Lifting the barriers to data sharing

Big Data and Transport Models
Roundtable 14-16 December 2020
Background

• “Big and open data” offer new opportunities for transport policy making

• Potential shift in regulatory function of governments
  – Data-driven policy making > monitoring and enforcement of legislation

• Private sector collect much of the “big data”
  – Access to these could improve planning and regulation while *reducing survey costs*

• Need for partnership between public and private sector
Working Group on Big Data and Open Data

- USA (Chair: Patricia Hu, US DOT), Canada, UK, France, Austria, Netherlands, Finland, Italy, Greece, UNECE, SEETO

- Assist member countries in leveraging and curating big data by:
  - Summarising quality standards that could be used to communicate the quality and statistical validity of big and open data
  - Developing a framework to facilitate data sharing between the private and the public sectors
## OECD Data Quality Standards and Big Data

<table>
<thead>
<tr>
<th>Standard</th>
<th>Big Data Benefits</th>
<th>Big Data Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevance</td>
<td>Allows the study of new subjects</td>
<td>Can lack information needed to respond to objectives (socio-demographic, metadata...)</td>
</tr>
<tr>
<td>Accuracy</td>
<td>More detail, no human error</td>
<td>Evaluation of validity of results have to be modified</td>
</tr>
<tr>
<td>Credibility</td>
<td>---</td>
<td>Can lack methods of scientific rigor and objectivity, although work is being done to improve this</td>
</tr>
<tr>
<td>Timeliness</td>
<td>More data at faster rates</td>
<td>Data becomes out-of-date faster, requiring new methods of analysis</td>
</tr>
<tr>
<td>Accessibility</td>
<td>Collected quickly / instantaneously</td>
<td>Metadata and collection methods can sometimes be confidential and costly to obtain, assuming proper documentation has been recorded.</td>
</tr>
<tr>
<td>Interpretability</td>
<td>If accessible, level of detail allow for more precise understanding of population</td>
<td>Can lack info on sample, variables of interest or those needed for extrapolation of findings. Less control over what is collected</td>
</tr>
<tr>
<td>Coherence</td>
<td>Potential to compare between countries (Google search etc)</td>
<td>Lacks consistent collection and documentation procedures</td>
</tr>
</tbody>
</table>
Case studies of data sharing

• Canada: Rail and trucking waybill, Fuel consumption survey

• Finland: NordicWay project, Aurora project

• Italy: PLUG-IN project

• France: E-Logistics database, Use of passive data sources to estimate traffic on regional trains, SNCF route planner, Use of GSM data for estimation of tourism traffic,

• Netherlands: KiM Mobility Panel

• Greece: Delivering driving analytics to insurance companies
Stakeholder interviews

• A structured interview
  – Experience with existing partnership (type, structure, benefits)
  – Attitudes towards partnership
  – Third party set-up
  – Value proposition
  – Data sets (ownership, data sharing approach, level of detail, conditions, re-sharing, dissemination, open data)

• Austria: ASFINAG
• Canada: Desjardins Insurance
• Greece: Attic Tollway
• Germany: PTV
• Norway: Scania, PostNord Logistics
Observations

• New models of public-private data sharing partnerships considered indispensable

• There are several types of partnerships (large-scale data-dumps as part of surveys, fixed term projects, commercial schemes)...

• with different settings (fixed research projects, business development activities, mandatory data sharing schemes)...

• and mostly temporary (limited duration, pilot projects, data access only for a period of time)

• Data typically only shared between parties involved due to data protection and privacy issues
Recommendations to facilitate data sharing

• Governments should try to make any data received more widely open
• Ensure compliance with privacy protection regulation
  – Before sharing - anonymisation, encryption of data
  – Location, trajectory data most vulnerable – apply the most robust protection methods
• Building trust important
  – Develop and endorse non-disclosure agreements, involve trusted “third party”, develop “safe answers approach” (only query results exchanged instead of raw data)
• Apply appropriate partnership model depending on type of data use
  – Offering to produce timely open data in exchange (can help business)
  – Financial compensation (recovers investment costs and addresses free-riding)
  – Mandating (limit to data of public interest, apply purpose specificity and data minimisation principles)
Thank you

Questions & comments to:

Jari Kauppila
E jari.kauppila@itf-oecd.org