Accident Analysis and Prevention 97 (2016) 19–27

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Investigating the gender differences on bicycle-vehicle conflicts at urban intersections using an ordered logit methodology

Joshua Stipancic[a,*], Sohail Zangenehpour[a], Luis Miranda-Moreno[b], Nicolas Saunier[c], Marie-Axelle Granié[d]

[a] Department of Civil Engineering and Applied Mechanics, McGill University, Room 391, Macdonald Engineering Building, 817 Sherbrooke Street West, Montréal, Québec, H3A 0C3, Canada
[b] Department of Civil Engineering and Applied Mechanics, McGill University, Room 268, Macdonald Engineering Building, 817 Sherbrooke Street West, Montréal, Québec, H3A 0C3, Canada
[c] Department of Civil, Geological and Mining Engineering, Polytechnique Montréal, C.P. 6079, succ. Centre-Ville, Montréal, Québec, H3C 3A7, Canada
[d] IFSTTAR-TS2-LMA, 304 Chemin de la Croix Blanche, F-13300 Salon de Provence, France

ARTICLE INFO

Article history:
Received 26 April 2016
Received in revised form 13 July 2016
Accepted 25 July 2016

Keywords:
Discrete choice
Cycling
Surrogate safety
Gender
Urban
Order logit

ABSTRACT

In the literature, a crash-based modeling approach has long been used to evaluate the factors that contribute to cyclist injury risk at intersections. However, this approach has been criticized as crashes are required to occur before contributing factors can be identified and countermeasures can be implemented. Moreover, human factors related to dangerous behaviors are difficult to evaluate using crash-based methods. As an alternative, surrogate safety measures have been developed to address the issue of reliance on crash data. Despite recent developments, few methodologies and little empirical evidence exist on bicycle-vehicle interactions at intersections using video-based data and statistical analyses to identify associated factors. This study investigates bicycle-vehicle conflict severity and evaluates the impact of different factors, including gender, on cyclist risk at urban intersections with cycle tracks. A segmented ordered logit model is used to evaluate post-encroachment time between cyclists and vehicles. Video data was collected at seven intersections in Montreal, Canada. Road user trajectories were automatically extracted, classified, and filtered using a computer vision software to yield 1514 interactions. The discrete choice variable was generated by dividing post-encroachment time into normal interactions, conflicts, and dangerous conflicts. Independent variables reflecting attributes of the cyclist, vehicle, and environment were extracted either automatically or manually. Results indicated that an ordered model is appropriate for analyzing traffic conflicts and identifying key factors. Furthermore, exogenous segmentation was beneficial in comparing different segments of the population within a single model. Male cyclists, with all else being equal, were less likely than female cyclists to be involved in conflicts and dangerous conflicts at the studied intersections. Bicycle and vehicle speed, along with the time of the conflict relative to the red light phase, were other significant factors in conflict severity. These results will contribute to and further the understanding of gender differences in cycling within North America.

1. Introduction

Road safety is a substantial concern for transportation professionals due to the high economic and social cost of traffic crashes (Abdel-Aty, 2003). In 2012, traffic crashes resulted in 2077 fatalities and over 165,000 injuries in Canada, where cyclists and pedestrians account for approximately 18% of both fatalities and injuries annually (Transport Canada, 2014). While the safety of motorists has commanded much attention, the protection of vulnerable road users has become common only recently (Kockelman and Kweon, 2002). In North America, cyclists are twelve times more likely to be killed than motor vehicle drivers (Moore et al., 2011; Strauss et al., 2014), and 269 cyclist fatalities occurred in Canada between 2008 and 2012 (Transport Canada, 2014).

