The Importance of Peers in Student Cycling Choice: A Discrete Choice Model with Endogenous Social Interactions for the Choice of Owning a Bike by a University Student in Toronto

Khandker Nurul Habib, Ph.D., PEng
Associate Professor
Department of Civil Engineering
University of Toronto
khandker.nurulhabib@utoronto.ca

Number of Words: 6184 + 2 Figure (500 words) + 2 Tables (500 words) = 7184

The paper submitted to the Committee ADB40 for consideration presentation at the 97th Annual Meeting of Transportation Research Board, January 7-11, 2018. Washington DC
Abstract

The paper employed a latent class discrete choice model with endogenous social interaction effect model to investigate the role of gender and age-cohort-specific peer or social interaction effect on the choice of owning a bike by the university students in Toronto. For the empirical investigation, it used a dataset collected through a web-based travel survey among the students of all 4 universities with 7 campuses in Toronto. Two latent classes of bicycle owners are identified: peer-conforming and peer-indifferent. Empirical model proves that the influence of age-cohort specific peer-group social interactions is real and very strong on the choices of bicycle ownership. Gender of the students play a defining role in partitioning a student into the latent classes of bicycle users, however, female students are influenced more by the choices of fellow female students than do the male students. Also, considerable heterogeneity exists in such age-cohort specific social interaction effects. The finding of this study has clear policy relevance and calls for policy initiatives. Policy initiatives that target female students of specific age groups will have greater potential in influencing bicycle usage. Overall, the choice of owning a bicycle is influenced by various variables in complex ways. This makes the potential of social engineering type policy initiatives highly potential for success in increasing bicycle ownership among the university students.

Key Words: Bicycle ownership, gender effect on bicycle choice, social interaction effect on bicycle choice, latent class discrete choice model, discrete choice model with endogenous variable
1. Introduction

The use of bicycles as a travel mode is considered one of the most sustainable and healthiest forms of transportation, yet cycling accounts for almost a negligible modal share in North American cities. Often the lack of safe and dedicated cycling infrastructure is cited for this (Akar and Clifton 2009), though little research has explored whether there may be other user-related factors at play. A better understanding of these factors could lead to more relevant and informed approaches to further develop cycling as a realistic transportation alternative for users in major cities.

Recent studies recognize that social interaction effects (e.g., how choices of the peers influence an individual’s decisions) are very strong on the choice of owning a bike (Maness and Cirillo 2016). Choice models have often found that the “gender” variable to be strong predictor variable in terms of bicycle ownership and use (Schoner et al 2015), yet little or no research has explored the “gender” variable in concert with the social interaction effects. There is even less information on cycling behavior of the “Millennial” population who could be expected to be the next cohort to shape transportation planning with respect to cycling.

This study employed a travel diary data set to explore the joint social and gender (female, male) effects on the choice of owning a bike by university students in Toronto who represent the post-secondary millennial population in Toronto. Social effects are defined as age-cohort specific effects by different gender type (male and female) by using an advanced discrete choice modeling framework. The hypothesis that the choices of people of similar age and gender have an influence on individual’s choice is investigated for the case of owning a bicycle in Toronto. Such gender-specific age-cohort effect is captured through the proportion of people owning a bicycle in a particular gender-age cohort as an explanatory variable in the choice model of an individual of the same gender-age cohort. Given the potential for an endogeneity issue (e.g., choice of one individual is influenced by the choices of other individuals), a new econometric model formulation is also proposed that can overcome this. Results reveal that cohort effects are real and have distinctive influences on how the two different genders, in this case, made choices of owning a bike.

The paper presents the study in the following ways: the next section presents a brief literature review on the suites that investigated social interaction effects on biking. This section is followed by the sections explaining the proposed econometric model; the section discussing the key features of the data set for empirical study and the section discussing the empirical model along with possible policy relevance. The paper concludes with a summary of key findings and recommendations for future studies.
2. Literature Review

Considering the main scope of this research, a review of studies/paper that presents investigations on gender and social interaction effects on choice of using or owning a bicycle is presented in this section.

Among few published studies that focused entirely on the effects of gender on bicycle choices, Krizek et al (2005) are the notable earliest ones. They focused on bicycle use and facility preferences to investigate difference influenced by gender in the US. Their empirical investigation identified a difference between males and females in terms of use and preference of bicycling. Overall, they found the females in their study were less likely to use a bicycle than males, but specifically, women were less likely to use a bicycle for commuting than men. The takeaway lesson of this study is that women perceive attributes of bicycling facilities (lanes, parking, safety etc.) differently than men and so even in the case of similar trip lengths, women tend not to use a bicycle as much as men do.

Emond et al (2009) also focused on the gender issue in bicycle choices in the US. They presented an empirical investigation by using data collected through an online (on the use of bicycle) in 6 western cities in the US (4 cities in California, 1 in Colorado and 1 in Oregon). They used the discrete choice model for investigating the factors influencing the choice of biking by women. They found that men and women are influenced by different factors in terms of the choice of using a bicycle. Among many factors, the perception of safety preference to bicycle facility types showed a strong gender difference in the case of the choice of biking. Handy et al (2010) used the same dataset to estimate a discrete choice model of the choice of owning and using a bike. Among many other variables, age and gender are found to have a strong influence on the choice of owning and using a bike.

