Modeling Bicyclists’ Injury Severity Levels in the Province of Nova Scotia, Canada using a Generalized Ordered Probit Structure

Muhammad Ahsanul Habib, PhD
(Corresponding author)
Assistant Professor
Dalhousie Transportation Collaboratory
School of Planning, and Department of
Civil and Resource Engineering
Dalhousie University, PO Box: 15000
Halifax, NS, B3H4R2, Canada
Phone: +1(902)494 3209, Fax: +1(902)423 6672
Email: ahsan.habib@dal.ca

And,

Justin Jamael Forbes
Graduate Research Assistant
Dalhousie Transportation Collaboratory
School of Planning, and Department of
Civil and Resource Engineering
Dalhousie University, PO Box: 15000
Halifax, NS, B3H4R2, Canada
Email: justin.forbes@dal.ca

Word Count: 6666 + (2 x 250) = 7166
Revised Paper Submission Date: November 15, 2013

93rd Annual Meeting of Transportation Research Board, January 12-16, 2014, Washington D.C.
**ABSTRACT**

This paper examines the factors affecting injury severity in bicycle collisions using a generalized ordered probit model. One of the unique features of the modeling approach adopted in this paper is its flexible threshold structure, which incorporates individual variations in the thresholds to account for heterogeneity likely present in the data, but not commonly accommodated in traditional ordered probit models. Additionally, previous research has focused primarily on the factors that affect injury severity for motorists; generally, less attention has been given to understanding factors that affect injury severity levels for cyclists. Furthermore, examination of neighborhood and land use attributes in association with injury severity is surprisingly limited in the existing literature. This study attempts to fill the gap, particularly in understanding how land use and neighborhood characteristics affect injury severity levels for bicyclists. The data covers 2007-2011 bicycle collisions taken from police collision reports from the Province of Nova Scotia, supplemented with Census tabulations, provincial land use information, and point of interest data specific to the individual collision locations. The results reveal that females, impaired cyclists, and persons aged 45-54 involved in bicycle collisions have an increased likelihood of sustaining more severe injuries. Road condition and configuration, bicyclists’ manoeuver, and lighting conditions also affect cyclists’ injury severity levels. Finally, characteristics of the neighborhood in which collisions occur, often ignored in previous collision studies, for instance land use mix, proximity to activity centers, and demographic attributes are found to be significant in explaining injury severity of bicyclists. The results suggest that neighborhood characteristics should be given more scrutiny and be an important consideration when evaluating and planning for cyclist safety.
INTRODUCTION
This paper presents the findings of a generalized ordered probit model that examines bicyclists’ injury severity levels. Bicycling is becoming an increasingly important element of sustainable transport systems; their pollutant and noise emissions, and the accident risks they pose for other road users are very low which contribute to a more attractive urban environment (1). Bicycling also offers economic benefits, such as reduced household expenditure on transportation, reduced work hours lost in traffic congestion, and reduced healthcare costs resulting from increased physical activity and reduced pollution (2). Furthermore, there is increased recognition of the health and wellness perspective of cycling as an effective way for people to cope with health problems and obesity (1). Given the benefits of bicycling, improving road safety for cyclists is an important consideration for encouraging people to cycle more. Effective injury reduction requires an understanding of the factors that affect the likelihood of a collision occurring as well as the characteristics that may mitigate or exacerbate the level of injury sustained (3). Many studies have applied various modeling frameworks to injury severity, although limited research has focused on bicycle collisions for improving road safety for the cyclists. Relatively little is known about the influence of the built environment and neighborhood attributes on injury severity, especially within the context of bicycling. This research attempts to fill the gap, specifically by investigating characteristics of the neighborhoods in which collisions occur. It is to our best knowledge that this study may be the first to solely model bicyclists’ injury severity levels, incorporating neighborhood characteristics in the model estimation. Furthermore, the results generated in this study provide an evidence based foundation for the implementation and evaluation of bicycle-related road safety strategies and campaigns in Nova Scotia, Canada.

We apply an ordered probit modeling approach to examine the factors affecting injury severity levels. A generalized ordered probit model structure is utilized to account for heterogeneity across individuals, particularly in relation to the thresholds parameters, something previously not attempted in bicycle collision research. The paper uses data drawn from the Nova Scotia Collision Record Database (NSCRD) at Service Nova Scotia and Municipal Relations (SNSMR) for the empirical application. The rest of the paper is organized as follows: first, we provide a review of the literature. Following the literature review, we discuss the data used in the empirical application. The next section describes the modeling approach used in the study, followed by a discussion of model results. The paper concludes by providing a summary of contributions and future research directions.

