4-1-2019

Who Do We Miss by Moving Travel Surveys Online? –Assessments from Vermont

Jonathan Dowds
University of Vermont, jdowds@uvm.edu

Lisa Aultman-Hall
University of Vermont

A Report from the University of Vermont Transportation Research Center

Who Do We Miss by Moving Travel Surveys Online? – Assessments from Vermont

Final Report

April 2019

Lisa Aultman-Hall and Jonathan Dowds

Who do we miss by moving travel surveys online? – Assessments from Vermont

April 2019

Prepared by:

Lisa Aultman-Hall
Jonathan Dowds

Transportation Research Center
Farrell Hall
210 Colchester Avenue
Burlington, VT 05405

www.uvm.edu/cems/trc/research-reports/

Acknowledgements

This paper was extracted from a project funded by the Vermont Agency of Transportation and the Chittenden County Regional Planning Commission. We gratefully acknowledge their support. We appreciate the opportunity to partner for data collection with RSG Inc. The conclusions drawn are those of the authors not our partners.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the UVM Transportation Research Center. This report does not constitute a standard, specification, or regulation.

This report content overlaps with but expands upon TRC Report 17-004, "Designing the All-in-One Vermont Transportation Survey," by Lisa Aultman-Hall and Jonathan Dowds.

Table of Contents

List of Tables

List of Figures

Abstract

Online travel surveys are increasingly common because of cost, user burden, and geocoding advantages. Consequently, it is important to ask how online survey samples compare to paper survey samples. This study compares paper and online responses to a 2016, state-wide, Vermont transportation planning survey. Internet and smartphone access were analyzed by socioeconomic characteristics as well as by residential location to assess rural coverage. Respondents' selection of the paper option was linked to lower population density. Online respondents showed significant spatial clustering. Crucially, the travel behavior and transportation attitudes of paper and online respondents differed even after weighting for demographic attributes. Smartphone ownership in Vermont is too skewed by age to be a primary travel survey method. Internet access is more widespread but does exclude some population segments. We recommend consideration of respondents by geographic location as well as socioeconomic characteristics when selecting survey mode and weighting, especially for state-wide surveys.

Keywords: Travel surveys, Survey modes, Survey design, Travel behavior, Bias

1. Introduction

Historically, paper mail-back and telephone surveys have been primary survey data retrieval methods for transportation agencies in the United States for household travel and public opinion surveys. Limitations related to response rates, sample representativeness, and the decreasing prevalence of landline telephones along with the simultaneous growth in Internet access and Global Positioning System (GPS)-enabled technology have spurred an rapid evolution in data collection methods. Numerous agencies are implementing web-based or mobile device-based data collection both to reduce costs and to take advantage of new opportunities for improving the quantity, accuracy and completeness of travel data collection. This trend toward online and mobile device-based surveys raises important questions about who may be excluded by this transition.

The need for new survey methods has been recognized for decades (Ampt and Bonsall 1997; Axhausen et al. 2007; Lee, Sener and Mullins 2016). In this rapidly changing environment, many challenges still exist, including sociodemographic and urban/rural difference in Internet and smartphone access. Urban/rural disparities may be particularly important for state-wide travel surveys since land use, travel demand patterns and transportation services vary between rural and more urbanized areas. In 2010, only 17.4% of Vermont's 625,000 residents lived in an urbanized area, 21.5% lived in an urban cluster and 61.1% lived in a rural area. Therefore, the question of who is excluded when Vermont uses online surveys includes not only questions about sociodemographic groups, but also rural versus urban residents, as would be the case for most state-wide survey efforts.

In 2016, the Vermont Agency of Transportation (VTrans) conducted a Long Range Transportation Planning Survey (LRTPS) of 2,232 respondents who used either paper (57.5%) or web-based retrieval (42.5%) methods. Respondents were recruited through address-based random sampling by Resource Systems Group (RSG) Inc. LRTPS data are analyzed here to generate insights about coverage by retrieval mode and community type. In order to consider responses by home location we made use of geo-coded home addresses of invitees. Additionally, data on smartphone ownership from a distinct 2017 telephone-based survey are reviewed in order to consider the viability of mobile device-based data collection.

These datasets offer an opportunity for an assessment of the characteristics of who may be missed when state-wide travel surveys move online. Following a background section, this paper considers the following research questions:

(1) Are responses and/or chosen retrieval mode spatially correlated?

(2) How do paper and online respondents differ from each other?

(3) How does Internet and smartphone access vary between different groups of Vermonters and by residential location?

2. Background

2.1 Transportation Survey Programs

Agencies responsible for both metropolitan area and state-wide travel surveys are currently implementing changes in their survey programs. Numerous recent surveys have been conducted using the Internet or mobile device apps. The United States Department of Transportation (USDOT) Federal Highway Administration's (FHWA) National Household Travel Survey (NHTS) is one of the most comprehensive one-day travel surveys conducted in the United States. The 2016 dataset includes travel data for all persons at least 5 years of age from approximately 129,000 households and was be collected using a web-based format for the first time (Transportation Research Board 2016; Westat 2015).

Many other agencies in the United States, mostly larger metropolitan areas or state Departments of Transportations (DOTs), also conduct their own travel surveys. For example, the New York Metropolitan Transportation Council collected their regional household travel survey in 2010-2011 from 19,000 households in 29 counties in 3 states using phone, mail, web, and wearable GPS (NYMTC and NJTPA 2013). California's 2010-12 survey was the largest outside of the NHTS program, including approximately 44,000 households (Kunzman and Daigler 2013). The Atlanta Regional Commission surveyed 10,000 households in 2011 and included 10% GPS collection (PTV NuStats 2011). Many of these cases are not scaled appropriately for replication in rural Vermont with a state population of 626,000. With a smaller sample size, weighting of responses is more challenging. Moreover, broadband Internet access and cell signal coverage are more significant issues in rural areas which are present in most state-wide study areas.

