CHANGING PRACTICES IN DATA COLLECTION ON THE MOVEMENT OF PEOPLE (CPiDC)

Summary of the study undertaken by Lee-Gosselin Associates Ltd. for Transportation Association of Canada

Consultant team:

- Eric Miller, University of Toronto (Principal Investigator)
- Martin Lee-Gosselin (Project Director)
- Khandker Nurul Habib, University of Toronto
- Catherine Morency, Polytechnique Montreal
- Matthew J. Roorda, University of Toronto
- Amer S. Shalaby, University of Toronto
SPONSORS

- Agence Métropolitaine de Transport
- City of Edmonton
- City of Mississauga
- Ville de Montréal
- City of Ottawa
- Halifax Regional Municipality
- Ministère des Transports du Québec
- Metrolinx
- TransLink
- Transport Canada
This study was carried out under the supervision of a project steering committee of volunteer members. The participation of these committee members throughout the project is gratefully acknowledged.

- Timothy Spurr, Agence Métropolitaine de Transport (Co-Chair)
- Lisa Salsberg, Metrolinx (Co-Chair)
- Stéphane Brice, Ville de Montréal
- William Hui and Lorie Srivastava, TransLink
- Ahmad Subhani, City of Ottawa
- Arthur Tai, Ontario Ministry of Transportation
- Rhonda Toohey, City of Edmonton
- Pierre Tremblay, Ministère des Transports du Québec
- Bob Sasaki, Norbert Orzel, Joe Perrotta, City of Mississauga
- David MacIsaac, Halifax Regional Municipality
- Eric Sevigny and Arif Husain, Transport Canada

- Katarina Cvetkovic, TAC (project manager)
Project Objectives

The objective of the Changing Practices in Data Collection on the Movement of People project is to prepare a practical, comprehensive framework for the coordination, collection, processing and management of data on the movement of people by all modes in Canadian urban areas that is implementable and addresses data needs across the range of Canadian transportation agencies.
Project Components

- Literature Review
- Data Integration/Fusion Methods
- Review of Data Sources for Urban Transportation Applications
- Design, Conduct & Report on a Survey of Canadian Transportation Data Collection Practice
- A Framework for Urban Passenger Transportation Data Collection & Management

=> 6 volume, 295 page Report
Recommended Framework about:

1. An institutional (and political) commitment to on-going data collection and management
2. Assessment of an agency’s data needs
3. A multi-method approach to comprehensively and cost-effectively meet an agency’s full data needs
4. Integration using data fusion and synthesis methods
5. Controlled experimentation to evolve methods
6. Growing of metropolitan travel survey territories
Focus of Literature Review

- **Population-based Surveys**
  - Definitions and Basic Concepts in Survey Design
  - Household Travel Surveys
  - Issues and Challenges of Telephone Interview Surveys
  - Methodological Advances

- **Choice-based Sample Surveys**
  - Roadside Intercept Surveys
  - Transit User Intercept Surveys

- **Technology-based Data Collection Methods**
Data Fusion Applications studied

- Using the combined datasets for analysis/modelling purposes (e.g., combined revealed/stated preference model estimation).
- “Filling in” missing data in a dataset (e.g., adding income estimates to a household survey).
- Creating a synthetic population for model forecasting purposes.
- Creating a “pseudo-panel” from repeated cross-section surveys.
Possible Ancillary Data Sources for Urban Transportation Applications

- Canadian Census.
- Other Statistics Canada (StatCan) datasets.
- Other federal (non-StatCan) datasets.
- Data collected by professional transportation organizations (TAC, CUTA, etc.).
- Municipal datasets.
- Provincial datasets.
- Commercial/private sector datasets.
- Open source datasets (typically web-based).
Types of Data Sources Used by Canadian Transportation Agencies

From our own survey:

• Over 50% used the Canadian Census, Municipality data, and Provincial data

• Over 25% used StatCan’s Labour Force Survey, Other StatCan data, and On-line, Open-source data
CPiDC Survey

