Route Choice: Models and Challenges

Shlomo Bekhor
Technion – Israel Institute of Technology

Swiss Transport Research Conference
September 1st, 2010
Presentation Topics

- Introduction
- Choice Set Generation
- Model Formulations
- Case Studies
- Accounting for Congestion
- Research Directions
Two types of choice behavior

1. **Pre–trip choice**: made before starting the trip
   - Continuous service systems (road and pedestrian networks) without unexpected events

2. **En-route choice**: made during the trip, to adapt to random or unknown events
   - Road systems with unexpected events
   - En-route information

Route choice models assume either pre-trip or mixed pre-trip/en-route choice behavior
   - Depending on the characteristics of the transportation service they are applied to.
Alternative modeling approaches

- Fuzzy logic (e.g., Lotan and Koutsopoulos, 1993; Lotan, 1997; Henn, 2000; Rilett and Park, 2001; Ridwan, 2004)
- Artificial neural networks (e.g., Yang et al., 1993; Dougherty, 1995; Yamamoto et al., 2002)
- Cognitive psychology (e.g., Nakayama and Kitamura, 2000; Nakayama et al., 2001)
- Random Utility – most common
- **This presentation** – route choice for a single mode (private car)
Route Choice Models: Two-stage Choice Process

1. Choice Set Generation

2. Route Choice Given a Choice Set
Route Attributes

- Travel Times: Time in Motion, Time at Stop Lights, Delay at Bottlenecks
- Cost: Out-of-Pocket, Long-Term
- Uncertainty or Variance of Travel Time
- Number of Stop Signs or Stop Lights on Route
- Volume of Conflicting Traffic or Pedestrian Movements
- Number of Turns on Route - Ease of Memory, Left Turns Against Traffic, Protected Lefts at Lights
- Street Width, Number of Lanes, Effort Required to Maneuver
- Circuitry of Route
- Safety, Roadway Condition
Path Utility Specification

- Link-Based Variables (such as travel time) versus Path-Based Variables (such as scenic route) - Affects Need for Enumeration

- Path-Based Variables (Including the “Path Size” or “Commonality Factor”) Require Enumeration
Choice Set Generation Models

- **exhaustive**
  - deterministic
    - K-shortest paths
    - Link penalty
    - Link elimination
    - Labeling (Ben-Akiva et al., 1984)
    - Constrained enumeration
  - stochastic
    - STOCH (Dial, 1971)
    - Simulation of link attributes
- **selective**
  - probabilistic
    - Formulation (Manski, 1977)
    - Captivity, Independent Availability (Ben-Akiva, 1977)
    - Choice Set Indicators (Ben-Akiva and Boccara, 1995)
    - Availability Model (Cascetta et al., 1998)
    - Random walk (Frejinger et al., 2009)
Simulation of Link Attributes

Network Topology → Random Link Costs → Shortest Path → Add to Choice Set

The same route may be found several times during the iterative process.

Nielsen (2000)
Bekhor et al. (2001)
Fiorenzo-Catalano and Van der Zijpp (2001)
Bierlaire and Frejinger (2005)
Bovy and Fiorenzo-Catalano (2006)
Link Elimination Method

- Network Topology
- Shortest Path
- Add to Choice Set
- Delete Link

The same route may be found several times during the iterative process.

- Schussler et al. (2010) Breadth-First Search

Azevedo et al. (1993)
Bekhor et al. (2001)
Prato and Bekhor (2006)
Frejinger and Bierlaire (2007)
Branch and Bound Method

- A link is inserted to the tree if and only if all the following conditions hold:
  - Directional constraint: (excludes from consideration links that take the driver significantly farther from the destination and closer to the origin)
  - Temporal constraint: (excludes paths with unrealistic travel times)
  - Loop constraint (remove paths with large detours)
  - Similarity constraint (remove high overlapping path segments)
  - Left turn constraint: (maximum number of left-turns per route)
Branch and Bound Tree

The algorithm processes a tree level before path segments of the next level are considered. The algorithm completes the connection search when all levels are processed and for all the branches the node corresponds to the destination.
Evaluation of Path Generation Algorithms

- Coverage: generated route matches the observed route at a specified threshold (Bovy, 2007):

\[
C_g = \frac{N_{cov}}{N_{obs}} = \frac{\sum_{n=1}^{N_{obs}} I(O_{ng} \geq \delta)}{N_{obs}} = \frac{\sum_{n=1}^{N_{obs}} \left( \frac{L_{ng}}{L_n} \geq \delta \right)}{N_{obs}}
\]