LITERATURE REVIEW
There is a wide body of safety literature examining the occurrence and outcomes of bicycle collisions. Several collision studies found that children and older individuals were the main groups that suffered from bicycle related injuries (4,5,6,7,8,9,10). Stone and Broughton (5) observed a higher incidence of fatalities in adults older than 50 years. About half of fatally injured bicyclists are 65 years or older in Sweden (11). A study by Kim et al. (12) support this finding; bicyclists older than 55 were found to be more susceptible to fatalities. An early study by Maring and van Schagen (10) noted that even though age itself is not a causal factor, age is strongly associated with relevant variables such as cognitive development and perception. Although not conclusive, gender has been found to play a role in injury outcomes. Males, in Kim et al.’s study (12), had a higher percentage distribution in bicycle-related injuries compared to females while Stone and Broughton (5) found no significant difference for males and females.

Several studies have investigated the relationship between bicycle safety and alcohol consumption. Kim et al. (12) found that alcohol consumption by both bicyclists and motorists increase the likelihood of fatal and incapacitating injuries for cyclists. Cyclist fatalities have been found to be significantly correlated with alcohol expenditure per capita (13). Comparing intoxicated and sober cyclists, some studies have found a greater risk of head and face injuries in intoxicated cyclists (14,15). Alcohol increases the bicyclists’ risk of injury from falling more than from collisions (15).
The literature is in general agreement that most fatal and serious bicyclist injuries are associated with higher speed limits (5,12,16,17,18,19) and that bicycle safety can be improved by reducing bicycle and vehicle speeds (20,21,22). High speeds make drivers pay attention to the most relevant direction and ignore the less relevant direction which modifies the driver's visual scanning pattern (23,24). This agrees with research that indicates the most frequent type of bicycle-motor vehicle collisions are related to a driver turning right and a bicycle coming from the driver's right (25,26).

Some research has shown that relative rates for falls and injuries is lower when cycling on-road compared to using an off-road path or sidewalk (27,28,29,8). However, Smith and Walsh (30) and Pucher (31) argue that bicycle bikeways and bike lanes make cycling safer. Eluru et al (17) recommend bicycle facilities be designed to be an off-roadway bicycle lane, physically separated from motorized vehicle traffic by an open space or barrier. Research has shown that accident risk for cyclists varies significantly.

There are safe, and unsafe bike paths, just as there will be areas where on-road riding is relatively safe. Rifaat et al. (32) studied effect of street pattern on the severity of crashes involving vulnerable road users and found loop and lollipop design to be associated with a higher likelihood of non-fatal injury in the event of a collision between a motor vehicle and a cyclist but a lower likelihood of non-fatal injury. In a study by Thom and Clayton (33), the most frequent contributing factor to bicycle-motorist accident risk for both bicyclists and drivers was the failure to yield to right away. Garder (19), and Kim and Li (34) observed that cyclists were more likely than drivers to violate traffic laws. Specific crash patterns and risk have been found to play a role in elevating bicyclist injury severity, for example: head-on and angle collisions, occurrence of running over bicyclists, roads without median/division, and heavy vehicle involvement (16).

Klop and Khattak (35) observed that injury severity increased in fog, on roads with both straight and curved grades and with higher speed limits. Klop and Khattak (35) also found that crashes occurring in higher average traffic result in less severe injuries. Wavnik (36) found the risk of injury accidents to increase in darkness. Stone and Broughton (5) studied the influence of lighting on fatalities and found that darkness with street lighting has the lowest fatality rate and found that a higher percentage of fatalities occur between 9 pm and 6 am. Some authors identified main causes of bicycle-traffic accidents and found that the main causes were excessive vehicle speed, lack of proper illumination during the afternoon peak period and at night, and a poor roadway design. Kim et al.’s study (12) found that collisions occurring during the AM peak (9-10 am) and weekends increase the likelihood of fatality.

