit will be surveyed more than once within a 5-year period. In the second stage, a sample of the addresses in the current year's subframe is selected. Addresses in different sampling strata are selected with different sampling rates. Selected addresses included in the resulting annual ACS sample are then sorted by their order of selection and assigned systematically to the 12 months of the year.

During the second phase of sampling, a sample of addresses for which neither a mail questionnaire nor a telephone interview has been completed is selected for Computer Assisted Personal Interviewing (CAPI). The CAPI sample for each month is selected systematically after sorting within county by CAPI sampling rate, mailable versus unmailable, and geographical order within the address frame (Hefter 2005; U.S. Census Bureau 2009).

Questionnaire design. The ACS questionnaire includes four sections. The first section verifies basic address information, determines the occupancy status of a housing unit, and identifies who should be interviewed. The second section of the questionnaire collects basic demographic data: sex, age, relationship, marital status, Hispanic or Latino origin[7], and race. The third section collects information on physical and financial characteristics of housing (e.g., value, rooms, year built), and the final section collects population data (e.g., citizenship, education, employment status). Different instruments are provided for the three data collection modes (mail, telephone, and in-person interviews). For instance, since mail interview by far is the least expensive mode of data collection, the instrument for this mode is designed accordingly to maximize the rate of mail response through multi-mail contacts (U.S. Census Bureau 2009). Additionally, to maximize responses, an Internet response option will be added to the mail data collection phase on the 2013 ACS (U.S. Census Bureau 2012). For telephone interview, questions with long or complicated response categories are broken up into separate questions.

To address current needs, a rolling sample survey accommodates questionnaire changes. However, one constraint is that the ACS must be accumulated over time, to provide acceptable levels of reliability for small geographic areas. One obvious outcome of introducing changes is that the data will not be released for small areas. Therefore, content changes have to be minimized in order to produce consistent estimates across geographical levels.

Data collection method. The ACS uses three modes of data collection, mail, telephone, and personal visit, in consecutive months. First, mail-back responses are collected; then, non-respondents are interviewed by telephone; and finally, about one-third of the still remaining non-respondents are randomly selected for personal visits by interviewers. The computer-assisted telephone interviewing (CATI) usually begins about 5 weeks after the first mail package is sent out. For the housing units that have not returned the completed questionnaires by the end of the first month and that have an available phone number, an attempt is made to interview them by CATI during the second month. The personal interview phase, or CAPI, takes place during the third month. A sample of those who have not been interviewed by the end of the second month (including the unmailable cases) is selected for CAPI.

[7] This is a term used by the Census Bureau.

Great Britain National Travel Survey (NTS)

Purpose. The National Travel Survey (NTS), conducted since 1988, provides up-to-date information about personal travel within Great Britain over one week. It is designed to monitor long-term trends in personal travel and to inform transportation policymaking.
Sample size and frame. Beginning in 1988, the survey had an annual sample size of 5,040 addresses, which increased to 5,796 by 2001. Due to the limited sample size, 3-year estimates were provided for most analyses. Since 2002 the annual sample size has increased to 15,048 addresses, which makes annual estimates available with a greater degree of precision (Rofique et al. 2011). The sampling frame is drawn from the Postcode Address File (PAF), the most up-to-date and complete address database in the UK, which is updated daily, monthly, or quarterly.

Sampling method. The NTS employs a stratified two-stage random probability sampling method (Rofique et al. 2011). The first-stage sampling selects Primary Sampling Units (PSUs), based on a quasi-panel design. According to the design, half of the PSUs in a given year's sample are retained for the next year's sample and the other half are replaced, which has been proved to reduce the variance of estimates of year-on-year change (Rofique et al. 2011). Before the selection, all the PSUs are stratified using a regional variable (numbered 1 to 40), car ownership (i.e., the proportion of households with no car) and population density to ensure that the different strata in the population are correctly represented. Random samples of PSUs were then selected within each stratum.

The second-stage sampling selects 22 addresses randomly within each selected PSUs. While the same PSU sectors might appear in the next survey year, no single addresses are allowed to be included in consecutive three-year periods. This is achieved by excluding all addresses selected for the previous three survey years from the sampling frame before the addresses for this year were selected.

Questionnaire design. The survey asks detailed information about each trip, including origin and destination, purpose, mode, time, traveled distance, vehicles used, etc. It uses a 7-day travel diary, which has two versions, one for adults (16 years and over), and one for children (younger than 16). The main difference between these two versions is that adults are also asked about any parking costs, road tolls or indicating whether they were a passenger or driver, and children are asked whether any time is spent in the street (e.g. playing, talking with friends, etc.) (Rofique et al. 2011).

Data collection method. The NTS uses two data collection methods: face to face interviewing using computer assisted personal interviewing and self-completion of a 7-day travel diary. Interviewers begin fieldwork at the start of each month. Advance letters are initially sent to the sampled addresses. Then interviewers contact the households to arrange placement interviews. A placement interview is conducted with all household members (for children under 11, proxy information is collected) recording information about the household, each individual member, vehicles and any long distance journeys recently made. At the end, a 7-day travel diary is provided to each household member. The Travel Weeks are assigned to selected households to ensure they are evenly spread over the days of the week as well as the weeks of each month.

Finally, within 6 days of the end of the Travel Week, interviewers pick up the diaries and also check whether any key factors have changed since the placement interview, such as the purchase of a new car (Rofique et al. 2011).

New Zealand Household Travel Survey (NZHTS)

Purpose. Beginning in 2003, NZHTS is an ongoing survey of people throughout the country to collect data about their travel over a two-day period. The purpose of this survey is to provide travel data to facilitate the development of transportation policy and evaluation of road use and safety programs.

Sample size and frame. The sample size was 2,200 households per year between 2003 and 2007, and has been increased to 4,600 households per year since 2008. The sampling frame is developed based on the most recent national Census taken by Statistics New Zealand. The sampling frame is updated by field interviewers who record the street address of each household during each visit.

Sampling method. The New Zealand Household Travel Survey employs a stratified, two-stage sampling method. First, within each Local Government Region (a total of 14 regions in New Zealand), blocks of houses used for the Census (called "meshblocks") are selected randomly.[8] Then within each selected meshblock, all houses are listed in order and every seventh household on the list is sampled; over seven or so years all households in the selected meshblock will be invited to take part in the survey, and no household will be surveyed more than once within the 7-year period. For the next cycle, the survey will move to different blocks and the process will start over again. The sample sizes per Region are largely proportional to the Census populations (New Zealand Ministry of Transport 2006).

Questionnaire design. There are two questionnaires used: one to gather basic information about the household, and the other to record a 2-day trip diary and alcohol usage for each household member.

Data collection method. The survey uses only personal interviews, which are conducted by a fieldwork team consisting of 20-25 interviewers nationwide. Each sampled household is assigned two consecutive days ("travel days") for collecting travel data. The travel days are assigned to make sure that samples are evenly distributed throughout the week (including weekends). Before the designated travel days, the household is sent an introductory letter outlining the survey and advising that an interviewer will call. During the week prior to the travel days, an interviewer visits the home and explains the survey to the household members. In addition, the interviewer leaves a "memory jogger" for each member of the household to record travel on the designated travel days. Soon after the travel days, the interviewer returns at an agreed time to conduct a personal interview with each household member. The interview includes questions about the

---

[8] The meshblock is the smallest geographic unit for which statistical data is collected and processed by Statistics New Zealand. It varies in size from a city block in urban areas to extensive tracts of land in rural areas.

travel in the memory jogger, crash history, alcohol consumption, and other personal characteristics.

At least four attempts (made at different times of the day) are made to reach people to minimize non-response bias. With the personal interview method, a fairly high response rate of 66% has been achieved (New Zealand Ministry of Transport 2006).

Sydney Household Travel Survey (HTS)

Purpose. The HTS is a continuous travel survey that has been conducted in the Greater Sydney Metropolitan Region since 1997. The main purposes of this travel survey are to understand the trends in travel behavior and collect data for modeling and forecasting travel patterns in the region (Battellino and Peachman 2003).

Sample size and frame. The population in this region is approximately 4.6 million and the survey is designed on a three-year cycle. To be representative of the target population, over 5,000 households are randomly chosen each year to participate in the survey. The cumulated sample size collected over three years is comparable to that of the previous decennial travel surveys (12,000 households, approximately 1 percent of population in 1991) (Battellino and Peachman 2003). Moreover, the HTS does not draw exactly rolling samples, and therefore some households may be interviewed more than one time within the three-year cycle. However, the number of such households is very small.

Sampling method. The HTS employs a stratified, three-stage cluster sampling method and the sample is drawn every quarter. In addition, all travel zones are sampled in 3 years. In the beginning, the entire metropolitan area is divided into 80 Statistical Local Areas (SLA) and samples are selected through the following stages. In the first stage, a sample of Census Collection Districts (CDs) is selected based on the Probability Proportional to Size (PPS) method that secures an approximately equal chance of being selected. In the second stage, those CDs are divided into blocks of 50 dwellings and one of those blocks is randomly selected. In the last stage, starting from a randomly selected point in the selected block, every 7th dwelling is chosen until 7 dwellings are included. Then selected households are randomly assigned to different days of a preselected week of the year.

Questionnaire design. The survey includes three questionnaires: household, person and vehicle information, and a 24-hour trip diary for all household members, which are comparable to Sydney 1991 home interview survey. Essentially, the questionnaire remains very similar over time to achieve a high level of comparability. However, sometimes researchers need information about new topics at certain time periods. For example, the increasing use of real time transit information is new and it will have a significant influence on people's travel behavior. Including this extra topic in the survey will provide substantial benefits to planners and policy makers. With a periodic survey, small modifications to collect supplementary data can be implemented at marginal costs (Ampt et al. 2009). Practically, supplementary surveys can be added to the core survey at different times to collect interested information at a particular time period while ensuring the comparability of the data in the main questionnaire. In the case of the HTS the questionnaires are reviewed at the end of each wave, which is defined as one year. If there are any new needs for other information, changes are made. The first supplementary survey was added in the fourth wave (2000) to investigate the impacts of the internet use on trip.

Data collection method. The HTS adopted the personal interviewing method to collect data. Transport Data Centre (1997) conducted an investigation on what survey method is the best based on diverse criteria including response rate, data quality, and cost. They compared six different options based on three survey administration modes and two diary types. The three modes are: face to face interviewing, drop off/mail back self-enumerated questionnaire, and mail out/mail back self-enumerated questionnaire; the two diary types are: activity diary and trip diary. A telephone survey method was not used since it provides a relatively low response rate when questions are long and complicated. They found that face to face interview method with a trip diary ("memory jogger") is the most suitable compared to other alternatives. Their recommendation, based on the research findings, was accepted by New South Wales Department of Transport (NSW DOT).

Summary and Comparison

The basic design features of these four rolling sample surveys are summarized in Table 1. To better understand rolling sample surveys, these existing applications are compared to highlight several similarities and differences.

A basic similarity, of course, is that all four are continuous surveys. They are all designed to obtain information about changes and trends, placing a major emphasis on the currency of the survey data. In addition, all four surveys employ sophisticated multi-stage sampling approaches, reflecting the complexities they commonly face in designing a rolling sample survey.

Furthermore, all four surveys produce data on some multi-year cycles, and a three-year cycle is commonly used. Also commonly observed in these surveys, but not shown in the summary table, a relatively small but well-trained professional staff is employed for field work, and the quality of field work is monitored.[9]

There are also significant similarities between sub-groups of these surveys. Notably, three of the rolling sample surveys, with the exception of ACS, are travel surveys. Also worth noting is that, in multiple cases, there is clear capacity for making questionnaire changes in responding to current issues. Additionally, it is particularly interesting that the New Zealand Household Travel

[9] In the Sydney case, the survey management staff holds a regular bi-monthly team meeting with interviewers, and any issues arising in the field and suggestions from interviewers are discussed and retraining is performed. In the New Zealand survey, all interviewers receive three-day training, and a supervisor checks data collected by each interviewer by visiting three households randomly selected from the interviewer's workload. In Great Britain, the NTS interviewers have to complete their own travel records for one week before attending a two-day training which covers all aspects of the survey (Rofique et al. 2011). For the ACS, in addition to training interviewers, supervisors are required to travel with interviewers to reinforce the procedures learned in the training session. Moreover, interviewers are randomly selected each month for supervisors to re-interview a sample of assigned cases, thus verifying that interviewers are conducting interviews and doing so correctly (U.S. Census Bureau 2009).

Survey and the Sydney Household Travel Survey share many features. Specifically, both the NZHTS and the HTS have a sampling design that is basically stratified two-stage cluster sampling, and employ face-to-face interview as the survey mode, with a travel diary ("memory jogger") playing a facilitating role.

There are major differences among these four surveys. One basic difference is in the purposes they serve: the ACS aims at obtaining data about the demographic, economic, and housing characteristics of the population, whereas the other three surveys narrowly focus on collecting data about people's travel patterns and trends. This difference also implies that the survey questionnaires for the ACS are likely to be quite different from those for the others.
A second significant difference is observed in the spatial and temporal accumulations of the resulting data. The ACS is quite distinct in that it produces at one-, three-, and five-year estimates, and the various temporal accumulations are associated with the different geographical levels for which data are published. ACS data are available across geographical levels from the whole country down to individual block groups, for which 5-year estimates are provided. The other three surveys have simpler data accumulation schemes, and cover smaller geographical areas. In the case of the HTS, the geographical coverage is a single metropolitan region. These differences partly explain the much larger sample size for the ACS.

A third major variation is seen in the sampling approaches. Even though all four surveys employ rolling samples, they employ different sampling designs and methods. The sampling for the ACS is especially distinct in its simultaneous use of strata for applying different sampling rates to different blocks and subframes for partitioning household units into samples for different years. The sampling design for the Great Britain National Travel Survey is also quite unique in that it is quasi-panel. However, the available literature provides no insightful information about the relative merits of the alternative sampling approaches.

Finally, different survey modes are employed in the four cases. Again, the ACS is quite distinguished in its heavy reliance on mail and telephone interview. For the three travel surveys, personal interview is the primary survey mode, with travel diary serving as an alternative or only playing a supporting role in personal interview. However, it is not quite clear how the different survey modes affect the survey results, although conceptually it can increase response rates by providing respondents with alternative survey modes.[10]

[10] Providing survey respondents with alternative modes of data collection and offering an incentive conditional upon the completion of every survey section are often suggested as means for increasing response rates. However, the available literature does not establish a clear relationship between survey mode mix and survey response rate. In Great Britain, the response rate of the NTS dropped from 80% in 1989 to 59% in 2011. In Sydney, the response rate of HTS decreased from 75% in 1997 to 63% in 2004. The New Zealand HTS has a relatively stable response rate of 66%, but it has been lower than in previous one-off surveys (Ortuzar et al. 2011).

Table 1 Basic Features of the Four Rolling Sample Surveys Reviewed

Survey	Purpose	Survey modes	Organization	Sampling design	Sampling methods	Starting year & Sample size (per year)	Estimates
American Community Survey (ACS)	To measure the changing demographic, socioeconomic, and housing characteristics of the U.S. population.	Mail, telephone, personal visit	U.S. Census Bureau	Two-stage, two-phase sample design; stratification with five sampling strata based on geographic entities	First phase: Systematically assign all addresses to five subframes (each subframe contains about 20% of total frame) and determine subframe for the current year. Select a subset of addresses from designated subframe. Second phase: A sample of nonmailable and nonresponding addresses are selected for personal interview.	Began in 2005. Annual sample includes 3 million addresses (approximately 3% annually, cumulating to 15% over 5 years).	One-, three-, and five-year estimates
Great Britain National Travel Survey (NTS)	To track long-term trends in personal travel and inform transportation policymaking	Personal interview and travel diary	U.K. Department for Transport	Stratified two-stage random probability sample (quasi-panel design)	Primary Sampling Units (postcode sectors) are stratified based on regional variable, car ownership, and population density. The primary sampling units within each stratum are randomly sampled. Addresses are then sampled within selected units.	Began in 1988. Annual sample currently includes about 8,000 households.	Three year cycle

Table 1 Basic Features of the Four Rolling Sample Surveys Reviewed (Continued)

Survey	Purpose	Survey modes	Organization	Sampling design	Sampling methods	Starting year & Sample size (per year)	Estimates
New Zealand Travel Survey	To facilitate the development of transportation policy and evaluation of road use and safety programs.	Face to face interviews and travel data recording using a "memory jogger"	Ministry of Transport	Stratified, two-stage sampling, with 14 Local Government Regions as strata	Meshblocks (basic geographic units) are sampled independently with probability proportional to size of the strata. Then addresses of houses are randomly sampled from within selected meshblocks. Households in selected meshblocks take part in the survey only once within a 7-year period.	Began in 2003. Annual sample initially included 2,200 households, but has increased to 4,600 households since 2008.	Three-yearly moving average
Sydney Household Travel Survey (HTS)	To understand trends in travel behavior and collect data for modeling and forecasting travel patterns in the region	Face to face interview and travel data recording using a "memory jogger"	NSW Department of Transport	Stratified, two-stage cluster sampling	Census collection districts are first selected, and then a block of 50 dwellings is randomly sampled from each census collection district. 7 dwellings are randomly drawn from each block and allocated to randomly selected survey week.	Began in 1997. About 5,000 households are interviewed per year.	Three year cycle

Chapter 2. Rolling Sample Surveys for MPOs

1 INTRODUCTION

For many Metropolitan Planning Organizations (MPOs), conducting household travel surveys regularly constitutes a major task. Households travel surveys serve two important purposes: 1) to provide a current picture of how people in a region (or nation) travel between places distributed in space; and 2) to provide data for travel demand modeling. As discussed in Chapter 1, household travel surveys face many challenges. Thus, an important question in the context of rolling sample surveys is whether it can address some of the existing challenges. This chapter attempts to answer this question. We do not intend to provide a comprehensive assessment on each aspect as such an assessment has been done at the conceptual level in Chapter 1. Instead, using the 2006 household travel survey in the Puget Sound Region as an example, we describe PSRC's specific challenges and discuss the implications if a rolling sample survey were to be conducted.

2 HOUSEHOLD TRAVEL SURVEYS: PROVIDING A SNAPSHOT OF THE CURRENT TRAVEL PATTERNS

Household travel surveys serve an important purpose, that is: to understand a region's travel patterns. This means answering the following five questions:

a) How many trips are people making?
b) Where are people going from and to?
c) When are people traveling?
d) What modes of transportation are being used?
e) What routes are chosen?

Answering the above five questions means obtaining estimates on the following variables:

a) Trip rates (often expressed as daily number of trips per person/household by purpose),
b) Spatial and temporal distributions of the region's trips, or O-D table (origin-destination table)
c) Mode splits, and
d) The loaded networks.

Estimates of these variables are typically obtained from a regional household travel survey during which a sample of the region's population is interviewed and activities and trips for each person in the selected households are recorded over a time period typically between 24 and 48 hours[11]. The levels of accuracy and precision of these estimates depend on a number of features

[11] Most of the travel surveys capture data for a 24-hour period. A few collect data for a 48-hour period, for example, the Puget Sound Regional Council.

associated with the survey including representativeness, sample size, and the amount of sample bias.

3 HOUSEHOLD TRAVEL SURVEYS: PROVIDING DATA FOR MODELING

In addition to providing an accurate snapshot of the region's travel patterns, household travel survey serves an important purpose of providing sufficient data for the development of travel demand forecasting models. In the Puget Sound region, PSRC develops, operates, and maintains two set of model systems: DaySim and UrbanSim. DaySim is an activity-based travel forecasting model system and UrbanSim is a land use model. Under each model system, there are various models and the needs and the requirements for each model may vary substantially. In this section, we discuss a number of near-term improvements that are needed on the two models.

3.1 Sensitivity of the Models

Travel demand forecasting models are often used to assess the sensitivity of the travel patterns in response to a change in the general environment, e.g., a new toll on a corridor. In order to evaluate how people may react to these changes, the models must be sensitive such that changes in travel patterns can be detected.

A critical requirement for a model to be sensitive is that there must be sufficient variations in the variable that measures changes in the general environment, which may be a new policy, a new transit alternative, or a change in the land use. For instance, in order to measure the responses to the new implementation of a toll on a corridor, the amount tolled must be both sufficient and varied: an amount too small may not be able to trigger a change in travel patterns or a fixed toll amount only allows the observation of a single behavioral change, rendering the sensitivity analysis impossible.

