Analysis of mode and walk-route choice in a downtown area considering heterogeneity in trip distance

Toshiyuki Yamamoto, Shinichi Takamura and Takayuki Morikawa

Nagoya Univ.
Analysis of mode and walk-route choice in a downtown area considering heterogeneity in trip distance

- Nested choice structure
- Large number of alternative routes
- Varying effect of difference in travel time among alternatives on route choice
Introduction (1)

• A critical problem with route choice models, especially in downtown areas, is the formation of the choice set

• Inappropriate choice set results in biased parameter estimation

Frejinger et al. (2009) proposed sampling of alternatives by random walk method
Introduction (2)

• Frejinger et al. (2009) investigated the effect of the random walk parameter on estimation efficiency
  – Using a hypothetical single origin-destination
  – Not clear the method provide efficient estimates with empirical data containing significant variations in trip distance
Purpose of the study

• The effect of heterogeneity in trip distance on sampled alternatives is investigated in this study

• A structured random walk parameter according to the trip distance is proposed to improve the efficiency of parameter estimates
Methodology (1)

• Random walk method (Frejinger et al. 2009)
  – At each node, a link is randomly selected based on the distance to the shortest path
  – Randomness is determined by $b_1$.
  – It includes the shortest path search when $b_1 = \infty$, and a simple random walk when $b_1 = 0$
  – Similar to stochastic assignment algorithm by Dial (1971)
Methodology (2)

Relative distance
\[ x_l = \frac{SP(v, s_d)}{C(l) + SP(w, s_d)} \]

\( SP(v, w) \): Shortest path from \( v \) to \( w \)

\( C(l) \): Cost of link \( l \)

Link weight
\[ \omega(l|b_1) = x_l b_1 \]

Probability of choosing link \( l \)
\[ q(l|E_v, b_1) = \frac{\omega(l|b_1)}{\sum_{l' \in E_v} \omega(l'|b_1)} \]

\( E_v \): Set of outgoing links from \( v \)

Probability of generating path \( i \)
\[ q(j) = \prod_{l \in \Gamma_j} q(l|E_v, b_1) \]
Methodology (3)

- Conditional probability of route choice
  - Lower level of nested logit model of mode and walk route choice
  - Identical to multinomial logit model with sampling of alternatives (Frejinger et al. 2009)

\[
P(i|C_n) = \frac{\exp\left\{\mu V_{in} + \ln(k_{in}/q(i))\right\}}{\sum_{j \in C_n} \exp\left\{\mu V_{jn} + \ln(k_{jn}/q(j))\right\}}
\]

\(k_{in}\): Number of times alternative \(i\) is generated
Methodology (4)

• Marginal probability of mode choice
  – Expanded logsum proposed by Lee & Waddell (2010)

\[
P(m) = \frac{\exp(\mu'V'_{mn})}{\exp(\mu'V'_{sn}) + \exp(\mu'V'_{wn})} \quad \text{for} \quad m = s, w
\]

\[
V'_{wn} = \frac{1}{\mu} \ln \left\{ \sum_{j \in C'_n} \left\{ k_{jn} / q(j) \exp(\mu V'_{jn}) \right\} \right\}
\]

Expansion of logsum
Methodology (5)

• Correlation among routes
  – Expanded path-size (Frejinger et al. 2009)

\[
EPS_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \sum_{j \in C_n} \frac{1}{\delta_{aj} \Phi_{jn}}, \quad \Phi_{jn} = \begin{cases} 
1 & \text{if } \delta_{jc} = 1 \text{ or } q(j)R_n \geq 1 \\
\frac{1}{q(j)R_n} & \text{otherwise}.
\end{cases}
\]

• Heteroscedasticity in route choice
  – Heteroscedasticity based on trip distance (e.g. Gliebe et al. 1999, Morikawa & Miwa 2006)

\[
\mu_n = \mu_0 d_n^\gamma
\]
Data

• Person trip survey data at Nagoya, Japan in 2008
• Mobile phone with GPS functions to track trajectories traveling within the city
• 76 subjects and 4 weeks of travel data
Survey area

Nagoya sta.  Downtown

3.55 km  4.46 km
Road network
Sample distribution of trip distance

- Walk
- Car
- Subway

Distance (km)

Cases

0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0
Number of alternatives by the trip distance

$b_1 = 30$
Number of alternatives by the trip distance

Higher number of alternatives gives more efficient parameter estimates

$b_1 = 10$
Length of alternative by the shortest path length

\[ b_1 = 10 \]
Length of alternative by the shortest path length

$b_1 = 30$  
Shorter alternatives give more efficient parameter estimates
Our proposal

• Structured random walk parameter

\[ b_1 = b_3 + b_4 SP(s_o, s_d) \]

- Stronger inclination to shortest path for longer trip distance
- More randomness for shorter trip distance
Number of alternatives by the trip distance

$\text{Number of alternatives}$

$\text{Shortest path (km)}$

$b_1 = 30$
Number of alternatives by the trip distance

\[ b_1 = 10 + 2d_n \]
Length of alternative by the shortest path length

\[ b_1 = 10 \]
Length of alternative by the shortest path length

\[ b_1 = 10 + 2d_n \]
## Route choice model (N = 91)