Schoner et al (2015) used regional travel survey datasets from Minneapolis to investigate gender gaps in bicycle use between 2000 and 2010. Their empirical model reveals that women are less likely to use a bicycle than males under same circumstances. They also found that the growth in the number of females using a bicycle is slower than that of the males. Their investigation alludes to the fact of greater household and socio-economic responsivity could be blamed for this. Singleton and Goddard (2016) used a statewide travel survey data set collected in Oregon to investigate gender gap in bicycling. They found that in the same household and social context females are less likely to prefer bicycle as a travel model than males. It is concluded that, perhaps, women's role and responsibilities in household maintenance activities might play the role in lower attraction to bicycle.

However, in many other studies, gender is considered as one of many covariates in investigating bicycle demand/choice behavior. For example, Sener et al (2009a) investigated bicycle route choice behavior in Texas and found that female bicycle users
have very different behavior than males. They used data collected through a web-based survey among a sample of bicyclists in Texas and found that female bicyclists have a different perception to bicycling parking availability as they prefer routes with bicycle parking facilities more than the males do. They also found that female bicyclists tend to avoid routes with steep hills. Sener et al (2009b) used the same dataset to investigate classification of bicyclists. They found that females, in general, have a lower rate of using a bicycle than the males do (regardless of trip purposes). Specifically, the females tend to have a worse perception of the same bicycle facilities than do the males and that could explain the lower rate of using bicycles by females. This study identified that both age and gender have combined effects on the choice of using a bicycle for travel. Ma et al (2014) investigated the effects of perceived environment on the choice of using a bicycle in Portland, Oregon. Their Structural Equation Model reveals that age and gender have compounded positive effect on the better perception of cycling environment.

Castillo-Manzano and Sa´nchez-Braza (2006) presented an empirical investigation on the satisfaction of using bicycle sharing system by using data from an opinion survey on bicycle satisfaction/need among university communities in Seville, Spain. Their empirical models (ordered probability model) show no significant effects of age and gender of students on perceived satisfaction of using a bicycle. Habib et al (2014) used a dataset collected through a web-based bicycle demand survey in the City of Toronto and estimated the hybrid discrete choice model. Among many other findings, this study revealed that females are less comfortable in biking in Toronto than are males. It was also found than younger females, in particular, had the lowest perception of comfortability in biking in the city. Ferna´ndez-Heredia et al (2016) also used data collected from university students in Madrid, Spain and used a structural equation modeling approach to investigate the factors that affect the perception and then their role in influencing the choice of using a bicycle. Unlike Seville study, they found the gender of the respondent in relevant in defining the choice of using a bicycle, but only when the bicycle modal share is low. In such case, females tend to have a lower preference to use a bicycle than the males largely due to the risk aversion nature of the females as well as their greater involvement in household maintenance activities.

Use of university student population for investigating bicycle choices in also common in North America. Akar and Clifton (2009) used a dataset collected from the students of a university in Maryland, US. They used the discrete choice model for investigating the effects of various personal and infrastructure attributes on the use of bicycle by the students. However, unlike in Seville, in Maryland, they found that being female has a strong negative influence on the use of the bicycle. Akar et al (2013) used a dataset collected from the students of Ohio State University and used the discrete choice model for investigating the use of bicycle by the university students. They also found that female students are, in general, less likely to prefer bicycle than do the male students. Also, they found that female
students are more positively influenced by the presence of bicycle infrastructures (e.g. bike trails, bike lanes, etc.) in choosing bicycle s travel modes than so the male students.

All of the above-mentioned studies (except one) identified females tend to have a lower preference to bicycle than males, but Fernández-Heredia et al (2016) pointed out the fact that, perhaps a lower share of female bicycle users has a role to play. Clearly, this recognizes the effects of social influence in the choice of bicycle usage by females. Maness and Cirillo (2016) is the only study that explicitly investigated the presence of social influence on bicycle choice. They used 2001 US National Household Travel Survey data to investigate the effects of social influence on the choice of owning a bike. They developed a latent class binary discrete choice model of bike ownership. They proposed a latent class model to capture two distinct (but unidentified to the necked eyes in the dataset) classes: socially-informed class and socially-indifferent class. The former class represents the tendency to follow what the people of similar characteristics do and the latter is indifferent about that. In their formulation, they used metropolitan statistical area (MSA) level average bicycle per household as the link to inform class membership choice the corresponding social influence. The idea is that the propensity of owning a bike of the socially informed class will increase with increasing average number of bike per household in the own MSC. Since MSA level household bike ownership rate is a function of the choices of individual household’s owning a bicycle, the model posed an endogeneity issue in model parameter estimation. They proposed a two-step control function approach to overcome the endogeneity issue. Nevertheless, their empirical investigation clearly reveals that the household’s choice of owning a bike is strongly influenced by the choices of similar households’ owning bikes.

The current study is inspired by Maness and Cirillo (2016) that social interaction takes place the best at a class level than individual choice level. However, it further contributes to the literature by focusing on individual cohort-specific social influence effects and by proposing one-step (full information) parameter estimation process. For the empirical investigation, this study used data from a travel survey collected from the students of all 4 universities in the City of Toronto. In terms of investigating the effects of social influence, it used age-specific gender cohorts as opposed to spatial location-based aggregated social influences. In terms of parameter estimation, instead of a two-step estimation process, this paper proposes a full-information control function approach of parameter estimation technique. The next section explains the proposed econometric modeling framework.

