Modeling shopping behavior in a neighborhood with endogenous representation of retail attractiveness

Hiroshima University
Makoto Chikaraishi, Akimasa Fujiwara, Fumihito Nishikawa, Hajime Seya, Junyi Zhang
Motivation (1/2)

• Our ultimate goal:
  – Describing collective decisions through modeling social interactions

• Some examples in our mind:
  – Why some communities do better collective decisions (e.g., maintaining small market in the neighborhood), while some others do not?
  – Why some (most?) slums are in a kind of lock-in situation, while some others are not?

• These questions are not very new (e.g., Schelling, 1978), but we could do a better empirical analysis on this topic
  – Progress on modeling social interactions
  – Advancement of observation/estimation methods

Our focus: shopping behavior in a neighborhood
Motivation (2/2)

- Many newtowns were built around 30-40 years ago in Japan
- Rapid aging population has been observed in these areas
- As decreasing mobility with aging, maintaining necessary facilities for daily life within the neighborhood becomes more important

- 59% of residents do shopping within NT
- 13% of residents do shopping within NT
- 27% of residents do shopping within NT

※ One-day travel diary survey in Hiroshima in 2008
Modeling social interactions

- Social interactions in choice modeling and its macro behavior (i.e., collective decision)

When the degree of interactions are higher than a certain level under non-cooperative scheme

Individual behavior (individual decision)

Macro behavior (Collective decision)
Classification of social interactions

Social interactions

Market interactions
Differences in final consequences are understood as differences in initial conditions (Harris and Wilson, 1978; Fujita et al., 2001; Guevara, 2010)

Market & non-market interactions
Differences in final consequences are understood as collective decisions (e.g., Putnam, 1993; Falk and Kilpatrick, 2000)

Economical point of view
Sociological point of view
Classification of social interactions

Social interactions

Market interactions

Market & non-market interactions

Local interactions

 Mostly cooperative scheme, usually not causing multiple equilibria

Global interactions

 Mostly non-cooperative scheme, potentially causing multiple equilibria

We could have in-between models, but our current knowledge may be largely restricted partly due to observation issues of social networks/interactions

\[ y_i = f(z_i, x_i, y_{-i}, x_{-i}, w_{i,-i}) \]

Environment factors

Individual attributes

Others’ behavior/attributes

Relations with others
Classification of social interactions

- Social interactions
  - Market interactions
    - Market & non-market interactions
      - Local interactions
        - Interactions through “average” behavior
          Individual decisions are affected by the “average” behavior of social group (e.g., Brock&Durlauf, 2001; Fukuda&Morichi, 2007; Walker et al., 2011)
      - Global interactions
        - Interactions through “aggregate” behavior
          Individual decisions are affected by the “aggregate” behavior of social group (e.g., Ueda, 1991)

Modeling interactions through “aggregate” behavior
Research purpose

• Developing a shopping destination choice model with social interactions
  – Modeling market/non-market interactions through “Aggregate behavior”
    • Similar with Brock and Durlauf’s (2001) model
  – Confirming in what condition multiple equilibria happen

• Implementing empirical analysis
  – 10 newtowns in Hiroshima, Japan
    • Distinguishing endogenous effects from contextual/correlated effects, by using a modified Rust’s (1987) Nested Fixed Point algorithm (binary logit \(\rightarrow\) a mixed binary logit)
    • Empirically confirm whether multiple equilibria exist or not
Model structure

Shopping destination choice model

Individual $i$'s choice probability of destination $j$

$P_{ij} = \frac{\exp(V_{ij})}{\sum_{j'} \exp(V_{ij'})}$

Retail attractiveness at destination $j$

$V_{ij} = \gamma S_j + \beta x_{ij} + \alpha_j$

Explanatory variables

Aggregation

Total shopping demand at destination $j$

$D_j = D \cdot P_j$

Total shopping trips

Retail attractiveness

$S_j = f(D_j)$

Feedback to utility

A fixed point problem:

$D_j = \sum_i TR_i \frac{\exp(\gamma S_j + \beta x_{ij} + \alpha_j)}{\sum_{j'} \exp(\gamma S_{j'} + \beta x_{ij'} + \alpha_{j'})}$

※ Similar with Ueda (1991)
Some assumptions

(a) There are two alternatives → inside and outside of the residential newtown

(b) Only newtown residents use shopping facilities in the newtown (assuming ③ is zero)