Some of the most recent, and most expensive, travel surveys were those with GPS components. Stand-alone, in-vehicle GPS device surveys, which showed promise to change the survey landscape in the mid-1990s, (Ampt and Bonsall 1997; Wolf et al. 1999; Murakami and Wagner 1999) have fallen out of favor quickly for the more straightforward mobile-device based GPS data collection (Janzen et al. 2018; Lee, Sener and Mullins 2016; Ritter and Greene 2017), pointing to exciting future options. Despite valuable efforts to standardize travel survey methods (Stopher et al. 2008), they remain dynamic and data comparing methods remain limited. This paper compares paper and web-based surveys as well as potential coverage limitations of smartphone surveys. Retrieval methods need to be considered carefully if they could exclude segments of the population due to access to technology, especially if access correlates with travel patterns or needs.

2.2 Survey Content

The data collected during a transportation survey falls into five typical categories:

- **socio-demographic and household context**;
- **general transportation and travel activity questions**;
- **attitudes about transportation issues** (including lifestyle factors);

- **customer satisfaction questions** (increasingly common to meet the performance measurement requirements in MAP-21 and the FAST Act for customer satisfaction; and

- travel behavior data, collected through a **travel diary or log.**

The **travel diary or log** is the most complicated transportation survey data to obtain. The need for the travel log influences the choice of survey data retrieval method. These sections of a survey seek to track all travel activities undertaken by a specific person/household for a given study period, typically one day. The data collected include the origin, destination, time of departure, mode(s), travel party, length and purpose of all travel. Under-reporting of trips has been a long-standing challenge (Bricka and Bhat 2006; Brog et al. 2982 and Hassounha et al. 1993; Wolf et al 2003). More extensive survey efforts have focused on recording a person's complete daily activities, recognizing that travel is most often a derived demand. In these cases, simulation of synthetic households might often be the intended model driver (Pritchard and Miller 2012 for example) but these techniques require very large datasets are more advanced than the methods typically used for state-wide modeling, including in Vermont.

2.3 Retrieval Methods

Every survey data retrieval method has specific advantages and disadvantages for transportation and travel surveys. Three established and two emerging methods are defined below. Each data retrieval method's performance on five key evaluation criteria are highlighted in Tables 1 and 2.

Table 1. Characteristics of Established Travel Survey Retrieval Methods

Retrieval Method	Paper Survey	Telephone Survey	Online Survey
Demographic Representation	No inherent limitations on demographic representation. Older participants and women have higher response rates in general.	Often limited to households with land lines which excludes cell-phone only households. Likely to over represent older individuals.	Limited to respondents with Internet access. May under represent older and/or low income respondents. Possible geographic variability given slower Internet in rural areas.
Completeness of Data	Methods that rely exclusively on respondent recall may not be as accurate as those that provide prompts based on automatically recorded location (see Table 2). Shorter trips, some legs of tours and non-motorized travel are most often missed. These methods easily facilitate collection of data for every person in s household (including children), either directly or by proxy reporting, thus creating complete household-based data.		
Spatial Accuracy of Location Data	Location data is limited to a street address or street intersection. Requires significant post-processing and generally has only moderate spatial accuracy.		Locations can be selected/confirmed on an interactive map, reducing the need for post-processing and increasing accuracy.
Participant Burden	Increasing question number and complexity create significant burden.	Increasing question number and complexity increase burden. Respondent can find it helpful to have interviewer assistance.	Survey burden may be lower as questions can be tailored to the specific respondent (e.g. skipping questions). Surveys may be stopped and continued later. Data may be auto-populated for repeat trips.
Cost	Mail and printing costs can be significant and are proportional to sample size.	Costs are proportional to sample size.	Low marginal costs for increasing sample size. Telephone support can be costly.

Table 2. Characteristics of Emerging Travel Survey Retrieval Methods

Retrieval Method	Mobile Device Survey App	Secondary Data Sources
Demographic Representation	Limited to respondents with smartphones although some agencies have tried loaner programs. Likely to underrepresent older and/or low income individuals to a greater degree than online surveys. Possible geographic variability due to variable cellular service.	Representativeness varies by source. Individual demographic data not included. Data are usually provided on an aggregate basis to protect confidentiality which is more challenging in rural zones.
Completeness of Data	Can improve trip recall, especially of shorter and discretionary trips including active travel, by location prompts for probable trips. Data may be missed when phone is off or has poor cell/GPS signal strength.	Proponents argue that this source produces large sample sizes that have more representative coverage than surveys.
Spatial Accuracy of Location Data	Locations are best auto-populated from phone GPS and can be confirmed on an interactive map, reducing the need for post-processing and increasing accuracy. Cell tower-based locations are less accurate than GPS.	Depends on data source and aggregation procedures.
Participant Burden	Survey burden may be especially low since questions can be tailored and auto-populated, including for repeat trips. Participants may incur data costs, device battery drain and have privacy concerns.	None.
Cost	Low marginal costs for increasing sample size. Telephone support can be costly.	Purchase prices tend to be substantial.

Note that survey recruitment, in which individuals or households are invited to participate, is distinct from data retrieval when the data are collected.

1. Paper Survey: Respondents fill-out and mail-back a hard copy paper survey. The VTrans LRTPS 2016 used a mixed web and paper retrieval method and had an 18.4% response rate (41.9% of respondents utilized the web-based survey and 58.1% completed the paper survey).

2. Telephone Survey: Respondents are contacted by phone (landline and cell phones are now both possible) and an interviewer records a respondent's answers and may prompt for additional details. The 2017 Vermonter Poll, part of the data used here, conducted by landline and cell phone had a 20.1% response rate. Telephone surveys are still viable in Vermont, perhaps due to the older and rural population.