In order to identify current Canadian urban passenger transportation data collection practice, issues & needs, a national web-based survey of Canadian transportation agencies was conducted.
## Respondent Characteristics

<table>
<thead>
<tr>
<th>Province</th>
<th>Planning Department</th>
<th>Transportation Department</th>
<th>Public Works Department</th>
<th>Transit Agency</th>
<th>Other</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alberta</td>
<td>2</td>
<td>7</td>
<td></td>
<td>3</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>British Columbia</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Manitoba</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>New Brunswick</td>
<td></td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Newfoundland &amp; Labrador</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Nova Scotia</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Ontario</td>
<td>4</td>
<td>8</td>
<td>9</td>
<td>12</td>
<td>6</td>
<td>39</td>
</tr>
<tr>
<td>Quebec</td>
<td>2</td>
<td>4</td>
<td></td>
<td>12</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>Saskatchewan</td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Northwest Territories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Yukon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>11</strong></td>
<td><strong>25</strong></td>
<td><strong>12</strong></td>
<td><strong>32</strong></td>
<td><strong>14</strong></td>
<td><strong>94</strong></td>
</tr>
</tbody>
</table>
Respondent Characteristics

Level of Government
- Local Municipality: 61 (65%)
- Regional Municipality: 10 (11%)
- Provincial: 7 (7%)
- Other: 16 (17%)

2011 Population Served by Agency
- <10K: 17 (18%)
- 10-50K: 24 (26%)
- 50-100K: 7 (7%)
- 100-500K: 12 (13%)
- 500K-1M: 10 (11%)
- >1M: 23 (25%)
Household Travel Surveys

Agency's procurement and use of household travel surveys

- Conduct: 24 (23%)
- Commission: 8 (7%)
- Do not use: 27 (26%)
- Do not use but plan to do so over the next five years: 10 (9%)
- Use 3rd party survey: 28 (27%)
- Would like to use the data, but these data are not available: 5 (15%)

How many different Household travel surveys has your agency conducted or commissioned over the past ten years?

- 0: 2 (6%)
- 1: 1 (43%)
- 2: 3 (9%)
- 3: 2 (6%)
- >3: 14 (43%)
Household Travel Surveys

Sampling Frame

Interview Method
Household Travel Surveys

Do you have any issues/concerns with your current home-interview survey methods?

- Yes: 21 (64%)
- No: 9 (27%)
- N/A: 3 (9%)

Do you have any plans for addressing these issues/concerns?

- Yes: 14 (43%)
- No: 10 (30%)
- N/A: 9 (27%)
### Technology-based Data Collection

<table>
<thead>
<tr>
<th>Technology</th>
<th>Use now</th>
<th>Plan to use in next 5 years</th>
<th>Do not use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global positioning systems (GPS)</td>
<td>57</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>Other distributed or remote sensing technologies</td>
<td>42</td>
<td>3</td>
<td>43</td>
</tr>
<tr>
<td>Smartcards</td>
<td>25</td>
<td>19</td>
<td>46</td>
</tr>
<tr>
<td>Other transit pass technologies</td>
<td>16</td>
<td>6</td>
<td>63</td>
</tr>
<tr>
<td>Debit/credit cards</td>
<td>23</td>
<td>11</td>
<td>52</td>
</tr>
<tr>
<td>Social media</td>
<td>29</td>
<td>22</td>
<td>35</td>
</tr>
<tr>
<td>Other internet</td>
<td>45</td>
<td>12</td>
<td>30</td>
</tr>
<tr>
<td>Mobile devices</td>
<td>27</td>
<td>24</td>
<td>38</td>
</tr>
<tr>
<td>Other technologies</td>
<td>9</td>
<td>5</td>
<td>63</td>
</tr>
</tbody>
</table>

*N = 93 public agencies*
Barriers to Improved Data Collection

- Over 50% indicated inadequate budget, too few staff and competing priorities
- Over 24% indicated lack of perceived importance, lack of political support and/or limited technical capabilities of staff
- 13% indicated lack of management support
- Little else…
Framework must face these issues and trends