<table>
<thead>
<tr>
<th>Random walk parameter</th>
<th>Structured $b_1 = 10 + 2d_n$</th>
<th>Constant $b_1 = 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance (100 m)</td>
<td>-5.89</td>
<td>-6.14</td>
</tr>
<tr>
<td>Street with department stores for the elderly (100 m)</td>
<td>7.34</td>
<td>8.76</td>
</tr>
<tr>
<td>Street with restaurants on holidays (100 m)</td>
<td>4.61</td>
<td>3.37</td>
</tr>
<tr>
<td>Street without stores (100 m)</td>
<td>1.58</td>
<td>1.55</td>
</tr>
<tr>
<td>InEPS</td>
<td>0.54</td>
<td>0.38</td>
</tr>
<tr>
<td>Heteroscedasticity of scale parameter ($\gamma$)</td>
<td>-0.56</td>
<td>-0.59</td>
</tr>
</tbody>
</table>
## Route choice model (N = 91)

<table>
<thead>
<tr>
<th>Random walk parameter</th>
<th>Structured $b_1 = 10 + 2d_n$</th>
<th>Constant $b_1 = 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>s.e.</td>
<td>s.e.</td>
</tr>
<tr>
<td>Distance (100 m)</td>
<td>1.37</td>
<td>3.83</td>
</tr>
<tr>
<td>Street with department stores for the elderly (100 m)</td>
<td>1.88</td>
<td>5.42</td>
</tr>
<tr>
<td>Street with restaurants on holidays (100 m)</td>
<td>1.78</td>
<td>2.35</td>
</tr>
<tr>
<td>Street without stores (100 m)</td>
<td>0.66</td>
<td>1.20</td>
</tr>
<tr>
<td>lnEPS</td>
<td>0.14</td>
<td>0.22</td>
</tr>
<tr>
<td>Heteroscedasticity of scale parameter ($\gamma$)</td>
<td>0.24</td>
<td>0.37</td>
</tr>
</tbody>
</table>
Route choice model (N = 91)

<table>
<thead>
<tr>
<th>Random walk parameter</th>
<th>Structured $b_1 = 10 + 2d_n$</th>
<th>Constant $b_1 = 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance (100 m)</td>
<td>t-stat.</td>
<td>t-stat.</td>
</tr>
<tr>
<td>Street with department stores for the elderly (100 m)</td>
<td>-4.30</td>
<td>-1.60</td>
</tr>
<tr>
<td>Street with restaurants on holidays (100 m)</td>
<td>3.91</td>
<td>1.62</td>
</tr>
<tr>
<td>Street without stores (100 m)</td>
<td>2.60</td>
<td>1.43</td>
</tr>
<tr>
<td>lnEPS</td>
<td>3.93</td>
<td>1.71</td>
</tr>
<tr>
<td>Heteroscedasticity of scale parameter ($\gamma$)</td>
<td>-2.37</td>
<td>-1.62</td>
</tr>
</tbody>
</table>
## Mode & walk route choice (N = 107)

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>t-stat.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Travel time (10 min.)</td>
<td>-0.93</td>
<td>-2.63</td>
</tr>
<tr>
<td>Waiting time for subway (10 min.)</td>
<td>-3.39</td>
<td>-2.42</td>
</tr>
<tr>
<td>Subway constant</td>
<td>-3.98</td>
<td>-3.07</td>
</tr>
<tr>
<td>Street with department stores for elderly (km)</td>
<td>1.49</td>
<td>2.50</td>
</tr>
<tr>
<td>Street with restaurants on holidays (km)</td>
<td>0.96</td>
<td>2.55</td>
</tr>
<tr>
<td>Street without stores (km)</td>
<td>0.35</td>
<td>2.10</td>
</tr>
<tr>
<td>lnEPS</td>
<td>1.38</td>
<td>2.90</td>
</tr>
<tr>
<td>Scale parameter ((1/\mu_0))</td>
<td>0.03</td>
<td>2.38</td>
</tr>
<tr>
<td>Heteroscedasticity of scale parameter ((\gamma))</td>
<td>-0.49</td>
<td>-2.85</td>
</tr>
</tbody>
</table>
Mode & walk route choice (N = 107)

<table>
<thead>
<tr>
<th>Trip distance</th>
<th>0.5km</th>
<th>1.0km</th>
<th>1.5km</th>
<th>2.0km</th>
<th>2.5km</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1/\mu$</td>
<td>0.019</td>
<td>0.026</td>
<td>0.032</td>
<td>0.037</td>
<td>0.041</td>
</tr>
</tbody>
</table>

The utility at the route choice level does not have a big effect on the mode choice.
Empirical findings

- Shorter routes are preferred
- Older pedestrians prefer main shopping streets with department stores
- Streets with restaurants are preferred on holidays (partly because more trips on weekdays are undertaken after 5 pm)
- Overlapping of paths significantly causes correlation of utility among routes
Conclusion

• Structured random walk parameter improves the efficiency of the parameter estimates with empirical data containing trips of various distance