Recent growth in bicycle activity and infrastructure improvement have increased awareness of bicycle safety issues in North America. In this context, the complex nature of cyclist-vehicle inter-

* Corresponding author.
E-mail addresses: joshua.stipancic@mail.mcgill.ca (J. Stipancic), sohail.zangenehpour@mail.mcgill.ca (S. Zangenehpour), luis.miranda-moreno@mcgill.ca (L. Miranda-Moreno), nicolas.saunier@polymtl.ca (N. Saunier), marie-axelle.granie@ifsttar.fr (M.-A. Granié).

http://dx.doi.org/10.1016/j.aap.2016.07.033
0001-4575/

actions must be understood by examining factors that contribute to cyclist safety (Klop and Khattak, 2007; Moore et al., 2011). One method for evaluating these factors is traffic crash modelling. Crash models evaluate collision frequency using count regression models, or injury severity using different techniques such as ordered logit, discrete choice regression models, and tree-based or neural network techniques. Many studies have considered the different statistical methods available. A recent summary of literature on crash frequency and severity modelling is provided by Mannering and Bhat (2014). Additional summaries were compiled by Miranda-Moreno (2006), Lord and Mannering (2010), and Savolainen et al. (2011). Traditional safety models calibrated with crash data are reactive, requiring crashes to occur before causes can be identified and countermeasures can be implemented. Additionally, crash-based methods require long observation periods, particularly when cyclists are involved. Given the low rate of crash occurrence for non-motorized users, many years of accident data are required to conduct a safety analysis. Finally, the study of dangerous behaviors and other human factors can often not be investigated.

As an alternative or complementary approach, surrogate safety techniques analyze interactions and conflicts rather than crashes. Conflicts are events that are physically and predictably related to traffic crashes, and are placed immediately below collisions in Hydén's model (Hydén, 1987), presented in Fig. 1. As the pyramid model suggests, traffic interactions are inherently ordered through their proximity to a potential collision, or severity, as measured by various indicators. Modelling interactions rather than collisions provides several benefits. Interactions occur much more frequently than collisions and statistically sufficient data can be collected in a shorter time period. The use of interactions and conflicts is proactive, rather than reactive, surrogate measures are insensitive to crash underreporting (Kockelman and Kweon, 2002), and human factors can be incorporated in the analysis. However, despite the benefits offered by a proactive surrogate approach, dangerous cyclist behaviour in vehicle-bicycle interactions, and their associated factors, have rarely been studied using video analysis, surrogate measures of safety, and regression models.

Existing efforts in injury severity modelling primarily concern motor vehicle occupants in single or multiple vehicle crashes. Multiple modeling techniques have been proposed to investigate the relationship between injury severity levels and associated factors. Mannering and Bhat (2014) provide a comprehensive literature review on the methods used in severity analysis in general, and Eluru et al. (2008) provides a summary of literature concerning analysis of cyclist injuries. Much empirical evidence has been reported in the literature identifying the key factors influencing injury severity, including characteristics of the roadway, environment, and road user, for passengers, drivers, and pedestrians. Moore et al. (2011) utilized a multinomial logit and mixed logit to model cyclist injury severity at intersections and non-intersection locations using seven years of crash data from Ohio. The study stated "the injury mechanisms are substantially different [...] at intersection and non-intersection locations". Klop and Khattak (2007) studied bicycle crash severity on rural roads in North Carolina using an ordered probit model and four years of crash data. Visibility and weather conditions were found to most significantly increase collision severity. Eluru et al. (2008) utilized data from the 2004 US national database to estimate a mixed generalized ordered response logit (MGORL) model of cyclist injury severity. The MGORL generalizes more standard ordered models, providing additional flexibility across observations. Cyclist age and vehicle speed were correlated with injury severity. In the existing literature, gender is often a contributory factor to injury severity.

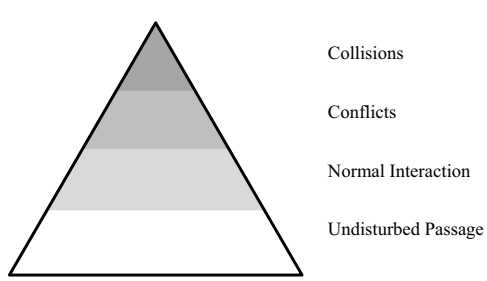

Fig. 1. Safety pyramid based on Hydén's classic model (Hydén, 1987).

