Understanding the Influence of Zonal Attributes on the Relationship between Public Transit and Ride Hailing Services

Patrick Loa, Sanjana Hossain, Khandker Nurul Habib
Outline

- Problem description
- Study area
- Sources of data
- Empirical model
- Results
- Conclusions and Future Work
Problem Description
Background

- The use of ride hailing continues to grow in North America

(City of Toronto – Big Data Innovation Team, 2019)
Background

- Transit ridership has stagnated and fallen below projections in some cities

![Average Weekday Transit Ridership](chart)

(City of Toronto, 2018)
Key Questions

Do zonal attributes influence the use of ride hailing and public transit?  
What is the extent of this influence?
Study Area
3,702 Dissemination Areas (DAs)
The City of Toronto

Public Transit:
- Subway, bus, streetcar, paratransit, commuter rail
- 521 million trips in 2018 (City of Toronto, 2018)

Ride Hailing:
- Uber: Since 2016
- Lyft: Since 2017
Sources of Data
Travel Information

- Number of trips that begin in each dissemination area (DA)

- Public Transit Trips → obtained from regional household travel survey
  Ride Hailing Trips → obtained from Uber, via the City of Toronto
Public Transit Trips

- **Source → Transportation Tomorrow Survey (TTS)**
- 5% sample of households in the Greater Golden horseshoe, every 5 years since 1986
- 2.1 million transit trips
- **2016 TTS: Sept 7 – Dec 16**
Ride Hailing Trips

- Sept 2016 – Sept 2018
- All pick-ups and drop-offs mapped to closest intersections

(City of Toronto – Big Data Innovation Team, 2019)
Zonal Attributes

- Land Use Classifications
- Enhanced Points-of-Interest
- 2016 Canadian Census
- GTFS Data
GTFS Data

Based on 400 and 800 m buffer around DA centroid

- Number of stops
- Length of bus, streetcar, subway lines
Empirical Model
The Bivariate Ordered Probit Model

- Model formulation:

\[
\begin{align*}
    y_{i,1}^* &= \beta_1'X_{i,1} + \varepsilon_{i,1} \rightarrow y_{i,1} = j \text{ if } \mu_{j-1} < y_{i,1}^* < \mu_j; j \in [0,J_1] \\
    y_{i,2}^* &= \beta_2'X_{i,2} + \varepsilon_{i,2} \rightarrow y_{i,2} = j \text{ if } \delta_{j-1} < y_{i,2}^* < \delta_j; j \in [0,J_2] \\
    \varepsilon_{i,1}, \varepsilon_{i,2} &\sim N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right]
\end{align*}
\]

- Response variables:
  - Level of transit trips generated by each DA
  - Level of ride hailing trips generated by each DA
Key Issues

Defining a “typical day” for ride hailing demand

Transforming trip generation values from continuous to ordered
Defining a “Typical Day”

- Weighted average of daily ride hailing trips based on TTS responses

\[
RH_j = \sum_{i=1}^{N} P(i) \times D_{ij}
\]

# of RH trips generated by zone j
Defining a “Typical Day”

- Weighted average of daily ride hailing trips based on TTS responses

\[ RH_j = \sum_{i=1}^{N} P(i) \times D_{ij} \]

- Number of ride hailing trips in zone j on day i
- % of survey responses on day i
Creating Ordered Responses

- DAs were classified based on the number of transit and ride hailing trips generated
- **Classification Strategies:**
  - Median (Below/ Above)
  - Tertiles (Low/ Medium/ High)
  - Quartiles (First/Second/Third/Fourth)
- Aimed for ~100 observations per pair of categories
# Models

<table>
<thead>
<tr>
<th>Model #</th>
<th>Ride Hailing Classification</th>
<th>Public Transit Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Median</td>
<td>Tertile</td>
</tr>
<tr>
<td>2</td>
<td>Tertile</td>
<td>Median</td>
</tr>
<tr>
<td>3</td>
<td>Tertile</td>
<td>Tertile</td>
</tr>
<tr>
<td>4</td>
<td>Quartile</td>
<td>Quartile</td>
</tr>
</tbody>
</table>
Results
Explanatory Variables