3. Modeling the Choice of Bike Ownership for Empirical Investigation

Modeling the choice of owning is bike is considered to be influenced by, varying degree, the same choices of the members of peer groups. So, as shown in Figure 1, two classes of student groups are assumed: ‘peer-conformable’ class and ‘peer-indifferent’ class. Both
classes are influenced by personal attributes, household attributes, transportation system attributes and land use attributes. In addition, the peer-conformable is influenced by the choices of similar peer groups, but that is not true for the peer-indifferent class. Considering that the presence/absence of classes is apparently hidden in observed data, the choice of owning a bike is influenced by the latent class membership.

Figure 1: Schematic Diagram of Bike Ownership Choice Process

Individual students derive utilities in being a member of the classes as follows:

\[ U_l = \Sigma (yz) + \Sigma_k (\delta_b \bar{y}_b)_k + \varepsilon_l \quad , \quad l = 1, 2 \]  

Here, \( U_l \) is the total utility of being a member of a latent class, \( l \). The subscript \( l \) has two values; \( C_C \) indicates the peer-conforming class and \( C_I \) indicates peer-indifferent class. \( z \) indicates a set of personal, land use and transportation variables and \( \gamma \) is corresponding covariates \( (\bar{y}_b)_k \) is the perceived proportion of students in the cohort \( k \) owning bikes and \( \delta_b \) is the corresponding coefficient \( \varepsilon_l \) is the random error component of the total utility
Considering that \( \varepsilon \) has Type I Extreme value distribution and following the random utility maximization assumption, the class membership function (probability) takes the logit formulation as follows:

\[
P[C_i] = \frac{\exp(\sum(yz) + \sum_k(\delta_b \gamma_b)k)}{1 + \exp(\sum(yz) + \sum_k(\delta_b \gamma_b)k)}
\]  

(2)

\[
P[C_i] = \frac{1}{1 + \exp(\sum(yz) + \sum_k(\delta_b \gamma_b)k)}
\]  

(3)

Conditional to the membership of a class, we consider that that choice of owning a bike is defined by random utility maximization theory.

Utility of bike owning given the class \( C_c \)

\[U_{b|C_c} = \sum(\beta x)_{b|C_c} + \varepsilon_{b|C_c}, \quad l = 1, 2\]  

(4)

Utility of bike owning given the class \( C_i \)

\[U_{b|C_i} = \sum(\beta x)_{b|C_i} + \varepsilon_{b|C_i}, \quad l = 1, 2\]  

(5)

\( x \) indicates a set of personal, land use and transportation variables and \( \beta \) is corresponding covariates.

\( \varepsilon_{b|C_c} \) and \( \varepsilon_{b|C_i} \) are Type I Extreme value random error components.

Corresponding conditional probability of owning a bike and being in class \( C_c \) is \( P_{b|C_c} \) and conditional probability of owning a bike and being in class \( C_i \) is \( P_{b|C_i} \).

Finally, the marginal probabilities of owning \( f[\text{own}] \) or not owning \( f[\text{not-own}] \) a bike are:

\[
f[\text{own}] = \frac{\exp(\sum(yz) + \sum_k(\delta_b \gamma_b)k)}{1 + \exp(\sum(yz) + \sum_k(\delta_b \gamma_b)k)} \cdot \frac{\exp(\sum(\beta x)_{b|C_c})}{1 + \exp(\sum(\beta x)_{b|C_c})} + \frac{1}{1 + \exp(\sum(yz) + \sum_k(\delta_b \gamma_b)k)} \cdot \frac{\exp(\sum(\beta x)_{b|C_i})}{1 + \exp(\sum(\beta x)_{b|C_i})}
\]  

(6)

\[
f[\text{not-own}] = \frac{\exp(\sum(yz) + \sum_k(\delta_b \gamma_b)k)}{1 + \exp(\sum(yz) + \sum_k(\delta_b \gamma_b)k)} \cdot \frac{1}{1 + \exp(\sum(yz) + \sum_k(\delta_b \gamma_b)k)} + \frac{1}{1 + \exp(\sum(\beta x)_{b|C_c})} \cdot \frac{1}{1 + \exp(\sum(\beta x)_{b|C_i})}
\]  

(7)

In these formulations, the perceived bike owning rate, \( (\bar{Y}_b)_k \) of any cohort \( k \), is unknown to us. So, we can fairly assume that this is a random function of modelled proportion of bike owning in the cohort:
\((\delta_b \bar{y}_b)_k = (\delta_b (\sum f[own] + \varepsilon_{bl}))_k\) \hspace{1cm} (8)

\((\sum f[own])_k\) is the modelled proportions of students owning bikes in the cohort \(k\)

\((\varepsilon_{bl})_k\) is a standard normal error term

Such assumption makes the Equation (5) and (6) the conditional probabilities of joint class membership and choice of bike owning. Both of the conditional probabilities are conditional to the distribution of \((\delta_b \bar{y}_b)_k\). Finally, the unconditional probabilities of owning a bike \((P[own])\) and non-owning a bike \((P[not-own])\) are (Guevara and Ben-Akiva 2010):

\[ P[own] = \int_{\varepsilon_{bl}} f(own)f(\varepsilon_{bl}) \, d\varepsilon_{bl} \] \hspace{1cm} (9)

\[ P[not-own] = \int_{\varepsilon_{bl}} f(not-own)f(\varepsilon_{bl}) \, d\varepsilon_{bl} \] \hspace{1cm} (10)

\(f(\varepsilon_{bl})\) indicates the distribution function of random error \(\varepsilon_{bl}\).