In many cases, the identification of a change in the general environment is difficult. Thus, the analysis is often conducted to a cross-sectional sample of the population who are exposed to different setups of the general environment. Under certain assumptions (Kitamura 1990), the estimates derived from such analyses can be interpreted as how travel patterns may respond to a change in the general environment. One example of such cases is the examination of the effect of the built environment on travel behavior: because it is in general difficult to identify a change in the built environment and be able to track people's behavioral changes over time, this built environment effect is often identified by comparing the behavioral patterns of people living in different environments (e.g., urban vs suburban). It is thus critical that the survey sample includes a sufficient number of people living in different types of areas with different characteristics.

In reality, the requirement that there is sufficient variation in the variable of interest (e.g., enough number of people living in different kinds of the built environment) is often not met in household travel surveys (Moudon and Sohn 2005; Moudon et al. 2009; Moudon et al. 2011; Moudon et al. 2011). In the case of the evaluation of the built environment effect, it is to some extent because of the way sampling is conducted. Respondents in household travel surveys are typically sampled through geographical locations and socio-economic and demographic characteristics, such as residence in county and household income. Because only a small portion of the population live in highly dense and diverse areas in the United States, sampling procedures through geographical and socio-demographic characteristics typically result in samples that contain too few people living in those areas. This means that behavioral differences between people living in different types of areas cannot be entirely observed. On the other hand, the use of geographical and socio-demographic attributes in sampling in a regional household travel survey is an important step that ensures the representativeness of the sample in those characteristics (geographic and socio-demographic), when the sample is compared to the census count.

3.2 Residential Self-selection

Amid the many studies that sought to understand the built environment effect on travel behavior, a central concern is residential self-selection. Residential self-selection refers to that people self-select into neighborhoods that would support their lifestyles. Thus, a household that prefers using transit will self-select into a neighborhood with transit services; a household who intends to drive will care little the availability of transit services in its residential location choice; and a bicyclist will want to live in a bicycling-friendly neighborhood. Because of the residential self-selection effect, the built environment effect observed by simply comparing the travel behavioral patterns of people living in different neighborhoods is likely over-estimated: one who prefers driving will continue to drive even if he/she lives in a transit-friendly neighborhood, not to mention the fact that he/she is unlikely to move into such neighborhood. This also suggests that it is important to account for residential self-selection when assessing the built environment effect.

Different methods have been proposed to account for the residential self-selection effect. These methods can be divided into two categories: 1) via experimental design; and 2) via statistical methods. The most ideal way to eliminate residential self-selection is via a randomized trial in which households are randomly assigned to different neighborhoods and their behavioral changes are recorded before and after the assignment. In reality, this randomized trial cannot be done in the context of housing choices. Therefore, probably the best way to guard against residential self-selection is via the use of mixed methods: via both experimental design and statistical methods. A combination of a natural experiment and some statistical model is probably the next best method after the “impossible” randomized trial (Shadish et al. 2002). This method could be ideally applied to a real-world scenario in which there is a change in the built environment (e.g., addition of a light rail), or so-called “natural experiment”. Here, two groups of households similar in their socio-economic and demographic characteristics can be selected: one group living close to the light rail and the other far from the light rail and their travel patterns before and after the light rail can be tracked and compared. Statistical models are then applied to both groups to estimate the built environment effect.

To make sure that the two groups differ in terms where they live, sampling must be done with a set of built environment characteristics, along with their socio-demographic attributes. Yet, the use of the built environment characteristics is not advisable to the entire sample for a regional household travel survey, since it complicates sampling matters and can result in unrepresentative samples. Furthermore, it is likely unnecessary since the interest here is to identify a suitable sample (with sufficient variation in built environment characteristics) to accurately estimate the effect of the built environment on travel behavior. In other words, while representativeness is required in accurately providing a snapshot of the travel patterns in the region (the first purpose of a regional household travel survey), it is not necessary for model estimation (the second purpose of a regional household travel survey).

3.3 Residential Location Choice Models

PSRC currently operates two residential location choice models: one is a relocation choice model that determines whether a household will relocate or not; the other is a location choice model that predicts where a household will decide to relocate. The last 2006 household travel survey has a set of questions relating to household relocation. During the interviewing process for the main sample, these questions are asked conditionally, only if a household indicated that it has relocated within the last 10 years. Answers to these survey questions are essential for the development on both residential relocation and location choice models.

The current data collection on variables related to residential location choices is confined within the main sample, thus it does not require additional effort for sampling. The extra survey administration effort on the recently relocated households is minimal, since it only requires administering a few extra questions within the main interview process. Consequently, the existing method for collecting variables for residential location choice models is simple and easy to implement. The downside is that one hardly has any control in the quality of sample as well as the variables collected. Since on average, only about 20% of the American families relocate within any year (Dieleman et al. 2000) and the selection of the main sample does not explicitly account for relocation-related factors, the resulting relocating sample from the main sample can be small. For example, the 2006 PSRC household travel survey has a total of 4,746 completed households in the main sample and 2,537 households indicated that they have relocated within the last 10 years and only 790 households (or 16.6%) moved into current locations less than 2 years ago.[12] With the current method, one also has little control over the quality of the variables collected. A relocation may happen between two very similar neighborhoods, or two very different neighborhoods; the latter case is not only much more interesting than the former but also provides better data in understanding people's relocation choices. Using the 2006 PSRC household travel survey, Table 2 shows location characteristics of the current and the previous homes for 780[13] households who moved within 2 years (between 2004 and 2006).

[12] If we consider households who moved less than 3 years ago, the total sample size becomes 1,132.
[13] 10 households were removed because they did not provide information on the current or previous homes.

Table 2 Neighborhood changes of households who relocated between 2004 and 2006 (statistics calculated based on the 2006 Puget Sound Regional Household Survey)

Previous Home Location	Current Home Location							
	Urban		Suburban		Rural/Exurban		Total	
	Number	%	Number	%	Number	%	Number	%
Urban	237	*30.38*	87	11.15	29	3.72	353	45.26
Suburban	83	10.64	210	*26.92*	33	4.23	326	41.79
Rural/Exurban	35	4.49	33	4.23	33	*4.23*	101	12.95
Total	355	45.51	330	42.31	95	12.18	780	100

From Table 2, we can see that the majority of these 780 households (62%) relocated to the same types of neighborhoods. This means that only 296 households were observed to have moved between neighborhoods of different types. A sample of 296 households is considered quite small in empirical data analysis, especially for a sophisticated discrete choice model for residential location choices. This simple frequency analysis shows that the current sampling method using geographic and socio-demographic attributes may not serve well for modeling households' residential location choices due to insufficient number of households who reside in different types of neighborhoods.

3.4 Pre- and Post Evaluations of Transportation Projects or Programs

Various transportation projects and programs being implemented today often need to be evaluated. A recent example in the Puget Sound region is the SR520 tolling project. While volumes on SR520 and its alternative routes such as I-90 and SR522 can be and are being closely monitored by WSDOT and guesses may be made about some behavioral changes, for example, an increase on the cross-bridge transit ridership[14] after tolling may be interpreted as people's making mode shifts. It is however unknown whether these guesses are right and how accurate these guesses are. Some of the behavioral changes cannot be observed by comparisons of aggregate-level traffic counts and transit ridership and such changes may include time of day changes, location changes, or simply eliminating trips altogether.

Behavioral changes may also occur as the result of land use (or built environment) changes, as shown in many studies (Chen et al. 2008; Chen and Chen 2009). An example in the Puget Sound region relates to the light rail project, in which case, people may shift to use transit and walk more since their accessibility to light rail or transit services in general is improved.

Understanding such behavioral changes is important not only from a model calibration point of view, but also from a policy recommendation perspective. To do this, one needs a panel dataset, in which the same households are tracked over time, from pre- to post-changes in the programs, policies, or land use. Such data is often hard to obtain. But with a continuous rolling sample

[14] Since tolling, there has been a 15% increase in cross-lake transit ridership according to Kevin Desmond (Nov. 14, 2012).

survey, the collection of the panel data may become easier since one may be able to tap on a sample of households who have already done the pre-change survey.

4 THE 2006 PSRC HOUSEHOLD TRAVEL SURVEY

4.1 Survey Design

There were three major data collection activities in the survey: (1) Activity and Travel Survey; (2) GPS Tracking; and (3) State-Preference (SP) Survey. The 48-hour Activity and Travel Survey collects basic demographics, activities, and tour and travel characteristics from each member of respondent households. The GPS tracking survey collects information on the travel paths of selected households in the same 48-hour period recorded in the Activity and Travel Survey. The SP survey is a follow-up survey after the main diary survey on people's attitudes and perception regarding travel in the Puget Sound region and a stated preference survey on choices between car and transit options, and between tolling and non-tolling options.

Approximately 133 households were recruited each weekday and weekend from 03/20/2006 to 06/04/2006. The total time is 11 weeks and the total number of households recruited is about 8,800.[15] This sample size was calculated based on an expected response rate of 56% and 7% additional for the replacement of households due to incomplete data (4,600*1.56*1.07 = 8,789).

4.2 Sample Size

The recommended sample size for the two-day household activity/travel survey is set at 4,600 households. This includes a main sample, a transit rider oversample, and a transit access oversample (

Table 3). Among these samples, a total of 220 household with vehicles were randomly selected for participation in the GPS tracking survey. Following the household survey, the SP survey recruited 800 individuals from the 4,600 households who completed the main travel survey.

Table 3 Recommended Sample Sizes by Survey Activity

Survey Type	**Final Household Sample Size**
Household Survey	4,600
RDD (Random Digit Dialing) Main Sample	3,600
Transit Rider & Transit Access Oversamples	1,000
Geographically Targeted Transit Access	800
Transit Rider Oversample – Park-and-Ride	150

[15] Every week, nearly 800 households (133 times 6 days is equal to 798, or nearly 800) are recruited. Six days includes 5 weekdays and 1 weekend day.

Transit Rider Oversample – Ferry Intercepts	50
GPS Tracking Survey (from main sample)	220
Stated Preference Survey (from main sample)	800

Source: Recreated from the PSRC 2006 Household Activity Survey Analysis Report

4.3 Sample Frames and Survey Administration

Four sampling frames are used in PSRC's household travel survey. They are described as follows:

Frame 1: RDD Main Sample

The main samples (3,600 households)[16] were collected from a geographically stratified RDD (random-digit-dial) household frame through two steps. First, the samples are allocated to each of the five PSRC service territories (see Table 4) , which are proportional to the total number of households as reproted in Census 2000. The sample sizes in the City of Seattle and Kitsap County (shown in bold) are disproportionally larger than other regions for precision purpose. Second, to ensure the representativness of the RDD sample, all households in each service territory are stratified by household size (1 to 4+), the number of household vehicles (0 to 3+), the number of workers (0 to 3+), and household income. All of the variables were obtained from the 2004 American Consumer Survey[17]. These variables were chosen because of their close association with travel behavior and travel patterns.

[16] The number of completed households is 3,937 households.
[17] This is indeed not the American Community Survey.

Table 4 Sample Allocations for the PSRC Service Territory RDD Sample

County	2000 Census Households	Percent of Households	Proportional Sample Allocation	Disproportional Sample Allocation
King	710,916	55.4%		
City of Seattle	258,499	20.1%	725	**900**
King without Seattle	452,417	35.3%	1,269	900
Kitsap	86,416	6.7%	242	**400**
Pierce	260,800	20.3%	732	750
Snohomish	224,852	17.5%	631	650
Totals	1,282,984	100.0%	3,600	3,600

Source: The PSRC 2006 Hsouehold Activity Survey Analysis Report.

MORPACE International Inc planned and conducted the 2006 Household Activity Survey for PSRC. First, a pre-recruitment letter was sent to selected households and then households were recruited by telephone. Households who agreed to participate in the survey received a package that included a cover letter, one or more activity diary, the instruction, and a postage-paid return package by mail. In the evening before the 1st travel day, each household received a reminder call. Person and diary information were collected over the phone in the evening following the assigned travel days. Retrieval interviews continued for five days until the Computer Assisted Telephone Interviewing (CATI) recorded all members' information. Respondents who did not want to proceed with the data collection phone interview could mail back their completed diaries. The information of respondents under 16 and members who could not be reached otherwise can be reported by proxy interviews (from adults in the household). For 20% of persons 18 or older, travel information was recorded by proxy interviews.

Frame 2:Transit Access Oversamples

To increase the number of transit riders in the sample, the sampling design oversamples households within certain zip+2 codes. Targeted zip+2 geographic areas are determined through the following steps:

1. Select zip+2 codes where transit options (access) are currently available;
2. Create a transit density map from PSRC's modeling network, based on the percentage of workers by each block group;
3. Overlay the selected zip+2 codes with the transit density map;
4. Select the Zip+2 areas based on density levels and the amount of overlap between the density geography and the Zip+2 geography.

A total of 1,724 Zip+2 areas are targeted. The distribution of the targeted Zip+2 geography across the Puget Sound region is presented in Table 5. Transit access oversamples are then allocated proportional to the size of targeted geographic areas within each county. For instance, King county has 80.2% of all targeted geographic areas in PSRC, the total sample size in King county is equal to 642 (800*0.802). Households are randomly sampled from this frame.

Table 5 Distribution of Targeted Zip+2 Areas

County	Total ZIP+2 Areas in a County	Zip+2 Areas In Targeted Transit Access Area	Percent of County	Percent of Total Targeted Zip+2 Areas
King	4,381	1,382	31.55%	80.2%
Kitsap	905	10	1.10%	0.6%
Pierce	2,882	183	6.35%	10.6%
Snohomish	2,218	149	6.72%	8.6%
Total	10,386	1,724		100.0%

Source: The PSRC 2006 Household Activity Survey Analysis Report.

Transit access and Transit rider oversamples: Households were randomly selected from targeted geographic areas (Areas were selected based on geographic proximity to specific transit-supported corridors).

Frames 3&4: Park-and-Ride Transit Rider and Ferry Rider Intercept Oversamples

Park-and-Ride transit users and ferry riders represent a unique subgroup from a sampling perspective, but their incidence is too low to expect enough for analysis purposes from the RDD or transit access frames. Therefore, to supplement this frame, intercept interviews were done at selected park-and-ride lots and ferry terminals, as presented in

Table 6 and Table 7. A total of 600 park-and-ride interviews and 200 ferry ride interviews were conducted. Those who agreed to participate were contacted by phone to participate in the standard recruitment interview process. The total completed sample sizes for transit rider and transit access oversamples were 809 households. It was implemented with the same procedure as the main sample but the respondents were recruited by interviewers at the parking/ferry lots.

Table 6 Park-and-Ride Intercept Sites

County	Location	Lot Name	Number Recruited
King	14200 SE Eastgate Way, Bellevue	Eastgate	80
King	I-5 & NE 65th St, Seattle	Greenlake	90
King	James St & Lincoln Ave., Kent	Kent Transit Center	90
King	32320 23rd Ave. S, Federal Way	Federal Way	90
Kitsap	SR 303 & McWilliams, Bremerton	McWilliams	50
Pierce	South 19th & Mildred, Tacoma	Tacoma Community College	100
Snohomish	202nd Street SW & 46th Ave. W, Lynwood	Lynwood Transit Center	100
Total			600

Table 7. Ferry Rider Intercept Sites

County	Location	Lot Name	Number Recruited
Kitsap	201 1st Street, Bremerton	Bremerton	50
King	Pier 52, 801 Alaskan Way, Seattle	Seattle Pier 52 Terminal	50
Snohomish	199 Sunset Ave. S, Edmonds	Edmonds Terminal	50
Kitsap	11264 State Route 104, Kingston	Kingston Terminal	50
Total			200

Park-and-ride Transit Rider Intercept Oversample: Interviewers collected contact information of commuters who were waiting for buses at park and ride lots. Those who agreed to participate were contacted by telephone following the recruitment at the Park-and-ride lots.

Ferry Rider Intercept Oversample: This oversample used the same procedure as the Park-and-Ride Transit Rider Intercept Oversample.

4.4 GPS Tracking Data

A total of 220 households with at least one vehicle were randomly selected from the RDD main sample frame. Up to three vehicles per household were equipped with GPS units. The objective was to compare the data results from GPS tracking of trips made with household vehicles with the diary trips reported by those household members. The GPS tracking data was analyzed to better understand the underreporting of trips and to provide insight into potential biases in the data.

The data for this sample was collected over 12 weeks between March 21st and June 16th in 2006. Households were randomly selected from the main sample. ECONorthwest (ECO) was subcontracted to conduct the GPS survey. It was responsible for programming the GPS devices and post-retrieval software. The survey was conducted over 12 weeks with four possible two-day travel periods (Mon-Tue, Tue-Wed, Wed-Thu, and Thu-Fri).

One package that included GPS devices, travel diaries, postage-paid return packaging, a thank you letter, and a single sheet of simple instructions was shipped to each participating household two days before their survey travel dates by FedEx's "Priority Overnight". That is, each household received a package one day before the first day of travel dates so that they could read the instructions and install the devices. The households can return the devices by dropping the package off at a FedEX location or contacting MORPACE to arrange a courier pickup. Sometimes participants did not return their diaries with devices. For this case, participants can verbally transmit all diary contents to a MORPACE representative via telephone or send back their diaries.

4.5 Stated Preference Survey.

The SP survey was conducted as a follow-up to the 2006 Household Activity/Travel Survey. Eight hundred households were targeted[18]. Respondents (age 16+) of the household survey were selected, based on trips reported in the household activity survey. In general, respondents were selected if their trips met certain criteria including the length of the trip and the location of origin and destination points, within defined geographic corridors of transit access and/or potential toll alternatives. For all the selected respondents, approximately one-third of the sample was provided with transit alternatives choice experiments; one-third received choice experiments related to toll usage, and the final third received choice exercises related to both. Data from the SP survey would provide input to activity choice models, and are used to better understand the market for transit services.

The SP survey was conducted between 07/12/06 and 09/25/06. Samples were selected based on their revealed trips (as noted earlier, the criteria included the length of the trips and the locations of origins and destinations, within defined geographic corridors of transit access and potential toll alternatives). The selected households were phoned to inform them of their selection for the follow-up survey. A survey packet containing attitudinal and perception related questions and four choice experiments was mailed to each participant and respondents mailed back their survey. Respondents who did not respond by mail were encouraged to respond by phone.

5. ADDRESSING MODELING ISSUES IN THE CONTEXT OF A CONTINUOUS SURVEY

5.1. Improving the Residential Location Choice Model

In the context of a continuous survey, both the quality of the sample and the variables collected for the residential location choice model development can be improved. Instead of administering these questions within the interview process for the main sample, special efforts may be made for sampling and survey administration. A sampling frame of all relocating households can be built

[18] The final completed number of households is 916.

by either accessing public records on real estate transactions or simply purchasing from commercial companies (e.g., Experian). Based on the addresses (previous and current addresses) contained in this sampling frame, one can calculate built environment characteristics associated with the previous and current addresses and classify households into different categories, for example, those who relocated from a suburban environment to an urban one or vice versa. An example of capturing relocating households using lists purchased from commercial companies can be found a number of studies: Chen and her students estimated multiple residential relocation models that assessed the impact of the built environment attributes at the previous addresses on current relocation choices (Chen et al. 2009; Chen and Lin 2011; Chen and Lin 2012). In the case that only current addresses are provided, one can also identify households living in neighborhoods with very different built environment characteristics. Identifying households living in different neighborhoods prior to the sampling will allow the researchers to sample a sufficient number of households from different categories and thus build a dataset that contains enough variations on key variables. Models built on datasets with large variations on key variables are more robust than those on datasets with little variation. In addition, data records offered by commercial companies often also include information on additional variables such as household size, income level, and marital status. These variables can be used to control potential influences brought by these variables.

In the actual survey administration, an optional module may be created containing only residential location choice related questions. Unlike the core module that is administered to a rolling sample continuously during the study period, an optional module is only administered to a special sample to which this optional module is specifically targeted. An optional module can be administered only once, multiple times to the same sample or different samples, depending on the purpose of the study at hand.

5.2 Timely Policy Evaluation

Probably the biggest disadvantage associated with the traditional regional household travel surveys is the fact that they are often conducted on a 10-year cycle and thus cannot provide answers to the effectiveness of many transportation programs and policies in a timely fashion. Nor can they be used to assess the impact on the region's travel patterns in response to a change in the general environment.