3. Online Survey: Respondents fill out a web-based, electronic survey on a computer, tablet or smartphone. This was one of the two retrieval methods used for the 2016 VTrans LRTPS. Some analysis has shown that online surveys tend to have lower response rates compared to mail-back surveys (Manfreda et al. 2008) but other research shows that a combination of mail and email recruitment contacts can yield comparable response rates (Millar and Morgan 2011).

4. Mobile Device App: Respondents use a smartphone App designed specifically for transportation data collection. Most Apps collect some data automatically and prompt the respondent to enter other data after a trip or day has been completed. Most infer data based on tracking location. Some passive Apps do not require any user input, and thus reduce the burden but cannot obtain information about trip purpose, attitudes, complete demographics or travel party. Unless specifically noted, references to mobile device data collection here refer to Apps that include both active and passive data collection.

5. Secondary Data Sources: Travel behavior data can be purchased from some 'big data' sources including blue-tooth, cell towers, or credit cards. Private companies sell this aggregated data, especially travel demand volume tabulated by Origin-Destination (OD) zones over a certain time period. In some cases, data providers can disaggregate these volume interchanges by crude trip purposes such as segregated work-based trips from others. Transportation agencies usually provide their traffic analysis zones (TAZ) which may range in size from a single block in an urban area to a good portion of a county in a rural area. An OD matrix is returned from the vendor (disaggregated by time of day or purpose). Home and work locations are often inferred but sociodemographic data for individuals is not known. Although the details are not discussed in this paper, our project included considering cell-tower based travel data for Vermont. In rural areas, low population density necessitated zone aggregation that was incompatible with the State DOT's travel demand model. Moreover, the data sources usually do not include international travelers, an important factor in border states including Vermont which is proximate to the Canadian Province of Quebec within which the metropolitan area of Montreal (population approximately 4 million) generates a significant number of trips.

The first four transportation planning survey components described above (all but the travel diary) usually involve multiple choice, Likert scales and limited open-ended questions. Transportation surveys with only these components could reasonably be conducted using any of the five data retrieval methods. Given cost and data accuracy considerations, the travel diary component of a transportation survey program can reasonably be conducted using a web-based or mobile device-based survey. Although travel diaries were historically collected by telephone, including proxy reporting for other household members, this practice is very limited at this time in the United States. Most transportation survey professionals now assess that it is not practical to collect accurate location data, and therefore a travel log, by paper and that the length of a telephone survey that includes a travel diary is an unreasonable burden. Data accuracy, as well as cost and demographic coverage, is a consideration in selecting a retrieval method. However, the research question in this paper is still pertinent: when we move surveys online, who do exclude?

One primary motivating factor for moving surveys online has been cost (Dijst et al. 2006). Cost can be highly variable and influenced by recruiting strategies, response rates and the length and/or complexity of the survey. It is challenging to meaningfully compare costs. Hartgen and San Jose (2009) report an average cost of $150 per completed survey with higher per unit costs for smaller sample sizes. The NHTS 2016 'add-on' was on the higher end ($225) due to its scope and scale. Conversations with survey managers nationwide during this project in 2016 indicate recent costs of $145-$225 per completed household for a travel survey. Many survey efforts include some GPS or mobile app data, the later often being much less costly than the former. In general, costs make online surveys attractive. Assuming solid recruiting, online surveys have predictable costs for large sample sizes and, as described above, offer lower burden and the ability to collect accurate geocoded information.

Travel survey response bias has long been recognized (Kam and Morris 1999 and Stopher and Greaves 2007). In the late 1990s experts were advocating for paid longitudinal panels as a means to gather better travel data (Ampt and Bonsall 1997; Stopher and Greaves 2007). At the time GPS-based surveys were recognized as adding spatial accuracy (Sharp and Murakami 2005) but still costly. The advent of GPS-enable smartphones is now hailed as the cost-effective solution specifically for smaller size agencies (Flake et al. 2017) and deemed to be appropriate for the reconsideration of panels. Many of the technical challenges related to battery life and participant privacy concerns are now addressed (Safi et al. 2015).

During the last 20 years, only a handful of travel studies have compared the attributes and representativeness of different retrieval modes. In 2006, using data from an e-shopping survey, Dijst et al. (2006) demonstrated that different people pick online and paper surveys, and that online data were comparable and in some cases better. An early Swiss comparison of online and paper studies, in 2008, revealed that better quality data were possible with well-designed web interfaces but that respondents expressed security concerns (Weiss et al. 2008). Agrawal et al. (2017) considered online and paper surveys for transit users and found differences that suggested that online surveys perform less well for transit riders. Findings by Cummins et al. (2013) were a bit more optimistic, suggesting that online methods could, in some cases, substitute well for onboard transit surveys. Xan and Handy (2014) compared online and telephone surveys for bicyclists and found that survey mode impacts resultant measures and models. While analysis by Lee and Pino (2012) leads to the suggestion that as the Internet continues to proliferate that online surveys will perform as well as telephone surveys. Recent work also indicates that smartphone surveys produce comparable data quality to web-surveys completed on a personal computer although care needs to be taken with sliders and other design elements that can be difficult to manipulate accurately on a small screen (Antoun, Couper and Conrad, 2017). Smith and Spitz (2010) consider several respondents with and without Internet access in two metropolitan areas and conclude that the bias introduced by lack of Internet access is limited. In the broader survey community, Sterret et al. (2017) compared the Internet access data from the General Social Survey (GSS) to assess changes in coverage bias for Internet surveys and found the coverage bias associated with education, income, race, and age had declined, but not been eliminated, between 2006 and 2014.