Numerous studies of bicycle collisions have focused on head injury and helmet usage. Generally, they found that head injuries were the most common type of bicycle-related injuries (9,37,38,39,40,41). Helmets are protective against head injury and brain injury (42,43,44,45,46,47). On the other hand, it has been argued that hemeted cyclists ride more recklessly as they feel more protected (48). Although helmets are effective for all cyclists, regardless of age, helmets are not always properly used leaving room to improve helmet design to mitigate improper use (49,50,45,51). Helmet usage has been linked to behavioral characteristics: for example, helmet usage rate in one study was correlated with time spent riding a bicycle each year (6). Neighborhood characteristics also have an influence on helmet use. One study found those who wear helmets are more highly educated (52). Another study found helmet use in rural areas to be lower than in urban areas across all age groups and for both genders (53). When legislative intervention is introduced to mandate helmet usage, the effectiveness is uncertain as injury and fatality rates may fall simply because the legislation produces a decline in bicycle use (54).

Socioeconomic factors, such as the percentage of poor households within a neighborhood, have been found to play an important role in the prediction of bicycle accident rates (55). Pless et al. (56) reported that higher risk of injury was related to fewer years of parent education, a history of accidents in the family, an environment judged as unsafe, and poor parental supervision. Macpherson et al. (38) found that children living outside urban centers had an increased risk of hospitalization due to bicycling-related injuries. Kim et al. (12) has noted a need for future research into the interplay between the built environment and bicycle collisions.

The majority of bicycle safety research concentrates on descriptive analysis of causes, occurrences and outcomes of collisions. Studies that employ multivariate models to pursue analysis of the
factors affecting injury severity at the level of individual collisions is limited (17). Even more limited, is
the amount of studies that examine the relationship between land use and neighborhood attributes, and
bicycle collisions. Therefore, this paper attempts to investigate the factors that affect injury severity levels
of bicyclists, including neighborhood characteristics using collision records of Nova Scotia, Canada.

DATA USED IN THE EMPIRICAL APPLICATION

Nova Scotia Collision Record Database
Collision records from 2007-2011 were the data source for model estimation. The collision records were
drawn from the NSCRD retained at SNSMR in Halifax, Canada. In Nova Scotia, all collisions involving
property damage over $1000 and injuries or fatalities occurring on a public road, as defined by the Motor
Vehicle Act, require reporting. The NSCRD consists of data representing collisions in 18 counties in
Nova Scotia. The 2007-2011 NSCRD data includes information on over 74,000 collisions involving about
208,700 individuals. Of these, about 470 collisions involved bicyclists. When a collision occurs, the
completed collision report forms (MV58A) record a number of accident-related attributes including the
characteristics of individuals involved, vehicle characteristics, roadway design attributes, environment
factors, and crash characteristics. The injury severity of each individual involved in the accident is
recorded on a five point ordinal scale: (1) not injured, (2) minor – no treatment, (3) moderate – treated &
released, (4) major – hospitalized, and (5) fatal. After cleaning the data for validity, consistency, and
uniformity, 425 bicycle collisions were deemed suitable and retained for further analysis.

Data Preparation for Modeling
The data preparation for modeling involved multiple stages. First, a database with all relevant attributes
was created. Second, collision locations were geocoded using the online service BatchGeo™. Third,
neighborhood characteristics were derived by means of the spatial join function in ArcGIS to combine the
collision location with dissemination area (DA) data from the 2006 Canadian Census. Joined data
included average household income, average number of rooms, housing stock, and dwelling type counts.
Population and dwelling densities were normalized by their respective fields with the DA area. Finally, a
land use mix index for Nova Scotia, originally proposed by Bhat and Gossen (57) and adapted by
Habib et al. (58), was spatially joined to the collision location. The land use mix index ranges from 0 to 1,
with 1 indicating perfect land use heterogeneity and 0 indicating perfect homogeneity. Other land use and
built environment measures were computed using geospatial Enhanced Point of Interest files obtained
from Desktop Mapping Technologies Inc. at a 250-meter (0.155 mile) buffer from each collision to
capture the context of the area where the collision occurred.

METHODOLOGY
This paper utilizes an ordered probit econometric structure in which the ordinal nature of the severity
levels of bicyclists are recognized at the level of individual collisions. The model assumes that there is a
latent continuous injury risk propensity metric underlying the observed ordinal responses. The continuous
variable $y_i^*$, albeit unobservable, can be written as a linear combination of predictors and an error term:

$$y_i^* = \beta X_i + \epsilon_i$$  \hspace{1cm} (1)