Finally, the likelihood function of any observation \(i, L_i\) becomes

\[ L_i = \left( \int_{\varepsilon_{bl}} f(own)f(\varepsilon_{bl}) \, d\varepsilon_{bl} \right)^{\omega} \left( \int_{\varepsilon_{bl}} f(not-own)f(\varepsilon_{bl}) \, d\varepsilon_{bl} \right)^{1-\omega} \] \hspace{1cm} (11)

\(\omega\) is an indicator that takes the value 1 if an individual is observed to own a bike and 0 otherwise.

This likelihood function is not of closed form and so we can use pseudo-likelihood (Simulated Likelihood) as an alternative:

\[ L_i = \left( \frac{1}{R} \sum_{r=1}^{R} f(own)|(\varepsilon_{bl})_r \right)^{\omega} \left( \frac{1}{R} \sum_{r=1}^{R} f(not-own)|(\varepsilon_{bl})_r \right)^{1-\omega} \] \hspace{1cm} (12)

\(f(.)|(\varepsilon_{bl})_r\) is the value of the function \(f(.)\) for a random draw of \((\varepsilon_{bl})_r\) from the standard normal distribution

\(R\) is the total number of random draws

However, the issue of this formulation is that the choice probability is a function of its own because of the incorporation of perceived bike owning rate as specified in equation (7).

This presents an endogeneity issue and makes the classical approach of Maximum Simulated Likelihood (MSL) method obsolete. However, as proven by Grange et al (2013 and 2015) for such mathematical problem, as per Banach Fixed-Point theorem, unique fixed points exist and an iterative estimation process can recover unbiased parameter estimates. The starting step (step 0) the algorithm uses the following formulas:
\[ f[own]^0 = \left( \frac{\exp(\Sigma(yz))}{1+\exp(\Sigma(yz))} \right) * \left( \frac{\exp(\Sigma(\beta x)b_{|C|})}{1+\exp(\Sigma(\beta x)b_{|C|})} + \frac{1}{1+\exp(\Sigma(yz))} \right) \] \hspace{1cm} (13) \\
\[ f[not - own]^0 = \left( \frac{\exp(\Sigma(yz))}{1+\exp(\Sigma(yz))} \right) * \left( \frac{1}{1+\exp(\Sigma(yz))} + \frac{1}{1+\exp(\Sigma(yz))} \right) \] \hspace{1cm} (14) \\

Step 0 does not have any endogenous variables and it gives estimates for step 1:

\[ (f[own])^1 = \frac{1}{R} \sum_{r=1}^{R} \left( \frac{\exp(\Sigma(yz)+\Sigma_k(\delta_p(\Sigma [f[own]^0]+(e_{bl})_r))_k)}{1+\exp(\Sigma(yz)+\Sigma_k(\delta_p(\Sigma [f[own]^0]+(e_{bl})_r))_k)} * \frac{\exp(\Sigma(\beta x)b_{|C|})}{1+\exp(\Sigma(\beta x)b_{|C|})} \right) \hspace{1cm} (15) \]

\[ (f[not - own])^1 = \frac{1}{R} \sum_{r=1}^{R} \left( \frac{\exp(\Sigma(yz)+\Sigma_k(\delta_p(\Sigma [f[own]^0]+(e_{bl})_r))_k)}{1+\exp(\Sigma(yz)+\Sigma_k(\delta_p(\Sigma [f[own]^0]+(e_{bl})_r))_k)} * \frac{1}{1+\exp(\Sigma(\beta x)b_{|C|})} \right) \hspace{1cm} (16) \]

Step 2 uses estimates step 1 as followings:

\[ (f[own])^2 = \frac{1}{R} \sum_{r=1}^{R} \left( \frac{\exp(\Sigma(yz)+\Sigma_k(\delta_p(\Sigma [f[own]^1]+(e_{bl})_r))_k)}{1+\exp(\Sigma(yz)+\Sigma_k(\delta_p(\Sigma [f[own]^1]+(e_{bl})_r))_k)} * \frac{\exp(\Sigma(\beta x)b_{|C|})}{1+\exp(\Sigma(\beta x)b_{|C|})} \right) \hspace{1cm} (17) \]

\[ (f[not - own])^2 = \frac{1}{R} \sum_{r=1}^{R} \left( \frac{\exp(\Sigma(yz)+\Sigma_k(\delta_p(\Sigma [f[own]^1]+(e_{bl})_r))_k)}{1+\exp(\Sigma(yz)+\Sigma_k(\delta_p(\Sigma [f[own]^1]+(e_{bl})_r))_k)} * \frac{1}{1+\exp(\Sigma(\beta x)b_{|C|})} \right) \hspace{1cm} (18) \]