It is clear that some bias is introduced by survey mode but the existing understanding of these differences is limited due to lack of comparative data. Concerns remain that new retrieval methods may exclude vulnerable populations that have limited or no access to technology. It is reasonable to hypothesize that the bias is greater in rural areas. Most states have rural areas, and this comparison of Vermont paper versus online survey respondents contributes to filling this data gap.

3. Data

The survey data used here consists of data collected RSG Inc. for VTrans in 2016 to support statewide planning. The LRTPS survey did not include a travel diary but did include the other four types of transportation survey questions described above. The random address-based sample was recruited from Vermont residents using two mailings. An initial postcard invited participants to complete a web-based survey. The second mailing included the survey web address as well as a paper version of the survey. An 18.4% response rate was achieved overall. Only 7.8% of invitees responded after the first mailing. The household member over 18 years old with the most recent birthday was instructed to complete the survey. Of the final sample of 2,232, 42% completed the survey online and 58% completed mailed-back paper surveys. In total 12,000 households were recruited to participate and the address information for 10,208 of these households (including 1,876 out of 2,232 responding households) was sufficient for geo-location.

A secondary source of data was used in this project. The annual Vermonter Poll survey conducted by the UVM Center for Rural Studies in February, 2017 used random digit dialing of landline and cell phones. The survey which aimed to generate a representative sample of the whole state, included the following question at the request of this project team: 'How many adults (including yourself) in your household have a data-enabled cell phone, that is a cell phone that can access the Internet?'. Information about data enabled cell phone ownership for 590 respondents was collected along with their sociodemographic data.

4. Results

4.1 Are responses and/or chosen retrieval mode spatially correlated?

Response rates and the proportion of respondents using online and paper surveys are tabulated by county in Table 3. The proportion of online respondents is statistically significantly different at the county level based on Chi-square tests. Moreover, the proportion paper decreases as county population density increases (Pearson's R -0.71, p=0.004). More rural places preferred paper at the county level. Response rates show no statistically significant differences at the county level.

Table 3. Response Rate and Retrieval Method by County (unweighted)

County	Response Rate	Retrieval Method		Population	Population Density (pop/mi²)
		Online	Paper		
Chittenden	19.2%	51.0%	49.0%	>100,000	219.7
Rutland	17.8%	41.6%	58.5%	>50,000	66.3
Washington	20.2%	49.6%	50.4%	>50,000	86.6
Windsor	17.2%	41.4%	58.6%	>50,000	58.5
Franklin	16.7%	40.0%	60.0%	30,000-50,000	75.3
Windham	17.9%	38.4%	61.7%	30,000-50,000	56.7
Addison	19.8%	43.9%	56.1%	30,000-50,000	48.0
Bennington	16.4%	30.8%	69.2%	30,000-50,000	55.0
Caledonia	20.0%	35.3%	64.7%	30,000-50,000	48.1
Lamoille	22.0%	37.9%	62.1%	10,000-30,000	53.3
Orange	18.9%	37.0%	63.0%	10,000-30,000	42.1
Orleans	16.6%	38.0%	62.0%	10,000-30,000	39.3
Essex	14.4%	33.3%	66.7%	< 10,000	9.5
Grand Isle	18.4%	38.1%	61.9%	< 10,000	85.2
Statewide Total	**18.4%**	**42.5%**	**57.5%**	**625,217**	**67.9**

Geo-located households as well as their response/non-response and mode choice status are shown in Figure 1. Unlike the more common Moran's I, which is designed for testing spatial autocorrelation in continuous variables, join-count statistical methods can be used to test for spatial autocorrelation in binary variables. As described in Morris and Doak (2002), Monte Carlo simulation was used to determine the probability of encountering a given distribution of outcome pairings between neighbors (e.g. response/response, response/non-response, non-response/non-response) within given radii thresholds. The simulation process randomized the outcome variable without replacement for each site resulting in random pairing distributions against which empirical data was compared. Empirical pairings outside of the 5^{th} – 95^{th} percentile simulation ranking are indicative of statistically significant spatial autocorrelation.