Where $y_i^*$ is the latent injury risk propensity for bicyclists $i$ in a given collision. $X_i$ corresponds to a set of
attributes associated with the collision, including personal, collision, and neighborhood characteristics. $\beta$
is a vector of unknown parameters to be estimated. $\epsilon_i$ is a random error term, which is assumed to follow
a normal distribution (i.e., a probit link), resulting in the ordered probit model examined in this paper.
The observed injury severity level, \( y_i \), takes on values 0 through \( m \) generating an ordered partitioning of the latent risk propensity into the observed severity categories according to the following scheme:

\[-\infty < \theta_1 < \theta_2 < \ldots < \theta_{m-1} < \infty \]  

(2)

Here, \( \theta \) represents threshold parameters in which \( \theta_0 = -\infty \) and \( \theta_m = \infty \). Hence, the observed injury severity levels can be represented as:

\[
\begin{align*}
  y_i^* &= 0 \text{ if } y_i^* \leq 0 \\
  &= 1 \text{ if } 0 < y_i^* \leq \theta_1 \\
  &= 2 \text{ if } \theta_1 < y_i^* < \theta_2 \\
  &\ldots \\
  &= m \text{ if } y_i^* > \theta_{m-1} \\
\end{align*}
\]

(3)

The estimation of this ordered probit model is straightforward. This model is an extension of a probit model for a binary outcome. Therefore, the probability of observing a particular ordinal outcome can be represented generically as:

\[
\text{Prob}(y_i = m) = \phi(\theta_m - \beta X_i) - \phi(\theta_{m-1} - \beta X_i)
\]

(4)

Assuming an indicator variable \( \psi_{im} \), which equals 1 if the bicyclist sustains an injury of level \( m \), and 0 otherwise, the log likelihood can be written as follows:

\[
\ln L = \sum_{i=1}^{n} \sum_{m=0}^{m} \psi_{im} \ln[\phi(\theta_m - \beta X_i) - \phi(\theta_{m-1} - \beta X_i)]
\]

(5)

This traditional ordered probit model, commonly used in accident research, restricts the thresholds \( \theta_m \) to be the same for every individual. Eluru et al. (17) argues that there will be several variables impacting injury risk propensity and several variables potentially influencing the thresholds in reality. Imposing a restriction of fixed \( \theta_m \) might lead to inconsistent injury risk propensity; thereby inconsistent effects of variables on the likelihood of severity level categories. Hence, several authors offered a generalized ordered probit model, allowing flexibility of varying thresholds (59,60). This generalized econometric structure assumes that threshold parameters can vary across collisions of different individuals due to both observed and unobserved factors. Let’s assume a specific functional form for the thresholds in order to constraint that all predicted probabilities are greater than zero and guarantee the ordering conditions (i.e., equation 2) for all data vectors. Then, the thresholds can be specified as:

\[
\theta_{im} = \exp(y_m + \delta' z_i)
\]

(6)

Where \( z_i \) is a set of exogeneous variables corresponding to the \( m^{th} \) threshold, \( \delta' \) represents parameters to be estimated, and \( y_m \) is a parameter associated with severity levels \( m = 1,2,\ldots,m \). Now, denoting \( F(\cdot) \) as the cumulative distribution of the standard logistic distribution, and \( \alpha_{im} \) as a dummy that exhibits the value 1 if the bicyclist \( i \) sustains an injury level of \( m \) and 0 otherwise, the log-likelihood function can be re-written for the \( i^{th} \) individual as:

\[
L_i = \int_\beta \int_\xi [F((\theta_{im} | \xi) - \beta X_i)] - F((\theta_{im-1} | \xi - \beta X_i)] \alpha_{im} \times g(\beta)g(\xi)d\beta d\xi
\]

(7)

Here, \( \beta \) and \( \xi \) are drawn from multivariate normal distributions \( g(\beta) \) and \( g(\xi) \). The overall log-likelihood function can be written as:
\[ L = \sum_i \ln L_i \] (8)

The parameters of this relaxed ordered probit formulation are estimated by maximizing the log-likelihood function of equation 8, and in relation to the moment parameters of the distributions \( g(\beta) \) and \( g(\xi) \).

Finally, the goodness-of-fit of the models are evaluated in terms of adjusted pseudo R-squared \((= 1 - (\text{logL}_{\text{constant only}} - Q)/\text{logL}_{\text{full model}}))\). Where \( Q \) is the number of parameters in the model.