- Transit Attributes
  - Based on buffers around zonal centroids
- Land Use Density Indicators
  \[ LUDI(k) = \ln \left( \frac{\text{Number of buildings of type } k}{\text{Area } [km^2]} + 1 \right) \]
- Zonal Attributes
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PTC</td>
<td>Transit</td>
<td>PTC</td>
<td>Transit</td>
</tr>
<tr>
<td>bus_400</td>
<td>0.0137</td>
<td>0.0496%</td>
<td>0.0129</td>
<td>0.0452*</td>
</tr>
<tr>
<td>streetcar_400</td>
<td>0.4983 #</td>
<td>0.2280 #</td>
<td>0.5266 #</td>
<td>0.1729 #</td>
</tr>
<tr>
<td>subway_800</td>
<td>0.2962 #</td>
<td>0.3707 #</td>
<td>0.3100 #</td>
<td>0.3242 #</td>
</tr>
<tr>
<td>stops_400</td>
<td>0.0271 #</td>
<td>0.0279 #</td>
<td>0.0262 #</td>
<td>0.0304 #</td>
</tr>
<tr>
<td>Transit Attributes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln_acco_food_density</td>
<td>0.1876 #</td>
<td>0.0686 ^</td>
<td>0.1544 #</td>
<td>0.1182*</td>
</tr>
<tr>
<td>ln_art_ent_rec_density</td>
<td>0.3653 #</td>
<td>0.2075 #</td>
<td>0.2892 #</td>
<td>0.2099 #</td>
</tr>
<tr>
<td>ln_edu_density</td>
<td>0.0906 *</td>
<td>0.2114 #</td>
<td>0.0995 %</td>
<td>0.1708 #</td>
</tr>
<tr>
<td>ln_health_density</td>
<td>0.1822 #</td>
<td>0.2544 #</td>
<td>0.1753 #</td>
<td>0.2318 #</td>
</tr>
<tr>
<td>ln_priv_dwelling_density</td>
<td>0.2714 #</td>
<td>0.2526 #</td>
<td>0.2621 #</td>
<td>0.3340 #</td>
</tr>
<tr>
<td>ln_retail_density</td>
<td>0.1569 #</td>
<td>0.0983 %</td>
<td>0.1977 #</td>
<td>0.0668 #</td>
</tr>
<tr>
<td>Land Use Density Indicators</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln_education_density</td>
<td>0.0012 #</td>
<td>0.0038 #</td>
<td>0.0011 #</td>
<td>0.0041 #</td>
</tr>
<tr>
<td>area_sq_km</td>
<td>1.2282 #</td>
<td>1.9333 #</td>
<td>1.2655 #</td>
<td>1.8684 #</td>
</tr>
<tr>
<td>avg_income ($1000's)</td>
<td>0.0017 #</td>
<td>-0.0008 #</td>
<td>0.0015 #</td>
<td>-0.0010 #</td>
</tr>
<tr>
<td>Zonal Attributes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>under_25</td>
<td>0.2726 #</td>
<td>0.2399 #</td>
<td>0.2445 #</td>
<td>0.2589 #</td>
</tr>
<tr>
<td>avg_income ($1000's)</td>
<td>0.0017 #</td>
<td>-0.0008 #</td>
<td>0.0015 #</td>
<td>-0.0010 #</td>
</tr>
<tr>
<td>avg_income ($1000's)</td>
<td>0.0017 #</td>
<td>-0.0008 #</td>
<td>0.0015 #</td>
<td>-0.0010 #</td>
</tr>
<tr>
<td>Correlation Parameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold Parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2.441 #</td>
<td>2.057 #</td>
<td>1.797 #</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>-3.276 #</td>
<td>2.959 #</td>
<td>-2.9552 #</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-3.276 #</td>
<td>-2.959 #</td>
<td>-2.9552 #</td>
</tr>
<tr>
<td>Model Fit Statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Observations</td>
<td>3473</td>
<td>3473</td>
<td>3473</td>
<td>3473</td>
</tr>
<tr>
<td>McFadden's Pseudo R²</td>
<td>0.2510</td>
<td>0.2388</td>
<td>0.2296</td>
<td>0.2022</td>
</tr>
<tr>
<td>Loglikelihood of Final Model</td>
<td>-4661</td>
<td>-4737</td>
<td>-5879</td>
<td>-7682</td>
</tr>
</tbody>
</table>

Significance Levels: #: >99.9%, %: 99.9%, *: 99%, ^: 95%
Key Findings

- Transit attributes have a positive impact on ride hailing and transit use.
- Land use density indicators also have a positive impact on ride hailing and transit use.
- Higher zonal income is associated with higher ride hailing use and lower transit use.
Model Performance

- Threshold parameters are all significant
- Correlation: 0.23 - 0.28
  - Correlation is significant in all models
- McFadden’s Pseudo $R^2$: 0.20 - 0.25
Partial Effects

- Based on Greene and Hensher (2010)

\[
\frac{\partial P(y = j | \bar{x})}{\partial x_k} = [f(\mu_{j-1} - \beta' \bar{x}) - f(\mu_j - \beta' \bar{x})] \beta_{x_k}
\]
Transit Attributes

Model 1

Model 2

Model 3

Model 4
Land Use Density Indicators

Model 1

Model 2

Model 3

Model 4
Zonal Attributes
Key Findings

The length of streetcar and subway lines are the two most influential transit attributes.

Zonal area has the greatest impact on the level of ride hailing and transit trip generation.

The density of private dwellings and commercial buildings drive transit and ride hailing use.
Conclusions and Future Work
Future Work

- Apply analysis to individual modes of transit
- Use different approaches to categorization
- Investigate the nature of the relationship between ride hailing and transit
- Determine the role of costs in the choice of the two services
Conclusions

- Higher density areas display higher levels of ride hailing and transit use
- Public transit network attributes influence both transit and ride hailing usage
- A positive correlation exists between the unobserved factors that drive ride hailing and transit use