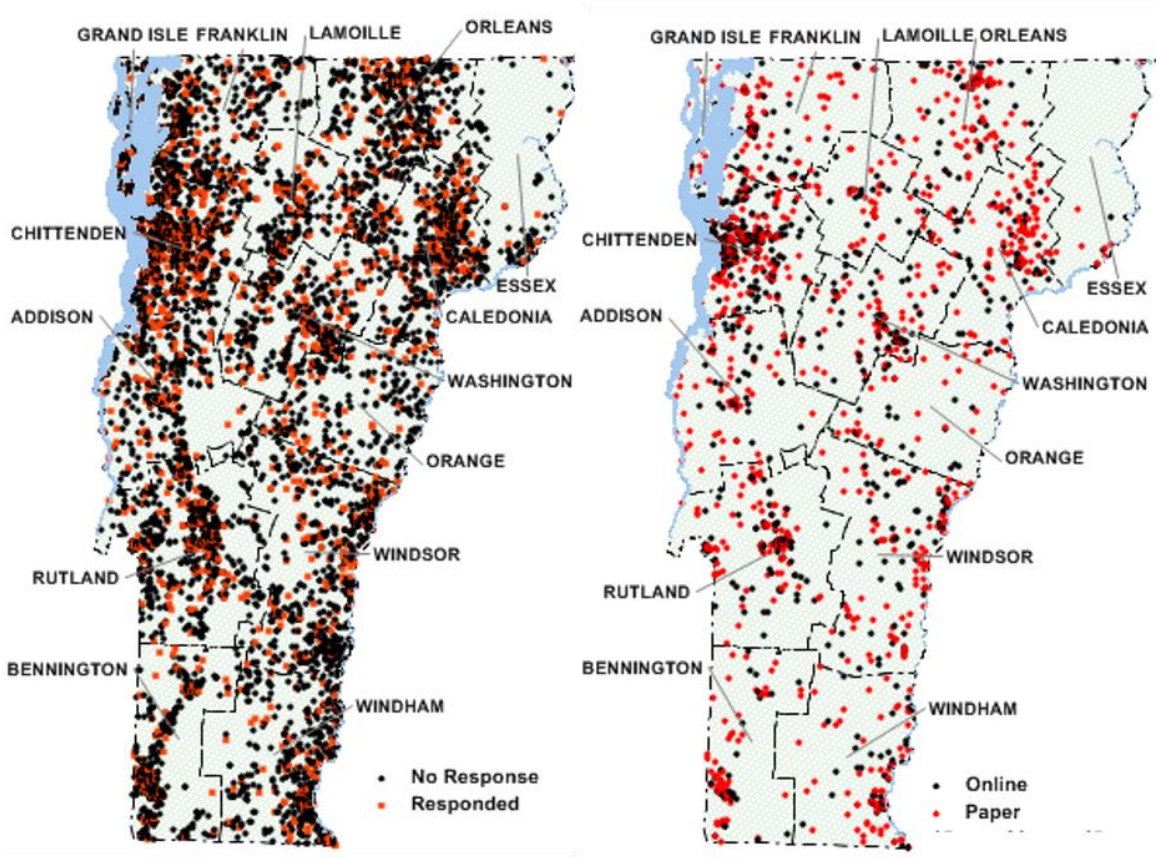

Figure 1. LRTPS recruits by response status (left) and respondents by retrieval method (right)

Significant spatial autocorrelation was indicated for online/online (more highly clustered than random) and paper/paper (less clustered than random) pairs at 1.6, 16.1, and 80.5 kilometer (1, 10 and 50 mile) distance thresholds[1]. Online/online pairing had a percentile rank of 96, 100 and 100 at the 1.6, 16.1, and 80.5 kilometer distance thresholds relative to 1,000 simulations. The corresponding paper/paper percentile ranks were 5, 0 and 0. Consistent with the correlation between retrieval method and population density, online respondents had a higher average number of pairs within a one mile radius than paper respondents (8.8 versus 7.6). Response and non-response show evidence of spatial autocorrelation at distant thresholds of 1.6 and 80.5 kilometers but not at 16.1 kilometers. These results support the data in Table 3 that rural residents are more likely to select paper surveys and also that certain neighborhoods/communities and regions were more likely to respond to the survey since responses are correlated in space.

4.2 How do paper versus online respondents differ from each other?

While potentially more accurate in capturing trips and geocodes, and less burdensome than traditional retrieval methods, online surveys also risk excluding individuals without access

[1] These distance thresholds were chosen to have a range of radii tested in the absence of any natural system association.

to or affinity for the Internet. In addition to capturing more rural residents, other differences exist among those who prefer paper surveys. To assess the possible differences between paper (57.5%) and online (42.5%) respondents, we considered the reported travel behavior, levels of customer satisfaction and issue prioritization after controlling for gender, age, income and region. Since the raw paper and online samples from LRTPS 2016 differed substantially in terms of demographic makeup, we weighted the online sample to match the paper sample in terms of gender, age, income and regional distribution using Izrael et al.'s (2004) Raking macro in SAS. The purpose of this weighting process was to facilitate comparisons between paper and online respondents while controlling for demographic variables rather than to replicate the demographics of the state population.

Table 4 shows the distribution of responses for the paper and weighted online samples for variables with statistically significant differences: household size, level of education and self-described neighborhood type. Paper surveys were more likely to be completed by smaller households, less educated individuals and individuals whose self-defined place of residence was a small village or town. Relative to the paper sample, the weighted online sample has fewer one person households and more two person households, is more highly educated, has higher rates of Internet access (especially smartphone access), is less likely to drive alone, and more likely to self-report living in a rural area than a small town/village.

Table 4. Demographic and Commute Mode Differences between Paper and Weighted Online Samples

Household Size**	Weighted Online survey	Paper survey
1 person	27.8%	34.3%
2 people	48.7%	42.2%
3 people	12.2%	12.6%
>4 people	11.4%	11.0
Highest level of education* **		
0-11 years, no diploma	2.2%	3.9%
High school graduate or GED	13.6%	25.1%
Some college, no degree	17.9%	17.6%
Associate's degree	9.8%	10.1%
Bachelor's degree	29.0%	23.1%
Graduate degree or higher	27.6%	20.3%
Neighborhood Type* **		
Urban/Suburban	34.9%	36.5%
Small village/town	26.7%	33.5%
Rural	38.4%	30.0%
Primary Commute Mode (last 12 months)**		
Drive alone	83.1%	90.5%
Carpool	3.6%	2.8%
Passenger in a private vehicle	1.9%	1.2%
Walk	3.1%	2.4%
Bicycle	1.5%	0.7%
Public transit bus	2.4%	1.4%
Specialized bus or van service	0.2%	0.0%
Ferry	0.2%	0.0%
Ride share service (e.g., Uber)	0.1%	0.0%
Vanpool	0.4%	0.2%
Other	3.6%	0.9%

* Significant at P = .1, ** Significant at P = .05, *** Significant at P = .01

Table 5 shows the differences in the customer satisfaction, issue importance and travel behavior, between paper and online respondents, weighted to account for differences in gender, age, income and region between paper and online survey respondents. The responses are significantly different for 7 of the 11 customer service questions. They are also significantly different for 10 of the 13 customer service questions. Finally, paper respondents are more likely to report both frequent trips to out-of-state destinations as well as to never or infrequently traveling outside of Vermont. Paper respondents were also more likely to report frequent unmet travel needs inside Vermont. (Table 6).