**DISCUSSIONS OF RESULTS**

Table 1 shows the summary statistics of the independent variables retained in the final model specification. Three alternative model specifications of bicyclists’ injury severity were estimated. The first model is a conventional ordered probit model that includes personal and collision characteristic variables only (Model 1). The second model is the generalized ordered probit model described in the earlier section, which allows flexibility in the model assumption to incorporate individual variations in the thresholds to account for heterogeneity likely present in the data, but not commonly accommodated in traditional ordered probit models (Model 2). Parameter estimates of this model are reported with the same variables that are used in the first model specification for consistent comparisons. The third model, the same generalized ordered probit model, retained all variables from previous specifications but was enhanced by the inclusion of variables reflecting neighborhood and land use characteristics (Model 3). The model results suggest that the signs and approximate values of the estimated coefficients in the previous models remain stable and generally improve with each new model specification. Table 2 reports parameter estimation results of the three models outlined above.

Overall, Model 3 results tend to exhibit stronger relationships between the explanatory variables and levels of injury severity. It also outperforms the previous specifications by demonstrating better model fit, evaluated in terms of adjusted pseudo R-squared. Most importantly, this model includes land use and neighborhood attributes, mostly ignored in previous bicycle collision modeling research. Therefore, Model 3 is selected as the final model in this study.

The parameter estimation results suggest that personal and collision characteristics are strong factors in explaining bicyclist injury severity outcomes. Neighborhood characteristics are also found to add to the explained variance in bicyclist collisions, and overall, inclusion of these variables in Model 3 improves the explanatory power of the model. The majority of the independent variables retained in the model are statistically significant at least at the 90% confidence interval. Some variables exhibit a lower \( t \)-statistic but have been retained in the final model, with the presumption that a larger dataset would result in statistically significant parameters.

The majority of personal characteristics of the bicyclists involved in the collision yield statistically significant associations with sustaining an injury. Cyclists aged 45-54 involved in a collision have a positive relationship with the injury severity levels, implying that these groups are more likely to suffer severe injuries compared to other age groups. This result compliments other studies that find older adults to be more positively associated with injury severity \((4,5,8)\). Age itself may not be a causal factor but may be strongly associated with other relevant variables correlated with age. For example, perception and reaction time, physical fragility, and likelihood of existing medical conditions which come with age; all of these might contribute to higher injury risk propensity. Additionally, it is likely that the 45-54 age group represents the oldest age cohort of cyclists. In our model, females are associated with higher levels of injury severity. The presence of alcohol or drugs (represented by a dummy of impairment at the time of collision) shows a strong positive relationship with severity, suggesting that persons affected by alcohol, drugs or other substances have a higher likelihood of being injured or dying in a collision.

The collision characteristic variables were also found to be significant factors. Not surprisingly, collisions reported with view obstructions are very strongly associated with a higher likelihood of injury severity, indicating that sightlines have an effect on the probability of serious injury for cyclists.
Collisions occurring when a cyclist is making a lane change positively influence the probability of a more severe injury. This is likely attributed to increased interaction with motor vehicles. The configuration of the road has an influence on the bicyclists’ injury likelihood. Specifically, a collision in an intersection has a higher injury risk. These findings are likely attributed to vehicles underestimating the speed of cyclists or perhaps not expecting bicycles to be on the road. Additionally, cyclists are subjected to maneuvering through conflicting vehicular movements if they need to make a turn at intersections. Road grades, particularly steep roadways, were found to increase the likelihood of a higher injury severity level outcome. It is possible that steeper grades allow riders, especially those who are inexperienced, to build up speeds on steep descents which may create hazardous conditions for stopping or staying in control.

The model also suggests that there is a positive, yet relatively weak relationship between street lighting and increased injury severity. As expected, a positive relationship with injury severity is found when cyclists are ejected from their bicycles during a collision. Weekend collisions were found to result in a greater likelihood of increased injury severity. This finding is consistent with the literature, which indicates bicycle use as a leisure activity increases on weekends which perhaps results in higher absolute injury severity. The model found inclement weather conditions to be a strong predictor of injury severity. The effect of weather on injury severity is likely a result of reductions in visibility and traction. Reduced visibility due to inclement weather can lead to a more severe collision since it can distract or reduce perception of both cyclists and drivers which reduces their ability to respond (e.g. brake or take an evasive maneuver). A positive and relatively strong relationship is found with collisions occurring after dark. Certainly lighting condition is directly related with visibility which primarily affects the risk of collisions, but also affects severity due to lack of evasive action (e.g. driver did not see bicyclist) which leads to greater impact and thus severity. Inclement weather also makes roads and trails more slippery which can lead to more severe injury since braking and steering are suboptimal, leading to greater impact speeds and possible worse impact angles.

