A Tour-Based National Model System to Forecast Long Distance Passenger Travel

Presented by:
Nazneen Ferdous
RSG

Authors:
Mark Bradley, Nazneen Ferdous, and Maren Outwater: RSG
Stephane Hess and Andrew Daly: University of Leeds
Chandra Bhat: The University of Texas at Austin
Ram Pendyala: Arizona State University
Jeff LaMondia: Auburn University

April 29, 2014
Overview

• Objectives of the project

• Model components

• Model estimation data

• Model estimation results
  ➢ Tour generation and scheduling models
  ➢ Tour mode and destination choice models
Objectives of the project
Objectives of the project

Create an operational national-level model to forecast long-distance person-trips (50 miles or longer) by car, air, bus, and rail.

This involves:

• Design an “ideal” model system,
• Estimate various types of models on available data sources,
• Calibrate and validate the initial model,
• Create a prototype model for application, and
• Advise on data needs to inform future long-distance travel surveys.
Objectives of the project (cont.)

Who will use the resulting model?

Federal planning agencies
• There is currently no national model they can use for policy analysis.

State planning agencies
• At a minimum, could provide background long-distance inter-state trips for a statewide model.
• At the maximum, could replace the need for a statewide model (although it does not predict shorter distance rural trips).

Regional planning agencies
• At a minimum, can provide background, long-distance inter-regional trips to supplement intra-regional forecasts
• At the maximum, could be integrated with a regional model to predict residents leaving the region instead of traveling within the region.
  ➢ Should be balanced by modeling visitor travel within the region. Our model will predict visitors entering the region, but not short distance trips that they make while there.
Model components
Model Components

Similarities with activity-based travel demand models:

• Travel is simulated at household-level.

• There is a similar set of “longer term” models predicting household auto ownership and workplace location for all workers.

• The proposed model uses a level of spatial detail for trip ends that is very fine compared to the size of the study area (in this case, all Census tracts in the US).

• The models are “vertically integrated”, meaning that mode choice, destination choice, and tour generation and scheduling are all mutually inter-related with causal effects in both directions.
Model Components (cont.)

Major differences with activity-based travel demand models:

• The long distance model predicts only tours to destinations 50 miles or more from home.

• The model does not predict all intermediate stop locations visited during long distance tours.
  ➢ It only predicts trips of 50 miles or more, but does not simulate shorter trips made while staying away from home.

• The long distance tours are scheduled across a year.

• The model jointly predicts destination and mode choice.
## Model Components (cont.)

<table>
<thead>
<tr>
<th>Model</th>
<th>What is predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Longer term choice models</strong></td>
<td></td>
</tr>
<tr>
<td>Synthetic population generation</td>
<td>The household and person characteristics of residents of each Census tract (Done using PopGen software).</td>
</tr>
<tr>
<td>Workplace location</td>
<td>The Census tract of the usual work location for each worker.</td>
</tr>
<tr>
<td>Auto ownership</td>
<td>The number of vehicles owned by the household.</td>
</tr>
<tr>
<td><strong>Tour level (“core”) models</strong></td>
<td></td>
</tr>
<tr>
<td>Tour generation, duration &amp; scheduling</td>
<td>The number of long distance tours made for each purpose category and duration category, scheduled across the months of the year. The tour purposes considered here include: Business, Commute, Visiting Friends and Relatives, Leisure, and Personal Business/Other.</td>
</tr>
<tr>
<td>Tour party size &amp; composition</td>
<td>The number of adults and children participating in each tour.</td>
</tr>
<tr>
<td>Tour primary destination and mode</td>
<td>The Census tract of the tour primary destination, and the main mode used for the tour (auto, air, rail or bus).</td>
</tr>
<tr>
<td><strong>Trip-level models</strong></td>
<td></td>
</tr>
<tr>
<td>Intermediate stop generation</td>
<td>The number of intermediate long-distance destinations visited during the tour, and the purpose of each stop.</td>
</tr>
<tr>
<td>Intermediate stop location</td>
<td>The Census tract of each intermediate stop visited.</td>
</tr>
<tr>
<td>Trip level mode choice</td>
<td>The mode for each trip, depending mainly on the main tour mode.</td>
</tr>
<tr>
<td>Trip level departure time</td>
<td>The day of week and time period of day for each trip, depending on the tour period and duration.</td>
</tr>
</tbody>
</table>
# Model Estimation Data

## Household Travel Survey: Suitability of the Datasets by Models to be Estimated

<table>
<thead>
<tr>
<th>Household travel survey</th>
<th>The dataset is suitable for the estimation of</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tour frequency</td>
<td>Scheduling</td>
</tr>
<tr>
<td>1995 ATS</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2001 NHTS – National</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>2001 NHTS - NY</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2001 NHTS - WI</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>2003 Ohio</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>2004/2009 Michigan</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>2010 Colorado</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012 CHTS</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

*New York does not currently have a state-wide model. So, for destination and mode choice model estimation, skims from the National Highway Planning Network (NHPN) would have to be used.*
## Model Estimation Data (cont.)