The subsequent steps use estimates of \( f[own] \) from the immediately prior step. The iterations converge to the fixed point within few iterations. However, because of the pseudo-likelihood function that Maximum Simulated Likelihood approach needs to be used after step 1. The algorithm is programmed in GAUSS (Aptech Systems 2017) and for simulating random draws Halton random number generation process is used. The proposed estimation process is of full-information approach and theoretically superior to any two-step approaches.
4. Data for Empirical Investigation

Data for empirical investigation of the paper came from an innovative travel diary survey of the students of all 4 universities (University of Toronto, York University, Ryerson University and OCAD University) from 7 campuses in Toronto. The survey was the result of collaborations of the administrations of the 4 universities to collect evidence on the travel behavior of university students in the city. These universities are hosts of 184,000 post-secondary students, who are often under-represented in regional household travel surveys. The survey was conducted in the fall of 2015 by using a web-based travel diary survey tool that resulted in a total of 8.3 percent response rate. Detailed description and summary statistics of the survey is available in StudentMoveTo (2015).

![Figure 1: Bike ownership by age group and gender](image)

Data from this survey are fused with zone-based population and land-use attributes for this investigation. After cleaning for missing and erroneous information, the final dataset that is used for this investigation has a total of 13600 individuals. Descriptions of the cleaned data set are reported in Habib et al (2017). The main focus of this paper is the bike ownership choice of the students in the context of peer-groups’ social influences. Figure 1 presents the distributions of bike ownership of males and females of different age groups in the sample. It is clear that male and female students have different patterns of bike ownership in the dataset. Also, for both male and females, bike ownership proportion changes distinctively by age groups.
5. Empirical Investigation and Policy Relevance

5.1 Empirical Model

The empirical model is a latent class bike ownership choice model with age-cohort-specific social interaction. It is presented in Table 1 (part 1 and part 2). The model has a binary membership choices for two latent classes (peer-confirming and peer-indifferent) and for each class, it has a binary choice model of bike ownership choice (owning or not owning a bike). A wide variety of possible specifications for the systematic utility functions are investigated and the final one (presented in the paper) is selected based on expected signs of the variables and statistical significance of the corresponding parameters. Although a 95 confidence limit (corresponding t-statistics of 1.64 for one-tailed or 1.96 for two-tailed test), some variables with lower than 95 confidence limit are retained in the model as these provide behavioral insights. Also, the expectation is that with a larger data set, these variables should show up with higher confidence limit. The presented model has a total 61 parameters and it passes the likelihood ratio test against a constant-only latent class model as well as a simple binary choice model. The final model also has higher AIC and BIC values than those of the same latent class model but without any social interactions. This justifies the accommodation social interaction effects in the model.

Figure 2: Endogenous Social Interaction Effects on Bicycle Ownership Choice
A wide variety of personal, household, land use and travel behavior related variables along with gender and cohort-specific social interactions influence a student to be peer-conforming or peer-indifferent. It is clear that age-cohort specific social interaction effects have positive effects on the choice of being peer-conformable and have higher odd ratio\(^1\) than any other variables except the constant. For clarity, Figure 2 presents comparisons of peer-group effects by plotting the odd-ratios that are reported in Table 1. Choices of owning a bike by the university students in Toronto are highly influenced by the choices of the fellow students of the same age and gender groups. Between male and female students, female students appear to be more influenced by such cohort-specific social interactions than the male students are. Also, heterogeneity in such social interaction effects exists for both males and females. Neither male nor female students of all age-cohort groups are influenced by such social interaction effects in the same ways. Cohort-specific social interaction effects are the highest for the age of 22 years that consistently increases from younger age groups (18 years or younger) and decreases afterward for the older than 22 years of ages. Perhaps, this is the end of the undergraduate-year effect. In general, students of age 22 are in their final year of undergraduate programs when their peer-influence reached the peak. After this age, students who linger in the undergraduate program or enroll in graduate programs have a gradually decreasing number of peers of similar age and similar decreasing peer-influence. However, in comparison to the male students, female students of age 22 are influenced by the choices of fellow females of the same age way more than do the male students of the same age group. This indicates more camaraderie with the peers that may exist more in females than the males.

In addition to, age-cohort specific social interaction effects, the peer-conformable bicycle ownership class utility function has a relatively large constant. This indicates that there are other variables/factors that influence a student to be peer-conformable in a choice of owning a bicycle, but are not available in the dataset. Among the other variables accommodated in the model, household car ownership has the highest influence. It seems the students who come from a home without a private car are more likely to be peer-conforming in the case of the choice of owning a bike. Personal possession of driving license and owning a car have opposite effects. As owning a car requires to possess a driving license, the final effect is that students who personally own a car are more likely to be peer-indifferent than peer-conforming. However, students who possess a driving license, but does not own a car, is even less likely to be peer-conforming. Students living with the family are more peer-conformable than the students living with a partner. Undergraduate students are more peer-conformable than are the graduate students. Full-time students are less peer-conformable than part-time students.