- Efficiency index: compares the path generation technique with an ideal algorithm that would replicate link-by-link all the observed routes and would produce for every origin-destination pair the chosen path and an alternative for modeling purposes Bekhor and Prato (2009):

\[
EI_G = \frac{1}{N_{obs}} \sum_{n=1}^{N_{obs}} \left\{ I \left( \frac{L_{G,n}}{L_n} \geq \delta \right) + \left( 1 - \frac{GR_{gen,n} - R_{rel,n}}{GR_{gen,n}} \right) \right\}/2
\]
Case Studies

- Boston (Ramming, 2001)
  - 900 Traffic Zones
  - 12,000 Nodes
  - 20,000 Links
  - 188 Observations

- Turin (Prato, 2005)
  - 182 O-D Pairs
  - 417 Nodes
  - 1,427 Links
  - 236 Observations
Screenshot of the Turin map with nodes
# Path generation techniques

<table>
<thead>
<tr>
<th>Technique</th>
<th>Boston</th>
<th>Turin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeling approach</td>
<td>( f(h) = \min(\text{distance}), \min(\text{free flow time}), \min(\text{travel time}) )</td>
<td>( pf = 1.05, I_h(L_{i,j}) = \text{travel time}, T = 15 \text{ iterations} )</td>
</tr>
<tr>
<td>Link penalty</td>
<td>( pf = 1.05, I_h(L_{i,j}) = \text{travel time}, T = 15 \text{ iterations} )</td>
<td>remove one link from shortest path, ( T = 10 \text{ iterations} )</td>
</tr>
<tr>
<td>Link elimination</td>
<td>remove one link from shortest path, ( T = 50 \text{ iterations} )</td>
<td>remove one link from shortest path, ( T = 10 \text{ iterations} )</td>
</tr>
<tr>
<td>Simulation</td>
<td>( f( I_h(L_{i,j})) ) are three normal distributions, ( I_h(L_{i,j}) = \text{travel time}, T = 16, 32, 48 \text{ draws} )</td>
<td>( f( I_h(L_{i,j})) ) are two truncated normal distributions, ( I_h(L_{i,j}) = \text{travel time}, T = 25, 35 \text{ draws} )</td>
</tr>
<tr>
<td>Branch and bound</td>
<td>( \Delta_D = 1.10, \Delta_T = 1.33, \Delta_L = 1.20, \Delta_O = 0.80, \Delta_{LT} = 7 )</td>
<td>( \Delta_D = 1.10, \Delta_T = 1.50, \Delta_L = 1.20, \Delta_O = 0.80, \Delta_{LT} = 16 )</td>
</tr>
</tbody>
</table>
## Coverage for different overlap thresholds

<table>
<thead>
<tr>
<th>Generation technique</th>
<th>Boston</th>
<th>Turin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100%</td>
<td>80%</td>
</tr>
<tr>
<td>Labeling approach</td>
<td>39.4</td>
<td>51.6</td>
</tr>
<tr>
<td>Link elimination</td>
<td>60.1</td>
<td>71.3</td>
</tr>
<tr>
<td>Link penalty</td>
<td>53.7</td>
<td>73.9</td>
</tr>
<tr>
<td>Simulation (16 draws)*</td>
<td>43.6</td>
<td>70.7</td>
</tr>
<tr>
<td>Simulation (32 draws)*</td>
<td>48.9</td>
<td>76.1</td>
</tr>
<tr>
<td>Simulation (48 draws)</td>
<td>50.0</td>
<td>78.7</td>
</tr>
<tr>
<td>Branch and bound</td>
<td>75.5</td>
<td>96.3</td>
</tr>
</tbody>
</table>

* Respectively 25 and 35 draws for the Turin network
# Efficiency Index for the Turin network

<table>
<thead>
<tr>
<th>path generation technique</th>
<th>efficiency index (%) for overlap threshold equal to:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100%</td>
</tr>
<tr>
<td>Labeling approach</td>
<td>43.0</td>
</tr>
<tr>
<td>Link elimination</td>
<td>48.2</td>
</tr>
<tr>
<td>Link penalty</td>
<td>42.5</td>
</tr>
<tr>
<td>Simulation (25 draws)</td>
<td>41.2</td>
</tr>
<tr>
<td>Simulation (35 draws)</td>
<td>36.2</td>
</tr>
<tr>
<td>Branch and bound</td>
<td>54.5</td>
</tr>
</tbody>
</table>
Comparison of Unique Routes Generated

