Travel Surveys – A Researcher’s View

KW Axhausen

IVT
ETH
Zürich

November 2014
Do we know the numbers? e.g. daily activities in Switzerland

![Graph showing the number of activities vs. average cohort age](image)
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, 2013

![Graph showing the relationship between A-priori estimate of the response burden and AAPOR response rate. Legend shows different categories: None, Only motivation call, Prior recruitment, Prior recruitment & incentive, and Predicted. The graph illustrates a negative correlation between the A-priori estimate of the response burden and the AAPOR response rate.]
Response is a non-random process

The graph shows the number of respondents over different waves and the target value. The data points represent the observed values, true mean, sample, and respondents. The graph indicates a non-random process as the number of respondents decreases over time.
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
Active/passive tracing: Many owners, locations, quality levels

Phone/SMS
Email
Smart card
GPS
Bluetooth
WiFi
CCTV
Filters imposed/suggested by the study: „Trips“

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>12</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out of home</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After “trip” filter:

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>12</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out of home</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Filters due to the respondent: Forgetting

Home

Out of home

Movement

After forgetting

Home

Out of home

Movement
Filters imposed by the respondent: Soft non-response

Home
Out of home
Movement

After soft non-response

Home
Out of home
Movement
Filters due to the respondent: Rounding

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>12</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out of home</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After rounding:

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>12</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out of home</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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
What happens next?
## Geocoding addresses

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideal</td>
<td>Street addresses identifying the entry to the network</td>
</tr>
<tr>
<td>Best-case</td>
<td>Unambiguous street addresses</td>
</tr>
<tr>
<td>State of the art</td>
<td>Street address</td>
</tr>
<tr>
<td>State of practice</td>
<td>Street address/mid-street block/street corners; missing conversion of facility names</td>
</tr>
<tr>
<td>Still seen in practice</td>
<td>Arbitrary zonal centroid, e.g. post offices</td>
</tr>
</tbody>
</table>
### Calculating distances & travel time

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideal</td>
<td>Complete GPS track for distance and times with pedestrian-networks added</td>
</tr>
<tr>
<td>Best-case</td>
<td>Minimal gaps, and state-of-the-art imputation of GPS tracks and modes</td>
</tr>
<tr>
<td>State of the art</td>
<td>SUE derived travel times and distances (navigation network)</td>
</tr>
<tr>
<td>State of practice</td>
<td>DUE derived travel times and distances (planning networks)</td>
</tr>
<tr>
<td>Still seen in practice</td>
<td>Shortest path on empty planning networks</td>
</tr>
</tbody>
</table>
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

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

www.ivt.ethz.ch