---
\(^1\) In binary logit model, odd ratio is the exponential of the estimated parameter of a variable and it presents the relative size effect of the variable on the choice probability
Table 1: Empirical Model (Part 1)

<table>
<thead>
<tr>
<th>Systematic Utility of Class Membership</th>
<th>Variable</th>
<th>Parameter</th>
<th>t-Stat</th>
<th>Odd ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer-indifferent class</td>
<td>Reference class</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peer-conforming class</td>
<td>Constant</td>
<td>1.71</td>
<td>2.86</td>
<td>5.54</td>
</tr>
<tr>
<td></td>
<td>Home to campus distance (in 100 km)</td>
<td>-0.11</td>
<td>-0.31</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>Graduate program</td>
<td>-0.05</td>
<td>-0.43</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>Full-time status</td>
<td>-0.18</td>
<td>-1.37</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>Having driving license</td>
<td>-0.47</td>
<td>-6.23</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>Having a car by own</td>
<td>0.27</td>
<td>2.84</td>
<td>1.31</td>
</tr>
<tr>
<td></td>
<td>Weekly frequency of school trips</td>
<td>-0.03</td>
<td>-1.24</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>Family household size (log transformed)</td>
<td>-0.04</td>
<td>-0.46</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>No car at the household</td>
<td>0.40</td>
<td>3.59</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>Living with family</td>
<td>-0.73</td>
<td>-3.61</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Living with partner</td>
<td>-0.97</td>
<td>-4.49</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>Living with roommates</td>
<td>-0.01</td>
<td>-0.07</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Home to nearest bus stop in m. (log of distance)</td>
<td>-0.17</td>
<td>-5.17</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>Average block length in m. (log transformed)</td>
<td>0.04</td>
<td>0.50</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>Walk buffer area (log transformed)</td>
<td>-0.93</td>
<td>-4.49</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>Population density (in 1000 per sq. km) in home zone</td>
<td>0.16</td>
<td>3.63</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>Employment density (in 1000 per sq. km) in home zone</td>
<td>0.31</td>
<td>5.47</td>
<td>1.37</td>
</tr>
</tbody>
</table>

Peer-Group Influence: Proportion of this group of students owning a bike

<table>
<thead>
<tr>
<th>Param</th>
<th>t-Stat</th>
<th>Odd ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 18 years or less and female:</td>
<td>0.78</td>
<td>3.11</td>
</tr>
<tr>
<td>Age 18 years or less and male:</td>
<td>0.00</td>
<td>---</td>
</tr>
<tr>
<td>Age 19 years and female:</td>
<td>1.32</td>
<td>3.30</td>
</tr>
<tr>
<td>Age 19 years and male:</td>
<td>0.34</td>
<td>0.88</td>
</tr>
<tr>
<td>Age 20 years and female:</td>
<td>1.49</td>
<td>4.78</td>
</tr>
<tr>
<td>Age 20 years and male:</td>
<td>0.97</td>
<td>2.54</td>
</tr>
<tr>
<td>Age 21 years and female:</td>
<td>1.85</td>
<td>5.26</td>
</tr>
<tr>
<td>Age 21 years and male:</td>
<td>0.86</td>
<td>1.87</td>
</tr>
<tr>
<td>Age 22 years and female:</td>
<td>2.46</td>
<td>4.15</td>
</tr>
<tr>
<td>Age 22 years and male:</td>
<td>1.48</td>
<td>2.75</td>
</tr>
<tr>
<td>Age 23 years or more and female:</td>
<td>2.06</td>
<td>5.07</td>
</tr>
<tr>
<td>Age 23 years or more and male:</td>
<td>0.55</td>
<td>1.53</td>
</tr>
</tbody>
</table>

Standard Deviation of bike ownership proportion

| Age 18 years or less and female: | 0.00 | --- | --- |
| Age 18 years or less and male: | 0.00 | --- | --- |
| Age 19 years and female: | 0.29 | 0.80 |
| Age 19 years and male: | 5.86 | 0.75 |
| Age 20 years and female: | 1.05 | 2.05 |
| Age 20 years and male: | 1.15 | 1.68 |
| Age 21 years and female: | 6.36 | 2.36 |
| Age 21 years and male: | 2.92 | 1.41 |
| Age 22 years and female: | 4.00 | 1.79 |
| Age 22 years and male: | 1.37 | 2.15 |
| Age 23 years or more and female: | 1.52 | 4.52 |
| Age 23 years or more and male: | 2.84 | 1.12 |

Living farther away from the university reduces the probability of being peer-conformable. In terms of land use attributes around the residence/home, it is clear that lower accessibility to transit (longer distance between home and the nearest bus stop) as well as lower walkability reduces the probability of being peer-conformable. Students living in dense neighborhoods are more likely to be peer-conformable than those living in suburbs.
Similarly, students living in higher employment density locations are more likely to be peer-conformable.

Given the class memberships, two separate models of the choice of owning a bicycle are estimated: one for peer-conformable students and one for peer-indifferent students. Bicycle ownership choice of peer-indifferent class is highly influenced by a couple of dominant variables. It seems that if a peer-indifferent student owns a transit pass, it is almost certain that the student also owns a bicycle. This is further assured by the effects of home to campus distance variables. Peer-indifferent students who live far from the campus are more likely to own a bicycle than those living close to the campus. It is understandable that living far from the campus makes the transit pass ownership more feasible and so does the choice of owning a bicycle. However, the complete opposite is true for the peer-conforming students. Peer-conforming students who live far from the campus and/or own a transit pass are less likely to own a bicycle than the peer-conforming students who live close to the campus and/or own a transit pass. It is clear that no single variable highly dominate the choice of owning a bike by the peer-confirming students as it is true for peer-indifferent students.