Table 5. Differences in Surveyed Factors by Retrieval Method

	Variable	Paper Sample vs. Weighted Online Sample (Chi-square except where noted)
Respondent Satisfaction with:	Highway Conditions	***
	Sidewalk Availability	--
	Bike Facility Availability	--
	Amtrak Service	***
	Park & Ride Availability	--
	Winter Maintenance	--
	Bus service convenience	***
	Specialized bus/van service	***
	Traveler info	***
	DMV Services	***
Importance Rating for VTrans Issues and Services:	Minimizing Cost to Taxpayers	*
	Supporting Job Creation	--
	Supporting Downtowns	*
	Protecting the Environment	***
	Ensuring Safety of Travelers	***
	Reducing Congestion	***
	Withstand extreme weather	--
	Roadway conditions	***
	Winter maintenance	***
	Bike/ped facilities	**
	Public Transit Services	***
	Passenger Rail Services	***
Travel Behavior	Primary commute mode	**
	Estimated weekday VMT	--Wilcoxon-Mann-Whitney
	Mode use Frequency	Mixed
	Unmet Travel Needs within Vermont	***
	Unmet Travel Needs outside of VT	***
	Frequency of trips outside Vermont	***

* Significant at P = .1, ** Significant at P = .05, *** Significant at P = .01

Table 6. LRTPS Means of Accessing the Internet

Internet Access***	Online survey	Paper survey
No Internet	0.1%	10.5%
Limited Internet	1.5%	6.6%
Home Access	30.3%	43.9%
Home and Mobile Access	64.2%	32.8%
Mobile Access	3.9%	6.2%

*** Significant at P = .01

4.3 How does Internet access vary?

Widespread access to the Internet and/or smartphones throughout a study population is necessary for web-based travel surveys to be successful. The LRTPS collected information about Vermonters' Internet access but not their smartphone ownership. Access to the Internet is considered here and smartphone ownership is covered in the next section.

Tables 6 and 7 summarize the levels of Internet access available to different populations as measured by the weighted sample which matches statewide gender, age, income and regional distributions (note these are different weights from the prior section). Respondents were characterized as having limited Internet access if they reported no access to the Internet at home or on a mobile device but could access the Internet in other ways (e.g. at work, school, or via public WIFI hotspots). Statewide, 84.9% had home-based Internet and 94.3% reported at least some form of Internet access. This compares to 79.1% of households in the 2015 ACS further suggesting bias in the LRTPS sample. According to the LRTPS approximately 52% had used a mobile device for Internet access and 5% used a mobile device primarily (i.e. did not have home access as well). Nationally, 64% of American adults owned a smartphone and 7% rely on these devices as their primary mode of Internet access (Smith and Page 2015). Internet access differs between online and paper survey respondents. Note in particular that 10.5% of paper respondents have no Internet access at all. Younger and higher income persons have better Internet access, as might be expected, but somewhat surprisingly, Internet access is better amongst rural residents than in other neighborhood types. Lack of access is most severe for lower income individuals, aged 75 years and older. These are meaningful differences to keep in mind as statewide survey programs are implemented. Even in the lowest income and highest age categories, over 60% or respondents had home or mobile-device based Internet access suggesting that it would be feasible to reach Vermonters in these groups with a web-based survey. However, one must consider if the travel patterns and needs of the group without access are the same as those with access. The attitude and issues ratings summarized in Table 6 would suggest important differences may be masked when we weight only for the standard sociodemographic variables.

Table 7. Internet Access Group (LRTPS)

Neighborhood Type	None	Limited	Home Access No Mobile	Home and Mobile Access	Mobile Access No Home
Urban/Suburban	5.7%	4.5%	32.8%	53.1%	4.0%
Small Town/Village	5.8%	4.5%	41.2%	41.8%	6.7%
Rural	4.8%	3.5%	41.1%	45.9%	4.8%
Household Income					
Less than $15,000	16.5%	14.1%	42.1%	13.3%	14.1%
$15,000 to $24,999	12.9%	9.3%	40.2%	30.5%	7.2%
$25,000 to $34,999	5.0%	3.1%	42.5%	41.6%	7.8%
$35,000 to $49,999	3.6%	3.4%	38.3%	50.9%	3.8%
$50,000 to $74,999	2.0%	1.2%	36.7%	54.0%	6.1%
$75,000 to $99,999	0.4%	2.0%	38.0%	58.1%	1.5%
$100,000 to $149,999	0.2%	0.0%	27.5%	70.1%	2.2%
$150,000 to $199,999	0.0%	0.0%	26.1%	73.9%	0.0%
$200,000 or more	0.0%	0.0%	18.8%	80.5%	0.7%
Age Category					
18-24 years	0.0%	9.6%	21.6%	50.3%	18.6%
25-34 years	0.0%	2.1%	21.4%	68.5%	8.0%
35-44 years	1.3%	1.4%	23.1%	69.3%	4.9%
45-54 years	2.5%	3.2%	38.5%	50.5%	5.4%
55-64 years	6.7%	3.8%	46.3%	40.3%	2.9%
65-74 years	9.6%	5.7%	59.2%	24.6%	0.9%
75 years or older	25.4%	9.3%	58.0%	6.1%	1.1%