In the context of a continuous survey, assessment on the changes in the region's travel patterns in response to a change in the general environment or a particular transportation program/policy can be made in a timely fashion. The continuity nature of the rolling survey will allow us to tap into the already established and ongoing recruitment, sampling and survey administration process. A simple way to do this is to utilize the existing sample and add additional relevant questions as an optional module to the existing questionnaire.

6. SAMPLING

6.1. Sampling Frame

As rolling sample surveys are conducted on an ongoing basis, an updated sampling frame is necessary. Updating the sampling frame is most ideally done at an annual basis, coinciding with the pace of rolling samples. Since ACS is also a rolling sample survey now and the ACS sampling frame is continuously being updated, it is desirable to directly use ACS' updated sampling frame. Other sources of data that can be used for updating include postal services and telephone and utility companies.

6.2. Sample Size Allocation

6.2.1. Main Sample

As noted earlier, a standard rolling sample design is one that selects k non-overlapping probability samples, each of which represents 1/F fraction of the population. Each sample is interviewed in a single time period until all samples are interviewed after k periods. Depending on the precision requirement as well as the area size, one or more samples together may provide estimates of a population. What should the k value be? The determination of k depends on our answer to the question: how long data can be assumed to be current and usable for model updates? An obvious answer is 10 years, since current household travel surveys are conducted about 10 years apart. However, in light of the five-year cycle that is adopted by ACS as well as the expected changes in the population and employment in the Puget Sound region, we also recommend a 5-year cycle instead of a 10-year one. Assuming a 5-year cycle and a fixed total sample size of 3,600 households for the main sample, Table 8 shows the annual targeted sample sizes for each of the five geographical areas in the Puget Sound region.

Table 8 Summary of annual targeted sample size allocation for the main sample in the PSRC region (assuming a five-year accumulation cycle)

County	Year 1 Count	Year 2 Count	Year 3 Count	Year 4 Count	Year 5 Count	Cumulative 5-year Total Count
King						
City of Seattle	180	180	180	180	180	900
King without Seattle	180	180	180	180	180	900
Kitsap	80	80	80	80	80	400
Pierce	150	150	150	150	150	750
Snohomish	130	130	130	130	130	650
Totals	720	720	720	720	720	3600

Given that the number of housing units is in general on the rise, a drawback of a fixed sample size is that over time, the standard error of the estimates from the accumulated samples will decrease. One way is fix this is to target for a fixed sampling rate, instead of a fixed sample size and this may account for about 1-2% increase in annual sample sizes. Assuming a 1% increase in annual sample size, Table 9 shows an example of those sample sizes for the five areas in the

Puget Sound Region. In this example, the total sample is 4,396 households, 796 more households than the 3600 households when the fixed sample rate is applied.

Table 9. Summary of annual targeted sample size allocation for the main sample in the PSRC region (assuming a five-year accumulation cycle)

County	Year 1 Count	Year 2 Count	Year 3 Count	Year 4 Count	Year 5 Count	Cumulative 5-year Total Count
King						
City of Seattle	180	198	218	240	264	1,099
King without Seattle	180	198	218	240	264	1,099
Kitsap	80	88	97	106	117	488
Pierce	150	165	182	200	220	916
Snohomish	130	143	157	173	190	794
Totals	720	792	871	958	1,054	4,396

A key attribute of a rolling survey that is different from a one-off survey effort is that there is an accumulation of data over time. In other words, even though there is data being collected continuously, only at the end of the study period, the sample size of the rolling sample survey is the same as or similar to that for the one-off survey. This potentially poses challenges for modeling. Some advanced models may require a large sample size that the first year of data collection cannot afford, if sample sizes are equally allocated across all years (e.g., see Table 8). One way to address this potential issue is via front-loading, or allocating a larger sample size for the first year than for latter years. Front-loading may result in additional complications in weighting if one wants to report on the data collected about a region. However, since the primary purpose of PSRC's household travel survey is to provide data for model development, front-loading is not expected to create a significant problem. Exactly what proportion of the sample should be front-loaded in the first year is determined by the sample requirements for the various models developed by PSRC.

When samples are being rolled from one year to another, the simplest way of sampling is just to randomly select a sample (the targeted sample size for year t times the inverse of the expected response rate) from the entire corresponding geographical area. For example, assuming a response rate of 20%, in year 1, 900 households will be sampled throughout the city of Seattle. In year 2, 990 households will be sampled throughout the city of Seattle. The potential downside is that a single household may be sampled more than once during the 5-year cycle. However, given the large population size, such probability is likely small.

An alternative way is to first divide the entire sampling frame into five subframes and each subframe will include parts of city of Seattle, King county without Seattle, Kitsap county, Pierce county, and Snohomish county. Then, the address of each household in the sampling frame (population) is randomly assigned to one of the five subframes. Each of these five subframes is randomly assigned to year t_1 through year t_5 of specified 5-year periods (for example, 2006-2010, 2011-2015, and so on). This guarantees that no household will be surveyed more than once within the 5-year period, since households for each yearly sample are selected from the corresponding subframe.

6.2.2. Transit oversample

In addition to the main sample, an additional 1,000 households living in geographically targeted transit access areas need to complete the survey. The same procedure as described above can be adopted: assuming a response rate of 20%, every year 1,000 households will be randomly sampled from these geographically transit access areas. To avoid a single household to be surveyed more than once in the 5-year cycle, one can also first create five subframes and randomly assign all households in the transit access areas to each of the subframe. Then, each subframe will correspond to a single year in the five-year cycle.

6.2.3. GPS survey and Stated Preference (SP) survey

In 2006, PSRC also conducted a GPS survey and a SP survey. 220 households participated in the GPS survey and 800 households participated in the SP survey. These 1020 households are not in addition to the main sample or the transit access oversamples. They are recruited from the main sample after satisfying certain criteria. In other words, after completing all survey requirements in the main sample, these 1,020 households, in addition, completed either a GPS survey or a SP survey. In this case, one can just simply identify 220 households and 800 households (assuming a 20% response rate) every year for the GPS survey and the SP survey, respectively.

7. SURVEY ADMINISTRATION

7.1. Survey Instrument

It is expected that PSRC will use the existing survey instruments as a starting point. This is particularly the case for the survey instruments used in the main sample, containing memory jogger/travel diary and that for the collection of socio-economic and demographic attributes of the households. For the purpose of both maintaining stability over time (such that sample sizes can be accumulated) and allowing additional timely questions to be answered, we propose a two-module survey instrument. One core module includes a core set of questions that remain exactly the same over time. Such questions should include those on the households' socio-demographics and the members' travel diaries. Other modules are optional, containing additional questions that can be asked to an annual cycle or to a limited number of cycles.

7.2. Survey Administration

7.2.1. The use of RDD

Unlike the one-off survey that only has a single sampling frame, a rolling survey has multiple sampling subframes to be used at different times throughout the study period. In other words, households listed on different subframes are to access the survey at different times. This means that the conventional sampling method of RDD (random digit dialing) is likely no longer applicable and a new sampling method (e.g., address-based) may be more appropriate. The primary advantage of RDD is the reduced reliance on having a sampling frame as complete as possible. In RDD, the complete sampling frame comprises all possible combinations of 0 to 9 (excluding certain kinds of numbers such as business numbers and cell phone numbers), given area codes or some prefix digits (which are used to identify a particular geographical area). The probability that a number is being selected is the same as that of any other number and the survey recruitment process stops when it reaches the desired number of households to contact. In other words, one cannot pre-allocate the numbers to be called at different times. Unless the sampling frame is sufficiently small, one may not be able to guarantee that a household will not be contacted at different times during a rolling survey. If the probability that a household may be contacted multiple times is sufficiently large, undesirable consequences may result—some households may be contacted repeatedly over time, resulting in biases in the data and possibly reduced trust by the public.

One may propose to generate a complete sampling frame that contains all possible numbers and then divide it into multiple subframes to be used at different times during a study period. This solution, though is theoretical feasible, also faces some practical difficulties. A major concern that the usability of such randomly generated numbers may decline significantly over time, since more households are switching from landlines to cell phones only.

7.2.2. A 5-year Schedule

Figure 1 shows a possible 5-year cycle that may be adopted by PSRC. Counting from the initial time for start-up and pilot survey to the time needed to analyze and report on the data collected in the 5th year, this 5-year cycle actually is about 5 years and 10 months. This 5-year cycle has 5 sub-cycles. Each sub-cycle is about 16 month long, including a planning period for the first two months, continuous data collection for 12 months, and the final two months for analysis and reporting. In other words, planning for the next sub-cycle's data collection coincides with the final two months of the data collection for the current sub-cycle; and the analysis of and reporting after the 12-month continuous data collection starts when the next sub-cycle of the data collection starts.

									Year 1												Year 2											
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Start-up months and pilot	Start-up				Pilot																											
Yr 1							Planning		Continuous data collection												Analysis reporting											
Yr 2																			Planning		Continuous data collection											
Yr 3																															Planning	

	Year 3												Year 4												Year 5												Year 6	
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2
Yr 2	Analysis Reporting																																					
Yr 3	Continuous data collection												Analysis reporting																									
Yr 4											Planning		Continuous data collection																									
Yr 5																							Planning		Continuous data collection												Analysis reporting	

Figure 1. A possible implementation schedule of the household travel survey for PSRC

7.3 Use of Technologies

The use of GPS technologies to accompany or even in place of paper surveys is an increasing trend in household travel surveys. In this section, we discuss a number of technology-related issues that may arise in the context of a rolling survey.

Household travel surveys have been traditionally relied on self-reporting facilitated by paper-based memory joggers or activity and travel diaries and telephone interviews. It has been recognized that self-reports under-report trips, in particular, non-motorized trips (Wolf et al. 1999; Wolf et al. 2003). There is also respondent fatigue, limiting the duration of most surveys to one or two days. The use of technologies in surveys, for example, GPS loggers or using smart phones, will potentially lengthen the survey period (Stopher et al. 2008d; Chen et al. 2010; Gong et al. 2011), reduce respondent burden (Stopher et al. 2008d; Chen et al. 2010; Gong et al. 2011), and improve data quality (Stopher et al. 2007; Stopher et al. 2009). Since a rolling survey is supposedly rolled over multiple samples over time, a complete paper-less household travel survey may result in a number of complications in survey administration logistics and costs.

Thus, in the context of a rolling survey, a number of technology-related questions may be posed and we list a few below:

- What if we completely eliminate the paper survey?
- How to carry out a paperless survey?
- What other information or what shall be done to ensure the completeness of the records to support the model development?"

The above questions need to be answered before one starts a rolling survey.

7.4. Cost considerations

Information on costs comparing that of a one-off household survey vs a rolling sample survey is mostly absent from the existing literature. Recently, the City of Calgary in Canada conducted a feasibility study of the rolling sample survey in which costs are considered. The study examined four options (Nustats 2012):

- Option 1: start and operate the survey program completely in-house, internal to the city;
- Option 2: contract out for data collection and cleaning, while data analysis and reporting are conducted internally by city staff;
- Option 3: start with a pilot implementation in which the data collection would be contracted out for the first two years and transitioned to a completely internal city program in the third and subsequent years if the program proves successful;
- Option 4: contract out a one-off household travel survey on the 10-year cycle and conduct separate, smaller surveys conducted every three years to provide transportation data for program monitoring.

To estimate the costs of the above four options, the study made a number of assumptions relating to response mode (e.g., CATI vs mail vs online), recruitment mode, retrieval mode, the percentage of the households that need reminder messages and call back to retrieve (better) data.

The estimated costs include staffing costs, other direct costs (printing, postage, incentives, and travel costs etc.), start-up costs that would only incur if the survey were done internally (e.g., cost of facilities and equipment, CATI software, geocoding costs, initial programming and testing costs), and costs unique to the external contractors (e.g., soft costs usually included in the contractors' bid such as contingency costs for unexpected occurrences). The study concluded that though the rolling sample survey option (the first three options) is more expensive than option 4, the four options studied are in fact quite similar with each other. The difference in costs between the first three options and option 4 ranges from $381,175 to $563,576. The largest cost difference within the first three options is $180,640. It is worthy to note that the sample size for the first three options is slightly larger (10,000 households accumulated in five years) than that for the one-off survey (9,000 households).

Chapter 3 Application: a Two-wave Pilot Survey

1 MOTIVATION

As part of this project, we conducted a two-wave pilot survey. The objectives of the survey are two-fold:

- To empirically test the feasibility and reliability of a rolling sample survey; and
- To understand the impact of the built environment on use of non-motorized transportation modes such as walking and cycling. To capture a sufficient number of non-motorized trips, we apply the parcel-based sampling technique.

In the rest of this chapter, we describe the sampling procedure, survey instrument design and administration, and results of the survey.

2 SAMPLING

2.1 Study area and sampling

Our study area, King County in which Seattle is located, is one part of the Puget Sound region in Washington State, U.S., and is surrounded by the Cascade mountains to the east, the Olympic mountains to the west, several large Lakes, and ocean bays and inlets. Its unique geographic characteristics with a moderate marine climate can easily trap air pollutants, causing a serious air quality problem. Based on the fact that transportation is one of the major sources of emissions, diverse land use and transportation planning strategies including Growth Management Act have been adopted to reduce air pollution and promote non-motorized transportation system.

Our survey was conducted from June 11, 2013 to October 9, 2013 with the combination of mail out/mail back (survey booklet) and Internet modes. Due to the limited budget, we decided to limit the geographical scope to around the SR520 bridge that crosses Lake Washington to connect Seattle and the east side where many technology companies including Microsoft are located. Based on the mailing lists purchased through USPS services, our target population comprises 6091 households in the study area (Figure 2).

A parcel-level database was used to develop the sampling frame for this study. For each of the 6091 households, we calculated residential density as the number of housing units within a half-mile buffer around the household's home location. Addresses and contact information were purchased through a commercial company. We then sorted these density measures into three groups by using two tertiles: low, median and high density neighborhoods. We dropped the middle group and retained the high and low density groups to compare the travel behaviors of those living in high-density area vs those living in a low-density area. Our low- and high-density areas are identified as those with a density lower than 7.95 units per acre and higher than 13.29

units per acre, respectively. Since there are only 150 single family households in the sample, we decided to include all of them by assigning households above and below the median density (9.88 units per acre) to the high and low density groups, respectively. This results in 102 single family households being assigned to the low density group and 48 households to the high density group. Finally, we randomly selected 1,500 households from high density neighborhoods and other 1,500 households from low- density areas and administered the survey. A total of 547 workers from 412 households completed our survey. This initial recruitment of 3000 households is obtained based on three assumptions: (1) our target sample size for this study is 100 households; (2) 90% of the USPS mailing addresses are correct; and (3) 4% of the households contacted will be willing to participate and finish the survey.

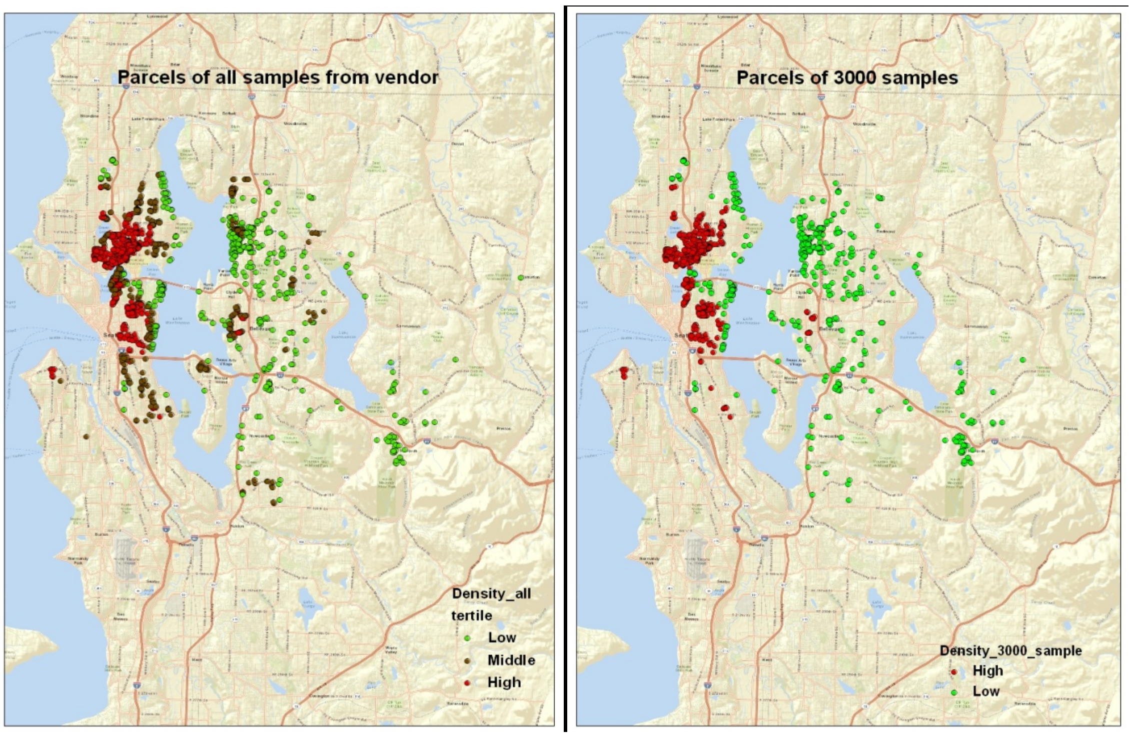

Figure 2 Spatial distribution of residential parcels based on high and low density tertiles. Left panel: all parcels obtained from the USPS mailing list (6,091unique households. Right panel: final sample for the Pilot Survey (3,000 unique households)

2.2 Sampling frames for the 2-wave survey

As discussed in Chapters 1 and 2, a rolling sample design requires fractions of the total sampling frame to be implemented continuously over time and each fraction is a random sample. Thus, for the 2-wave survey, we divide the sampling frame of 3000 households randomly into two sub-frames, each containing 1500 households. Note that each of this 1500-household sub-frame is expected to contain an approximately equal number of households living in high-density and low-density areas. In principle, statistically similar results from the two waves would demonstrate the validity of a rolling sample survey.

3 SURVEY INSTRUMENT DESIGN

3.1 Identifying important variables

Selection of variables and survey questions is crucial to the success of our survey. In order to identify relevant variables for our pilot survey, we have first reviewed empirical studies to examine possible factors affecting travel behaviors. We also reviewed empirical survey instruments that have been validated in the field.

Table 10 provides a list of factors and variables found important in previous studies involving trip frequency and mode choice, which are directly relevant to our study.

In addition to the above studies, Olsson (2003) conducted a comprehensive literature review and classified six factors that have critical impacts on mode choice: 1) Transportation-specific factors (i.e., travel time, fare, parking cost and availability, punctuality, information, travelling time, timetable, frequency of service, proximity to stops and stations, congestion charges, and service level); 2) Environment-specific factors (i.e., topography and access to other facilities); 3) Individual-specific factors (i.e., age, gender, attitudes, lifestyle, security, status, and habits); 4) Trip-specific factors (i.e., purpose of trip and trip chain); 5) Policy factors (i.e., taxes and restrictions); and 6) Quality factors (i.e., safety, reliability, and security).

For the purpose of our study, we focus on three categories of factors: socio-economic and demographic factors, perception/attitudes toward neighborhoods and transportation modes, and travel activities.

3.2 Survey Instrument

Questions related to *socio-demographic characteristics* are mostly built based on three sources such as the Pedestrian and Bicycling Survey (PABS), American Community Survey (ACS) and 2006 PSRC Household Activity Survey since their survey instruments have either been validated with the test-retest method (Forsyth et al. 2012) or is viewed as one of the authoritative surveys in the U.S. The questions are slightly modified to fit in our survey. For example, we ask the number of days when respondents are involved in certain kind of trip instead of the number of such trips. This makes it easier for the respondents to recall their experiences. Also, our questions are all on a past-7-day basis to capture more non-motorized travel activities given that only a small number of people ride bicycle or walk every day. Questions about how much time people generally engage in physical activities per day, week or month are also included.