### Land use, level-of-service, and other input data

<table>
<thead>
<tr>
<th>Data type</th>
<th>What is included</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Land use data</strong></td>
<td>Population, employment by sector, university enrollment, parks and land coverage at Census tract level.</td>
</tr>
<tr>
<td><strong>Synthetic population</strong></td>
<td>Household and person records for the entire US synthetic population, located to Census tract.</td>
</tr>
<tr>
<td><strong>Rail network</strong></td>
<td>Passenger rail station-to-station matrices of-vehicle time, distance, transfers, frequency and fares, based on Amtrak data.</td>
</tr>
<tr>
<td><strong>Air network</strong></td>
<td>Airport-to-airport matrices of distance, in-vehicle time, transfers, frequency, fare, and on-time reliability, from the DB1B ticket and On-time databases.</td>
</tr>
<tr>
<td><strong>Auto network</strong></td>
<td>A base national auto network, with connectors added for each Census tract, airport and rail station.</td>
</tr>
<tr>
<td><strong>Auto skims</strong></td>
<td>Distance, time and toll for each TAZ-TAZ pair. TAZ’s are an intersection of PUMAs and counties, to avoid having too much population or land area in any one TAZ.</td>
</tr>
<tr>
<td><strong>Airport and station access</strong></td>
<td>Auto travel distance and time from each Census tract to all airports and stations within a 100 mile radius.</td>
</tr>
<tr>
<td><strong>Air and rail skims</strong></td>
<td>Best tract-to-tract air and rail paths are found “on the fly” via all reasonable airport or station pairs.</td>
</tr>
<tr>
<td><strong>Mode/destination accessibility logsums</strong></td>
<td>Calculated for each Census tract / income / car ownership / purpose / party size combination, for tours to destinations in various distance bands.</td>
</tr>
</tbody>
</table>
Model estimation results
Tour Generation and Scheduling Models

• Separate long distance tour generation and scheduling model for employer’s business, visit friends and relatives, vacation/leisure, and personal business.

• For employer’s business travel, a combination of negative binomial model, linear regression, and hazard duration model was used.

• MDCEV models were estimated for non-business tour purposes.
  ➢ simultaneous modeling of long distance tour generation, purpose, duration, and scheduling (season of the year).
  ➢ Captures “satiation” in the number and duration of long distance tours.
  ➢ Represent effects of time budget.

• 1995 American Travel Survey (ATS) was used.
  ➢ Dated, little spatial details, the date of travel is recorded at 3 months intervals.

• Additional tour frequency models were estimated to address the limitations of ATS data.
Tour Destination and Mode Choice Models

• Separate models for Vacation/Leisure, Visit Friends and Relatives, Employer’s Business, and Personal Business (includes school, shopping, and medical).

• 20,764 choice alternatives.
  ➢ 4 modes: car, air, rail, and bus.
  ➢ 5,191 destination zones.

• 2012 California Statewide Household Travel Survey.
  ➢ ≈ 30,000 tour records.
  ➢ Not used to model commute tours.

• Tested MNL, nested, and cross-nested logit model structure.
Tour Destination and Mode Choice Models (cont.)

Nested logit model
• We tested both destinations nested under modes and modes nested under destinations.

• The best order of nesting appears to vary by tour purpose:
  ➢ Destination nested above mode: Employer’s Business and Vacation/Leisure.
  ➢ Mode nested above destination: Visit Friends and Relatives and Personal Business.

• There is significant heterogeneity in the model scale for destinations of different distances from the tour origin.
  ➢ Long distance models can cover a very wide range of distances, from 50 miles up to more than 5,000 miles.

• We are testing alternative specifications where the error scale for each destination alternative is a non-linear function of the distance to reach that alternative.

Cross-nested logit model
• CN structure appears to work better than either nesting order.
• More tests require.
Contacts

Nazneen Ferdous, RSG
Email: nazneen.ferdous@rsginc.com
Phone: 240-283-0634

Brad Gudzinas, FHWA
Email: brad.gudzinas@dot.gov
Phone: 202-366-5024