4.4 How does smartphone access vary?

Of the respondents reached by landline or cell phone in the Vermonter Poll, 79% of individuals reported that a least one adult in their household owned a data-enabled cell phone. Table 8 indicates household cellphone penetration levels by county. Household smartphone penetration is considered to be 'full' if there are at least as many data-enabled cell phones as adults and partial if there are fewer cell phones than adults. Cell penetration is highest in the more urban northwest region of the state. However, none of the smartphone penetration measures are correlated with population density. Full smartphone penetration is relatively high, ranging from 40-70% of households but a significant number of households still lack even a single smartphone. Tables 9 also confirms that smartphone ownership is highly correlated with income with nearly 60% of households with income below $25,000 having no data-enabled cellphones. In addition, Table 9 shows the breakdown of data enabled cellphone ownership by education and age. Because these variables were only collected for the individual respondent and cellphone ownership was collected at the household level, these variables only include respondents with full cellphone penetration or with no cellphone penetration (n = 496). Smartphones are more prevalent for younger and more educated individuals. A smartphone-based survey might systematically limit data from older and/or less educated Vermonters. The barriers to access to a smartphone-based survey

are considered to be more significant than access to an Internet-based survey based on the results here. It should be noted that some other States have experimented with loaner smartphone travel data collection, but the cost may be prohibitive.

5. Conclusions

In their international review of survey methods, Inbakaran and Kroen (2011) stated that 'there is not the one perfect method and the advantages and disadvantages of the different methods with regard to representativeness, response rates, data accuracy and costs have to be weighed against each other.' Since that time, forces beyond the control of a transportation agency have continued to alter the transportation survey landscape. Given that surveys continue to be an essential source of the data that are needed for system planning and operation, effective methods to collect survey data must be found. These methods must balance advantages and disadvantages but, most relevant here, they must balance cost with coverage. While many travel surveys are undertaken in predominantly metropolitan areas, state-wide surveys must address concerns about adequate coverage of rural residents. This analysis of Vermont data suggests that rural-based limitations in survey coverage do exist as well as the perennial issues of income and age biases.

At this time, our analysis suggests that Smartphone access in Vermont is too restrictive to be a primary survey method. This may or may not be true in other rural regions but this finding suggests pause before moving to quickly to smartphone-only based surveys. Internet access is more widespread in Vermont but does exclude some population segments. Smartphone and Internet access are not strongly related to urban versus rural home locations.

Within the raw sample, the choice of the paper survey over the online survey was linked to lower population density and online respondents showed significant spatial clustering. More importantly, after weighting for basic demographic variables, the travel behavior and transportation attitudes of paper versus online survey participants differed. The demographic weighting that is typically undertaken for transportation surveys may be insufficient to address the full range of coverage issues, especially with low response rates and small sample sizes. Weighting for demographic variables is unlikely to create truly representative samples in terms of transportation needs and priorities. Better understanding of what drives these differences between people is needed. It is clear from this analysis that when we move surveys online we are missing some portion of the population with lower incomes, less access to the Internet, and those living in more rural locations. These individuals are more likely to drive and live alone. It may be possible to realistically conduct travel surveys that are more inclusive and weighted by location especially with address-based recruiting (Bradley et al. 2015), but we should take care to understand that weighting based on sociodemographic variables alone is likely still leaving us with an understatement of travel needs.

References

Agrawal, A., S. Granger-Bevan, G. Newmark, and H. Nixon. 2017. "Comparing Data Quality and Cost from Three Modes of on-Board Transit Surveys." *Transport Policy*, Vol. 54, pp. 70–79.

Ampt, E., P. Bonsall, and PTRC Education and Research Services Limited. 1997. *Current Issues in Travel and Transport Demand Surveys*. PTRC European Transport Forum, Brunel University, England.

Antoun, Christopher, Mick P. Couper, and Frederick G. Conrad. 2017. "Effects of Mobile versus PC Web on Survey Response Quality A Crossover Experiment in a Probability Web Panel." *Public Opinion Quarterly* 81 (S1): 280–306. https://doi.org/10.1093/poq/nfw088.

Axhausen, K. W., M. Löchl, R. Schlich, T. Buhl, and P. Widmer. 2007. "Fatigue in Long-Duration Travel Diaries." *Transportation* 34 (2): 143–60. https://doi.org/10.1007/s11116-006-9106-4.

Bradley, M., A. Bergman, M. Lee, E. Greene, and S. Childress. 2015. "Predicting and Applying Differential Response Rates in Address-Based Sampling for a Household Travel Survey." *Transportation Research Record*, No. 2526 pp. 119–125.

Bricka, Stacey, and Chandra Bhat. 2006. "Comparative Analysis of Global Positioning System-Based and Travel Survey-Based Data." *Transportation Research Record: Journal of the Transportation Research Board* 1972 (January): 9–20. https://doi.org/10.3141/1972-04.

Brög, Werner, Erhard Erl, Arnim H. Meyburg, and Manfred J. Weymuth. 1982. "Problems Of Nonreported Trips In Surveys Of Nonhome Activity Patterns." *In Transportation Research Record*. https://trid.trb.org/view.aspx?id=189338.

Cummins, B., G. Spitz, T. O'Malley, and M. Campbell. 2013. "How Close Is Close Enough? Statistical Equivalence of Onboard versus Online Surveys of Transit Customers." *Transportation Research Record,* No. 2351 pp. 23–29.

Dijst, M., S. Farag, and A. de Blaeij. 2006. *Effects of Data Collection Methods in Travel Behavior Surveys: Comparing Internet and Mail Sample*. Transportation Research Board Annual Meeting.

Flake, L., M. Lee, K. Hathaway, and E. Greene. 2017. *Smartphone Panels as Viable and Cost-Effective GPS Data Collection for Small and Medium Planning Agencies*. #17-01359.

Hartgen, D., and E. San Jose. 2009. *Costs and Trip Rates of Recent Household Travel Surveys*. Retrieved from http://www. hartgengroup. net/Projects/National/USA/household_travel_summary/2009-11-11_Final_Report_Revised. pdf.