Table 10 Variables found important in travel behavior and built environment studies

Attributes	Study	Variables used in studies
Socio-economic factors	Khattak and Rodriguez (2005)	Household size, # of vehicles
	Bento et al. (2005)	Age of worker, gender, # of children, income, education, race, weather, gasoline cost of driving per mile
	Cervero and Duncan (2002)	# of vehicles, driver license, age, race
	Kitamura et al. (1997)	Household size, persons over 16yr, # of vehicles, driver license, age, professional dummy, student dummy, education, income
	Grazi et al. (2008)	Gender, # of workers, age, income, # of weekly hours worked
	Van Acker and Witlox (2011)	Household size, # of children, income, car availability, age, gender, marital status, employment, car needed during work hours (car work often, always), commuting distance, parking difficulty
Attitudinal factors: Use data reduction techniques (factors analysis) to identify each factor from various questions	Kitamura, Mokhtarian, and Laidet (1997)	Pro-Environment, Pro-Transit/Ridesharing, Suburbanite, Automotive mobility, Time pressure, Urban villager, TCM, Workaholic (39 questions)
	Schwanen and Mokhtarian (2005)	Neighborhood type preferences, personality (adventure seeker, organizer, loner, and calm), lifestyle factors (status seeker, workaholic, family/community-oriented, frustration), attitudes toward traveling (travel dislike, pro-environmental policy, commute benefit, travel freedom, pro-high density, travel stress)
	Outwater et al. (2003)	Desire to help the environment, need for flexibility, sensitive to travel stress, insensitivity to transport cost, sensitivity to personal travel experience
Attributes of modes	Cervero (2002)	Total travel time differential between modes, Direct travel cost differential
	Frank et al. (2008)	Time and cost (auto and transit cost, auto and transit-in-vehicle time, walk and transit-walk time, transit wait time, transit-transfers, bike-travel time)
	Bhat (2000)	Travel cost, in-vehicle time, out-of-vehicle time
	Outwater et al. (2003)	Travel cost and in-vehicle travel times, Out-of-vehicle travel time
Experience	Aarts et al. (1997)	habit (from survey)

Questions related to *perceptions towards neighborhoods* are from the Neighborhood Environment Walkability Survey. This survey aims at assessing people's perceptions towards neighborhoods that are associated with physical activities. It includes 98 questions covering from land use characteristics to residents' satisfaction. In addition, the test-retest reliability method was applied to each question to evaluate the reliability of survey instrument. We select questions that have intra-class correlation coefficient (a measure of correlation between answers for the same question from the same person at different times) greater than 0.59. For example, if a question has a reliability score higher than 0.59 in either two studies (Active Living Research 2003; Brownson et al. 2004) then we assume the question is substantially acceptable. From these questions, we can measure perceived safety and built environment factors for our analyses. There are also questions measuring *people's attitudes towards various travel options*. For example, we add questions such as "I prefer to walk rather than drive whenever possible" and "I need a car to do many of the things I like to do" to measure peoples' preference towards different transportation modes. These questions can be helpful in removing the self-selection bias.

Lastly, questions related to *travel activities* in our survey are built based on the PABS because it specifically focuses on non-motorized travel activities and it has been recently tested for its reliability through the test-retest analysis, which showed that the majority of the questions were acceptable based on statistics such as Parson's correlations and Kappa statistics (Forsyth et al. 2012). Physical activities including walking and cycling with different purposes (i.e., to work, transit stations or leisure) are recorded on the past 7 day basis, and the number of days are asked instead of the frequency because the former is easier to remember and could be more accurate.

4 SURVEY ADMINISTRATION

4.1 Mail out/mail back and Internet approaches

Typical approaches to reach survey participants include mail out/mail back, drop off/mail back, Internet, door to door, or telephone. A combination of mail out/mail back (survey booklet) and online questionnaire (using Survey Monkey https://www.surveymonkey.com) was used in our survey to meet the preferences of different age groups. The content of the online questionnaire is customized to be consistent with that of the survey booklet.

4.2 Survey administration

The two waves had an interval of about 3 weeks in between. Wave 1 began on June 11, 2013 with the printing of recruit packages containing a cover letter, a mail-back form that asks basic information of a potential participant (to determine their eligibility) and his/her preference on the use of a paper or online survey, and a business reply envelope in which the respondent, if willing to participate, can simply drop at a post office mailbox. On June 14, 2013, the addresses in the sampling frame were checked against a national address database provided by USPS for accuracy. Subsequently, packages were mailed to those identified correct addresses on June 17, 2013 with an expected delivery date of June 19, 2013. The same process for wave 2 began three weeks later on July 8, 2013 when the recruit packages were mailed. A total of 1,473 and 1,448 packages were mailed for recruitment for wave1 and wave2, respectively.

In the meantime, a reminder postcard was designed and sent to potential participants one week after the mail-out of the recruit packages; this was done on June 24 for wave 1 and July 15 for wave 2, respectively. Follow-up letters were sent two weeks after the reminders, on July 8 for wave 1 and on July 29 for wave 2, respectively. The time between the first and the second reminders is set to two weeks mainly to avoid accounting for the travel activities during the Independence Day week, a likely deviation from the normal days. Participants who finished our survey questionnaires received a 5$ Starbucks giftcard or a 5$ Amazon giftcard based on their preferences.

4.3 Final Sample and Results

4.3.1 Final sample size

Table 11 shows the calculations of response rates for wave 1 and wave 2. The percentage of accurate addresses in our sample exceeds our initial assumed rate of 90%. There were 371 and 332 returned mails for wave 1 and wave 2, respectively, accounting for about 25% and 23% of the addresses obtained, more than two times larger than our expectation. In addition, there are other cases such as “no working adults” (since our survey targeted at least one working adult in a household), “respondents moved out of state”, “respondents passed away”, or “respondents not in WA during the study period). The numbers reported in Table 11 are calculated because the respondents sent back our recruitment letter with the specific reason noted or called us to tell us know the reason why he/she cannot participate. The real numbers are likely much higher than those reported numbers.

Table 11 Response rates for Wave 1 and Wave 2

	Wave 1 numbers	Wave 2 numbers
Number of household addresses mailed out	1473	1448
No working adults	86	8
Respondents moved out of state	11	4
Not in WA during study period	1	0
Respondents passed away	9	2
Wrong addresses	371	332
Number of households who completed the survey	206	207
Response rates	21%	19%

Our reported response rates are 21% for wave 1 and 19% for wave 2. These response rates greatly exceed the current response rates for activity and travel surveys in WA and in the nation. Because we did not account fully those who cannot participate (mainly no working adults), the actual response rates are higher than these reported ones. We attribute this to a very careful survey administration process.

4.3.2 Comparing between wave 1 and wave 2

Figure 3 to Figure 9 show the distributions in residential density, race, education level, household income, commuting modes, and commuting distance for wave 1 and wave 2 households respectively. The distributions are quite similar overall. We also conducted t tests for continuous variables and chi-square tests for discrete or qualitative (e.g., race) variables. All test results except the one on commuting time suggest that there is no discernable difference between wave 1 and wave 2 households. On commute time, wave 2 respondents appear to have lower values than wave 1 respondents. It is possible that this may be related to that the survey period for wave 1 crosses over the independence day week, even though we have tried to avoid that. Our test results are shown in Table 12 and Table 13.

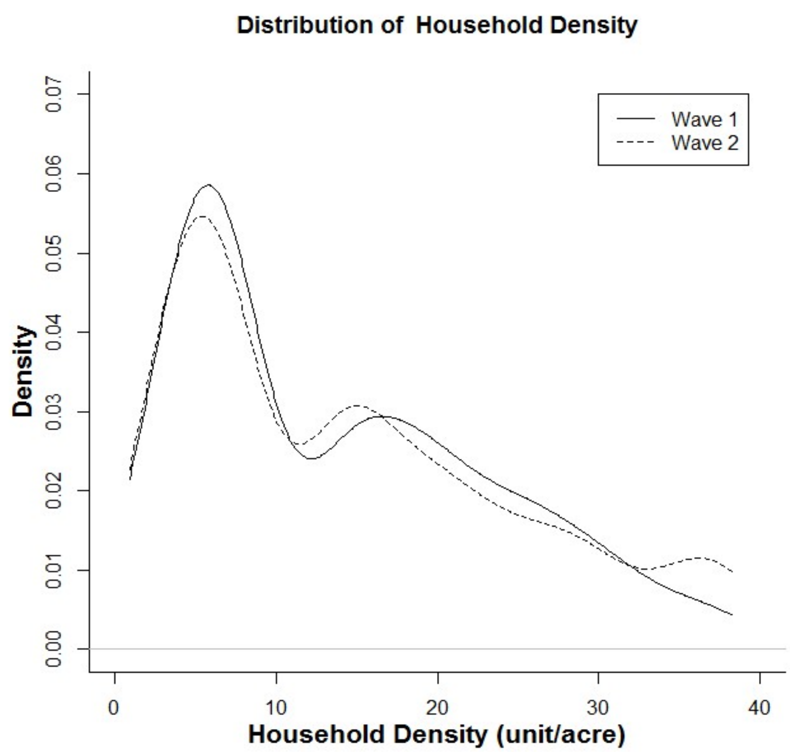

Figure 3 Density distributions between wave 1 and wave 2

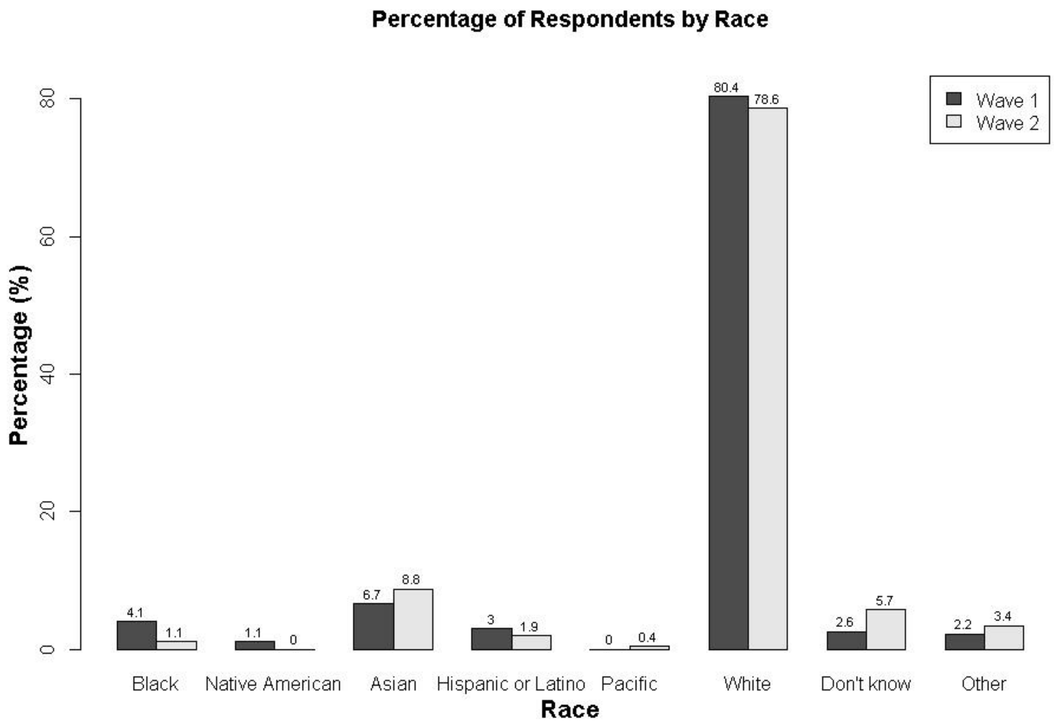

Figure 4 Distribution of race for wave 1 and wave 2

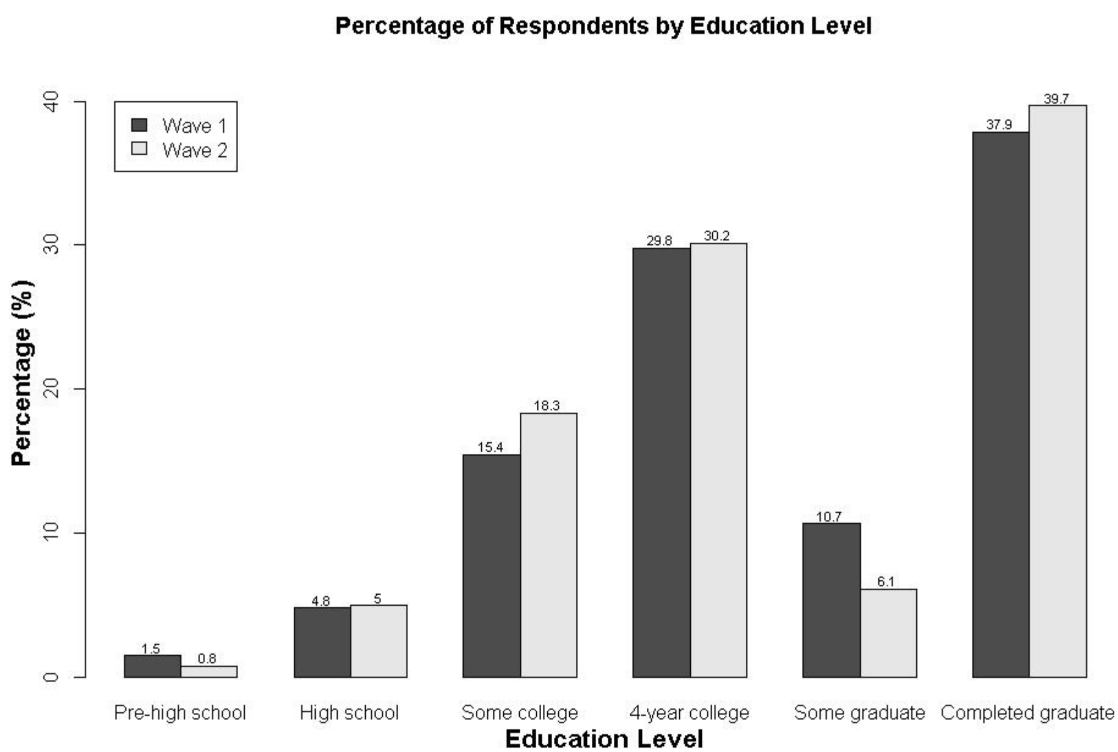

Figure 5 Distribution of education level for wave 1 and wave 2

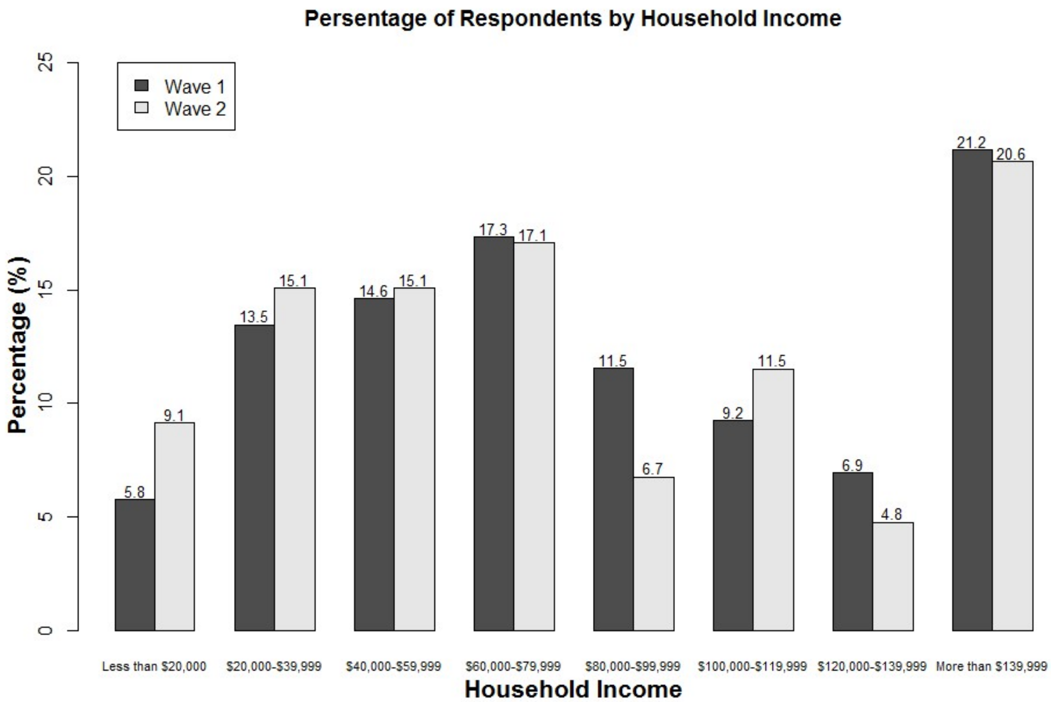

Figure 6 Distribution of household income for wave 1 and wave 2

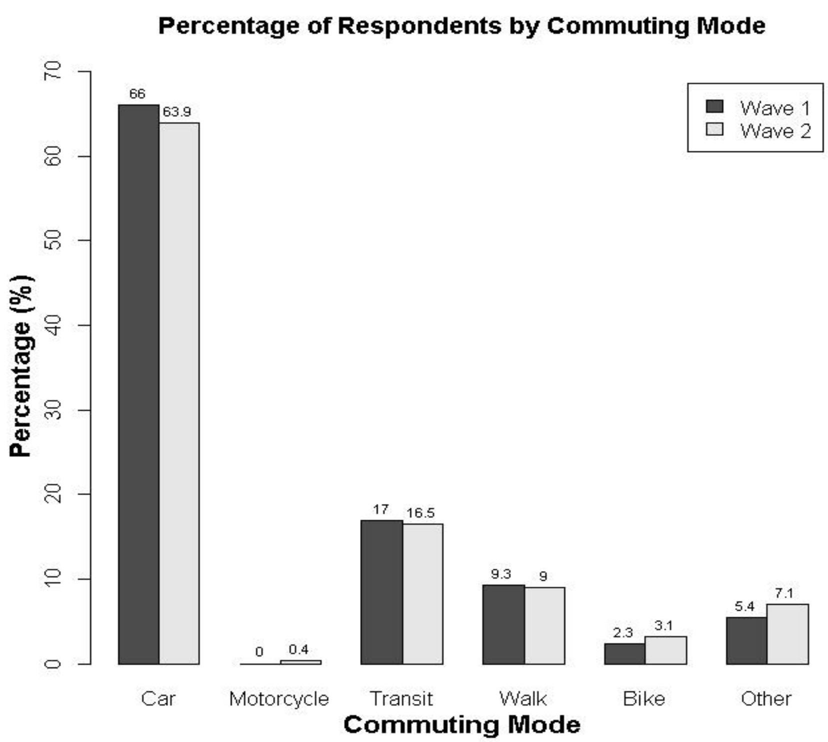

Figure 7 Distribution of commuting modes for wave 1 and wave 2

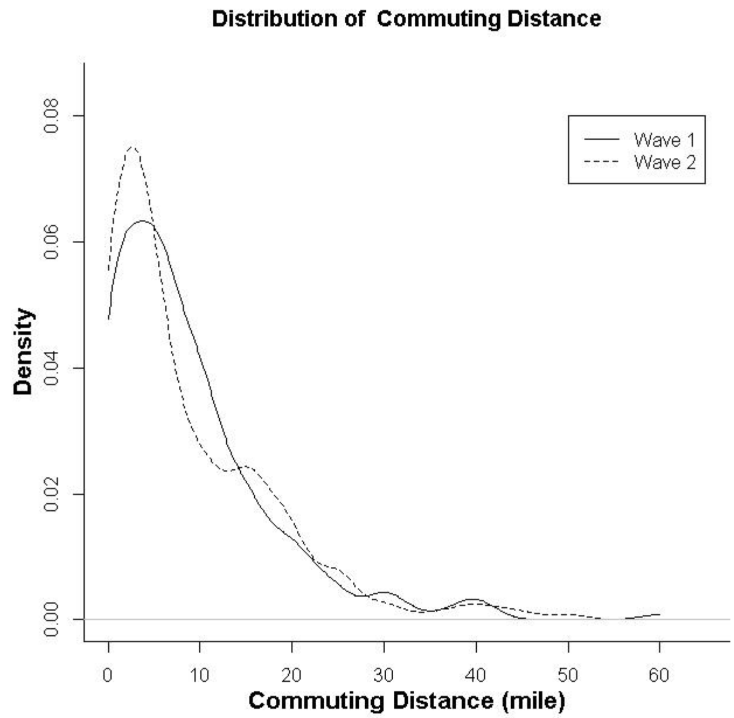

Figure 8 Distribution of commuting distance for wave 1 and wave 2

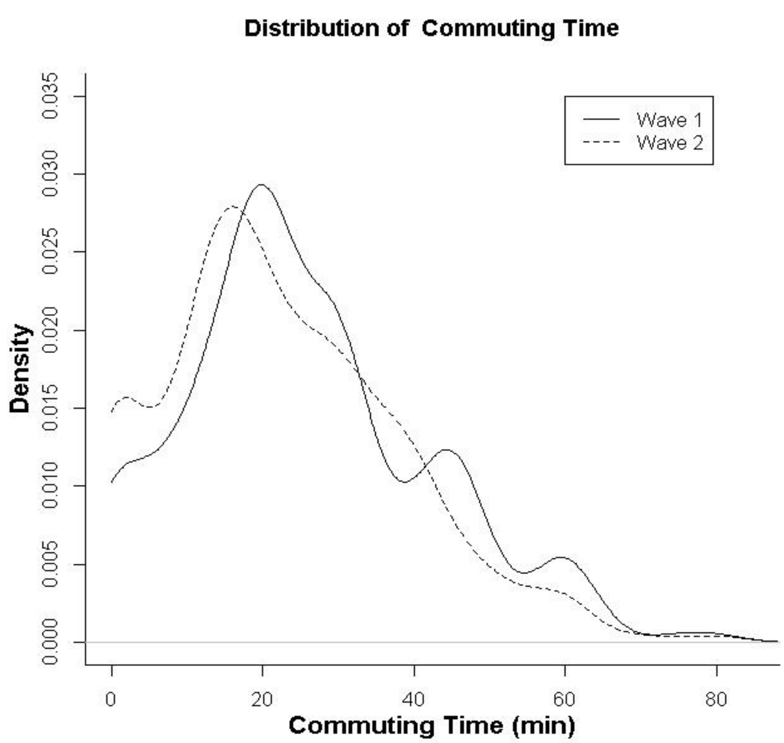

Figure 9 Distribution of commuting time for wave 1 and wave 2

Table 12 T-test results

Variable	Wave 1 mean	Wave 2 mean	t value	P value
Household size	1.98	2.59	-0.94	0.347
Race	5.56	5.76	-1.85	0.066
Education	4.57	4.55	0.17	0.862
Commute mode	1.98	2.08	-0.74	0.460
Commute time	26.43	22.87	2.26	0.024
Commute distance	8.95	8.72	0.27	0.788
Household income	4.77	4.54	1.10	0.270
Residential density	13.92	14.66	-0.86	0.389

