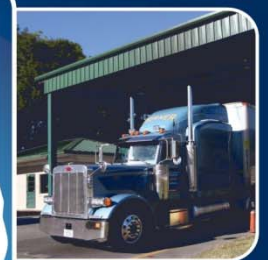


Assessment of GPS Travel Survey Data

presented by
Thomas Rossi, Cambridge Systematics

May 7, 2015



Everybody's Doing It...



- Household survey technology has evolved over time...
 - Web based data retrieval
 - Use of GPS for either a subset or for all households
- Most current/recent surveys include some GPS based component
- (But you already knew that, and I'm not here to talk about the advantages of GPS in surveys)

Common Knowledge?



- GPS results show 15%-20% more trips than daily surveys
- Short trips/activities are often unreported in diary surveys
- GPS data do not provide information about trip purpose or mode
 - But these aspects of travel can be imputed to some extent

It's Still Evolving



- Battery life is sometimes an issue (especially with older devices)
- Signals can still get lost (though this has improved)
- People sometimes forget to bring the devices

What lessons have been learned?



- What needs to be done before embarking on a GPS survey?
- What needs to be done to process the raw data to make it usable?
- What are the limitations of the imputation process?

Before the Survey...



- Know your data needs!
- Keep the survey streamlined
- Consider a **prompted recall** approach
- Prepare respondent training materials
- Get location information during recruitment
- Recognize limitations...
 - Of data collection itself (e.g., young children not tracked on their own)
 - Of the data imputation process

Respondent Training Materials



- Make sure that:
 - All surveyed respondents have devices
 - Devices are brought every time the respondent leaves the home
 - Devices are brought when making subtours (e.g., going to lunch from work)
 - Respondents know how to keep batteries charged

How Accurate Are the GPS Results?



- Ideally, one would have both diaries and GPS data for the same respondents
 - But this is usually impractical for an entire survey
- Diaries specifically define activity locations, GPS imputes them
- Diaries require users to accurately recall/report everything
- GPS requires users to carry the devices everywhere
- GPS collects time information passively
 - Diaries require user recall, which can be inaccurate and usually involve rounding

Cincinnati and Twin Cities Surveys



- Cincinnati (Ohio-Kentucky-Indiana Regional Council of Governments (OKI))
 - **GPS only** (no diary)
 - Up to 3 days of travel
 - 2,591 households, 2,051 complete
- Minneapolis-St. Paul (Met Council)
 - Diaries for 14,000 households (1 day)
 - 262 households with GPS data (up to 7 days)
 - 125 households, 187 persons with matched GPS and diary data

Twin Cities Findings



- A “wrong” day is sometimes reported in the travel diary
- There are GPS trips that are missing in the diary survey
- There are trips in the diary survey that are missing in the GPS dataset
- There are incomplete days of travel in the GPS dataset or an excessive number of days in which the GPS units were carried around.

Twin Cities Findings



- After initial processing, about 8% fewer activities reported in diary
 - Similar % difference for trips
 - “Missing trips” nearly all < 5 minutes
- Activity duration frequency distributions similar
- Processed data showed zero-duration activities (5%)

Trip Length Frequency Comparison



Trip Length (min)	GPS	Diary
1-4	183	92
5-10	334	340
11-20	284	276
21-30	144	139
31-60	71	85
>60	16	15
Total	1,032	947

Comparing GPS Data to Diaries #1



- Diary

Trip Number	Origin Location	Destination Location	Start Time	End Time	Mode	Access-Egress
1	Home	Drop-Off	7:30 AM	7:40 AM	Auto	N/A
2	Drop-Off	Work	7:45 AM	8:30 AM	Bus	Drive-Walk
3	Work	Eat	3:00 PM	3:15 PM	Walk	N/A
4	Eat	Work	3:45 PM	3:53 PM	Walk	N/A
5	Work	Pick-Up	5:02 PM	5:33 PM	Transit	Walk-Drive
6	Pick-Up	Home	5:46 PM	5:53 PM	Auto	N/A

- GPS

Trip Number	Origin Match Location	Destination Match Location	Start Time	End Time	Origin Matched with Trip Diary	Destination Matched with Trip Diary
1	Home	Drop-Off	7:32:26 AM	7:39:33 AM	Yes	Yes
2	Drop-Off	Work	7:44:41 AM	8:12:25 AM	Yes	Yes
3	Work	Unknown	5:11:37 PM	5:14:14 PM	Yes	No
4	Unknown	Pick-Up	5:14:14 PM	5:34:02 PM	No	Yes
5	Pick-Up	Home	5:40:30 PM	5:50:22 PM	Yes	Yes



Comparing GPS Data to Diaries #1



- Diary

Trip Number	Origin Location	Destination Location	Start Time	End Time	Mode	Access-Egress
1	Home	Drop-Off	7:30 AM	7:40 AM	Auto	N/A
2	Drop-Off	Work	7:45 AM	8:30 AM	Bus	Drive-Walk
3	Work	Eat	3:00 PM	3:15 PM	Walk	N/A
4	Eat	Work	3:45 PM	3:53 PM	Walk	N/A
5	Work	Pick-Up	5:02 PM	5:33 PM	Transit	Walk-Drive
6	Pick-Up	Home	5:46 PM	5:53 PM	Auto	N/A

- GPS

Trip Number	Origin Match Location	Destination Match Location	Start Time	End Time	Origin Matched with Trip Diary	Destination Matched with Trip Diary
1	Home	Drop-Off	7:32:26 AM	7:39:33 AM	Yes	Yes
2	Drop-Off	Work	7:44:41 AM	8:12:25 AM	Yes	Yes
3	Work	Unknown	5:11:37 PM	5:14:14 PM	Yes	No
4	Unknown	Pick-Up	5:14:14 PM	5:34:02 PM	No	Yes
5	Pick-Up	Home	5:40:30 PM	5:50:22 PM	Yes	Yes