(c) Total shopping demand (number of trips) is fixed

(d) The relationship between shopping demand and retail attractiveness follow a linear function

\[ S_j = kD_j \ (k \geq 0) \]
Self-consistent equilibrium (1/2)

With the above assumptions, equilibrium properties are very similar with Brock and Durlauf (2001), in which the basic mathematical idea is the same with the mean field approximation of Ising model in Physics.

\[ P_{im1} = \frac{e^{\gamma kD_{m1} + \beta x_{im1} + \alpha_1}}{e^{\gamma kD_{m1} + \beta x_{im1} + \alpha_1} + e^{\gamma k(D-D_{m1}) + \beta x_{im2}}} \]

Assuming that \( \omega_i \) represents the choice results of shopping destination where \( \omega_i = 1 \) when inside newtown is chosen and \( \omega_i = -1 \) when outside newtown is chosen (i.e., \( \omega_{im} = 2P_{im1} - 1 \)). Under this specification, expectation of \( \omega_i \) is:

\[ E[\omega_{im}] = \frac{e^{h_{im} + \gamma kD_{m1}}}{e^{h_{im} + \gamma kD_{m1}} + e^{-(h_{im} + \gamma kD_{m1})}} - \frac{e^{-(h_{im} + \gamma kD_{m1})}}{e^{h_{im} + \gamma kD_{m1}} + e^{-(h_{im} + \gamma kD_{m1})}} \]

\[ = \tanh \left( \frac{(\beta x_{im1} + \alpha_1) - (\gamma kD + \beta x_{im2})}{2} + \gamma kD_{1m} \right) \]

\[ h_{im} = \frac{(\beta x_{im1} + \alpha_1) - (\gamma kD + \beta x_{im2})}{2} \]
Self-consistent equilibrium (2/2)

Based on the assumptions (b) “Only newtown residents use shopping facilities in the newtown”, and (c)” Total shopping demand is fixed”,

\[ D_{m1} = TR_{m1} \bar{p}_{m1} = 1/2 TR_{m1} + 1/2 TR_{m1} E[\omega_i] \]

\[ \rightarrow \ E[\omega_{im}] = \tanh \left( \frac{(\beta(x_{im1} - x_{im2}) + \alpha_1 - \gamma kD + \gamma kTR_{m1}}{2} + \frac{\gamma}{2} kTR_{m1} E[\omega_{im}] \right) \]

Since \((\gamma \text{ and } k), \text{ and } (\alpha_1 \text{ and } \gamma kD)\) are cannot be distinguished,

\[ \gamma' = \gamma k \]

\[ \alpha' = \alpha_1 - \gamma' D \]

are assumed, and finally we obtain:

\[ \omega^*_m = \tanh \left( \frac{(\beta(x_{im1} - x_{im2}) + \alpha' + \gamma' TR_{m1}}{2} + \frac{\gamma'}{2} TR_{m1} \omega^*_m \right) \]
Behavior of equilibria

Hypothetical settings: $x_{im1}, x_{im2}$: Generalized travel cost ($x_{im2}$ is fixed as 1200 JPY), $\beta = -0.1, \alpha' = -3.0, \gamma' = 0.9$. Average number of shopping trips: 3 (i.e., total demand in NT: $TR_{m1} = \text{pop. in NT} \times 3$) Note that ⬤: stable equilibria, ○: unstable equilibrium

<table>
<thead>
<tr>
<th>Pop. in NT = 6,000</th>
<th>Travel cost in NT:100JPY</th>
<th>Travel cost in NT:300JPY</th>
<th>Travel cost in NT:600JPY</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
<td><img src="image3" alt="Graph" /></td>
<td></td>
</tr>
<tr>
<td><img src="image4" alt="Graph" /></td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
<td></td>
</tr>
<tr>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
<td><img src="image9" alt="Graph" /></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pop. in NT = 9,000</th>
<th>Travel cost in NT:100JPY</th>
<th>Travel cost in NT:300JPY</th>
<th>Travel cost in NT:600JPY</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image10" alt="Graph" /></td>
<td><img src="image11" alt="Graph" /></td>
<td><img src="image12" alt="Graph" /></td>
<td></td>
</tr>
<tr>
<td><img src="image13" alt="Graph" /></td>
<td><img src="image14" alt="Graph" /></td>
<td><img src="image15" alt="Graph" /></td>
<td></td>
</tr>
<tr>
<td><img src="image16" alt="Graph" /></td>
<td><img src="image17" alt="Graph" /></td>
<td><img src="image18" alt="Graph" /></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pop. in NT = 12,000</th>
<th>Travel cost in NT:100JPY</th>
<th>Travel cost in NT:300JPY</th>
<th>Travel cost in NT:600JPY</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image19" alt="Graph" /></td>
<td><img src="image20" alt="Graph" /></td>
<td><img src="image21" alt="Graph" /></td>
<td></td>
</tr>
<tr>
<td><img src="image22" alt="Graph" /></td>
<td><img src="image23" alt="Graph" /></td>
<td><img src="image24" alt="Graph" /></td>
<td></td>
</tr>
<tr>
<td><img src="image25" alt="Graph" /></td>
<td><img src="image26" alt="Graph" /></td>
<td><img src="image27" alt="Graph" /></td>
<td></td>
</tr>
</tbody>
</table>
Technical issues in model estimation