- b%b 100%
- b%b 80%
- Simulation 100%
- Simulation 80%

Coverage vs. Number of Unique Routes
Characteristics of the Choice Set

The diagram illustrates the cumulative percentage of observations against the maximum number of routes for OD pairs for Turin and Boston. The cumulative percentage increases as the number of routes increases, showing that a smaller proportion of the observations have a higher number of routes. The graph distinguishes between Turin (black dots) and Boston (pink squares).
The Sampling Problem

- Stochastic route choice set generation procedures may sample routes with unequal selection probabilities.
- The selection probability of a route depends on the properties of the route itself, such as length or travel time.
- The systematic utility of the routes should be corrected for the unequal selection probabilities.
- Frejinger et al. (2009): Sampling correction using random walk as generation method
Route Choice Model Formulations

- Deterministic choice models
  - Shortest path (generalized cost)
  - Used in most transportation packages

- Probabilistic choice models
  - Multinomial Logit (MNL) – Dial’s algorithm
  - Modified Logit (C-Logit, Path-Size Logit)
  - GEV Models (CNL, LNL, PCL)
  - Probit / Logit Kernel
The Overlapping Problem

- Introduces Correlation - Violates IIA - MNL is Unsuitable
- Traditional Example with Three Paths, Two Overlapping
  - As overlap approaches 100 percent (b→T), expect close to 50/25/25 shares
  - As overlap approaches 0 percent (b→0), expect close to 33/33/33 shares
- MNL Predicts 33/33/33 Shares for Any Value of b
The C-Logit model (Cascetta et al., 1996)

\[ P(i) = \frac{\exp \mu(V_i - CF_i)}{\sum_{j \in C} \exp \mu(V_j - CF_j)} \]

CF - Commonality Factor - Several possible specifications:

(i) \( CF_i = \ln \sum_{j \in C} \left( \frac{L_{ij}}{\sqrt{L_i L_j}} \right) \)

(ii) \( CF_i = \sum_{a \in \Gamma_i} \left( \frac{L_a}{L_i} \sum_{j \in C} \delta_{aj} \right) \quad \Gamma_i - \text{Set of links included in route } i \)
The Path-Size Logit model

\[
P(i) = \frac{\exp \mu(V_i + \ln PS_i)}{\sum_{j \in C} \exp \mu(V_j + \ln PS_j)}
\]

Ben-Akiva and Bierlaire (1998)

\[
PS_i = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \sum_{j \in C} \delta_{aj}
\]

Ramming (2001)

\[
PS_i = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \frac{1}{\sum_{j \in C} \left( \frac{L_i}{L_j} \right)^\gamma} \delta_{aj}
\]

Bovy et al. (2008) – Path Size Correction

\[
P(i) = \frac{\exp \mu(V_i + PSC_i)}{\sum_{j \in C} \exp \mu(V_j + PSC_j)}
\]

\[
PSC_i = -\sum_{a \in \Gamma_i} \frac{L_a}{L_i} \ln \sum_{j \in C} \delta_{aj}
\]
Adapting the CNL to route choice

Example Network

Route 1: Link a
Route 2: Links b-c
Route 3: Links b-d

Model Structure

$$\alpha_{km} = \frac{L_m}{L_k} = \frac{\text{link "length"}}{\text{route "length"}}$$
Adapting the Multinomial Probit model to Route Choice Situation

- **Problem**: define a variance-covariance matrix \( \Sigma \)
- **Solution**: Daganzo (1980), Sheffi and Powell (1982)
  - variances are proportional to the mean travel time
- **Example**: structured covariance matrix: Yai et al. (1996)

\[
\Sigma = \sigma^2 \begin{bmatrix}
L_1 & L_{12} & \cdots & L_{1J} \\
L_{12} & L_2 & \cdots & L_{2J} \\
\vdots & \vdots & \ddots & \vdots \\
L_{1J} & L_{2J} & \cdots & L_J
\end{bmatrix}
\]

Transit route choice (small number of alternatives)

The probability was computed using numerical integration
The Logit Kernel model

\[ U = \beta X + F \xi + \nu \]
\[ \xi = T \zeta \]

\[ \text{cov}(U) = F T T^T F^T + (g/\mu^2) I \]

\(\beta\) - (K*1) vector of unknown parameters
\(X\) - (J*K) matrix of explanatory variables
\(F\) - (J*M) factor loadings matrix
\(T\) - (M*M) lower triangular matrix of unknown parameters
\(\xi\) - (M*1) vector of unknown factors
\(\nu\) - (J*1) vector of i.i.d. Gumbel variables
Logit Kernel Probability Calculation

