G. Global Positioning System Travel Survey

2006 PSRC GPS-Assisted Household Activity and Stated Preference Survey

Technical Memorandum
GPS Travel Survey
12/21/06

Introduction

ECONorthwest (ECO) was subcontracted by MORPACE International Inc. (MORPACE) to conduct the GPS component of the Puget Sound Regional Council's 2006 GPS-Assisted Household Activity and Stated Preference Survey. The 12-week survey period began on 3/21/2006 and ended on 6/16/2006. During that time, ECO was responsible for programming the GPS devices and post-retrieval software to provide the GPS time stamping and travel measurement capabilities required of the study. Specifically, it was ECO's responsibility to:

- Collect GPS waypoints relevant to the survey time frame
- Identify trip starts and ends using algorithms that distinguish between periods of motion and periods of no recoded motion
- Match trips logged by the GPS devices with travel reported in the travel diaries using algorithmic and semi-manual methods.

This memorandum presents the methodology and findings from the GPS component of the household travel survey.

Methodology

ECO deployed a total of 150 GPS devices to 285 households with 518 different vehicles over the course of the 12-week survey. GPS data were collected, downloaded, and processed to compile a trip-level database. That database was used to match GPS trips to those reported by in the travel diaries. ECO provided a single Excel workbook that included the original trip-level GPS data, the original trip-level diary data, and several tables of matched trips.

GPS Data Logger

ECO identified a one-piece GPS data logger that can be easily installed by participants. The device can be powered by batteries or cigarette lighter. ECO elected not to include the cigarette lighter adapter for this study to reduce the risk of data loss resulting from mishandling or vehicle electrical problems. The hardware, GPS antenna, and batteries are encased in a single plastic container. The container used for this study measures approximately 4 x 2 x 1 inches in size. The device attaches to the vehicle's dashboard with special double-sided tape. The tape is easy to remove and will not damage surfaces it is affixes to.

Current GPS technology works by using radio frequency to communicate with satellites directly overhead. No GPS device will function without establishing an unobstructed line-of-sight to those satellites. Because this survey was conducted with dashboard-mounted devices, it was necessary to screen out vehicles with metallic or vertical windshields.

The GPS device is programmed with firmware that logs successive waypoints at a specified time interval. That interval was set to 10 seconds so the individual data points could be accurately translated to sets of continuous trips. Positions are only recorded when the vehicle is moving to manage memory and power constraints. Each position includes: date and time, speed, direction, latitude and longitude coordinates, and altitude.

Logistics

ECO and MORPACE conducted a pilot study to confirm data quality, device reliability, and the appropriate timeline for shipment and retrieval of the GPS devices. The findings from that study were combined with assumptions about device and data loss due to theft, mishandling, or malfunction to establish the total number of GPS households needed to successfully retrieve 230 with complete GPS data. It was assumed that five percent of households would never return the devices and 20 percent of households would return one or more devices with missing or corrupt data. Based on those loss rates, it was determined that a 300 GPS household deployment would be adequate to meet the survey's GPS quota.

The total number of GPS household deployments and survey timeline were analyzed to determine the appropriate deployment schedule and necessary inventory of GPS devices. The survey was set to take place over a 12-week period with four possible combinations of two-day travel periods or "waves" per week (Monday-Tuesday, Tuesday-Wednesday, Wednesday-Thursday, and Thursday-Friday). Four waves per week over the course of 12 weeks translates to a total of 48 waves. Therefore, slightly more than six households per wave are needed to deploy 300 households. Assuming that 25 percent of the devices are returned in one week, 50 percent are returned in two weeks, 20 percent are returned in three weeks, and five percent are never returned, the appropriate inventory is 100. Fifty additional devices were programmed and placed on standby in case return rates were slower than expected.

Packaging and Shipping

Each participant household was sent one package that included GPS devices, travel diaries, postage-paid return packaging, a thank you letter, and a single sheet of simple instructions. Those instructions were designed to be understandable for even the most technologically challenged users. The instructions explained device installation (affix tape to the dashboard), activation (pull the plastic tab), and removal (pull on the tape).