Hassounah, Mazen I., Loy-Sai Cheah, and Gerald N. Steuart. 1993. "Underreporting Of Trips In Telephone Interview Travel Surveys." *Transportation Research Record*, no. 1412. https://trid.trb.org/view.aspx?id=385056.

Izrael, D., D. C. Hoaglin, and M. P. Battaglia. 2004. *To Rake or Not to Rake Is Not the Question Anymore with the Enhanced Raking Macro*. 29th Annual SAS Users Group International Conference, Montreal.

Inbakaran, C., and A. Kroen. 2011. *Travel Surveys–Review of International Survey Methods*. Australasian Transport Research Forum.

Janzen, Maxim, Maarten Vanhoof, Zbigniew Smoreda, and Kay W. Axhausen. 2018. "Closer to the Total? Long-Distance Travel of French Mobile Phone Users." *Travel Behavior and Society* 11 (April): 31–42. https://doi.org/10.1016/j.tbs.2017.12.001.

Kam, B., and J. Morris. 1999. *Response Patterns in Travel Surveys: The VATS Experience*.

Kunzman, M., and V. Daigler. 2013. *2010 - 2012 California Household Travel Survey Final Report*.

Lee, C., and J. Pino. 2012. "Hang Up the Phone and Get Online: Measuring the Effectiveness of Web-Based Surveys." *Transportation*.

Lee, Richard J., Ipek N. Sener, and James A. Mullins. 2016. "An Evaluation of Emerging Data Collection Technologies for Travel Demand Modeling: From Research to Practice." *Transportation Letters* 8 (4): 181–93. https://doi.org/10.1080/19427867.2015.1106787.

Manfreda, Katja Lozar, Jernej Berzelak, Vasja Vehovar, Michael Bosnjak, and Iris Haas. 2008. "Web Surveys versus Other Survey Modes: A Meta-Analysis Comparing Response Rates." *International Journal of Market Research* 50 (1): 79–104.

Messer, Benjamin L., and Don A. Dillman. 2011. "Surveying the General Public over the Internet Using Address-Based Sampling and Mail Contact Procedures." *Public Opinion Quarterly* 75 (3): 429–57. https://doi.org/10.1093/poq/nfr021.

Millar, Morgan M., and Don A. Dillman. 2011. "Improving Response to Web and Mixed-Mode Surveys." *Public Opinion Quarterly* 75 (2): 249–69. https://doi.org/10.1093/poq/nfr003.

Morris, W., and D. Doak. 2002. *Quantitative Conservation Biology*. Sinauer Associates, Inc., Sunderland, Massachusetts.

Murakami, E., and D. Wagner. 1999. "Can Using Global Positioning System (GPS) Improve Trip Reporting?" *Transportation Research Part C: Emerging Technologies*, Vol. 7, No. 2 pp. 149–165.

NYMTC, and NJTPA. 2011. *2010/2011 Regional Household Travel Survey Final Report*.

Ritter, Chloe, and Elizabeth Greene. 2017. *Have Smartphone, Will Travel: Long Distance Travel Surveys with Smartphone GPS*. Transportation Research Board. https://trid.trb.org/view/1437465.

Pritchard, D., and E. Miller. 2012. "Advances in Population Synthesis: Fitting Many Attributes per Agent and Fitting to Household and Person Margins Simultaneously." *Transportation*, Vol. 39, No. 3 pp. 685–704.

PTV NuStats. 2011. *Regional Travel Survey Final Report*.

Safi, H., B. Assemi, M. Mesbah, L. Ferreira, and M. Hickman. 2015. "Design and Implementation of a Smartphone-Based Travel Survey." *Transportation Research Record*, No. 2526, pp. 99–107.

Sharp, J., and E. Murakami. 2005. "Travel Surveys: Methodological and Technology-Related Considerations." *Journal of Transportation and Statistics*, Vol. 8, No. 3, pp. 97–113.

Smith and Page. *U.S. Smartphone Use in 2015*. Pew Research Center, 2015.

Smith, C., and G. Spitz. 2010. "Internet Access: Is Everyone Online yet and Can We Survey Them There?" *Transportation Research Record*, No. 2176, pp. 35–41.

Stopher, P. R., and S. P. Greaves. 2007. "Household Travel Surveys: Where Are We Going?" *Transportation Research Part A*, Vol. 41, No. 5, pp. 367–381.

Stopher, P., R. Alsnih, C. Wilmot, C. Stecher, J. Pratt, J. P. Zmud, W. Mix, M. Freedman, K. W. Axhausen, M. Lee-Gosselin, A. E. Pisarski, and W. Brog. 2008. *Standardized Procedures for Personal Travel Surveys*. *NCHRP Report*, No. 571.

Transportation Research Board. 2016. *Exploring New Directions for the National Household Travel Survey. Phase Two, Report of Activities. National Academies*.

Weis, C., A. Frei, K. Axhausen, B. Fell, and T. Haupt. 2008. A Comparative Study of Web- and Paper-Based Travel Behaviour Surveys. *Working paper/Institute of Transportation, Traffic, Highway-and Railway-Engineering (IVT)*, Vol. 519, pp. 1–19.

Westat, Inc. 2015. *National Household Travel Survey Task C: Sample Design, Final Report*. FHWA.

Wolf, Jean, Marcelo Oliveira, and Miriam Thompson. 2003. "Impact of Underreporting on Mileage and Travel Time Estimates: Results from Global Positioning System-Enhanced Household Travel Survey." *Transportation Research Record: Journal of the Transportation Research Board*, no. 1854: 189–198.

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*, No. 1660, pp. 66–74.

Xing, Y., and S. Handy. 2014. "Online versus Phone Surveys: Comparison of Results for a Bicycling Survey." *Transportation Planning and Technology*, Vol. 37, No. 6, pp. 554–567.