If the factors $\zeta$ are known:

$$\Lambda(i|\zeta) = \frac{\exp(\mu(X_i\beta + F_T\zeta))}{\sum_j \exp(\mu(X_j\beta + F_T\zeta))}$$

Since the factors are unknown, the unconditional probability is given by:

$$P(i) = \int_\zeta \Lambda(i|\zeta) \prod \phi(\zeta) d\zeta$$

This probability function can be estimated by simulation:

$$P(i) = \frac{1}{D} \sum_{d=1}^{D} \Lambda(i|\zeta^d)$$
Adaptation of LK to route choice situation (as in Probit)

- Link specific factors are iid Normal
- Variance proportional to the link “length”
- The T matrix is the link factors variance matrix (diagonal matrix)
- Bekhor et al. (2002): The F matrix is the link-path incidence matrix
- Frejinger and Bierlaire (2009): Subnetwork approach
Adapting the LK for route choice

Route 1: Link a
Route 2: Links b-c
Route 3: Links b-d

F matrix (J*M)

\[
F = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 \\
\end{bmatrix}
\]

T matrix (M*M)

\[
T = \begin{bmatrix}
\sigma_a & 0 & 0 & 0 \\
0 & \sigma_b & 0 & 0 \\
0 & 0 & \sigma_c & 0 \\
0 & 0 & 0 & \sigma_d \\
\end{bmatrix}
= \alpha \begin{bmatrix}
\sqrt{t_a} & 0 & 0 & 0 \\
0 & \sqrt{t_b} & 0 & 0 \\
0 & 0 & \sqrt{t_c} & 0 \\
0 & 0 & 0 & \sqrt{t_d} \\
\end{bmatrix}
\]

Utility vector

\[
U_i = \begin{bmatrix}
\beta X_1 + \sigma_a \zeta_1 + \nu_1 \\
\beta X_2 + \sigma_b \zeta_2 + \sigma_c \zeta_3 + \nu_2 \\
\beta X_3 + \sigma_b \zeta_2 + \sigma_d \zeta_4 + \nu_3 \\
\end{bmatrix}
\]

Covariance matrix

\[
FTT^T F^T = \begin{bmatrix}
\sigma_a^2 & 0 & 0 \\
0 & \sigma_b^2 + \sigma_c^2 & \sigma_b^2 \\
0 & \sigma_c^2 & \sigma_b^2 + \sigma_d^2 \\
\end{bmatrix}
\]
Example 1: Red Bus – Blue Bus Network

![Graph showing the probability of choosing route 1 against the length of common link b. The graph includes lines for MNL, C-Logit, PS-Logit, CNL_0, CNL_0.5, CNL (mi var.), and Probit. Each line represents a different model, and the probability increases as the length of the link increases.]
Integration of Latent Variables

Latent variables: structural equations

- Latent variables as function of travelers’ characteristics
- L latent variables, M explanatory variables

\[ X^*_l = S_l \gamma_l + \omega_l \quad \omega_l \sim N(0, \sigma_{\omega_l}) \]

Latent variables: measurement equations

- Indicators as functions of latent variables
- R indicators, L latent variables

\[ I_m = X^*_l \alpha_r + \nu_m \quad \nu_m \sim N(0, \sigma_{\nu_r}) \]
Integrated Model

Route choice: structural equations

- Utilities as functions of observable and unobservable variables
- J alternatives, K observable variables, L latent variables

\[ U_{jn} = Z_{jn} \beta_{obs} + X^*_{jn} \beta_{lat} + \varepsilon_{jn} \quad \varepsilon_{jn} \sim Gumbel \]

Route choice: measurement equations

- Choice of alternative i as function of utilities
- J alternatives, N observations

\[ y_{in} = \begin{cases} 1 \text{ if } U_{in} \geq U_{jn} \quad \forall j \neq i, 0 \text{ otherwise} \end{cases} \]
Integrated Model

Case study assumptions

- Latent variables are orthogonal (null covariances in Sw)
- Indicators are independent (estimate variances in Su)
- Choice model is a Path Size Correction Logit

Choice probability

\[
P(y_n, I_n | Z_n, S_n, \alpha, \beta, \gamma) = \\
\int_{X^*_n} P(y_n | X^*_n, Z_n, \beta) g(I_n | X^*_n, \alpha) f(X^*_n | S_n, \gamma) dX^*_n
\]
Integrated Model