During the study, MORPACE periodically furnished electronic lists of new GPS participants. The devices were programmed, packaged, and shipped to those participants two days before their survey travel dates. A participant would generally receive the package one day prior to the first day of recorded travel. That provided enough time to read the instructions, install the device, and activate the power supply before beginning recorded travel. The participants were instructed to return materials after the survey by either dropping the package off to a FedEx location or contacting MORPACE so that a courier pickup could be arranged. The packages were sent to participants using FedEx's "Priority Overnight" shipping service and they were returned using FedEx's "Second-day Air."

MORPACE was forwarded a spreadsheet on a daily basis that tracked the survey's progress and the disposition of each unit and diary. That "GPS Status Update" contained a summary of all survey activities and device/vehicle-level information. MORPACE used that information to adjust the deployment schedule and follow up with delinquent households or households that failed to return a portion of the survey materials. The devices and travel diaries were returned to ECO in one return package and then ECO forwarded the diaries to MORPACE.

Data Processing

Each package was returned directly to ECO's Portland office. Upon receipt, packages were immediately opened and inspected to ensure that all GPS devices and travel diaries were accounted for. The GPS data were immediately downloaded and inspected to confirm completeness (i.e. data were not corrupt and movement was logged during the household's travel days). However, it was not uncommon for a household to retain the diaries. The instructions offered participants the option to verbally transmit all diary contents to a MORPACE representative via telephone instead of sending back their diaries. When travel diaries were returned with the GPS devices, ECO also visually confirmed that travel reported in the diaries seemed reasonably consistent with movement logged in the GPS devices. Once the data were successfully downloaded, the device's memory was cleared and the unit was programmed for redeployment.

Results

Completed Households

There are two definition of a completed household. The "all or nothing" definition stipulates that a household is completed only if data were successfully retrieved from the devices in every vehicle from that household. The "partial complete" definition applies to multi-vehicle households where data were retrieved from some, but not all of the household vehicles. A partial GPS account of a household's travel is still useful in trip-level travel modeling. This definition credits a percentage that's equal to the share of devices with complete data in a given household (i.e. if data were retrieved from one out of two devices in a household, that household would be considered 50 percent complete).

ECO deployed a total of 285 GPS households with the goal of retrieving complete data from 230 households, or approximately 400 vehicles (assuming an average of 1.75 vehicles per household). Travel data were ultimately retrieved from 415 vehicles. Under the "partial complete" definition of a retrieved household, the survey returned exactly 230 households. Under the "all or nothing" definition, the survey returned 218 households.

Obstacles

There are numerous potential failure-points in the process of deploying GPS devices for travel survey purposes. The devices, and ultimately the data, are in peril at virtually every step of the logging process. To account for those hazards, ECO assumed the aforementioned non-retrieval rate of 25 percent. The actual non-retrieval rate turned out to be very close to that assumption. Again, it was anticipated that five percent of households would never return the devices and 20 percent would return one or more units with missing or corrupt data. In the end, 5.3 percent of households never returned the devices and 18.6 percent of households returned one or more devices with bad data. The majority of those non-retrievals resulted from incorrect installation or non-participation.

It was initially assumed that 230 complete households could be confidently obtained by deploying a total of 300 households. With actual loss rates comfortably below expectation, the only problem came with the fact that many devices were staying in the field too long. Even after deploying all 50 of the backup GPS devices, the number of households in several waves had to be reduced because there were not enough devices on hand. Table 1 summarizes the expected and actual return rates of survey materials.

Table 1: Number of Weeks it Took Participant Households to Return Survey Materials

Duration	Expected % of Households	Actual % of Households
<1 Week	25%	9%
1-2 Weeks	50%	5%
2-3 Weeks	20%	70%
>3 Weeks	0%	11%
Never Returned	5%	5%
Total	**100%**	**100%**

Vehicles

A total of 518 GPS devices were shipped to the survey participants during this study and 415 were successfully retrieved. Of those shipped, the average number of vehicles per household was 1.82. The average number of vehicles per household in households where complete data were retrieved was 1.78. Nineteen percent of the partially complete households had two vehicles and three percent had three vehicles. Table 2 summarizes the distribution of this sample according to the number of vehicles in each household:

Table 2: Distribution of Sample by Number of Household Vehicles

Household Type	Total Shipped	Shipped % of Total	Total Retrieved	Retrieved % of Total
One Car	98	34%	81	37%
Two Car	141	49%	102	47%
Three Car	46	16%	35	16%
Total	285	100%	218	100%

Trips

Trips are compiled by collapsing the waypoint data around periods of non-movement. Two parameters are used to convert waypoint data into trips. The first parameter, which registers a trip's beginning by defining the distance that must be traveled before a waypoint is logged, was set to 110 feet. That parameter is known as "GPS wobble" and is necessary because disruptions in a satellite's signal (usually due to weather) can be incorrectly interpreted as movement. The second parameter, which registers a trip's end by defining the amount of time a device must be motionless before a trip is broken, was set to two minutes.

It is difficult, if not impossible, to draw inferences from trip-level GPS data. Travel behavior is very complicated and it does not necessarily correspond to the movement of an individual's car. In many cases, individuals did not even use the family vehicles during their survey period. Table 3 summarizes some descriptive statistics pertaining to the volume of trip-level data from the GPS devices. These data, however, do not begin telling a story until they are aligned with information from the travel diaries.

Table 3: Trip-Level Descriptive Statistics

	Households	Vehicles
Total	225	343
Total Trips	2,942	2,942
Average Trips Per	13.1	8.6
Maximum Trips Per	35	27
Minimum Trips Per	1	1

Trip Matching

ECO's final task was to match trips from the GPS data with the corresponding diary data. The first step in the matching process was algorithmic and the second step was semi-manual. Some households and vehicles could not be matched because of incompleteness in the diary data. Other households never traveled in their household vehicles, so there were no to match. When the GPS and diary data were aligned, 183 households were present in both datasets. Out of the 2,942 total GPS trips, 1,748 were matched algorithmically and 104 were matched semi-manually.

Matching the GPS and diary trips begins with a computer algorithm developed by ECO that analyzes the timing of each trip. The program's primary focus is on finding trips with similar duration among trips that start on or around the same time. The program loops through both datasets, each time permitting an increasingly lax difference between trip start-time and duration. It also matches across vehicles to find situations where the participant(s) misreported which household vehicle was used for a trip. The program is capable of complex matching such as chaining GPS or diary trips to find one-to-many and many-to-one relationships in the data. It does not, however, consider origin or destination locations, mostly because those are often missing from diary data.

Trips that could not be matched algorithmically were separated and, to the extent possible, matched semi-manually. The semi-manual matching is focused on finding trips in the diary data that do correspond to GPS trips, but were severely misreported and not picked up by the program. There are a number of ways trips can be misreported in a way that allows them to be easily matched through visual inspection. The most common is to either report a trip on the wrong day or to report a long trip but make one or more pronounced stops over the course of that trip. The focus of this semi-automated matching was not to force matches on data that are fundamentally different, but to find obvious reporting mistakes. The remaining trips from both datasets were separated and classified as unmatchable.

Conclusion

This GPS travel survey was conducted through a process of generating GPS travel data, arranging that data, and matching it with diary trips. That process was executed as planned and the expected outcome was strikingly similar to the actual survey results.

Lessons Learned

A small number of devices from the first waves suffered damage while in transit. Despite the use of heavily cushioned packages, something occurred during FedEx's shipping process that subjected packages to one or more severe physical shocks. It is impossible to determine exactly what happened to those packages, but shipping preparations in future efforts should account for this. Specifically, medium to large-sized FedEx boxes lined with layers of bubble wrap proved sufficient to keep the problem from persisting.

The process of prequalifying GPS survey candidates is very important because deploying this technology is quite costly. GPS survey participants must cooperate with the survey process so that devices are returned promptly with useful data. The opportunity cost of the time devices are in the field is consequential, especially if the household decides not to participate and no data are recovered. To any extent possible, future emphasis should be placed on stringent screening of these candidates to promote better results and mitigate unnecessary costs.

The only significant obstacle encountered over the course of this survey was the length of time that it took households to return the devices. On average, 86 percent of devices were still in the field two weeks after being deployed. ECO and MORPACE responded by increasing the frequency of post-survey follow-up calls. Those calls proved quite effective and return rates did ultimately improve. Future efforts should adhere to a strict schedule of following up with participant households from the very beginning.