• Distinguishing endogenous effects from contextual/correlated effects (e.g., Manski, 1993)
  – BLP approach (Berry et al., 1995)
  – Control function approach
  – Full maximum likelihood approach

• Introducing the endogenous variable as an exogenous variable may not be self-consistent
  – Nested Fixed Point (NFXP) algorithm (Rust, 1987)
  – A bit more efficient NFXP algorithm (Aguirregabiria and Mira, 2002)
  – Bayesian estimation (Imai et al., 2009)
  – MPEC (Su and Judd, 2012)
Model estimation method

- NFXP (Nested Fixed Point) algorithm (Rust, 1987)

1. Given the observed choice probability \( P_{ij}^{(0)} \), estimate a set of parameters \( \theta^{(1)} \)

2. Repeat the following steps until convergence
   (a) Inner fixed point algorithm:
      Calculate choice probability \( P_{ij}^{(k)} \) with parameters \( \theta^{(k)} \)
   (b) Outer hill climbing algorithm:
      Given \( P_{ij}^{(k)} \), estimate parameters \( \theta^{(k+1)} \)

When introducing random parameters in the 2(b) step, individual parameters (Train, 2003) are calculated and put in the 2(a) step
Empirical analysis

• One-day travel diary survey in Hiroshima in 2008
  – Sample size in the target area: 5,620 (1,453 shopping trips)

• Study area
  ※ Because the available spatial information is limited, surrounding areas are also included in some Newtowns.

<table>
<thead>
<tr>
<th>List of newtown</th>
<th>Population in NT</th>
</tr>
</thead>
<tbody>
<tr>
<td>あさひが丘</td>
<td>8155</td>
</tr>
<tr>
<td>美鈴が丘</td>
<td>10086</td>
</tr>
<tr>
<td>五月が丘</td>
<td>6925</td>
</tr>
<tr>
<td>藤の木</td>
<td>4948</td>
</tr>
<tr>
<td>向洋</td>
<td>14847</td>
</tr>
<tr>
<td>井口台・井口鈴が台</td>
<td>22544</td>
</tr>
<tr>
<td>山田新町・鈴が峰</td>
<td>11886</td>
</tr>
<tr>
<td>矢野</td>
<td>18617</td>
</tr>
<tr>
<td>高陽</td>
<td>39423</td>
</tr>
<tr>
<td>高陽第一</td>
<td>20437</td>
</tr>
</tbody>
</table>
## Basic statistics

<table>
<thead>
<tr>
<th>Sample size (persons)</th>
<th>Origin = residential area</th>
<th></th>
<th>Origin = non-residential area</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Shopping trips</td>
<td>Freq. per day per person</td>
<td>Share of inside newtown (%)</td>
<td>Shopping trips</td>
</tr>
<tr>
<td><strong>あさひが丘</strong></td>
<td>312</td>
<td>48</td>
<td>0.154</td>
<td>39.58</td>
</tr>
<tr>
<td><strong>美鈴が丘</strong></td>
<td>447</td>
<td>56</td>
<td>0.125</td>
<td>23.21</td>
</tr>
<tr>
<td><strong>五月が丘</strong></td>
<td>259</td>
<td>29</td>
<td>0.112</td>
<td>34.48</td>
</tr>
<tr>
<td><strong>藤の木</strong></td>
<td>143</td>
<td>11</td>
<td>0.077</td>
<td>18.18</td>
</tr>
<tr>
<td><strong>向洋</strong></td>
<td>485</td>
<td>64</td>
<td>0.132</td>
<td>28.13</td>
</tr>
<tr>
<td><strong>井口台・井口鈴が台</strong></td>
<td>728</td>
<td>102</td>
<td>0.140</td>
<td>50.98</td>
</tr>
<tr>
<td><strong>山田新町・鈴が峰</strong></td>
<td>408</td>
<td>55</td>
<td>0.135</td>
<td>21.82</td>
</tr>
<tr>
<td><strong>矢野</strong></td>
<td>656</td>
<td>80</td>
<td>0.122</td>
<td>28.75</td>
</tr>
<tr>
<td><strong>高陽</strong></td>
<td>1394</td>
<td>274</td>
<td>0.197</td>
<td>71.17</td>
</tr>
<tr>
<td><strong>高陽第一</strong></td>
<td>788</td>
<td>150</td>
<td>0.190</td>
<td>52.67</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>5620</td>
<td>869</td>
<td>0.155</td>
<td>48.68</td>
</tr>
</tbody>
</table>
Travel cost (accessibility) calculation