Table 13 Chi-square test results

Variable	Degree of freedom	P value
Household size	7	0.071
Race	7	0.060
Education	5	0.458
Commute mode	5	0.847
Household income	7	0.421

4.3.3 Comparing between High-density Residents vs Low-density Residents

Figure 10 to Figure 16 show differences between households living in low-density and high-density areas in a number of aspects. The number of respondents living in high-density areas is slightly more than those in low-density areas (Figure 10). People living in high-density areas live in smaller households (Figure 11), and more educated (Figure 12) than those in low-density areas. They also tend to use more alternative modes of transportation (Figure 13), have shorter commuting times and distances (Figure 14 and Figure 15).

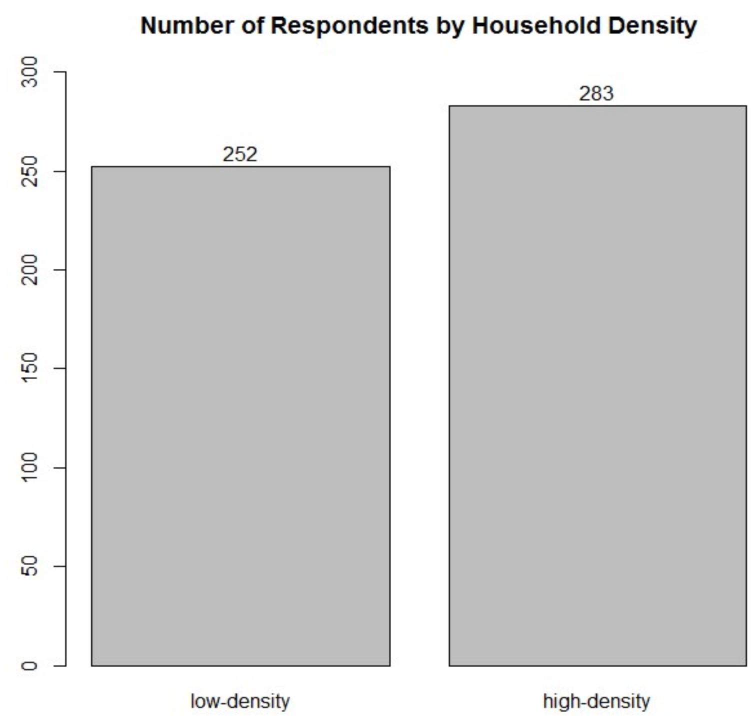

Figure 10 Number of respondents living in low-density and high-density areas

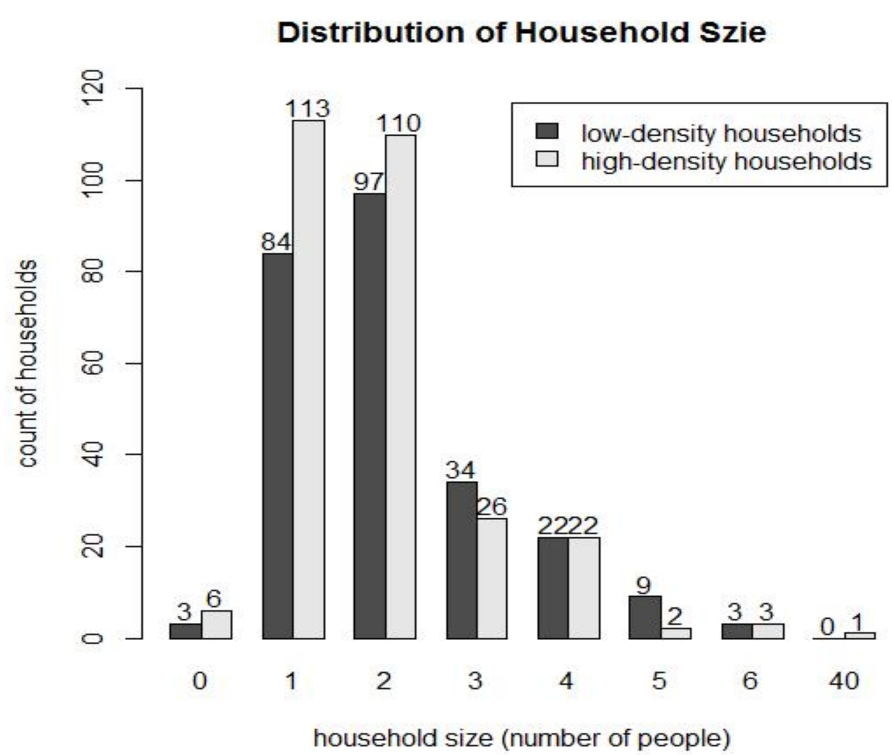

Figure 11 Household size for households living in low-density and high-density areas

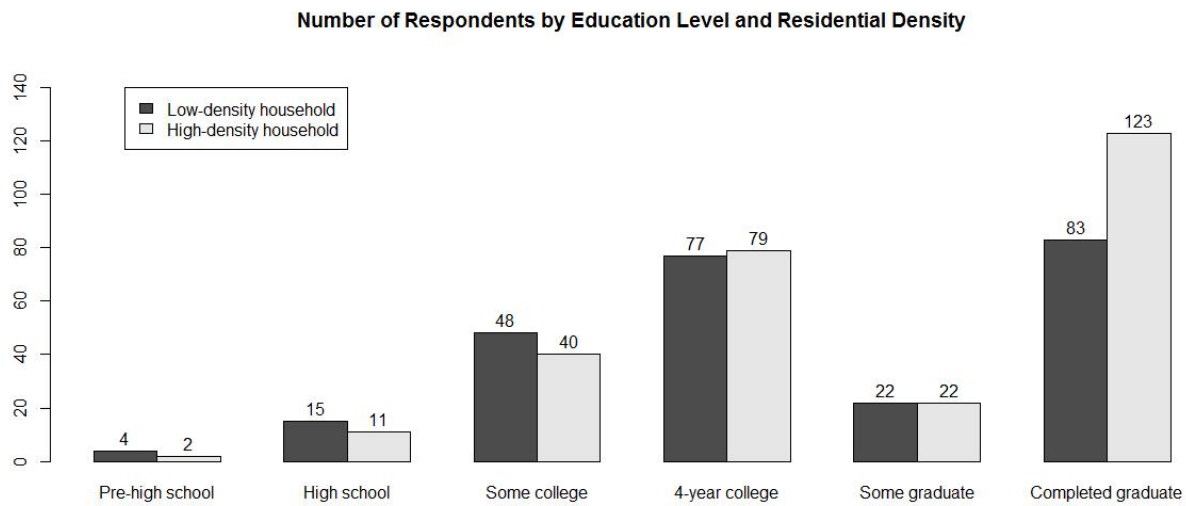

Figure 12 Education level for households living in low-density and high-density areas

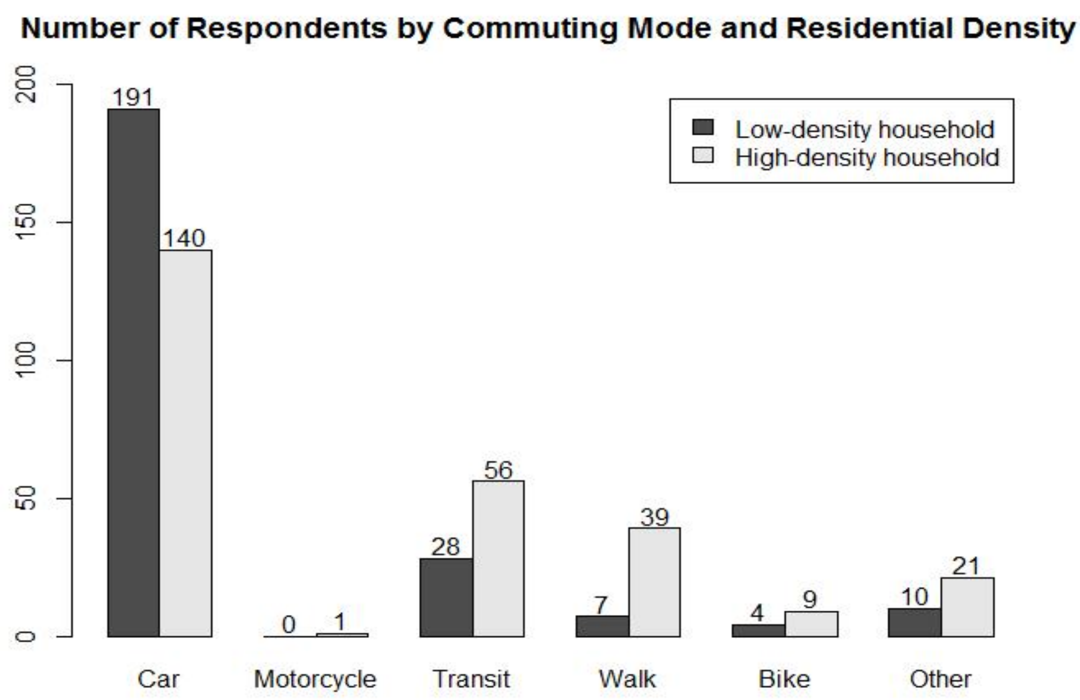

Figure 13 Use of commuting modes for households living in low-density and high-density areas

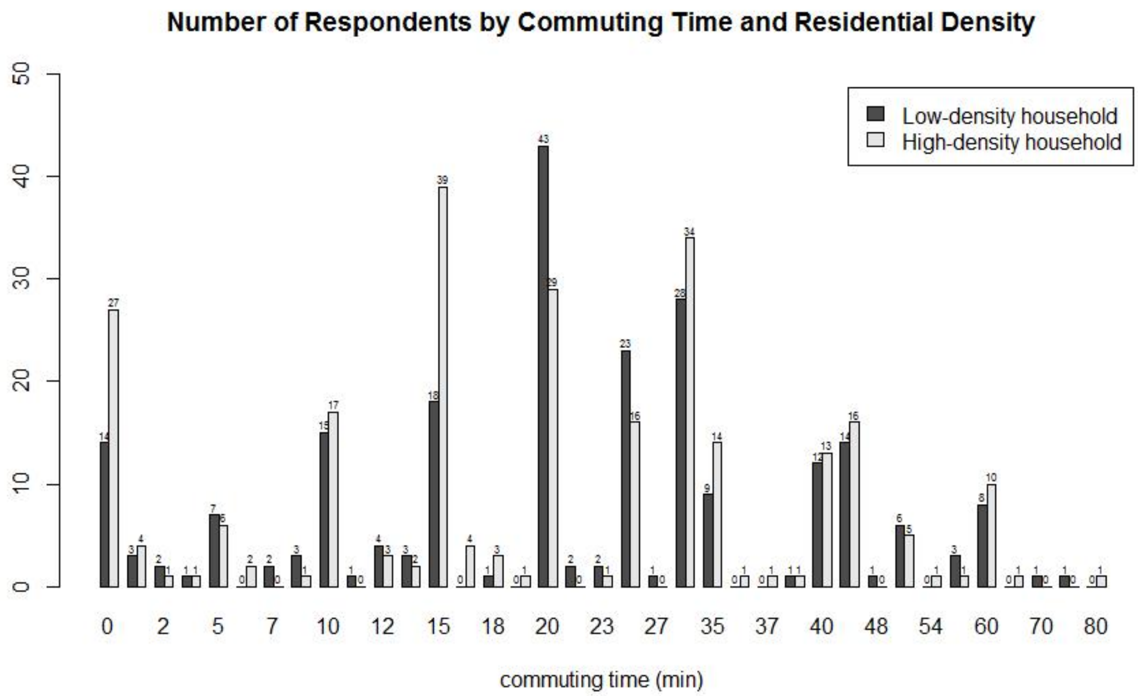

Figure 14 Commuting times for households living in low-density and high-density areas

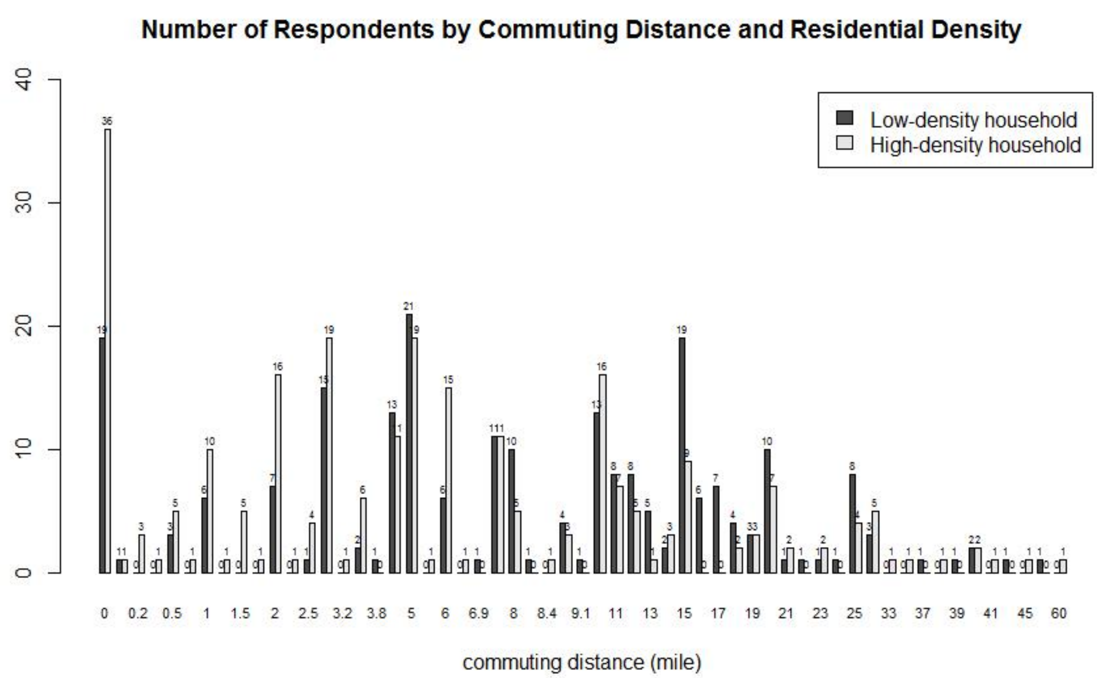

Figure 15 Commuting distances for households living in low-density and high-density areas

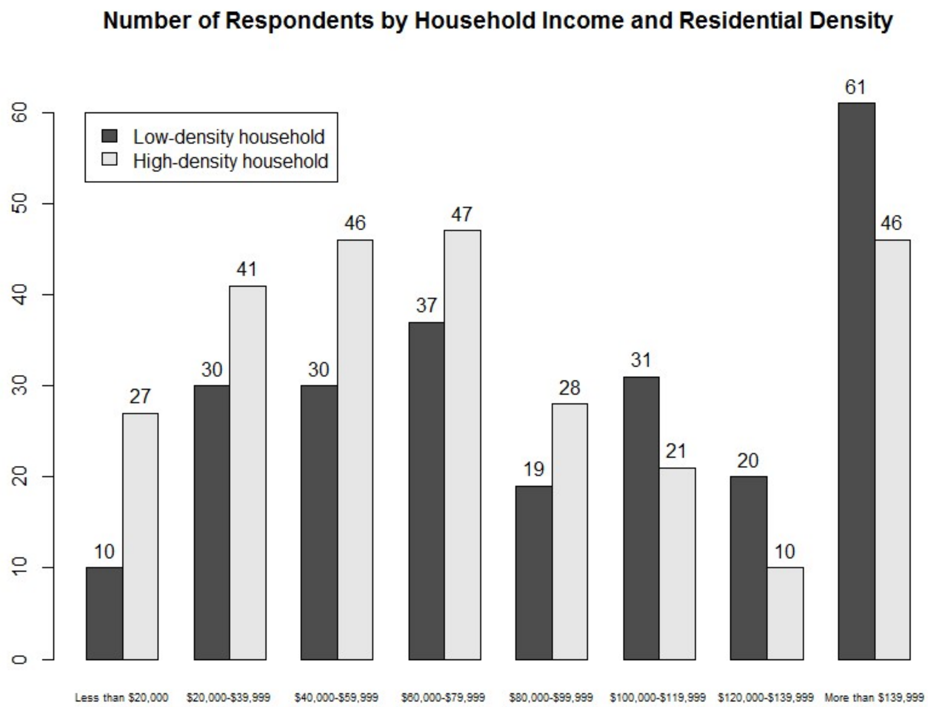

Figure 16 Household incomes for households living in low-density and high-density areas

Table 14 Survey Administration Schedule

	June		July					August			
	3 week	4 week	1 week	2 week	3 week	4 week	5 week	1 week	2 week	3 week	4 week
1st recruitment package for wave 1											
Reminder post card for wave 1											
Survey links and booklets for wave 1											
2nd follow-up letter for wave 1											
1st recruitment package for wave 2											
Reminder post card for wave 2											
Gift card											
Follow-up call & mail for persons who did not finish survey instrument in wave 1											
Survey links and booklets for wave 2											
2nd follow-up letter for wave 2											
Follow-up call & mail for persons who did not finish survey instrument in wave 2											

Chapter 4 Conclusions

1. OVERVIEW

Nationwide, household travel surveys are facing increasing challenges. The current practice of conducting one large survey every ten years requires a large amount of dollars reserved for a one-time survey. This budgetary requirement puts pressure on Metropolitan Planning Organizations (MPOs) that conduct the surveys. A more plausible budgetary scenario is having a smaller, but continuous budget over time, to support survey efforts that span over space and time continuously. Hence, **an innovative survey design with rolling samples** is proposed to cope with the changing budgetary constraints in the future.

A rolling sample design potentially offers many advantages to a traditional one-time cross-sectional survey. It can address a number of well-recognized problems associated with the current practice, including non-representativeness (Stopher and Greaves 2007; Murakami 2008), declining sample sizes, and increasing non-response rates (Wilson 2004; Stopher and Greaves 2007). A rolling-sample design can potentially reach rare or hard-to-reach populations (Rust 2010), reduce costs (Rust 2010), provide better estimates for small communities (Kish and Verma 1983), generate frequent updates on the population (Kish and Verma 1983; Kish and Verma 1986), facilitate the adjustment of seasonal effects (e.g., bicycle use, walking, and long distance trips for holidays) (Ortúzar et al. 2011), and allow a policy topic to be addressed in a timely fashion (Kish 1986).