Simulated choice probability

\[
\hat{P}(y_n, I_n | Z_n, S_n, \alpha, \beta, \gamma) = \\
\frac{1}{D} \sum_{d=1}^{D} \sum_j \exp \left( Z_{in} \beta_{obs} + X_{in}^d \beta_{lat} + PSC_i \beta_{PSC} \right) \exp \left( Z_{jn} \beta_{obs} + X_{jn}^d \beta_{lat} + PSC_j \beta_{PSC} \right) g(I_n | X_{n^d}^*, \alpha)
\]

Sequential estimation

- SEM estimator for the latent variables model
- Maximum simulated likelihood for the route choice model

\[
X_{in}^* = S_{ln} \gamma_l + \omega_{ln}^d = S_{ln} \gamma_l + \sigma_{\omega_l} \tilde{\omega}_{ln} \quad \text{where} \quad \tilde{\omega}_{ln} \approx N(0,1)
\]
Measurement equations of the latent variable model

<table>
<thead>
<tr>
<th>Variable</th>
<th>MEM</th>
<th>Variable</th>
<th>HAB</th>
<th>Variable</th>
<th>TSAV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>est.</td>
<td>t-stat.</td>
<td>est.</td>
<td>t-stat.</td>
<td>est.</td>
</tr>
<tr>
<td>Memroute</td>
<td>1.000</td>
<td>-</td>
<td>1.000</td>
<td>-</td>
<td>1.000</td>
</tr>
<tr>
<td>Memhome</td>
<td>0.838</td>
<td>5.74</td>
<td>0.975</td>
<td>2.46</td>
<td></td>
</tr>
<tr>
<td>Memmind</td>
<td>0.712</td>
<td>5.35</td>
<td>1.382</td>
<td>2.84</td>
<td></td>
</tr>
<tr>
<td>Memlayout</td>
<td>0.690</td>
<td>4.96</td>
<td>1.668</td>
<td>2.86</td>
<td></td>
</tr>
<tr>
<td>Memway</td>
<td>1.398</td>
<td>7.56</td>
<td></td>
<td>-0.962</td>
<td>-2.05</td>
</tr>
<tr>
<td>Mempark</td>
<td>1.372</td>
<td>7.62</td>
<td></td>
<td>-1.266</td>
<td>-2.87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>FAM</th>
<th>Variable</th>
<th>SPAB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>est.</td>
<td>t-stat.</td>
<td>est.</td>
</tr>
<tr>
<td>Dscfamrt</td>
<td>1.000</td>
<td>-</td>
<td>1.000</td>
</tr>
<tr>
<td>Dscrthom</td>
<td>0.612</td>
<td>5.61</td>
<td></td>
</tr>
<tr>
<td>Evalroute</td>
<td>0.650</td>
<td>5.33</td>
<td></td>
</tr>
<tr>
<td>Navhome</td>
<td>0.150</td>
<td>2.59</td>
<td></td>
</tr>
<tr>
<td>Drvmain</td>
<td>-0.116</td>
<td>-2.17</td>
<td>-0.446</td>
</tr>
</tbody>
</table>
Structural equations of the latent variable model

<table>
<thead>
<tr>
<th>Variable</th>
<th>MEM</th>
<th>HAB</th>
<th>FAM</th>
<th>SPAB</th>
<th>TSAV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>est.</td>
<td>t-stat.</td>
<td>est.</td>
<td>t-stat.</td>
<td>est.</td>
</tr>
<tr>
<td>Male</td>
<td>0.228</td>
<td>1.93</td>
<td>-0.080</td>
<td>-0.92</td>
<td>0.260</td>
</tr>
<tr>
<td>Age&lt;35</td>
<td>0.198</td>
<td>2.83</td>
<td>-0.220</td>
<td>-2.22</td>
<td>-</td>
</tr>
<tr>
<td>Age&gt;55</td>
<td>-0.134</td>
<td>-1.73</td>
<td>0.233</td>
<td>2.04</td>
<td>-</td>
</tr>
<tr>
<td>Education</td>
<td>0.292</td>
<td>2.37</td>
<td>0.214</td>
<td>2.35</td>
<td>-</td>
</tr>
<tr>
<td>Single</td>
<td>-</td>
<td>-</td>
<td>-0.316</td>
<td>-2.63</td>
<td>-</td>
</tr>
<tr>
<td>Children</td>
<td>0.311</td>
<td>2.37</td>
<td>-</td>
<td>-</td>
<td>0.263</td>
</tr>
<tr>
<td>Family</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Stops</td>
<td>-0.360</td>
<td>-2.63</td>
<td>0.287</td>
<td>2.96</td>
<td>-0.360</td>
</tr>
<tr>
<td>City Resid.</td>
<td>-</td>
<td>-</td>
<td>0.256</td>
<td>2.84</td>
<td>0.231</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.215</td>
<td>-1.87</td>
<td>-0.131</td>
<td>-1.48</td>
<td>-0.293</td>
</tr>
</tbody>
</table>
Structural equations of the latent variable model