1. Estimating travel mode choice
2. Calculating logsum

- For inside of newtown, LOS variables are prepared for each 50m grid cell, and then take a population-weighted average (since we do not have detailed information of residential location)
- For unchosen travel modes’ LOS, we put an average value of others who live in the same newtown

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant (Car)</td>
<td>2.239</td>
<td>6.73</td>
</tr>
<tr>
<td>Constant (Bicycle)</td>
<td>-1.204</td>
<td>-3.23</td>
</tr>
<tr>
<td>Constant (Walk)</td>
<td>0.540</td>
<td>1.78</td>
</tr>
<tr>
<td>Travel time (100 min)</td>
<td>-3.255</td>
<td>-2.28</td>
</tr>
<tr>
<td>Travel cost (1000 JPY)</td>
<td>-2.463</td>
<td>-1.78</td>
</tr>
<tr>
<td>Age&gt;=75 dummy (car)</td>
<td>-0.071</td>
<td>-0.17</td>
</tr>
<tr>
<td>Initial log-likelihood</td>
<td>-1053.2</td>
<td></td>
</tr>
<tr>
<td>Final log-likelihood</td>
<td>-539.0</td>
<td></td>
</tr>
<tr>
<td>Adjusted rho-square</td>
<td>0.483</td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>864</td>
<td></td>
</tr>
</tbody>
</table>
Estimation results of shopping destination choice model

<table>
<thead>
<tr>
<th></th>
<th>No social interactions</th>
<th>With endogenous effects</th>
<th>With contextual/correlated effects</th>
<th>With all effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parameter</td>
<td>t-value</td>
<td>Parameter</td>
<td>t-value</td>
</tr>
<tr>
<td>Constant</td>
<td>4.585</td>
<td>8.22**</td>
<td>1.068</td>
<td>1.03</td>
</tr>
<tr>
<td>Travel cost (logsum)</td>
<td>-0.019</td>
<td>-1.98*</td>
<td>-0.018</td>
<td>-1.81+</td>
</tr>
<tr>
<td>Age &lt;=30</td>
<td>-0.600</td>
<td>-1.78+</td>
<td>-0.598</td>
<td>-1.75+</td>
</tr>
<tr>
<td>Age &gt;=65</td>
<td>-0.035</td>
<td>-0.20</td>
<td>0.057</td>
<td>0.31</td>
</tr>
<tr>
<td>Male</td>
<td>-0.202</td>
<td>-1.06</td>
<td>-0.271</td>
<td>-1.39</td>
</tr>
<tr>
<td>Housewife</td>
<td>-0.293</td>
<td>-1.47</td>
<td>-0.302</td>
<td>-1.50</td>
</tr>
<tr>
<td>Non-worker</td>
<td>-0.249</td>
<td>-1.12</td>
<td>-0.301</td>
<td>-1.34</td>
</tr>
<tr>
<td>ln(# of shops)</td>
<td>0.960</td>
<td>9.25**</td>
<td>0.356</td>
<td>1.96*</td>
</tr>
<tr>
<td>Retail attractiveness</td>
<td></td>
<td></td>
<td>1.014</td>
<td>3.93**</td>
</tr>
<tr>
<td>Contextual/correlated</td>
<td></td>
<td></td>
<td>0.095 [7.96]</td>
<td>3.2.E-15 [0]</td>
</tr>
<tr>
<td>Random effect</td>
<td></td>
<td></td>
<td>864</td>
<td>864</td>
</tr>
<tr>
<td>Sample size</td>
<td></td>
<td></td>
<td>864</td>
<td>864</td>
</tr>
<tr>
<td>Initial log-likelihood</td>
<td>-598.88</td>
<td>-598.88</td>
<td>-598.88</td>
<td>-598.88</td>
</tr>
<tr>
<td>Final log-likelihood</td>
<td>-546.40</td>
<td>-538.51</td>
<td>-542.42</td>
<td>-538.51</td>
</tr>
<tr>
<td>AIC</td>
<td>1108.80</td>
<td>1097.03</td>
<td>1102.84</td>
<td>1099.03</td>
</tr>
</tbody>
</table>