The presence of schools within the collision location is associated with an increase in injury severity. This variable may represent environments that are associated with higher levels of cyclist activity. Distance to the nearest shopping center was found to have a strong, positive relationship with injury severity. As distance increases, the likelihood of higher injury severity increases, indicating that collisions near shopping centers are less likely to result in higher levels of severity. Arguably, there are better cycling facilities located near shopping malls compared to outlying areas.

This paper examines a land use mix variable, defined as an index that ranges from 0 to 1, with 0 indicating perfect land use heterogeneity and 1 indicating perfect homogeneity. In general terms, land use heterogeneity is a spatial phenomenon in which a given area contains a high mix of land uses. An area with land use heterogeneity blends residential, commercial, industrial, government, park, and open space land uses. The model results reveal that the land use mix variable has a negative association with injury severity, indicating that severity risk is lower in relatively higher mixed land uses. Land use heterogeneity is touted as a viable planning option as it promotes greater housing variety and density, reduced travel distances, more compact development, and strong neighborhood character. Based on our observations, these environments typically include pedestrian and bicycle supportive design, which may be a reason why the likelihood of serious injury decreases in these locations. As expected, speed limits over 50 km/hour (31.1 miles/hour) are associated with an increase in the probability of injury severity. Collisions in areas with high population densities are associated with greater levels of injury severity. One likely explanation may be the high numbers of people who cycle in these areas. The model shows that average gross rent in the neighborhood is associated with a lower likelihood of injury severity. Presumably, affluent neighborhoods typically have better quality street lighting, roadway markings, and traffic calming measures that may contribute to the association of lower injury severity levels in these areas. On the other hand, average persons per household in the neighborhood negatively affect injury severity levels. This finding may be attributed to greater numbers of people cycling in these areas; or may also represent environments that are less supportive of safe cycling.

As discussed earlier, this paper examines several variables for the threshold parameter specification. Two major variables: cyclist wearing helmets at the time of collision (dummy) and cyclist
at fault in the collision (dummy) which exhibits 44% and 79% respectively for the sample collision records. The helmet variable exhibits a negative sign, indicating a downward shift of the threshold parameters. On the other hand, the cyclist at fault variable shows a positive relationship, an upward shift of the threshold parameters. These results confirm the hypothesis that the location of the set of thresholds varies across individuals.

Finally, several other variables were tested during model estimation but those hypotheses could not be confirmed due to lack of reasonable statistical significance and unexpected signs. For example, some neighborhood characteristics, such as intersection density, and institutional and commercial land use densities, yielded counter-intuitive results. These counter-intuitive results could be due to high correlations between the other built environment variables.

CONCLUSION

This paper presents the findings of a generalized ordered probit model that examines bicyclists’ injury severity levels, capturing the ordinal nature of injury severity and allowing flexibility by incorporating individual variations in the thresholds to account for heterogeneity. The analysis reveals the patterns of cyclist injury severity relative to personal, collision, land use, and neighborhood characteristics. The findings generally are consistent with the existing literature but offer some interesting insights to the role of neighborhood and land use attributes in explaining bicycle injury severity levels, which are limited in the existing literature.

There are many important empirical findings. The results reveal that females, impaired cyclists, and persons aged 45-54 involved in bicycle collisions have an increased likelihood of sustaining more severe injuries. The estimation results show that there are a number of important collision characteristics that increase the probability of more severe injuries: cycling manoeuvres (lane change), road conditions and configurations (at intersections and on steep road grades), and lighting conditions (after dark and when street lights are off) to be significant in explaining cyclist injury severity. Characteristics of the neighborhood in which collisions occur, specifically land use characteristics (heterogeneous land use), accessibility measures (presence of schools and distance to nearest shopping center), and household characteristics (average person per household and average gross rent) were found to be important predictors of bicyclists’ injury severity.