Challenges also abound. While the concept of a rolling sample design was raised decades ago, it has only been recently implemented at a large scale, for example, the 2010 American Community Survey, which has replaced the decennial long-form census. For its implementation in household travel surveys, many questions and concerns exist, for example questionnaire design, sampling, inference and weighting, and survey administration. Some are technical issues while others are cost and procedural-related. **The primary purpose of this project is to understand these issues and provide recommendations for a future household travel survey with rolling samples**.

It is also expected that a rolling sample design can potentially address some critical questions in helping us understand travel behavior better and thus improve travel behavior modeling. With mounting concerns for congestion, air pollution, and climate change, land use and infrastructure related strategies (e.g., Transit-oriented neighborhoods) are increasingly being proposed as a set of promising solutions for VMT (Vehicle Miles Traveled) reduction. Accurate assessment of these land use and infrastructure strategies requires a targeted sample of people who live and work in places that differ sharply from each other, a requirement that even a large cross-sectional survey is unable to fulfill. A rolling sample design has the potential to meet this requirement because of its capability to sample participants living in very different neighborhoods. In addition, a modified rolling sample design may also allow a subset of subjects to overlap across periods, such that those subjects can be followed over time. This design can provide important information on the net behavioral changes in response to a policy and potentially address the self-selection issue. **The second purpose of this project is to conduct a pilot two-wave rolling survey, evaluate its reliability, and use the collected data to address these modeling related questions, in particular, the potential of land use and infrastructure related strategies for VMT reduction**.

2 BETTER SAMPLING, BETTER UNDERSTANDING OF VMT REDUCTIONS

Policies aiming to reduce VMT need to be supported by an understanding of the circumstances within which travel takes place. Numerous studies have shown that land use in both the home and the work neighborhoods have an effect on travel behavior, including mode choice, number of daily trips taken, and distances travelled (Frank and Pivo 1994; Litman and Steele 2012). This effect remains after adjusting the effects for personal and household characteristics. In the home neighborhood, land use indicators related to travel behavior include residential density and the presence of utilitarian destinations. In the neighborhood where people work, employment density and parking costs and policy relate to mode choice. These effects of land use on travel are not surprising since higher development density (measured in terms of the number of residential units and jobs within a defined area) is associated with land use mix, which in turn reduces distances between activities and therefore decreases the need to drive between activities (Cerin et al. 2007; Frank 2008). As well, higher development density also supports transit service.

Land use patterns in US urban and suburban areas are such that only 20% to 30% of the population lives in areas that provide alternatives to SOV travel and hence support lower VMT (Moudon et al. 2011). These lower VMT areas have been called “transportation efficient” or TE areas (Moudon and Sohn 2005; Moudon et al. 2011). Past and current travel surveys, which aim to capture a representative sample of the population, usually provide a good basis for understanding mode choice. Sampling for these surveys is typically stratified to enable observation of the characteristics of the transit user population. However, most travel surveys do not include a sufficient number of observations in TE areas to identify the factors (personal and land use) that are related to trip length and frequency (Krizek et al. 2005; Levinson and Krizek 2008; Steiner and Srinivasan 2010). For example, less than 6% and 1% of the trips are walking or bicycling trips in the 2006 Puget Sound Household Survey. Also, the trips take place in too few settings to find the significant relationships with land use patterns that are known to exist in studies using local data (McCormack et al. 2001).

Effective policies to reduce VMT require data on a sufficient sample of the population that generates low VMT. In this project, we applied an innovative parcel-based sampling approach for our two-wave rolling survey. We show that we can capture sufficient number of respondents living in both very low- and very high-density areas effectively.

3. CONCLUSIONS AND DISCUSSIONS

Compared to the traditional one-off household travel surveys, there are many advantages offered by the rolling surveys. We discussed those advantages in Chapter 1 and researched its feasibility and provided recommendations to MPOs in Chapter 2 (using PSRC as an example). To demonstrate, we conducted a pilot two-wave rolling survey which was described in Chapter 3. In addition to demonstrating the practice of a rolling survey in the field, we also use this opportunity to collect data to evaluate the impact of the built environment on non-motorized travel behaviors. In the process, we applied an innovative parcel-based sampling procedure to show that through this method, we are able to reach households who are living in very low-density areas and those who are living in very high-density areas. These households, especially the latter, are often very few in a traditional household travel survey effort, resulting in difficulties in estimating various models related to the impact of the built environment. Our two-wave survey was a success, reporting 21% and 19% response rates for wave 1 and wave 2, respectively, far exceeding the current response rates for activity and travel surveys in the Seattle Metropolitan area and in the nation. When comparing the characteristics between wave 1 and wave 2, we show that they are not statistically different, demonstrating the validity of a rolling survey in place of a traditional one-off survey. We also compare the characteristics between respondents living in low-density areas and those living in high-density areas; they are statistically different, suggesting the impact of the built environment. Based on the results of the project, we conclude that a rolling survey is a promising survey method that can potentially replace the traditional one-off travel surveys.

Appendix 1 Rolling Sample Design of the American Community Survey (ACS)

1. ACS SAMPLING

Starting from 2006, the ACS collects samples from both housing units (HU) and group quarters (GQ). As the GQ sample only accounts for 6.5% of the whole ACS sample, and household travel surveys do not typically include GQ, our review is on the housing unit sample selection process only.

1.1 Sample Size

We follow Kish's definition for rolling sample surveys here—a "rolling sample" design jointly selects k non-overlapping probability samples, each of which constitutes 1/F of the entire population. One sample is interviewed each time period until all samples have been interviewed after k periods (Kish and Verma 1983).

In ACS, a monthly sample size of 500,000 HUs was initially suggested to produce multi-year accumulations for small areas at the same level of sampling reliability as the long-form census sample (Alexander, 1993), but budget constraints limit the size of the ACS sample to 250,000 addresses monthly, or 3 million per year. The same address will be not be sampled again within 5 years (U.S. Census Bureau, 2009). This corresponds to a monthly rolling sample with an average rate of approximately F = 480 or an annual sample with F = 40. The survey uses k = 60. In any given year about 2.5% (1 in 40) of U.S. households will receive the survey. Over any 5-year period about 1 in 8 households should receive the survey (as compared to about 1 in 6 that received the census long form in the 2000 census).

1.2 Sampling Frame

The sampling frame used for the American Community Survey (ACS) is an extract from the national Master Address File (MAF), which contains mailing and location addresses, geocodes, and other attribute information for all known living quarters in the U.S. Addresses in the MAF are maintained dynamically through addition, deletion, or revision, based on information collected by the U.S. Postal Service's (USPS) Delivery Sequence File (DSF), the Community Address Updating System (CAUS), and the Census Bureau's other household surveys (U.S. Census Bureau 2009). All MAF HUs addresses are arranged in each county by geographical location.

Given that the number of housing unit addresses is on the rise, the fixed sample size (i.e., 3 million per year) results in a decrease in sampling rates at different levels of geography (e.g., county and tract), initiating concerns that the standard errors of the ACS estimates for each estimate period (1-, 3-, and 5-year) may increase over time. To account for growth in the population, one suggested option is making the shift from a fixed target sample size to a constant, target sampling rate, which would entail an approximately 1.6% increase in the annual sample size (Citro and Kalton 2007).

1.3 Survey Administration

The ACS depends on mail-back responses, with non-response follow-ups by telephone, and about one-third of the remaining non-respondents are randomly selected for personal visits by interviewers. Mail is the least expensive method of data collection, and the success of the survey program depends on a high level of mail responses. Sample addresses are reviewed to determine whether the available information is adequate for mailing. An address is considered unmailable if it includes "only physical descriptions of an HU and its location, or post office (P.O.) box addresses", or address missing place names and ZIP Codes" (U.S. Census Bureau 2009). Unmailable HU addresses are skipped in the mail and telephone phases, but will be sampled at the personal interview phase, at a rate of 2 in 3. This reduces possible coverage biases that can be created by the inability to sample unmailable addresses. Once an address is determined as mailable, it will be sent out to the household in the first phase.

The second phase of the ACS data collection is the computer-assisted telephone interviewing (CATI), which usually begins about 5 weeks after the first mail package is sent out. For the HUs that have not returned the completed questionnaires by the end of the first month and that have an available phone number, an attempt is made to interview them by CATI during the second month. If a household's questionnaire is received at any time during the CATI operation, that case is removed from the CATI sample and is considered a mail response.

The last phase is the computer-assisted personal interview phase, or CAPI. A sub-sample of those who have not been interviewed by the end of the second month (including the unmailable cases) is selected for interview by CAPI in the third month. This phase typically lasts for the entire month.

In general, all addresses mailed a questionnaire can return a completed questionnaire, through mail, telephone, or personal visit, during the three-month time period.

1.4 Sampling

The ACS sampling is operated via two phases, using systematic sampling so that each sample is spread throughout the United States in an unclustered way.

During **first-phase** sampling, each census block is assigned to one of the seven sampling strata, each with a unique sampling rate, and the initial sample is selected from the MAF. During the **second-phase** sampling, a subsample of addresses for which neither a mail questionnaire nor a telephone interview has been completed is selected for computer-assisted personal interviewing (CAPI). Figure A1 provides a visual overview of the HU address sampling process (U.S. Census Bureau 2009).

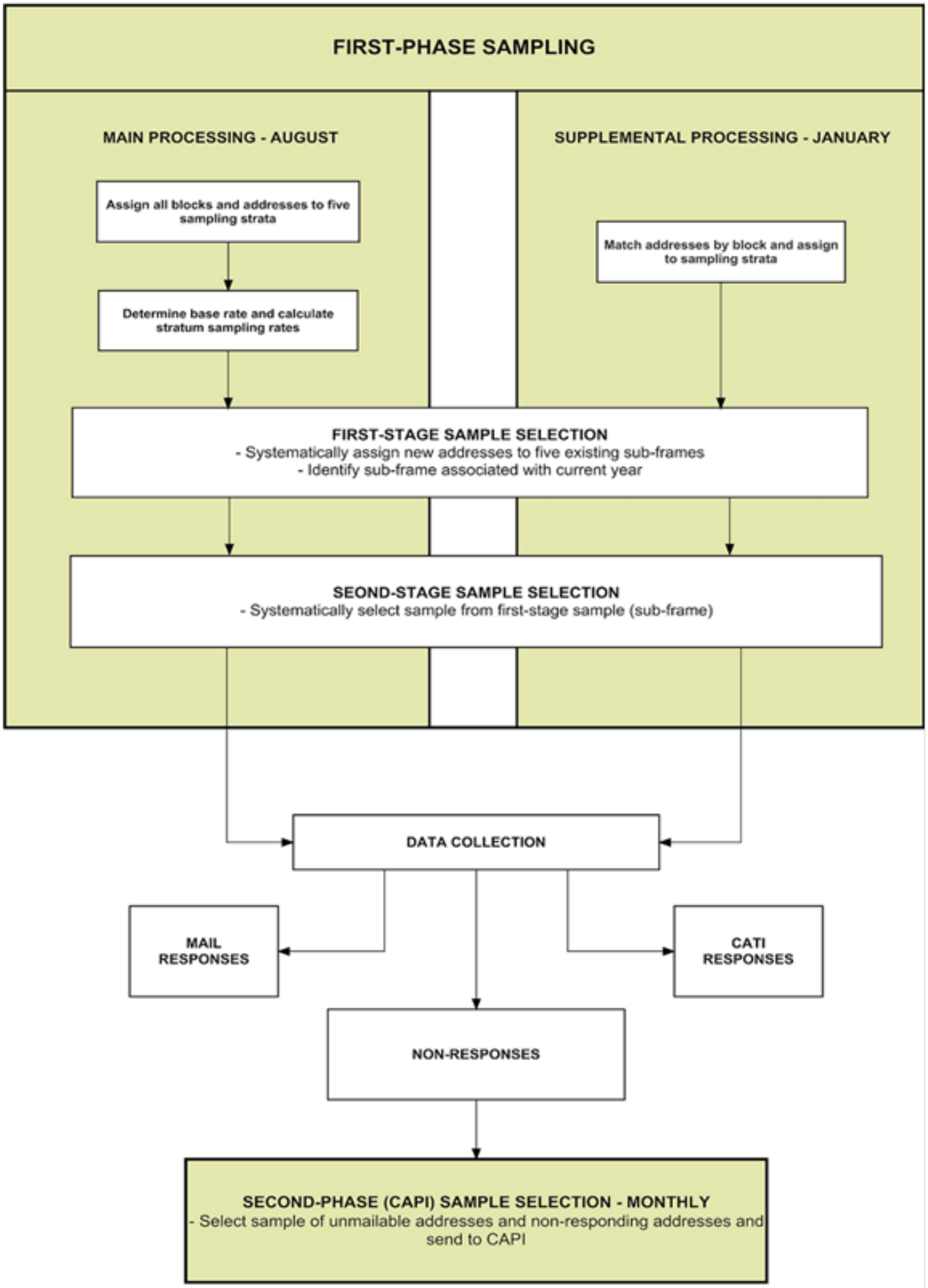

Figure 17. Selecting the samples of housing unit addresses (U.S. Census Bureau 2009).

1.4.1 First-Phase Sampling

In the first phase, the sample selection is conducted in two stages. First, each HU address on the MAF is randomly assigned to one of the five subframes. Each of these five subframes is randomly assigned to year t_1 through year t_5 of specified 5-year periods (for example, 2006-2010, 2011-2015, and so on). In other words, sampling within any year is only conducted from one of the five subframes. This guarantees that no HU will be surveyed more than once within the 5-year period. Second, HUs addresses for each yearly ACS sample are selected from the corresponding subframe.

Before constituent addresses can be selected, each block must be assigned to one of the five sampling strata. The goal of the ACS sample design is to produce small area estimates similar to those provided by the census long-form sample. To accomplish this, each block and its constituent addresses within each stratum are sampled at a unique sampling rate. These rates are set so that the sample sizes for the smallest geographic areas are selected with higher rates, while the sample sizes for the largest areas are selected with lower rates.

Earlier studies revealed that estimates in smaller geographic areas (e.g., cities, places, minor civil divisions) have lower level of reliability than those for larger geographic areas, which means the goal of the ACS sample design is not fully realized. One major reason for the unequal reliability is the very small number of discrete initial sampling rates, which creates diverse sampling errors for areas with very little difference in population size because they may fall into different sampling rate groups. To achieve the desired estimates, an alternative sampling design with 16 sampling strata was developed, by minimizing the differences of the coefficients of variation for tract level estimates by size class (Sommers & Hefter, 2010). The new strata shift samples from the very largest tracts to the smallest tracts, resulting in small area estimates with approximately equal reliability. The Census Bureau began to implement the new ACS sample design beginning from 2011.

Main and supplemental sampling

At the block level, samples are selected twice a year and these two annual processes are referred to as main and supplemental sampling. This allows new addresses to be potentially selected during supplemental sampling. Main sampling occurs in August preceding the sample year and accounts for 99% of the total annual ACS sample, with the most recently updated MAF obtained in the July preceding the sample year. Supplemental sampling occurs in February of the sample year and accounts for 1% of the total sample, with the MAF obtained in January of the sample year. Systematic sampling method, which regularly samples units over a geographically sorted list, is used in both sampling processes.

For the main sample, addresses are selected from the subframe assigned to the corresponding sample year. These sample addresses are then allocated systematically to each month of the sample year. During supplemental sampling, addresses new to the frame are systematically assigned to the five subframes. The new addresses in the current year's subframe are sampled and are systematically assigned to the months of April through December of the sample year for data collection.

1.4.2 Second-Phase Sampling

Mailable addresses with responses to neither mailout surveys nor CATI interviews, are eligible for the personal interview phase (CAPI), along with the unmailable addresses. The CAPI sample is selected systematically by sorting within county by CAPI sampling rate, mailable versus unmailable, and geographic order within the address frame.

The CAPI subsampling uses three different rates in order to approximately equalize the precision of estimates for areas with higher and lower mail/CATI response rate, and uses 66.7% for unmailable addresses and addresses in Remote Alaska (See Table A1.1). CAPI samples, due to sampling for nonresponse follow-up, generally have larger weights than those completed by mail or CATI. The variance of the estimates for an area will tend to increase as the proportion of mail and CATI responses decreases. Large differences in these proportions across areas of similar size may result in substantial differences in the reliability of their estimates. To minimize this possibility, tracts in the U.S. that are predicted to have low levels of responses completed by mail and CATI have their CAPI sampling rates adjusted upward from the default 1-in-3 rate for mailable addresses. This tends to reduce variances for the affected areas both by potentially increasing their total numbers of completed interviews and by decreasing the differences in weights between their CAPI cases and mail/CATI interviews (Hefter 2005).

Table 15 Second-Phase (CAPI) Sampling Rates (Hefter 2005)

Address and tract characteristic	**CAPI sampling rate (%)**
Unmailable addresses and addresses in Remote Alaska	66.7
Mailable addresses in tracts with predicted levels of completed interviews prior to CAPI subsampling between 0 percent and 35 percent	50.0
Mailable addresses in tracts with predicted levels of completed interviews prior to CAPI subsampling greater than 35 percent and less than 51 percent	40.0
Mailable addresses in other tracts	33.3

1.5 Developing Sampling Weights

For one-year estimates, ACS applies weights in three steps:

1. calculation of the base weights,
2. adjustments for non-response, and
3. application of controls (Beaghen et al. 2012)

In the first step, base weights are calculated as the inverse of the sampling rate. Then, the sampling weight is calculated as a multiplication of the base weight and a CAPI subsampling factor (SSF), which is the inverse of the CAPI subsampling rate (see Table A1.1).

In the second step, adjustments for nonresponse are conducted. There are three factors used for this step, which are associated with housing unit responses (building type, census tract, and month of data collection). If all factors are used at the same time, there will be too many cells that should be filled out with a sufficient number of sample HUs (i.e., 10). To solve this problem, ACS designed two steps separately: one based on building type and tract, and the other based on building type and tabulation month. Nonetheless, there may be cells containing fewer than 10 households, in which case the cells are collapsed with neighboring tracts until meeting the minimum size criterion. In addition, possible mode related noninterview bias is adjusted since three modes (mail, telephone, and personal interview) are employed in the survey.

In the last step, weights are controlled by housing units and population estimated by the Population Estimates Program (PEP). PEP provides a set of estimates by sex, age, race and Hispanic origin on an annual basis based on the previous decennial census data (i.e., 2000 census for the ACS). Subcounty controls are applied for each subcounty area by matching the ACS estimates of totals by sex, age, race and Hispanic origin with the PEP estimates. Different from Census, ACS uses samples and therefore estimates may deviate from true population. Generally, sample surveys cannot capture a whole population so that the weighted population without control adjustments tends to be smaller than the census population estimate. In addition, population estimates from sample surveys can vary from sample to sample, producing sampling error for estimates such as the number of population living in urban areas. If weights are controlled by fixed values of controls, it will reduce sampling errors of other estimates by eliminating the sampling errors for the controlled population.

For 3-year and 5-year samples, weights for the multi-year estimates (MYE) are applied in four steps:

1. reweighting,
2. adjustments,
3. application of controls, and
4. model-assisted estimation.

First, all samples are pooled over data collection periods (i.e., 3- or 5-year) and combined into a single file. Then, new weights with adjustments are calculated to make samples representative for the population. The simplest way is to take the one-year base weights divided by the number of multiple years, so that each year contributes proportionally to the multiyear estimate.

Second, compared to one-year estimates, multi-year estimates make adjustments to nonresponse using two more factors: geography and monetary values. As geographic boundaries (e.g., census place and county subdivision) are likely to change over time to reflect new political boundaries or new developments, the most recent year’s geography is used in MYEs. Similarly, all income and dollar value estimates are adjusted for inflation to the final year of the estimate period.

Third, different from one-year estimates, weights for MYEs are controlled using the average of the annual independent population estimates over the period. For example, the 2008-2010 controls are the average of the 2008, 2009, and 2010 independent population estimates.

