Smartphones: Silver Bullet for Mobility Data Collection?

KW Axhausen

IVT
ETH
Zürich

January 2016
Survey challenges
What do we want to know?

Who travels

when?

where?

with whom?

how?

for how long (space and time)?

for what purpose?

and spends how much?
Protocols and response
Surveys, observations are „talk“

Two speakers

managing their „image“
staying within the rules of talking
staying within their socially allocated/identified role
fulfilling social expectations

talk and report with/to each other

=>

„Maintaing the willingness of the respondent to report“
Response as a function of response burden @IVT, 2015

TRB 2016
Known „error“ generating processes
Activities, movement and traces: A full example record

Home
Out of home
Movement
Phone/SMS
Email
Smart card
GPS
Bluetooth
WiFi

TRB 2016
Active/passive tracing: Many owners, locations, quality levels

- Phone/SMS
- Email
- Smart card
- GPS
- Bluetooth
- WiFi
- CCTV

TRB 2016
What is left?

**True**

- 5 at home
- 9 Out of home
- 26 Stages,
  - 11 trips,
  - 1 subtour,
  - 2 tours

**After all processes**

- 3 at home
- 2 Out of home
- 4 trips,
  - 2 tours

TRB 2016
What do we know?
### What do we know?

<table>
<thead>
<tr>
<th>Variable</th>
<th>Diary</th>
<th>GPS (logger or mobile) (no prompted recall)</th>
<th>Mobile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation</td>
<td>Self-selected</td>
<td>Self-selected</td>
<td>(Random)</td>
</tr>
<tr>
<td>Duration</td>
<td>1 day (- 6 weeks)</td>
<td>1 day (- 6 weeks)</td>
<td>1 day (Unlimited)</td>
</tr>
<tr>
<td>Stage</td>
<td>Yes, underreported</td>
<td>(Yes)</td>
<td>No</td>
</tr>
<tr>
<td>Trip</td>
<td>Yes, underreported</td>
<td>Yes</td>
<td>(Yes)</td>
</tr>
<tr>
<td>Journey</td>
<td>Yes</td>
<td>Yes</td>
<td>(Yes)</td>
</tr>
<tr>
<td>Time</td>
<td>Rounded</td>
<td>Exact</td>
<td>Imputed</td>
</tr>
<tr>
<td>Location</td>
<td>Rounded</td>
<td>Exact</td>
<td>Imputed</td>
</tr>
<tr>
<td>Mode</td>
<td>Yes</td>
<td>Imputed</td>
<td>Imputed</td>
</tr>
<tr>
<td>Purpose</td>
<td>Yes</td>
<td>Imputed</td>
<td>Imputed</td>
</tr>
<tr>
<td>Group</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Expenditure</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
## What do we know?

<table>
<thead>
<tr>
<th>Variable</th>
<th>Diary</th>
<th>GPS (logger or mobile) (no prompted recall)</th>
<th>Mobile</th>
</tr>
</thead>
<tbody>
<tr>
<td>$/reported day</td>
<td>High</td>
<td>High-medium</td>
<td>Low</td>
</tr>
<tr>
<td>Data availability</td>
<td>Months</td>
<td>Week</td>
<td>Daily</td>
</tr>
<tr>
<td>Corrections</td>
<td>Needed</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Imputations</td>
<td>Needed</td>
<td>Needed</td>
<td>Needed</td>
</tr>
<tr>
<td>Choice models</td>
<td>Yes</td>
<td>Yes</td>
<td>Difficult</td>
</tr>
<tr>
<td>Socio-demographics</td>
<td>Yes</td>
<td>Yes</td>
<td>Imputed</td>
</tr>
</tbody>
</table>
French mobile data: long distance journeys over 5 months

![Graph showing the average number of long-distance tours by city size level.]

Legend:
- Orange GSM Data
- National Survey

City size levels:
- Rural
- <5k
- 5k–10k
- 10k–20k
- 20k–50k
- 50k–100k
- 100k–200k
- 200k–900k
- Paris

TRB 2016
What should we do?
Next steps

• Query what we really need for
  • Cost-benefit analysis
  • Planning of prices and services
  • Planning for the slow modes
  • Social accounting

• High-quality multi-modal surveys to establish the measurement errors (add bluetooth and wifi senders, noise profile)
• Error correction models
• Cross check against third party sources
Next steps

• Treat survey data as indicators in a measurement model
• Treat traces as indicators in a measurement model

• Improve imputation methods (mode, purpose) (socio-demographics)

• Address the self-selection issue of the survey/tracing participation
• Address the self-selection into different service providers