**Table 1: Empirical Model (Part 2)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Peer-Conformable class</th>
<th>Peer-Indifferent class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Param</td>
<td>t-stat</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.81</td>
<td>-5.75</td>
</tr>
<tr>
<td>Home to campus distance (km)</td>
<td>-0.17</td>
<td>-6.04</td>
</tr>
<tr>
<td>Graduate program</td>
<td>-0.03</td>
<td>-0.27</td>
</tr>
<tr>
<td>Having a driving license</td>
<td>0.71</td>
<td>5.21</td>
</tr>
<tr>
<td>Having a transit pass</td>
<td>-2.55</td>
<td>-4.07</td>
</tr>
<tr>
<td>Family household size</td>
<td>0.03</td>
<td>0.48</td>
</tr>
<tr>
<td>More than 1 car in Family</td>
<td>0.26</td>
<td>1.30</td>
</tr>
<tr>
<td>Not living on-campus</td>
<td>0.15</td>
<td>0.96</td>
</tr>
<tr>
<td>Number of women in family</td>
<td>0.09</td>
<td>1.25</td>
</tr>
<tr>
<td>Family age ratio:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age difference (oldest and youngest) by Age of the oldest member</td>
<td>-0.31</td>
<td>-0.96</td>
</tr>
<tr>
<td>Age (log transformed) of female</td>
<td>1.21</td>
<td>4.44</td>
</tr>
<tr>
<td>Age (log transformed) of male</td>
<td>1.43</td>
<td>5.13</td>
</tr>
</tbody>
</table>

There are some variables that have similar influences (magnitude and signs) on the choices of bike ownership by both classes of students. Both of these cases, the unexplained systematic effects (constant) are relatively low compared to the effects of other variables.
Possession of driving license increases the choice of owning a bicycle for both classes. Students of both classes who belong to a large family are more likely to own a bike.

There are some variables that have a similar effect in terms of sign, but very different in terms of magnitude on the choice owning a bicycle. Higher car ownership by the household of a student increases the probability of owning a bicycle, but its effect is almost 10 times higher for the peer-indifferent class than that for the peer-conforming class. Similarly, a student of both classes, who lives off-campus is more likely to own a bicycle than those living on-campus. However, this effect of living off-campus is almost 10 times higher for the peer-indifferent class than that for the peer-conforming class. Family configuration a student (in terms of combinations of age and gender of the family members) play important role in the choice of owning a bicycle in both classes. For the peer-indifferent class, students from the families with large differences between the ages of the youngest and the oldest members are highly likely to own a bicycle. However, this variable does not have a significant effect on the students of the peer-conforming class. Age and gender of a student seem to have systematic, but opposite effects on the choice of owning a bicycle for members of both classes of bicycle owners. For the peer-conforming class, older female and male students are more likely to own a bicycle than the younger ones. However, the opposite is true for the members of the peer-indifferent class.

5.2 Policy Relevance of the Empirical Findings

The empirical investigation of the paper clearly shows that the choices of bike ownership by the university students are influenced by various factors and two distinct classes of students are identified: influenced by the choices of the peers (peer-conforming) and not influenced by the choices of the peers (peer-indifferent). This study reveals the fact that the policy makers should consider such complex interactions of a various demographic and socioeconomic factors in evaluating the effectiveness of any bicycle promotion initiatives. The focus of this investigation on the university students has particular policy relevance that they represent the millennial. Any policy that can influence the choice of bicycle owning (and likely to use) will have lasting effects in the society as they will enter into the work force in near future if their preference to bicycle follows with their life stages. These data show that bicycle ownership in the sample peaks at age 20, yet “peer conformity” peaks at age 22 (typically a student’s last year), suggesting that the greatest opportunity to increase bicycle ownership would be among female students in their final year of undergraduate studies. The data suggest ownership by this cohort would be expected to have the greatest impact on encouraging others to pursue cycling. Previous studies revealed that bicycle Infrastructure investments have a positive effect on increasing bicycle ownership (Gorrard et al. 2008; Dill et al. 2015). It is difficult to target age and gender specific groups through infrastructure options only. Perhaps, application social engineering approaches e.g. extensive advertisements, public promotions, bicycle awareness week, etc.
catered for different age groups will have a greater return on investment if are undertaken along with bicycle infrastructure improvements. Specifically, if such programs/approaches target females, the return can be higher than that of gender-indifferent programs/approaches. Habib et al (2014) also called for similar policy initiatives of targeting young females to reduce the gender gap in bicycle mode share following an empirical investigation in the same study area but on the general population and using different data set.