The relationship between biological gender and cyclist safety is a particular issue that remains to be investigated using microscopic conflict data. More specifically, it is not that gender itself is necessarily a deciding factor, but rather it is associated with unobserved differences in behaviour, physiology, and experience (which themselves are gender-related) that can be captured using video data. Important behavioral differences have been identified for both pedestrians and motorists of different genders at intersections (Holland and Hill, 2007; Santamarina-Rubio et al., 2014; Tom and Granie, 2011), yet very little is known about this for cyclists. For now, studies on gender differences among cyclists have focussed largely on variation in behaviour or preference. Johnson et al. (2011) studied red-light compliance of cyclists in Australia and found that male cyclists were more frequently non-compliant than females. Bernhoft and Carstensen (2008) used a survey approach, finding male cyclists in Denmark tended to act less cautiously, but also felt safer than their female counterparts. Similarly, French studies show that male cyclists tended to overestimate their ability more so than females whereas females tend to overestimate their carefulness more than males (Felonneau et al., 2013) and that cyclist risk-taking behaviour seems to be gender specific even at an early age (Granié, 2011). Gender differences and their associated risks may help to explain the disparity in ridership currently experienced in cities (Garrard et al., 2012), as well as the implementation of the appropriate designs and facilities.

However, while gender differences are evident in terms of behaviour, behavioural differences do not necessarily create disparity in risk, and recent research has only shown minor differences in actual crash risk for males and females (Kaplan et al., 2014; Martinez-Ruiz et al., 2014). Kaplan et al. (2014) evaluated injury severity using a generalized ordered logit model on Danish crash data collected over a 5-year period, and found no correlation between gender and injury risk. Martinez-Ruiz et al. (2014) used 17 years of crash data from Spain to calculate the crash rate ratio by gender for different ages. Without adjusting for cycling exposure, males were more likely to be involved in a collision than females. When controlling for exposure, crash rates were approximately equal for males and females with age being a stronger determinant of crash risk. Despite parity in risk, in most countries, females still cycle less than males, particularly in English speaking countries that are less bicycle-friendly, including Canada, the U.K., Australia, and the U.S. (Garrard et al., 2012). Route conditions and vehicle interactions greatly influence individuals' likelihood to cycle (Winters et al., 2011) and female cyclists, much more than males, prefer to use routes with maximum separation from motorized traffic (Garrard et al., 2008) and signalized crossings (Bernhoft and Carstensen, 2008), whereas males prefer the fastest routes (Bernhoft and Carstensen, 2008). Studies in the city of Montreal show that routes with separated cycle tracks attract higher cyclist volumes than those without (Strauss and Miranda-Moreno, 2013) and that streets with cycle tracks have a lower injury risk (Lusk et al., 2011). As female cyclists seem to favour sites with cycle tracks (Garrard et al., 2008) a possible explanation for this lack of difference in the risk of accidents could be found in the behavior of male and female cyclists on roads with cycle tracks.

The purpose of this study is to estimate a segmented ordered logit model for bicycle-vehicle interactions at urban intersections with cycle tracks. Intersections, being one of the most dangerous locations in the urban network, deserve specific attention (Wang and Abdel-Aty, 2008). The objectives of this research are to investigate bicycle-vehicle conflicts and contributory factors using an ordered-logit modelling approach and to examine the influence of gender and other cyclist, vehicle, and environmental factors on conflict severity. While research on severity modelling includes consideration for cyclists and advanced formulations, few studies have attempted to apply modeling techniques to surrogate measures for bicycle safety (Zangenehpour et al., 2013, 2015b). Analysis of differences between male and female cyclist risk by surrogate safety techniques has also been limited. Rather than using independent models for males and females, a segmented modelling approach is preferred in order to increase the number of observations available for estimation and to make differences between genders instantly comparable.