※ [] indicates the result of χ² test

Based on AIC comparison, the model only with endogenous effects are selected.
No multiple equilibria

Only if when \( B = \frac{\beta T R_1}{2} > 1 \), there is possibility to have multiple equilibria

<table>
<thead>
<tr>
<th></th>
<th>NT pop.</th>
<th>Freq. per day per person</th>
<th>( B )</th>
</tr>
</thead>
<tbody>
<tr>
<td>あさひが丘</td>
<td>8,155</td>
<td>0.269</td>
<td>0.046</td>
</tr>
<tr>
<td>美鈴が丘</td>
<td>10,086</td>
<td>0.268</td>
<td>0.035</td>
</tr>
<tr>
<td>五月が丘</td>
<td>6,925</td>
<td>0.201</td>
<td>0.026</td>
</tr>
<tr>
<td>藤の木</td>
<td>4,948</td>
<td>0.231</td>
<td>0.011</td>
</tr>
<tr>
<td>向洋</td>
<td>14,847</td>
<td>0.225</td>
<td>0.049</td>
</tr>
<tr>
<td>井口台・井口鈴が台</td>
<td>22,544</td>
<td>0.249</td>
<td>0.165</td>
</tr>
<tr>
<td>山田新町・鈴が峰</td>
<td>11,886</td>
<td>0.270</td>
<td>0.042</td>
</tr>
<tr>
<td>矢野</td>
<td>18,617</td>
<td>0.227</td>
<td>0.069</td>
</tr>
<tr>
<td>高陽</td>
<td>39,423</td>
<td>0.285</td>
<td>0.470</td>
</tr>
<tr>
<td>高陽第一</td>
<td>20,437</td>
<td>0.277</td>
<td>0.185</td>
</tr>
</tbody>
</table>

The empirical results indicate that there are no newtowns which potentially have multiple equilibria (though the significant social interactions exist)
Conclusions

• **Model development**
  – Modeling shopping destination choice with social interactions
    • Representing retail attractiveness as endogenous
      ➢ Brock and Durlauf (2001) → “average” behavior of others
      ➢ This study → “aggregate” behavior of others
    • Model estimation
      – NFXP algorithm with a distinction between endogenous and contextual/correlated effects

• **Empirical analysis**
  – 10 newtowns in Hiroshima city
    • Confirming the existence of endogenous effects
    • But, the degree of endogenous effects are not big enough to cause multiple equilibria
      – The situation is not like “collective decisions” through equilibrium selection
Future works

• Separating market interactions and non-market interactions
  – Interpretation of interactions, and possible policy discussions

• Modeling residents’ shopping behavior from outside of newtown
  – Modeling trip chaining

• Modeling frequency of shopping trips and money they spent
  – Elderly may less participate in shopping activities

• Applying the model to evaluate the impacts of improving transportation system in newtown
Future tasks

For each 50m grid cell

Pop data by age group

Sharing participation

The current travel cost
Reducing travel cost within newtown
Recalculate travel cost

Trip freq.
Mode choice

Individual attributes, etc.

LOS of PM sharing within a newtown
Location of station, price, reservation system, # of vehicles, etc.

PM sharing participation model
Non-market interactions
Market interactions

Sharing participation

LOS for Public transport (Navitime)
LOS for road transport (road network data)

The proposed model

Service level of shops
Variety, open/close time, price, etc.

Individual attributes, etc.

Shopping destination
Shopping demand
Non-market interactions
Market interactions

Non-participants
Participants

Profit

Donation

Profit
Initial trial
References

- 上田孝行 (1991) 交通改善による生活機会の増大が人口移動に及ぼす影響のモデル分析. 土木計画学研究・論文集 No. 9, 237-244.