6. Conclusions and Recommendations for Future Research

The paper presents empirical investigations on the choice of owning a bicycle by the university students in Toronto. It starts with a hypothesis that the choice of owning a bicycle is influenced by the same choices of people of similar age and gender group (social interaction effects) along with many other socioeconomic and land use variables. It used a latent class discrete choice modeling approach following Maness and Cirillo (2016) considering that social interaction effects influence the membership of latent classes than the choices of owning a bike that is conditional to the latent class memberships. However, the paper extends the modelling approach further and contributes to the literature on bicycle choice related studies in two ways. First, it extends the econometric formulation of latent class bicycle choice model by proposing a simultaneous full-information modeling approach that is more efficient and consistent than the two-stage model proposed by Maness and Cirillo (2016). Further, as opposed to considering spatial aggregate social interaction effects, this study developed a latent class model for age and gender specific peer groups. Though the empirical application of on the university students in Toronto, the model can be extended and applied to any such socio-economic peer class/group definition. In addition, the empirical application of the model for university students has particular policy interested as these group of people represent the Millennials and are expected to have influences on the future demands that they would carry from the current level of the life cycle.

The empirical model clearly reveals that university students are highly influenced by the choices of their peers in case of the choice of owning a bicycle. In terms of peer groups, clear distinctions are visible by age and gender combinations. Between male and female students, female students are more influenced by the choices of fellow female students of the same age group than do the male students. Car ownership, either by the household or by the student by his/her own, have negative effects on being peer-conformable in bicycle ownership choices. Peer conformability reduces for the students do not live close to the university campus and also live in the neighborhood with poor transit accessibility and poor walkability. Living in suburban areas with lower home and employment density seems to lower the probability of being peer-conformable of the university students in case of the choices of owning a bicycle.
It is interesting to note that, peer-indifferent students are highly influenced by a couple variables in case of the choice of owning a bicycle. Peer-indifferent students who live off-campus, far from the university and own a transit pass are mostly likely to own a bicycle. The opposites are true for peer-conformable students. Also, a wide variety of variables plays a comparable role in influencing the choice of owning a bicycle by the peer-conformable students. Overall, peer-conformability seem to be advantageous from policy initiative point of view as a large number of factors can be manipulated to influence the choice of owning a bike. These findings of empirical investigations call for soft policies (e.g. peer campaign, social marketing, educational campaign and gender-specific bicycle promotion campaigns etc.) to accompany hard infrastructure investments to promote bicycle in urban areas.

The proposed latent class discrete choice model is developed for the binary choice of bicycle ownership. However, the formulation can easily be extended to multinomial or ordered choices of a number of bicycle ownership. A similar framework can also be used for investigating bicycle ownership and usage. The dataset available for this study did not allow for these investigations, but these are plausible future investigations for the different dataset. Finally, the empirical investigation presented in this paper can be extended to the general population and for varieties of socio-economic peer groups by using different data set.

Acknowledgement

The study was funded by an NSERC Discovery Grant. The author wishes to thank StudentMoveTO survey team for making the dataset available for current research. The author also wishes to thank Dr. Steven Farber for kindly sharing the land use and transit accessibility related variables that are used in the empirical investigation. The paper is also benefitted from the comments Dr. Trevor Henson. However, the author has the sole responsibility for all comments and conclusions made in this paper.

References

Akar, G., Clifton, K., 2009. Influence of individual perception and bicycle infrastructure on the decision to bike. Transportation Research Record 2140: 165-1712
Castillo-Manzano, J.I., Sañchez-Braza, A. 2013. Managing a smart bicycle system when
demand outstrips supply: The case of university community in Seville.
Transportation 40:459-477
Dill, J., Goddard, T., Monsere, C.M., McNeil, N. 2015. Can protected bike lanes help close the
gender gap in cycling? Lessons from five cities. Paper presented at the 94th Annual
Emond, C.R., Tang, W., handy, S.L. 2009. Explaining gender difference in bicycling
behaviour. Transportation Research Record 2125: 16-25
The role of perceptions. Transportation 43:1-23
Gorrard, J., Rose, G., Lo, S.K. 2008. Promoting transportation cycling for women: The role of
bicycle infrastructure. Preventive Medicine 46: 55-59
Transportation Research Part A 53:1-9
Grange, L.d., González, F., Vargas, I., Troncoso, R. 2015. A logit model with endogenous
explanatory variables and network externalities. Network and Spatial Econometrics
15:89-116
Guevara, C.A., Ben-Akiva, M. 2010. Addressing endogeneity in discrete choice models:
Assessing control-function and latent-variable methods, in Stephane Hess, Andrew
370
City of Toronto: Investigating the effects of perception, consciousness, and
comfortability on the purpose of biking and bike ownership. Transportation
Research Part A 70:67-80
in mobility tool ownership choices of post-secondary students in Toronto”. Paper
Handy, S., Xing, Y., Buehler, T.J. 2010. Factors associated with bicycle ownership and use: a
study of six small U.S. cities. Transportation 37: 967-985
Krizek, K. J., Johnson, P.J., Tilahun, N. 2005. Gender differences in bicycling behavior and
facility Preferences. In Conference Proceedings 35, Research on Women’s Issues in
Transportation Volume 2: Technical Papers, Transportation Research Board of the
bicycling? Transportation 41: 1135-1152
Maness, M. Cirillo, C. 2016. An indirect latent informational conformity social influence
choice model: Formulation and case study. Transportation Research Part B 93: 75-101


Sener, I., Eluru, N., Bhat, C.R. 2009b. Who are bicyclists? Why and how much are they bicycling? Transportation Research Record 2134: 63-72