- Males have higher mnemonic capability and better spatial abilities and time saving skills
- Young respondents have better memory and time saving skills
- Educated respondents have higher abilities in terms of memory, spatial orientation, time saving, and also habitual behavior
- Singles have less habitual behavior and lower time saving skills, while having children relates to higher memory and familiarity
- Residents in the city relate with familiar and routine behavior
# Model Estimation Results

## Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>w/o latent variables</th>
<th>with latent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>t-stat.</td>
</tr>
<tr>
<td>Dist</td>
<td>-0.680</td>
<td>-4.59</td>
</tr>
<tr>
<td>Time</td>
<td>-0.346</td>
<td>-6.84</td>
</tr>
<tr>
<td>Delay %</td>
<td>-0.458</td>
<td>-3.92</td>
</tr>
<tr>
<td>Time on major roads %</td>
<td>0.530</td>
<td>3.74</td>
</tr>
<tr>
<td>Path Size Correction</td>
<td>0.681</td>
<td>3.35</td>
</tr>
<tr>
<td>Memory</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Habit</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Familiarity</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Spatial ability</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Time saving skill</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Estimated parameters</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Null log-likelihood</td>
<td>-1298.38</td>
<td></td>
</tr>
<tr>
<td>Final log-likelihood</td>
<td>-1062.81</td>
<td></td>
</tr>
<tr>
<td>Adjusted rho-bar squared</td>
<td>0.178</td>
<td></td>
</tr>
</tbody>
</table>
Route Choice Model

**Interpretation**

- Expected and logical effects of distance, travel time, congestion and attractiveness of major roads
- Mnemonic, spatial and time saving abilities seem have a positive effect: most likely reflect that individuals tend to look for better alternatives and use them since they remember them
- Habit and familiarity appear to have a negative effect: most likely reflects that individuals do not tend to search for better alternative routes even if their choice is not optimal
Route Choice and Equilibrium models

- Congestion Effect
- Stochastic Effect
- MNL SUE Assignment
- Similarity Effect
- Route Choice Models
Example 2: Grid Network

\[ Q = 1000 \]

\[ c_a = t_{0a} \left( 1 + 0.6 \left( \frac{x_a}{1000} \right)^4 \right) \]
Example 2: Grid Network

Deterministic UE Assignment

MNL Assignment

CNL Assignment

PCL Assignment
Example 2: Grid Network

- Flow on link 2 (percentage of total demand)
- Demand

Graph showing the flow on link 2 for different values of demand with curves for C-Logit, MNL, GNL, and PSL models.

- V = -0.5t
- V = -1.0t

Legend:
- C-Logit
- MNL
- GNL
- PSL

45
Example 3: Winnipeg Network

948 nodes, 2535 links, 154 traffic area zones
4345 OD pairs with positive demand
Total demand 54459 trips.
Winnipeg Network – Choice Set Generation

Maximum 50 routes for each OD pair
Total 174491 paths generated (40.1 paths on average)
Influence of Path-Set Size – Winnipeg

Objective Function vs. Maximum number of routes for OD pair.

- The graph shows the relationship between the maximum number of routes for an OD pair and the objective function, which decreases as the maximum number of routes increases.
- The y-axis represents the objective function values, ranging from 900000 to 1150000.
- The x-axis represents the maximum number of routes, ranging from 0 to 50.

The curve indicates a decreasing trend, suggesting that increasing the maximum number of routes for an OD pair decreases the objective function.
Iterations Needed to Reach Convergence

![Graph showing iterations needed to reach convergence for Maximum Number of Routes per OD pair for MNL and CNL models.](image)
Research Directions

- Data Collection Issues
  - Passive methods
  - Smart phones
  - Map-matching

- Choice Set Generation
  - Inclusion of behavioral variables
  - Sampling correction

- Accounting for congestion
  - Path-based algorithms
  - Choice set composition