Finally, model assisted weighting step is only implemented for multiyear weighting and aims at reducing variances for subcounty estimates. Previous studies found that the variances of ACS tract level estimates are higher than those of Census Long form estimates due to the absence of subcounty controls at the tract level or a lower level. Administrative records, which represent a census-like portrayal of the population for the period between decennial censuses, are available on an annual basis by the PEP program. Each record contains the basic demographic characteristics (i.e., age, sex, race and Hispanic origin) of a person, and an identifier linking to each HU address in the MAF. Model assisted weights are calculated through three steps: 1) linking the administrative records to the MAF for creating estimate totals at the subcounty area for the 3-year data, but at the tract level for the 5-year data; 2) linking the administrative records to the ACS sample to create totals for the same geographic areas; and 3) using generalized regression estimation (or GREG) method to calibrate the ACS final weights so that the ACS sample-based totals match the MAF-based totals.

2. INSTRUMENT DESIGN

2.1 Questionnaire Structure and Instruments

The ACS questionnaire includes four sections. The first section verifies basic address information, determines the occupancy status of a HU, and identifies who should be interviewed by applying the interview rule and the residence rule. The second section of the questionnaire collects basic demographic data: sex, age, relationship, marital status, Hispanic origin, and race. The third section collects information on physical and financial characteristics of housing (e.g., value, rooms, year built), and the final section collects population data (e.g., citizenship, education, employment status).

Different instruments are provided for the three data collection modes (mail, telephone, and in-person interviews). For instance, since mail interview by far is the least expensive mode of data collection, the instrument for this mode is designed accordingly to maximize the rate of mail response through multi-mail contacts (i.e., pre-notice letter, initial questionnaire package, reminder postcard, and potential second questionnaire package due to nonresponse to the initial package) (U.S. Census Bureau 2009). Additionally, to maximize responses, an Internet response option will be added to the mail data collection phase on the 2013 ACS (U.S. Census Bureau 2012). For telephone interview, CATI, questions with long or complicated response categories are broken up into separate questions. As an aid in answering questions, flash cards are provided to provide respondents needed information during the CAPI phase.

One challenge, however, faced by a rolling sample surveys is that on one hand, subjects and questions in the questionnaire should remain stable for samples to be accumulated; on the other hand, topics of interests from data users would invariantly change over time.

2.2 Content Change Requirements

To address current needs, a rolling sample survey allows questionnaire changes in a timely manner, which is unachievable for the previous decennial census. For example, it can provide the effects on per capita income for areas of large natural disasters like the mid-west flood or Hurricane Andrew (Scarr 1994). However, one constraint is that the ACS must be accumulated over time, to provide acceptable levels of reliability for small geographic areas. One obvious outcome of introducing changes is that the data will not be released for small areas, if a question changes significantly, or has not been surveyed for long enough to accumulate three or five years' worth of data (U.S. Census Bureau 2009). For this reason, content changes have to be minimized, and consistency must be maintained throughout all ACS data collection operations.

Certain limitations apply to the timing of implementing the new questionnaire content for the future ACS. First, new content will be incorporated i(U.S. Census Bureau 2009)nto the ACS only after a content test has been completed. The test typically consists of new versions of current questions (e.g., with high missing data rates, or low reliability) as well as questions on new topics (U.S. Census Bureau 2009). Second, because the ACS accumulates data over time to provide reliable estimates, e.g., 1-year, 3-year, 5-year, ACS content can be added to or revised at most once a year to maintain consistency. That is to say, any content change occurring within a year will result in only partial releases of data for that particular year . For instance, the Census Bureau made changes to the questionnaire wording of a series of "disability" questions from 2008, which means that there are no small area disability data available until the 2008-2012 accumulated ACS records are available. Third, the ACS questionnaire has been same as the Census since its inception until 2008. 2008 was selected as the first year of implementing a new questionnaire, because 2008 ACS marks the first year of a three-year aggregated data product, the 2010 decennial census (2008 - 2010). Similarly, the year 2013 is selected for implementing new content changes allowing for 5-year estimates to be released which used the same questionnaire since 2008.

The ACS Content Test uses a similar data collection methodology as the current ACS, though cost and time constraints could result in some deviations such as only HUs addresses being sampled in the Content test. With a smaller sample size (e.g., 62,900 for the 2006 Content test; 70,000 for the 2010 Content test), the Contest Test assigns approximately half of the sample addresses to a test group, which got the test version of the questions and the other half to a control group, which received the current questions. Different from the ACS sample selection, the Content Test sample is first selected systematically from a geographically sorted list, and then the next address in the list is selected as its pair. For each pair, one member is randomly assigned to the control group, and the other assigned to the test group (2010 ACS content test for property income).

Certain indicators are used for evaluating the quality of the test questions relative to current ACS questions, which include response variance, gross difference rates, and other data quality measures, such as item nonresponse rates and measures of distributional changes. As recommended by these tests, the 2008 ACS questionnaire included new questions on the subjects of marital history, health insurance and coverage, and veterans' service connected disability ratings. Additionally, two new subjects will be implemented on the 2013 ACS and they relate to computer and Internet usage and parental place of birth.

2.3 The Mandatory Nature of the ACS

Due to increasing concerns on public privacy, the Census Bureau conducted a test to address the effects of a voluntary ACS on response rates, data quality, workloads and costs in 2003. The results of the voluntary sample (March – April 2003) were compared to the mandatory ACS results (March – April 2002) (U.S. Census Bureau 2003; U.S. Census Bureau 2004). The major impacts of shifting to the voluntary data collection methods include: 1) reductions in response rates across all three modes of data collection (e.g., more than 20% drop in mail response); 2) reductions in the reliability of estimates; and 3) cost increases of more than $59 million annually if reliability was maintained. The ACS has been implemented as a mandatory survey since its inception.

2.4 Residence Rules

Residence rules define who should be interviewed at a sample address. The ACS uses different residence rules from those that have been used in decennial censuses. Decennial censuses and most other surveys use the usual residence concept—most often the place where they spend the most time. The usual residence rule does not count people who are staying somewhere other than their usual residence as occupants of that place, for example, “snowbirds”.

The ACS, in contrast, is based on the current residence concept and uses the Two-Month Rule. Under this rule, anyone who is living in a survey unit for more than two months as of the date of interview (either by mail, telephone, or personal visit) is considered to be a current resident of that unit (U.S. Census Bureau 2009). The current residence concept suits the ACS, because the ACS continuously collects information from monthly samples throughout the year. The current residence concept accounts for the fact that people can live in more than one place throughout a year (Citro and Kalton 2007).

2.5 Measurement Error

Measurement error, common to all surveys, is defined as the inaccuracy in responses recorded on survey instruments, which indicates how reliable and precise the estimates are. Measurement error may arise from: 1) interviewers’ effects on the respondents’ answers to survey questions, for example, the tone used in reading questions; 2) the inability of respondents to answer questions, lack of requisite effort to obtain the correct answer, or other psychological or cognitive factors; 3) vague or ambiguous questions easily misinterpreted by respondents; and 4) data collection mode effects (U.S. Census Bureau 2009).

Measurement error is minimized in the ACS by improving the questionnaire and instrument design in several ways (U.S. Census Bureau 2009):

1. Including a questionnaire instruction booklet in the mail questionnaire package to help respondents interpret and answer specific questions.

2. Providing a toll-free telephone questionnaire assistance (TQA) line for respondents to speak with trained interviewers when having questions.

3. For the CATI/CAPI phase, instruments could be improved by providing extensive training for the interviewing staff (e.g., reading questions, answering respondent questions), automating skips to show the interviewer only relevant questions, providing multi-language support, and implementing a quality re-interview program with CAPI respondents to minimize falsification of data.

3. INFERENCE

3.1 One-Year vs. Multiyear Estimates (MYEs)

ACS provides reliable annual estimates for large areas whose population size is greater than 65,000. To obtain similarly detailed estimates for mid-size areas with populations between 20,000 and 65,000, three years of data are utilized. Finally, for smaller areas with a population less than 20,000, ACS recommends five years of data for estimation. Overall, the estimates of ACS have greater sampling errors compared to those of census long-form, mainly due to the differences in sample sizes (for the same area, the total sample size for ACS is still smaller than that for the census long-form). However, one can determine the length of data cumulation for estimates for various applications by keeping the sampling error close to that of census long-form (Chand and Alexander 2000). For example, ACS provides several statistics for MYEs to check the magnitude of sampling errors such as standard error (SE) and margins of error (MOE) (U.S. Census Bureau 2008). Specifically, ACS provides MOE at 90% CL for the published data, which indicates the precision of estimate at 90% confidence level. MOE is calculated by multiplying SE and Z score (1.645 for 90% confidence level) and represents the maximum difference between the estimate and upper/lower bounds. SE is the variability of an estimate caused by using samples not the population. If SE is small, it indicates that the estimate is very similar to the population value. These estimates are directly associated with sample size and this is the reason why ACS only uses multiple years of data for small regions. Similarly, the regional household travel survey can adopt the same idea if one is not ready to directly apply ACS's criteria (based on population size) for one-year or multi-year estimates. In other words, keeping the sampling errors similar to each other (between the rolling sample survey and the original cross-sectional regional household travel survey) would provide clues for the length of data accumulation needed for specific estimates and a particular geographic scale.

Table 16 One-Year, Three-Year, and Five-Year Estimates

Length of average	Population size	Features
1 year estimate	≥ 65,000 population	Collected during 12 months Currency is more valuable than precision Large population, geographic scales.
3 years estimate	≥ 20,000 population	Collected during 36 months Intermediate between 1 year and 5 year estimates Medium size population where 1 year estimate is not available
5 years estimate	All geographic areas (down to the tract and block group levels)	Collected during 60 months Precision is more valuable than currency

Extra caution should be noted, especially for small domains (Beaghen and Weidman 2007). From rolling samples, annual estimates can be derived to allow the detection of changes over time. However, if we try to use annual estimates for smaller domains, there exists larger sampling variability due to their small sample sizes. As an alternative, asymmetrical cumulations from rolling sample surveys have been proposed. It refers to the cumulation of different numbers of years for different geographic levels or domains. For instance, ACS provides 3-year and 5-year estimates for small geographic scales to obtain sufficient sample sizes while it provides annual estimates at the national scale. This issue is critical for the transportation modeling because Traffic Analysis Zones (TAZ) often employed as an analytical unit is relatively small, requiring certain levels of sample size to obtain reliable estimates.

3.2 Calculating MYEs

In theory, MYEs are period estimates, that is, the average over time (Beaghen and Weidman 2007). For clarification, a period can be 1 year, 3 years, or 5 years. These estimates are not the simple averages of 12 or 36 monthly values. In addition, the MYEs are not the average of 3 or 5 single year estimates. Data are collected continuously every day with evenly distributed samples across the entire period and results are aggregated over particular periods. For example, assuming equal weights over time and no corrections, the MYE of 3-year data can be simply expressed as follow:

$$\bar{y}= \frac{\sum_{i=1}^{n} \pi_i y_i}{N}, \quad \text{Equation 1}$$

where y_i is our interested variable for household i, π_i is the sampling weight for household i, N is the population size, and n is the sample size. In practice, several corrections including nonresponse adjustments, geography, and monetary values are implemented to estimate multi-year estimates (Beaghen et al. 2012).

In addition, MYEs are generally better estimates for small subpopulations than single year estimates since its increased sample size results in smaller standard errors. Obviously, 3-year estimates are based on samples that are three times greater than those for 1-year estimates and this greater sample size results in smaller standard error. However, we also should recognize that there is a trade-off between currency and precision. If single year estimates give enough precision, they are better than MYEs for depicting current conditions since they include more recent information. For example, travel patterns across the whole region can be better compared using 1 year estimates with relatively good precisions. To evaluate the precisions of single year or multiyear estimates, coefficient of variation (CV) is often employed, which is the ratio of the standard error of the estimate and the estimate itself. This index represents the relative amount of sampling error with respect to estimate. If a single year estimate has a small enough CV it will be more useful than a multiyear estimate in terms of currency (Beaghen and Weidman 2007).

However, it should be noted that the multiyear estimates can be misleading under certain circumstances. In practice, simplified assumptions are often made. It is assumed that the MYE is the simple average of 1-year estimates. This assumption is only reasonable for areas that have no significant changes in population over time and the variances of yearly estimates are the same (Citro and Kalton 2007). Under this assumption, we can adopt an equal weighting scheme. As an example, we can calculate n-year estimates of trip frequency per household of the same region at tth year (see Equation 2 below).

$$\boldsymbol{Average}\ (y_t) = \frac{\sum_{i=0}^{n-1} y_{t-n+1+i}}{n}. \qquad \text{Equation 2}$$

For example, 3-year estimate of trip frequency at 3rd, 4th, or 5th year can be written as follows:

$$\widetilde{y_3} = \frac{y_1 + y_2 + y_3}{3} \qquad \text{Equation 3}$$

$$\widetilde{y_4} = \frac{y_2 + y_3 + y_4}{3} \qquad \text{Equation 4}$$

$$\widetilde{y_5} = \frac{y_3 + y_4 + y_5}{3} \qquad \text{Equation 5}$$

Then, it seems obvious that the difference between MYEs depends on non-overlapping years. That is, the difference between $\widetilde{y_4}$ and $\widetilde{y_3}$ is $(y_4$-$y_1)$/3, canceling out the overlapping one year estimates. As we can see, it does not indicate the difference between the most recent years and it is hard to interpret. For example, let's assume that the one year estimates of trip frequency per household are 8.2, 8.7, 8.6, and 9.1 from 2003 to 2006. Then, the averages of 2003-2005 and 2004-2006 are 8.5 and 8.8, respectively. If users consider the difference between these two MYEs as a result of changes in the most recent years, 2005 and 2006, one could reach an incorrect conclusion by saying that the trip frequency increases by 0.3 between 2005 and 2006 even though the actual change is 0.5. This estimate truly means the difference between 2003 and 2006 divided by 3.

In addition, the standard error of the difference is related to the correlations in sampling errors because they are not independent due to the two overlapping years. The standard error of the difference is calculated by $\sqrt{1\text{-C}}\sqrt{\mathrm{SE}_{\widetilde{y_4}}^2 + \mathrm{SE}_{\widetilde{y_3}}^2}$, where C is the fraction of overlapping years, representing the approximation of the sampling correlation. Here, C is used to account for the covariance between $\widetilde{y_4}$ and $\widetilde{y_3}$ due to the lack of knowledge.

On the other hand, MYEs can be interpreted relatively easily by comparing different time periods that are not overlapped across different regions. The standard error for the difference between two estimates of non-overlapping time periods can be approximately calculated with the square root of the sum of the variances of estimates ($\sqrt{SE_1^2 + SE_2^2}$) since those are nearly independent (Beaghen et al., 2012). For example, if the estimates of 2000-2002 and 2003-2005 are 8.2 with a SE of 0.11 and 8.5 with a SE of 0.13, the difference between these two estimates is 0.4 with a SE of 0.17 ($\sqrt{0.11^2 + 0.13^2}$).

3.3 Inference for a Given Year with Several Years of Data

Multi-year period estimates are better ones than single-year estimates when some areas have experienced significant changes in interested characteristics (Citro and Kalton 2007). As we discussed, however, it also results in complications in interpretation. For this reason, strategies to derive an estimate of a single year have been studied: essentially, it can be viewed as the weighted sum of sample averages of multiple samples if the sample size and design for annual samples are assumed to be the same.

Let's think about n-year cumulated samples. $\bar{y} = \sum w_i \bar{y}_i$, where $\bar{y}_i$ is the mean for year i, w_i is the weight applied to year i, and $\bar{y}$ can be viewed as either an estimate of the current level or a multiyear estimate. In addition, $\sum w_i = 1$. The choice of w_is implies different interpretations. Setting w_1=1 and all other w_i=0 represents the practice of using the sample in the base (first) period to infer on the population. On the other hand, setting w_n=1 indicates that only data in the last sample are utilized. This weighting method may work well when one is more interested in the current year than over several years; this is especially the case when an area experiences large fluctuations in recent years (Kish 1999), for example, New Orleans after Hurricane Katrina. However, the downside is the large sampling error when only data for the current year is used, resulting in a lower degree of precision for the estimates. Obviously, 1-year estimates for large domains are reliable due to their relatively larger sample sizes. If there are small changes in trend, researchers can employ equal weights for all years, that is, w_i=1/n. It also means that variations of estimates from different years are simply due to sampling variation not because of systematic errors. If we assume that the population does not change significantly over 3 or 5 years and the variance of 1-year estimates are the same, and ignore adjustments for non-response and calibration, multiyear estimates from ACS are simply simple average of the 1-year estimates with the equal weighting scheme.

Lastly, one may allow weights to vary with samples. For example, one may set $w_1 \leq w_2 \leq \ldots \leq w_i \leq w_{i+1} \ldots$, such that samples in later periods are weighted more heavily than the early ones. When each year estimates may change to some extent, it is reasonable to assume that the estimates of recent years would reflect the current condition better than estimates of early years. This weighting scheme indicates that recent data are more representative than older data and these weights can be simply determined by models, empirical data, or their combination. The possible methods for determining weighting schemes are briefly described in the next section. However, it should be also noted that the variance of this non-decreasing weighting scheme increases substantially compared to that of equal weighting (Citro and Kalton 2007). For example, two weighting schemes can be compared, one for equal weights (0.2 for 5-year estimate) and the other for increasing weights (0.05, 0.15, 0.2, 0.25, and 0.35 for 5-year estimate). The variance of $\bar{y}$ is calculated as $\sigma^2 \sum w^2$ when the sample size of each year and the variance of the each year estimate are the same. The variances of two schemes are $0.2\sigma^2$ and $0.25\sigma^2$, respectively.

In general, an equal weight scheme provides the minimum variance compared to other weighting schemes. In the case of ACS, Census Bureau has decided to use period estimate with the equal weighting scheme to reduce serious biases due to rapid changes over period and obtain the lowest variance.

3.4 Possible Methods for Determining Non-Decreasing Weight Schemes

As we discussed, ACS recommends using annual estimates only for large areas where population size is greater than 65,000 people. For the smaller areas, the Census Bureau proposes to combine information across time depending on the population size. However, given multiple years of data, there are methods that can be employed to infer on the current population with time-series and smoothing modeling approaches. These methods take different weighting schemes. Below, a simple procedure about how the non-decreasing weights to the 1-year estimates can be determined with statistical models is described. Let's assume a signal-plus-noise model with time-series annual estimatesY_t.

$$\boldsymbol{Y_t = \theta_t + \varepsilon_t} \qquad \text{Equation 6}$$

where θ_t is the true population characteristic and ε_t includes both sampling and nonsampling errors. To simplify the computation, let's assume that ε_t is uncorrelated over time. Technically, the random vector of the annual estimates and unobserved population during m years can be written as follow:

$$Y_t = (Y_{t-m+1}, \dots, Y_t)^\intercal \text{ and } \theta_t = (\theta_{t-m+1}, \dots, \theta_t)^\intercal$$

For instance, the random vector of 3 annual estimates at 2012 are $Y_{2010\,(2012-3+1)}, Y_{2011}, and\ Y_{2012}$. Then, we can consider linear estimators of θ_t as the multiplication of a m*m matrix, W, and Y_t, that is, $\widehat{\theta_t} = WY_t$. This matrix W is the weighting schemes we are interested in and can be chosen in several ways. For example, the principle of filter design commonly employed in the time-series analyses or smoothing splines can be used. Specifically, stochastic (e.g. Kalman filter) or deterministic (e.g. nonparametric regression) models can be employed to determine the weighting schemes. Besides using statistical models, another simple way is to use weights proportional to period. For example, one can use 1/6, 2/6, and 3/6 for weights of three 1 year estimates and expand this principle with several domains (Friedman et al. 2002). After we calculate diverse versions of weighting schemes, estimates can be compared by using a certain standard like mean squared error (Friedman et al. 2002; Citro and Kalton 2007).

3.5 Comparing Estimates between Areas

In principle, the same period estimates should be used to compare different areas. For Instance, it would be undesirable to compare a 1-year estimate for a large city and a MYE for a smaller city because the potential difference between these two areas is confounded by the potential difference in trend.