- Household surveys will remain a data collection priority
- Current methods face major challenges & must adapt:
  - Problems with land-line-based sampling frame
  - Problems with contacting & recruiting respondents
  - Problems with retrospective/proxy reporting
- New technology offer opportunities:
  - Multi-instrument surveys
  - Web-based surveys
  - Positioning (GPS/other) increasingly found on portable devices
- Social/Travel Behaviour Trends
- Technology Trends
- Trends in planning, policy issues, etc.
1. An institutional (and political) commitment to on-going data collection and management, that is not ad hoc or fragmented, of secure, high-quality data to support evidence-based planning.

2. A careful assessment of an agency’s data needs, leading to a comprehensive, ideally object-oriented, model of its data requirements that guides the extent and timing of data collection, while protecting privacy.

3. A multi-method approach to comprehensively and cost-effectively meet an agency’s full data needs through the wise and efficient utilization of a requisite variety of relevant data, including core and satellite transport surveys, and co-opted data from other sources.
FRAMEWORK: main recommendations 4-6

4. Integration using **data fusion** and synthesis methods that are *designed in* at the database planning stage to help operationalize the data model framework.

5. Controlled experimentation given changing needs and technological opportunities

6. Growing of metropolitan travel surveys to the provincial level
Towards a complete view of needed data

Data Collection Systems

Population, Employment, Car ownership, etc.
Road network, Transit network, Schedules, etc.
OD demand by mode, by purpose, by time of day, etc.
Travel times, speeds, pollution, safety
Fuel prices, interest rates, weather,
Census, Inventories, surveys, counts, passive data, impact monitoring, etc.

Transport/Activity Systems Operations

Activity System
Transportation Supply
Travel Demand
Performance and Impacts
System Characteristics
System Behaviour
External Factors
A Standardized Model of Transportation Data

- To capture this “complete view”, an “object-oriented” (OO) model of transportation systems and the data characterizing these systems is recommended. Uses:
  - Provides a systematic, comprehensive representation of important data elements & their relationships: determination of data needs for various applications.
  - Starting point for organizing data management systems.
  - Tool for model development
Example: OO Data Model for Urban Travel Demand Modelling
Core-Satellite Multi-Instrument Survey Design

Adapted from: Goulias, et al., (2011)
Core Survey

- Large-sample survey that gathers primary information concerning respondents & their behaviour.

- Characteristics:
  - Key data fundamental to agency primary needs.
  - Common variables that link to satellite data for joint use (integration/fusion).
  - Sample size permits valid statistical inferences.
  - Expandable to full population.
  - Large geographic area.
  - Stable (but not necessarily static) over time: consistent time-series.
  - Relatively short (minimize response burden; cost-effective for large samples).
Satellite Surveys

- Smaller sample, more focused and detailed - enrich/augment the core for:
  - Special, detailed models
  - Analysis of special behaviours of interest
  - Analysis of small sub-populations
- Must be linkable to the core
- Examples:
  - Extra questions and/or instrumentation of a subset of the core.
  - Additional survey of a core sub-sample
  - Increased (stratified) sampling of specialty populations
  - Surveys conducted on different samples but with common data items for linking/fusion
Core-Satellite Paradigm

- Very flexible/generalizable approach.
- Applicable to a wide variety of contexts.
- Defined by content, not method.
- Permits controlled experimentation, evolution over time:
  - Use satellites to test new methods.
  - “Grow” data over time by adding satellites as need and opportunity permit.
  - Respond to “hot button”, new issues/needs in a timely, flexible manner.
CORE SURVEY: Home Interview Survey
• Large sample
• Key/core variables
• Key household & person variables
• Trips by mode, purpose & time of day

Satellite 1
HOV Usage

Satellite 2
Bicycle Usage

Satellite 3
Elderly Travel Needs

Satellite 4
Auto Ownership & Usage

Example Core-Satellite Data Collection Design

- Screenline counts, all vehicle types
- Transit boarding counts
- Roadway speed-time studies
- Transit line headways, speeds, etc.; transit fare policies
- Parking supply & price
- Auto operating costs, including tolls
- Road segment capacities, speeds, etc.
- Census data

- Vehicles by type, vintage, fuel type, etc.
- Capital & operating costs of vehicles by type, vintage, fuel type, etc.
NB: Unlike most industrialized countries, Canada has no National Travel Survey to measure everyday travel.