Comparing GPS Data to Diaries #1



- Diary

Trip Number	Origin Location	Destination Location	Start Time	End Time	Mode	Access-Egress
1	Home	Drop-Off	7:30 AM	7:40 AM	Auto	N/A
2	Drop-Off	Work	7:45 AM	8:30 AM	Bus	Drive-Walk
3	Work	Eat	3:00 PM	3:15 PM	Walk	N/A
4	Eat	Work	3:45 PM	3:53 PM	Walk	N/A
5	Work	Pick-Up	5:02 PM	5:33 PM	Transit	Walk-Drive
6	Pick-Up	Home	5:46 PM	5:53 PM	Auto	N/A

- GPS

Trip Number	Origin Match Location	Destination Match Location	Start Time	End Time	Origin Matched with Trip Diary	Destination Matched with Trip Diary
1	Home	Drop-Off	7:32:26 AM	7:39:33 AM	Yes	Yes
2	Drop-Off	Work	7:44:41 AM	8:12:25 AM	Yes	Yes
3	Work	Unknown	5:11:37 PM	5:14:14 PM	Yes	No
4	Unknown	Pick-Up	5:14:14 PM	5:34:02 PM	No	Yes
5	Pick-Up	Home	5:40:30 PM	5:50:22 PM	Yes	Yes



Comparing GPS Data to Diaries #1



- Diary

Trip Number	Origin Location	Destination Location	Start Time	End Time	Mode	Access-Egress
1	Home	Drop-Off	7:30 AM	7:40 AM	Auto	N/A
2	Drop-Off	Work	7:45 AM	8:30 AM	Bus	Drive-Walk
3	Work	Eat	3:00 PM	3:15 PM	Walk	N/A
4	Eat	Work	3:45 PM	3:53 PM	Walk	N/A
5	Work	Pick-Up	5:02 PM	5:33 PM	Transit	Walk-Drive
6	Pick-Up	Home	5:46 PM	5:53 PM	Auto	N/A

- GPS

Trip Number	Origin Match Location	Destination Match Location	Start Time	End Time	Origin Matched with Trip Diary	Destination Matched with Trip Diary
1	Home	Drop-Off	7:32:26 AM	7:39:33 AM	Yes	Yes
2	Drop-Off	Work	7:44:41 AM	8:12:25 AM	Yes	Yes
3	Work	Unknown	5:11:37 PM	5:14:14 PM	Yes	No
4	Unknown	Pick-Up	5:14:14 PM	5:34:02 PM	No	Yes
5	Pick-Up	Home	5:40:30 PM	5:50:22 PM	Yes	Yes



Comparing GPS Data to Diaries #2



- Diary

Trip Number	Origin Location	Destination Location	Start Time	End Time	Mode	Access-Egress
1	Home	Work	7:10 AM	7:50 AM	Walk	N/A
2	Work	Home	5:35 PM	6:38 PM	Walk	N/A

- GPS

Trip Number	Origin Match Location	Destination Match Location	Start Time	End Time	Origin Matched with Trip Diary	Destination Matched with Trip Diary
1	Home	Work	6:57:00 AM	7:49:04 AM	Yes	Yes
2	Work	Work-Vicinity	9:19:21 AM	9:22:55 AM	Yes	Yes
3	Work-Vicinity	Work-Vicinity	9:22:55 AM	9:25:25 AM	Yes	Yes
4	Work	Home	5:34:23 PM	6:36:33 PM	Yes	Yes
5	Home-Vicinity	Home-Vicinity	6:58:46 PM	7:12:27 PM	Yes	Yes

Comparing GPS Data to Diaries #2



- Diary

Trip Number	Origin Location	Destination Location	Start Time	End Time	Mode	Access-Egress
1	Home	Work	7:10 AM	7:50 AM	Walk	N/A
2	Work	Home	5:35 PM	6:38 PM	Walk	N/A

- GPS

Trip Number	Origin Match Location	Destination Match Location	Start Time	End Time	Origin Matched with Trip Diary	Destination Matched with Trip Diary
1	Home	Work	6:57:00 AM	7:49:04 AM	Yes	Yes
2	Work	Work-Vicinity	9:19:21 AM	9:22:55 AM	Yes	Yes
3	Work-Vicinity	Work-Vicinity	9:22:55 AM	9:25:25 AM	Yes	Yes
4	Work	Home	5:34:23 PM	6:36:33 PM	Yes	Yes
5	Home-Vicinity	Home-Vicinity	6:58:46 PM	7:12:27 PM	Yes	Yes

What Is Happening Here?



- GPS processing apparently adds stops/activities
 - Some may be “within site”
 - Some apparently are transit stops
- GPS travel times appear to be more accurate (no rounding/estimating/recall)
 - Note: Usually we use times from skims in model estimation/validation
- Apparent “missing trips” in GPS (forgot device?)

How to Handle These Issues?



- We did not adjust diary results to reflect missing trips in either GPS or diary
 - Substantial disaggregate analysis would be required to determine the relative effects
- Assume zero-duration activities are invalid
- 8% more activities total, 5% are zero-duration
→ **3% more activities in GPS**
- These findings are not generalizable to other surveys but provide insights into what to look for

Acknowledgments



- Anurag Komanduri, Cambridge Systematics
- Jonathan Ehrlich, Metropolitan Council
- Andrew Rohne, Ohio-Kentucky-Indiana Regional Council of Governments
- Kimon Proussaloglou, Cambridge Systematics