This study has certain limitations associated with the characteristics of the data. Collisions resulting in no or minor injury are likely to be underrepresented in the results which may skew injury severity levels toward more severe collisions. Unreported collisions could also be an issue. Furthermore, the study employs a relatively small sample size. Nevertheless, this study contributes in many ways. Previous research mainly focuses on motorists’ injury severity. Moreover, studies on injury severity levels that consider characteristics of the neighborhood in which they occur are surprisingly limited. This study contributes in understanding how land use and neighborhood characteristics influence injury severity levels for bicyclists. Our findings have important implications for engineering, enforcement, and education safety interventions. For example, older adults need to be educated about safety risks of cycling as they have been found to be more vulnerable to severe injuries when bicycle collisions occur. The findings can also inform the direction and prioritization of policy interventions. For example, one takeaway is the need to focus on visibility of cyclists (sight lines and lighting) and reduce the need and ability of cyclists to maneuver when traveling. Possible interventions could include bike boxes, shared roadway markings, cycle tracks, and colored bicycle lanes; facilitates which increase cyclist visibility and maneuverability. Finally, the contributions of this research are timely given the increased awareness and emphasis on the use of alternative modes of transportation, including bicycling in recent years. The finding of this paper will be valuable in road safety planning and policy discussions that would likely to encourage bicycle use in different regions.
ACKNOWLEDGEMENTS

The authors would like to thank the Natural Sciences and Engineering Research Council (NSERC), Nova Scotia Road Safety Advisory Committee (RSAC), and the Nova Scotia Department of Energy for their contribution in supporting the research.
REFERENCES


List of Tables

TABLE 1  Summary Statistics of Explanatory Variables used in the Generalized Ordered Probit Model for Bicyclists’ Injury Severity
TABLE 2  Parameter Estimation Results from Generalized Ordered Probit Model for Bicyclists’ Injury Severity
### TABLE 1 Summary Statistics of Explanatory Variables used in the Generalized Ordered Probit Model for Bicyclists’ Injury Severity

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Mean / Proportion</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 45-54</td>
<td>Cyclists between age of 45 and 54 (dummy)</td>
<td>11%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Gender - female (dummy)</td>
<td>23%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclist impairment</td>
<td>Cyclist impaired at time of collision (dummy)</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Collision characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intersection</td>
<td>Collision occurred in intersection with parking lot entrance/exit, private driveway or laneway (dummy)</td>
<td>14%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Road slope</td>
<td>Collision occurred on steep road grade (dummy)</td>
<td>25%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane change</td>
<td>Lane change manoeuver by cyclist at time of collision (dummy)</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>View obstructed</td>
<td>Cyclist view was obstructed at time of collision (dummy)</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclist ejected</td>
<td>Cyclist was ejected from bicycle (dummy)</td>
<td>31%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Street light off</td>
<td>Street lights not on at time of collision (dummy)</td>
<td>62%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After dark</td>
<td>Collision occurred after dark (dummy)</td>
<td>19%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekend</td>
<td>Collision occurred during the weekend (dummy)</td>
<td>22%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weather</td>
<td>Collision occurred during inclement weather (dummy)</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Neighborhood characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schools</td>
<td>At least one school present within the 250 meter buffer (0.155 mile) (dummy)</td>
<td>7%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shopping Centre</td>
<td>Distance to nearest shopping center (m)</td>
<td>12,531</td>
<td>30,373</td>
<td>36</td>
<td>197,725</td>
</tr>
<tr>
<td>Land use mix</td>
<td>An index, ranges from 0 to 1, with 1 indicating perfect land use heterogeneity and 0 indicating perfect homogeneity.</td>
<td>0.24</td>
<td>0.2</td>
<td>0</td>
<td>0.66</td>
</tr>
<tr>
<td>Road facility type</td>
<td>Collision occurred on a collector road (dummy)</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed limit</td>
<td>Collision occurred on road with speed limit &gt;50 km/hour (31.1 miles/hour)</td>
<td>10%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population density</td>
<td>Population density of the dissemination area where the collision occurred (persons per sq. km)</td>
<td>15.69</td>
<td>14.69</td>
<td>0.001</td>
<td>55.55</td>
</tr>
<tr>
<td>Average person per household</td>
<td>Average person per household (log) in the dissemination area where the collision occurred</td>
<td>2.81</td>
<td>0.26</td>
<td>2</td>
<td>4.5</td>
</tr>
<tr>
<td>Average gross rent</td>
<td>Average gross rent (log) in the dissemination area where the collision occurred ($)</td>
<td>442.69</td>
<td>321.44</td>
<td>0</td>
<td>1643</td>
</tr>
<tr>
<td><strong>Threshold covariates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helmet</td>
<td>Cyclist wearing helmet at time of collision (dummy)</td>
<td>44%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclist fault</td>
<td>Cyclist was at fault during time of collision (dummy)</td>
<td>79%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 2 Parameter Estimation Results from Generalized Ordered Probit Model for Bicyclists’ Injury Severity