2. Methodology

The methodology consists of four steps; video data collection and processing using video tracking and classification methods; definition and computation of surrogate safety indicators; model formulation and estimation, and; site selection. The details of each step are provided in the following sections.

2.1. Video data collection and processing

The methodology for data collection and processing uses a similar approach to the one implemented in several past studies (Jackson et al., 2013; Saunier et al., 2010; Sayed et al., 2013). Video data was collected using an inexpensive and commercially available video camera which stores video and is powered internally. The camera is mounted using a telescoping fibreglass mast to ensure a clear view above the intersection and approaches. The camera system is introduced in detail by Zangenehpour et al. (2015a). Once collected, the video data was processed using an open-source road user tracking software, Traffic Intelligence (Jackson et al., 2013; Saunier et al., 2010; Saunier, 2015). Road user trajectories (position and speed at each frame) were automatically extracted from the video footage. Road users were then classified as pedestrians, cyclists, or vehicles using a technique developed by Zangenehpour et al. (2014), which is now available in Traffic Intelligence, and the trajectories were filtered to isolate the desired interactions. Filtering is completed by designating a start and end zone for each road user type, and considering only those trajectories which begin and end in the desired zones. Any erroneous interactions were removed manually from the data set. In this study, approximately 15% of all interactions identified automatically were deemed unclear or non-existent.

2.2. Computation of surrogate safety measures

Popular surrogate measures of safety include time-to-collision (TTC) and post-encroachment time (PET). TTC is "the time required for two vehicles to collide if they continue at their present speed and on the same path" (van der Horst et al., 2014) or more generally if their movements remain unchanged, which can include variations in speed and direction. PET is the difference in time between two road users occupying the same location in space, or the potential conflict point (Peesapati et al., 2013). TTC is measured continuously and, depending on the choice or motion prediction method, will yield several measurements over time when there is a collision course (when some predicted trajectories would lead the road users to collide). PET is based on observed trajectories and can be computed only if trajectories intersect. Both measures can usually be computed for the same interaction and are complementary in the analysis of bicycle conflicts (van der Horst et al., 2014). Surrogate safety data has been successfully collected using video-based detection systems and extracted using computer vision techniques (Jackson et al., 2013; Sayed et al., 2013). PET was selected as the surrogate measure of safety in this research as all interactions involve intersecting trajectories. PET is better suited to interactions involving turning movements than common TTC with assumption of constant velocity and it is simpler and faster to compute than TTC with more realistic motion prediction methods.

Once trajectories were extracted, PET calculation for each interaction was automated using a Python script to count the number of frames (converted to seconds) between consecutive road users occupying the conflict point. The dependent choice variable was created by classifying PET as:

- ◯ Alternative 0: Normal interaction with PET greater than 5 s;
- ◯ Alternative 1: Conflict with PET between 3 and 5 s, and;
- ◯ Alternative 2: Dangerous conflict with PET less than 3 s.

These thresholds have been used successfully in previous work, such as Zangenehpour et al. (2015b), where readers are referred for more details on the determination of threshold values. This formulation yields results that are easy to interpret, as positive coefficients indicate an increase in the severity of conflicts (decrease in PET) as the variable associated with that coefficient increases (Kockelman and Kweon, 2002).