The existence of well-established, large (population & geography) urban surveys in many provinces provides the potential to expand these surveys to the provincial level.

Significant potential advantages:
- Incremental, “bottom-up” approach to evolving an national data collection program (possibly the only way this is likely to happen).
- Provides uniform/standardized data across the province.
- Provides data to smaller/medium-sized areas very cost-effectively.
- Eliminates “urban boundary” effects: very important for fast-growing regions.
- Provides a framework for gathering long-distance & rural travel data.
Where do we go from here?

- Good case for creating a small and stable network of Canadian transportation data collection experts and major clients

- Possible first steps should focus on collaboration and sharing:
  - A Data Strategy Workshop for representative stakeholders from municipal, metropolitan, provincial and federal agencies, which as an output could specify....
  - ...the funding and creation of an all-Canada Working Group to design a coherent program of pilots and trials of the main components of the Data Framework, and translating its objectives (coordination, collection, processing and management) into action
  - The creation of a permanent Clearinghouse for data methodology and innovation
THANK YOU!

QUESTIONS?

CHANGING PRACTICES IN DATA COLLECTION ON THE MOVEMENT OF PEOPLE
SPONSORS

- Agence Métropolitaine de Transport
- City of Edmonton
- City of Ottawa
- Halifax Regional Municipality
- Ministère des Transports du Québec
- Metrolinx
- TransLink
- City of Mississauga
- Transport Canada
- Ville de Montréal
Extra slides for question period
FRAMEWORK: main recommendations

1. An institutional (and political) commitment to on-going data collection and management, that is not ad hoc or fragmented, of secure, high-quality data to support evidence-based planning.

2. A careful assessment of an agency’s data needs, leading to a comprehensive, ideally object-oriented, model of its data requirements that guides the extent and timing of data collection, while protecting privacy.

3. A multi-method approach to comprehensively and cost-effectively meet an agency’s full data needs through the wise and efficient utilization of a requisite variety of relevant data, including core and satellite transport surveys, and co-opted data from other sources.

4. Integration using data fusion and synthesis methods that are designed in at the database planning stage to help operationalize the data model framework.

5. Controlled experimentation given changing needs and technological opportunities.

6. Growing of metropolitan travel surveys to the provincial level.
Data Management

- Sound, systematic data management procedures essential
- Key elements include:
  - Storage
  - Documentation (metadata)
  - Completeness (contextual data)
  - Access
  - Dissemination
  - Quality Control – cumulative lessons reported by analysts
Privacy/Security

- Also essential; increasingly important:
  - Increasingly disaggregate datasets
  - Increasingly strict legislation, public concern
- Elements include
  - Gathering/storing private information
  - Access to confidential data
  - Identification of respondents
  - Secure, confidential archiving/storage
- Needs to be addressed at every point in the data collection process:
  - Survey design
  - Accessing respondents & gaining consent
  - Analysis of data
  - Storage/maintenance of data

“Privacy by design”
Types of Data Sources Used by Canadian Transportation Agencies

Source: Web-based survey of 93 public agencies conducted as part of project
Barriers to Improved Data Collection

Barriers for developing improved data collection & management methods

- Inadequate budget: 61
- Too few staff: 60
- Limited technical capabilities of staff: 22
- Competing priorities: 51
- Lack of management support: 12
- Lack of political support: 23
- Union/contract rules: 3
- Lack of perceived importance: 27
- Other: 6

(Bar chart showing the distribution of barriers to improved data collection and management methods)