<table>
<thead>
<tr>
<th></th>
<th>Model 1: Conventional Ordered Probit</th>
<th>Model 2: Generalized Ordered Probit without Neighborhood Characteristics</th>
<th>Model 3: Generalized Ordered Probit with Neighborhood Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coefficient</td>
<td>t-stat.</td>
<td>coefficient</td>
</tr>
<tr>
<td><strong>Personal characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 45-54</td>
<td>0.36850114</td>
<td><strong>2.072</strong></td>
<td>0.39367546</td>
</tr>
<tr>
<td>Female</td>
<td>0.10762794</td>
<td>0.826</td>
<td>0.10580487</td>
</tr>
<tr>
<td>Cyclist impairment</td>
<td>0.6799049</td>
<td><em>1.864</em></td>
<td>0.64417965</td>
</tr>
<tr>
<td><strong>Collision characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intersection</td>
<td>0.25052195</td>
<td>1.562</td>
<td>0.25815015</td>
</tr>
<tr>
<td>Road slope</td>
<td>0.24521584</td>
<td>1.93</td>
<td>0.23255537</td>
</tr>
<tr>
<td>Cyclist manoeuvre (lane change)</td>
<td>0.67397524</td>
<td><em>1.811</em></td>
<td>0.67366255</td>
</tr>
<tr>
<td>View obstructed</td>
<td>1.33739608</td>
<td><strong>3.482</strong></td>
<td>1.3311095</td>
</tr>
<tr>
<td>Cyclist ejected from bicycle</td>
<td>0.44713461</td>
<td><strong>3.779</strong></td>
<td>0.43204158</td>
</tr>
<tr>
<td>Street light off</td>
<td>0.16827643</td>
<td>1.407</td>
<td>0.16697159</td>
</tr>
<tr>
<td>After dark</td>
<td>0.4206568</td>
<td><strong>2.755</strong></td>
<td>0.4177134</td>
</tr>
<tr>
<td>Weekend</td>
<td>0.26242587</td>
<td><em>1.958</em></td>
<td>0.26413158</td>
</tr>
<tr>
<td>Weather</td>
<td>0.6276363</td>
<td>1.287</td>
<td>0.59065724</td>
</tr>
<tr>
<td><strong>Neighborhood characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence of schools</td>
<td></td>
<td></td>
<td>0.00014964</td>
</tr>
<tr>
<td>Distance to nearest shopping center</td>
<td></td>
<td></td>
<td>0.19936163</td>
</tr>
<tr>
<td>Land use mix</td>
<td></td>
<td></td>
<td>-0.00022238</td>
</tr>
<tr>
<td>Road facility type (collector road)</td>
<td></td>
<td></td>
<td>0.53980227</td>
</tr>
<tr>
<td>Speed limit &gt;50 km/hour (31.1 miles/hour)</td>
<td></td>
<td></td>
<td>0.29969097</td>
</tr>
<tr>
<td>Population density</td>
<td></td>
<td></td>
<td>0.00005584</td>
</tr>
<tr>
<td>Average gross rent (log)</td>
<td></td>
<td></td>
<td>-0.0436924</td>
</tr>
<tr>
<td>Average person per household (log)</td>
<td></td>
<td></td>
<td>0.04402381</td>
</tr>
<tr>
<td><strong>Threshold parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theta(1)</td>
<td>0.80498049</td>
<td>12.89</td>
<td>-0.27855624</td>
</tr>
<tr>
<td>Theta(2)</td>
<td>2.77491176</td>
<td>23.692</td>
<td>0.97111263</td>
</tr>
<tr>
<td>Theta(3)</td>
<td>3.79534857</td>
<td>15.981</td>
<td>1.28331608</td>
</tr>
<tr>
<td><strong>Threshold covariates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helmet</td>
<td></td>
<td></td>
<td>-0.08069972</td>
</tr>
<tr>
<td>Cyclist fault</td>
<td></td>
<td></td>
<td>0.11989971</td>
</tr>
<tr>
<td>Constant</td>
<td>0.3851968</td>
<td>3.026</td>
<td>0.39059542</td>
</tr>
<tr>
<td>Pseudo R-squared</td>
<td>0.0590564</td>
<td>0.0615906</td>
<td>0.0780812</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>Number of observations</td>
<td>425</td>
<td>425</td>
<td>425</td>
</tr>
</tbody>
</table>

**95% confidence interval; *90% confidence interval