2.3. Modelling framework definition and model estimation

Ordered response models are the most widely used statistical models in crash severity analysis. Unlike other multiple choice models that include utility variables for each alternative, ordered models have only one propensity, or latent variable (O'Donnell and Connor, 1996). The ordered logit model assumes that the error term, representing the unobserved component of the latent variable, is logistic distributed with a mean of zero and a variance of $\pi^2/3$. The latent variable (outcome), y^*, is defined by

$$y* = \beta X + \varepsilon (1)$$

where $\boldsymbol{X}$ is the vector of explanatory variables (in this case, user, traffic, and built environment factors), β is the vector of unknown parameters, and ε is the logistic distributed random error term. This propensity is bound by unknown thresholds, τ_i, which delineate alternatives. For the case presented herein, the probability of each choice alternative (in this case, three conflict types by severity) is:

$$\begin{aligned} P(y=0) &= CDF\left(\tau_1 - \beta\boldsymbol{X}\right)(\text{for normal interactions}) \\ P(y=1) &= CDF\left(\tau_2 - \beta\boldsymbol{X}\right)(\text{for conflicts}) \\ P(y=2) &= 1 - CDF\left(\tau_2 - \beta\boldsymbol{X}\right)(\text{for serious conflicts}) \end{aligned} \quad (2)$$

where CDF is the cumulative distribution function of the logistic distribution defining ε.

Exogenous segmentation is a method that provides heterogeneity in the coefficients across multiple segments, working particularly well when segments are few (Bhat, 1997). For this study, segmentation allows for variation in the value of the parameters across cyclist gender. Segmentation is achieved through the generation of new independent segmented variables, which are the product of the existing independent variables and some discrete variable (gender). When the model is estimated, the parameter on the basic variables represents a contribution to both genders, while the parameters on the segmented variables represent the variation

between genders. The propensity for the segmented model can be represented by

$$y* = \beta_1 X + \beta_2 wX + \beta_3 w + \varepsilon (3)$$

where $\boldsymbol{X}$ are the unsegmented variables and β_1 are their coefficients, $\boldsymbol{wX}$ and β_2 are the segmented variables and their coefficients, respectively, and $\boldsymbol{w}$ and β_3 are the segmentation variable (gender) and its coefficient, respectively.

The conflict severity model was developed in three steps, with each step incorporating additional gender heterogeneity. An ordered logit model (OL1) was estimated assuming homogeneity between genders. This model included all available variables without consideration for gender. With the addition of the gender, OL2 was created. The inclusion of this variable introduced simplistic heterogeneity, by allowing the latent propensity to change according to cyclist gender. The final segmented ordered logit model (SOL) included all available explanatory variables with segmentation according to gender. This allowed not only the propensity, but also the value of each parameter, to vary between genders. While similar results can be achieved through the estimation of separate models, segmentation increases the number of observations for estimation and makes differences between genders instantly comparable.

2.4. Site selection

Seven test sites were selected along Maisonneuve Boulevard, a one-way urban arterial in downtown Montreal, Canada. The sites, shown in Fig. 3, featured two travel lanes and a fully separated bidirectional cycling facility, or cycle track, along Maisonneuve. In Canada, motorists operate on the right side of the road, and turning right on a red light is prohibited on the island of Montreal. The sites along Maisonneuve were selected specifically because of the present cycling infrastructure. The two-way separated cycle track is a relatively new design that has only been studied more recently. Maisonneuve is also one of the busiest cycling corridors on the island of Montreal, and the specific intersections were chosen because of their geometric uniformity. On all the observed intersections, the cycling lane is on the left of the two other travel lanes. Video data was collected using a single camera at each site with a clear view above the intersection. Approximately 4 h of video was collected at each site in the afternoon peak period from June 16 to June 20, 2014, ensuring that traffic volume, weather, time of year, and time of day remained consistent.

Similar to Wang and Abdel-Aty (2008), a single traffic interaction was analyzed. Conflicts were limited to those between cyclists traveling straight through the intersection and left-turning vehicles. An example of a dangerous conflict is provided in Fig. 2. Given the site geometry and design of the cycling infrastructure, these interactions represent the majority of safety concerns since cyclists wishing to go through the intersection should consider vehicles located on their right wishing to turn left. Any erroneous interactions were removed manually from the data set. Several interactions were also missed by the tracking algorithm, which could not be included as part of this study. Once filtered, 1514 useable observations remained for analysis.