3.6 Comparing Estimates over Time

Similar to the comparisons between areas, ACS recommends using the same length of time periods estimates to compare estimates over time (e.g. two 1-year estimates or two 5-year estimates). One question is whether or not we should compare two estimates with overlapping time periods. As we discussed above, using non-overlapping multiyear estimates will be simpler and straightforward to interpret. It is still possible to compare estimates of overlapping time periods but the users should remember that the difference between these two estimates is not the change between the last 2 years. In addition, the correlation between two estimates due to the overlapping time periods should be considered in the estimation of standard errors.

3.7 Possible Influences of Irregularity in Collecting Data on Multiyear Estimates

Let's think about ACS multiyear estimates. The weights are developed starting with sample selection probabilities and then a number of adjustments are made for corrections (U.S. Census Bureau, 2011). Among the adjustments, non-interview adjustment uses three characteristics including census tract, building type, and month of data collection. Based on a cross-classification table made by two of the three factors, samples are placed into cells. For example, selected housing units are placed into cells based on the cross-classification table made by building type and month. If the cell contains too small a sample, that is, fewer than 10 households, that cell should be collapsed into a nearby cell to obtain a sufficient sample size. Because of this, pooling data over multiple years will guarantee less collapsing due to increased sample size, which therefore, will better preserve seasonal trend effects, if there are any. This implies that if we collect data irregularly some months will be over-represented while other months may be under-represented, affecting the adjustment process. Therefore, it is possible that the estimates will be less able to preserve seasonal trends. In sum, if data are collected irregularly it will likely influence the weighting process and the estimates may not be good as the estimates from samples collected regularly.

4. IMPLEMENTATION ISSUES

To implement a rolling sample survey in the real world, several issues need to be addressed. These issues include: management issues, staffing, costs, and future developments.

4.1 Management Issues

Decentralizing the data collection organization

The ACS is managed by the Census Bureau. It is recommended that roles within the survey organization should be split for better efficiency, for instance, separating the management of the fieldwork team from other tasks of the survey (Ortuzar et al. 2011). For the ACS, the Census Bureau headquarters are responsible for the design and management of all primary survey activities, but all the National Processing Centers (NPC) of the Census Bureau are charged with survey operation. Within the headquarters, the roles are also split, for instance, the Content Council manages all requests for content changes to existing ACS content, providing guidelines for pretesting and implementing the changes (U.S. Census Bureau 2009). It will also be beneficial to the survey if the data collection agency should seek feedbacks from stakeholders (e.g., data users) on the survey approach and any potential pitfalls and obstacles (U.S. Census Bureau 2009).

Piloting and testing

Rolling sample surveys often require testing of early prototypes for a small number of sites before being fully implemented. The pilot test is essentially a rehearsal of the main survey, used to test the interviewing procedures, questionnaire wording, adequacy of the training, the field work, as well as the response rates (U.S. Census Bureau 2009). Furthermore, if there are any requests for content changes, a content test needs to be completed before the changes are made (U.S. Census Bureau 2009).

4.2 Staff Issues

A rolling sample survey conducted on an ongoing basis requires a permanent, professional team for both project management and interviewing. It is also essential to provide team members with high quality training to keep them motivated continuously (Ortuzar et al. 2011). For the ACS, besides the training session which provides interviewers appropriate interviewing procedures, their supervisors are required to travel with them to reinforce the procedures learned in training. In addition, interviewers are randomly selected each month for supervisors to re-interview a sample of their cases, thus verifying that interviewers are conducting interviews, and doing so correctly (U.S. Census Bureau, 2009).

4.3 Cost Issues

Budgetary implications are an important consideration in moving from a one-time survey to a continuous survey, but budget constraints still apply. For the ACS, a monthly sample size of 500,000 HUs was initially suggested to produce multi-year accumulations for small areas at the same level of sampling reliability as the long-form census sample (Alexander 1993) but the budget only allows a half of it, 250,000 addresses monthly. Furthermore, to obtain sufficient sample sizes, data users from small geographic areas have to wait for one to several years to have access to the aggregated data products, although samples are collected yearly, monthly, or even daily in a rolling sample survey.

4.4 Future Development

Continuous rolling sample surveys, compared to one-time surveys, provide smaller samples, but collect richer, more in-depth data over a longer period. One challenge in rolling sampling surveys is how to reduce the respondents' self-reporting burdens without losing details. Since the past decade, various technologies have been increasingly used to assist in data collection.

For the ACS, staring from 2013, an internet response option, aiming to maximize responses, will be added to the mail data collection phase as guided by the 2011 ACS internet test (U.S. Census Bureau 2012). In previous studies (Brady 2004; Bentley 2006), the strategy of pushing households to respond by the Internet has not proved effective. However, the internet instrument was retested in the ACS content test. This is largely driven by its potential cost savings associated with printing, mailing, data capture, and nonresponse follow-up costs, if the instrument is successful in maintaining or even increasing response (U.S. Census Bureau 2012).

REFERENCES

Aarts, H., B. Verplanken and A. V. Knippenberg (1997). "Habit and information use in travel mode choice." Acta Psychologica **96**: 1-14.

Active Living Research. (2003). "Test-retest reliability coefficeints for individual items of the Neighborhood Environment Walkability Scale." from http://www.activelivingresearch.org/files/NEWS_Item_Reliability_0.pdf.

Alexander, C. H. (1993). Three general pototypes for a continuous measurement system. Internal Census Bureau Reports CM-1. Washington, DC, U.S. Census Bureau.

Alexander, C. H. (2002). "Still rolling: Leslie Kish's "rolling samples" and the American Community Survey." Survey Methodology **28**(1): 35-42.

Alexander, C. H., Navarro, A (2003). The quality of estimates from the American Community Survey for small population groups. The Joint Statistical Meetings - Section on Survey Research Methods, San Francisco, California.

Ampt, E. S. and J. d. D. Ortúzar (2004). "On best practice in continuous large−scale mobility surveys." Transport Reviews **24**(3): 337-363.

Ampt, E. S., A. Sinclair Knight Merz and A. J. de Dios Ortúzar (2009). On large scale on-going mobility surveys: The state of practice. Transport Survey Methods: Keeping up with a Changing World. M. L.-G. P. Bonnel, J. Zmud, J.-L. Madre. Bingley, Emerald**:** 505-532.

Battelle (1997a). Global Positioning Systems for Personal Travel Surveys: Lexington Area Travel Data Collection Test. A Report to Office of Highway Information Management (HPM-40). Columbus, Ohio.

Battelle (1997b). Lexington area travel data collection test final report. Washington, DC, Federal Highway Administration.

Battellino, H. and J. Peachman (2003). "The joys and tribulations of a continuous survey." Transport Survey Quality and Innovation: 49-68.

Beaghen, M., T. McElroy, L. Weidman, M. Asiala and A. Navarro (2012). Interpretation and use of American Community Survey multiyear estimates. Research Report Series (Statistics #2012-03). . Washington, DC, U.S. Census Bureau.

Beaghen, M. and L. Weidman (2007). Statistical issues of interpretation of the American Community Survey's one-, three-, and five-year period estimates. The Joint Statistical Meetings, Salt Lake City, UT.

Bentley, M., Tancreto, J. (2006). 2005 National Census Test: self-response options analysis. 2010 Census Test Memorandum Series: 2005 National Census Test, No. 26. Washington, DC, U.S. Census Bureau.

Bento, A. M., M. L. Cropper, A. M. Mobarak and K. Vinha (2005). "The effects of urban spatial structure on travel demand in the United States." The Review of Economics and Statistics **87**(3): 466-478.

Bhat, C. R. (2000). "Incorporating observed and unobserved heterogeneity in urban work travel mode chocie modeling." Transportation Science **34**(2): 228-238.

Brady, S., Stapleton, C., Bouffard, J. (2004). 2003 National Census Test: response mode analysis. DSSD 2003 Memorandum Series # B-02. Washington, DC, U.S. Census Bureau.

Brownson, R. C., J. J. Chang, A. A. Eyler, B. E. Ainsworth, K. A. Kirtland, B. E. Saelens and J. F. Sallis (2004). "Measuring the environment for friendliness toward physical activity: A comparison of the reliability of 3 questionnaires." American Journal of Public Health **94**(3): 473-483.

Cerin, E., E. Leslie, L. d. Toit, N. Owen and L. D. Frank (2007). "Destinations that matter: Associations with walking for transport." Health & Place **13**(3): 713-724.

Cervero, R. (2002). "Built environments and mode choice: toward a normative framework." Transportation Research Part D **7**(4): 265-284.

Cervero, R. and M. Duncan (2002). Residential self selection and rail commuting: A nested logit analysis. California, University of California Transportation Center Berkeley.

Chand, N. and C. H. Alexander (2000). Multi-Year Averages from a Rolling Sample Survey. Washington, DC, U.S. Census Bureau.

Chen, C. and J. Chen (2009). "What is responsible for a response lag of a significant change in discretionary time use: the built environment, family and social obligations, temporal constraints, or a psychologicaly delay factor?" Transportation **36**: 27-46.

Chen, C., J. Chen and H. Timmermans (2009). "Historical deposition influence in residential location decisions: a distance-based GEV model for spatial correlation." Environment and Planning A **41**(11): 2760-2777.

Chen, C., H. Gong, C. Lawson and E. Bialostozky (2010). "Evaluating the feasibility of a passive travel survey collection in a complex urban environment: lessons learned from the New York City case study." Transportation Research Part A **44**(10): 830-840.

Chen, C., H. Gong and R. Paaswell (2008). "Role of the built environment on mode choice decisions: additional evidence on the impact of density." Transportation **35**(3): 285-299.

Chen, C. and H. Lin (2011). "Decomposing Residential Self-selection via a Life Course Perspective." Environment and Planning A **43**(11): 2608-2625.

Chen, C. and H. Lin (2012). "How far do people search? Analyzing the roles of housing supply, intra-household dynamics, and the use of information channels." Housing Studies **27**(7): 898-914.

Citro, C. F. and G. Kalton (2007). Using the American Community Survey: Benefits and Challenges. Washington, D.C., National Academies Press.

Dieleman, F. M., W. A. V. Clark and M. C. Deurloo (2000). "The geography of residential turnover in twenty-seven large US metropolitan housing markets, 1985-95." Urban Studies **37**(2): 223-245.

Edmonston, B. (2001). "The case for modernizing the US census." Society **39**(1): 42-53.

Forsyth, A., K. Krizek and A. W. Agrawal (2012). "Reliablity testing of the PABS (Pedestrian and Bicycling Survey) Method." Journal of Physical Activity and Health **9**(5): PMID: 21946264.

Forsyth, A., K. J. Krizek, A. W. Agrawal and E. Stonebraker (2012). "Reliability testing of the Pedestrian and Bicycling Survey(PABS) method." Journal of Physical Activity & Health **9**(5): 677-688.

Frank, L., Mark Bradley, Sarah Kavage, James Chapman, and T. Lawton (2008). "Urban form, travel time, and cost relationships with tour complexity and mode choice." Transportation **35**(1): 37-54.

Frank, L. D., M. Bradley, S. Kavage, J. Chapman and T. K. Lawton (2008). "Urban form, travel time, and cost relationships with tour complexity and mode choice." Transportation **35**(1): 37-54.

Frank, L. D. and G. Pivo (1994). Relationship Between Land Use and Travel Behavior in the Puget Sound Region. Olympia, WA, Washington State Department of Transportation.

Friedman, E. M., D. Jang and T. V. Williams (2002). Combined Estimates from Four Quarterly Survey Data Sets. The American Statistical Association, Alexandria, VA.

Gong, H., C. Chen, E. Bialostozky and C. Lawson (2011). "A GPS/GIS method for travel mode detection in New York City." Computers, Environment, and Urban Systems **36**(2): 131-139.

Grazi, F., J. C. van den Bergh and J. van Ommeren (2008). "An empirical analysis of urban form, transport, and global warming." The Energy Journal **29**(4): 97-122.

Griffiths, R., A. Richardson and M. Lee-Gosselin (2000). Travel surveys. Transportation in the New Millennium. Washington, DC, Transportation Research Board.

Hato, E. and Y. Asakura (2001). New approach for collection of activity diary using mobile communication systems. 80th Annual Meeting of the Transportation Research Board, Washington, DC.

Hefter, S. (2005). American Community Survey: specifications for selecting the computer assisted personal interview samples. Washington, DC, Census Bureau.

Khattak, A. J. and D. Rodriguez (2005). "Travel behavior in neo-traditional neighborhood developments: A case study in USA." Transportation Research Part A **39**: 481-500.

Kincannon, C. L. (2003). American Community Survey witnesses (testimony).

Kish, L. (1986). "Timing of surveys for public policy." Australian Journal of Statistics **28**(1): 1-12.

Kish, L. (1990). "Rolling samples and censuses." Survey Methodology **16**(1): 63-79.

Kish, L. (1999). "Cumulating/combining population surveys." Survey Methodology **25**(2): 129-138.

Kish, L. and V. Verma (1983). "Census plus samples: Combined uses and designs." Bulletin of the International Statistical Institute **50**(1): 66-82.

Kish, L. and V. Verma (1983). "Censuses plus samples: combined uses and designs." Bulletin of the International Statistical Institute **50**(1): 66-82.

Kish, L. and V. Verma (1986). "Complete censuses and samples." Journal of Official Statistics **2**(4): 381-395.

Kitamura, R. (1990). "Panel analysis in transportation planning: an overview." Transportation Research Part A **24A**(6): 401-415.

Kitamura, R., P. L. Mokhtarian and L. Laidet (1997). "A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area." Transportation **24**(2): 125-158.

Krizek, K., Y. Li and S. Handy (2005). "Spatial Attributes and Patterns of Use in Household-Related Information and Communications Technology Activity." Transportation Research Record: Journal of the Transportation Research Board **1926**(1): 252-259.

Levinson, D. M. and K. J. Krizek (2008). Planning for Place and Plexus: Metropolitan Land Use and Transport. New York, Routledge.

Litman, T. and R. Steele (2012). Land Use Impacts on Transport: How Land USe Factors Affect Travel Behavior. Victoria, Canada, Victoria Transport Policy Institute.

MacDonald, H. (2006). "The American Community Survey: warmer (more current), but fuzzier (less precise) than the Decennial Census." Journal of the American Planning Association **72**(4): 491-503.

McCormack, E., S. G. Rutherford and M. G. Wilkinson (2001). "Travel impacts of mixed land use neighborhoods in Seattle, Washington." Transportation Research Record: Journal of the Transportation Research Board **1780**: 25-32.
Moudon, A. V., B. Saelens, S. Rutherford and M. Hallenbeck (2009). A Report on Participant Sampling and Recruitment for Travel and Physical Activity Data Collection. Seattle, WA. Seattle, WA, Transportation Northwest.
Moudon, A. V. and D. W. Sohn (2005). Transportation-Efficient Land Use Mapping (TELUMI): Phase 3 of Integrating Land Use and Transportation Investment Decision-Making. Olympia, WA, Washington State Department of Transportation, Transportation Research Center (TRAC).
Moudon, A. V., D. W. Sohn, S. E. Kavage and J. E. Mabry (2011). "Transportation-efficient land use mapping index (telumi), a tool to assess multimodal transportation options in metropolitan regions." International Journal of Sustainable Transportation **5**(2): 111-133.
Moudon, A. V., D. W. Sohn, S. E. Kavage and J. E. Mabry (2011). "Transportation-Efficient Land Use Mapping Index (TELUMI), a Tool to Assess Multimodal Transportation Options in Metropolitan Regions." International Journal of Sustainable Transportation **5**(2): 111-133.
Murakami, E. (2008). Hard to Reach Populations (Presentation at the NYMTC Survey Workshop). New York.
New Zealand Ministry of Transport. (2006). "New Zealand Travel Survey 2006." from http://www.transport.govt.nz/research/travelsurvey/.
Nustats (2012). Feasibiilty Study for a Continuous Household Activity Survey Program (CHASP). Task 3 Implementation Plan. Calgary, City of Calgary.
Olsson, A.-L. L. (2003). Factors that Influence Choice of Travel Mode in Major Urban Areas, Royal Institute of Technology.
Ortúzar, J. D., J. Armoogum, J.-L. Madre and F. Potier (2011). "Continuous mobility surveys: the state of practice." Transport Reviews **31**(293-312).
Ortuzar, J. D. D., J. Armoogum, J. L. Madre and F. Potier (2011). "Continuous mobility surveys: the state of practice." Transport Reviews **31**(3): 293-312.
Outwater, M. L., S. Castleberry, Y. Shiftan, M. Ben-Akiva, Y. S. Zhou and A. Kuppam (2003). "Attitudinal market segmentation approach to mode choice and ridership forecasting: structural equation modeling." Tranportation Research Record **1854**: 32-42.
Parkany, A. E. and E. Murakami (2004). Seasonality of transportation data. The Eirass, Moving Activity Travel Approaches to Practice, Maastricht, the Netherlands.
Rofique, J., A. Humphrey, K. Pickering and S. Tipping (2011). National Travel 2010 Technical Report. Washington, D.C., Department of Transportation.
Rust, K. (2010). Potential Role of the American Community Survey in Sampling Rare Populations. CNSTAT Workshop on the Future of Federal Household Surveys. Washington DC.
Rust, K. (2010). Potential Role of the American Community Survey in Sampling Rare Populations. Presented at the CNSTAT Workshop on the Future of Federal Household Surveys. Washington D.C., Westat.
Scarr, H. A. (1994). Continuous Measurement - Harry A. Scarr to Association of Public Data Users (APDU). Washington, DC.

Schwanen, T. and P. Mokhtarian (2005). "What affects commute mode choice: neighborhood physical structure or preferences toward neighborhoods?" Journal of Transport Geography **13**: 83-99.

Shadish, W. R., T. D. Cook and D. T. Campbell (2002). Experimental and Quasi-experimental Designs for Generalized Causal Inference. New York, Houghton Mifflin Company.

Steiner, R. and S. Srinivasan (2010). VMT-based Traffic Impact Assessment: Development of a Trip Length Model. Gainesville, FL, Center for Multimodal Solutions for Congestion Mitigation.

Stopher, P., E. Clifford, J. Zhang and C. FitzGerald (2008b). Deducing mode and purpose from GPS data. Working Paper of the Austrian Key Centre in Transport and Logistics. Sydney, Australia, University of Sydney.

Stopher, P., C. FitzGerald and M. Xu (2007). "Assessing the accuracy of the Sydney household travel survey with GPS." Transportation **34**: 723-741.

Stopher, P., K. Kockelman, S. P. Greaves and E. Clifford (2008d). Reducing burden and sample sizes in multi-day household travel surveys. 87th Annual Meeting of the Transportation Research Board. Washington, D.C.

Stopher, P., Y. Zhang, J. Zhang and B. Halling (2009). Results of an evaluation of TravelSmart in South Australia. The 32nd Australasian Transport Research Forum (ATRF), Auckland, New Zealand.

Stopher, P. R. and S. P. Greaves (2007). "Household travel surveys: where are we going?" Transportation Research Part A **41**: 367-381.

Stopher, P. R. and S. P. Greaves (2007). "Household travel surveys: Where are we going?" Transportation Research Part A: Policy and Practice **41**(5): 367-381.

U.S. Census Bureau (2003). Testing the Use of Voluntary Methods - Additional Results. U.S. Census Bureau. Washington, DC.

U.S. Census Bureau (2004). Testing the Use of Voluntary Methods - Additional Results. U.S. Census Bureau. Washington, DC.

U.S. Census Bureau (2008). A Compass for Understanding and Using American Community Survey Data: What General Data Users Need to Know. U.S. Government Printing Office. Washington, DC.

U.S. Census Bureau (2009). Design and Methodology: American Community Survey. U.S. Census Bureau. Washington, DC.

U.S. Census Bureau (2012). Design of the American Community Survey Internet Instrument. United States Department of Commerce. Washington, DC.

Van Acker, V. V. and F. Witlox (2011). "Commuting trips within tours: how is commuting related to land use? ." Transportation **38**(3): 465-486.

Wilson, J. (2004). Measuring Personal Travel and Goods Movement (Transportation Research Board Special Report 277). Washington, D.C., Transportation Research Board.

Wolf, J., S. Hallmark, M. Oliveira, R. Guensler and W. Sarasua (1999). "Accuracy issues with route choice data collection by using Global Positioning System." Transportation Research Record **1660**: 66-74.

Wolf, J., M. Oliveira and M. Thompson (2003). "Impact of underreporting on mileage and travel time estimates--results from Global Positioning System-enhanced household travel survey." Transportation Research Record **1854**: 189-198.