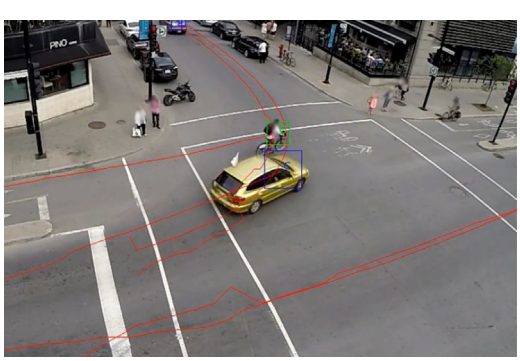

Fig. 2. Example of a dangerous bicycle-vehicle conflict.

3. Results

3.1. Data exploration

Attributes of the cyclist, vehicle, and environment were compiled as independent variables. Cyclist and vehicle speed at the conflict point were automatically computed from the trajectories. Additional explanatory variables, including cyclist gender helmet use, whether the vehicle was a truck, van, or SUV, and whether the vehicle was the first in a platoon, were extracted manually from the video. As in previous studies (Johnson et al., 2011), gender was classified based on physical appearance. The environmental attributes included variables representing if the interaction occurred immediately after a red light phase (both cyclist and vehicle were simultaneously waiting at a red light before the interaction occurred) and site-specific constants for each site, with Crescent used as the reference. Dummy variables were used to indicate whether the cyclist was the first object to reach the conflict point and whether pedestrians (or contra-flow cyclists) were present in the intersection at the time of the conflict, capturing the behavioural freedom of the road users. A summary is provided in Table 1. Site specific data has been omitted for brevity.

A preliminary exploration of the raw variable data was conducted with descriptive statistics. Of the 1514 observed interactions, only 496 involved female cyclists, reflecting the disparity in number of cyclists by gender. The data set contained 33% female cyclists, which is close to the share of 37% female cyclists reported by Vélo Québec in 2010 (Vélo Québec, 2010). The distribution of conflicts was determined for males and females, and is presented below in Fig. 4.

The proportion of conflicts, dangerous conflicts, and normal interactions were determined for males and females and are provided in Fig. 5. As the distributions seem to be drawn from the same underlying distribution given the result of the chi square test at 95% confidence (test statistic of 1.23 compared to the critical value from the chi-square distribution of 5.99), these results point towards equivalency in terms of safety for male and female cyclists. The data exploration provided little evidence of differences between genders in the considered variables.

3.2. Model estimation

The models were estimated using all available parameters and systematically eliminating statistically insignificant variables. In the results that follow, variables retained were significant at 80% confidence or greater. Variables significant at 90% confidence are italicized, and variables significant at 95% confidence are bolded. Results for the ordered logit models are presented in Table 2. In Model OL1, which assumed gender homogeneity, both cyclist and vehicle speed were found to be significant, along with the red light variable and several site specific constants. A log-likelihood ratio test was used to compare OL1 to the constants-only model. The null hypothesis was that none of the variables help to explain conflict severity. LR for Model OL1 was 55.72, compared to the critical value from the chi-square distribution of 14.07 (7 ◦ of freedom at 95 % confidence). The null hypothesis was rejected, and OL1 was superior. Adding the gender variable in Model OL2 produced only minor changes in the model results. Although the male dummy variable

Fig. 3. Study sites along Maisonneuve at Crescent (a), Stanley (b), Peel (c), Mackay (d), Metcalfe (e), St Denis (f), and Union (g).

Table 1
Variables and Statistics for All Captured Interactions.

	Units	Description	Mean	Min.	Max.	Std. Dev.
Cyclist						
Speed	km/h	Speed of the cyclist at the conflict point	13.16	1.71	43.51	5.39
Male	Dummy	Cyclist was male (1) or female (0)	0.67	0	1	0.47
Helmet	Dummy	Cyclist wearing a helmet (1) or not (0)	0.51	0	1	0.50
Vehicle						
Speed	km/h	Speed of the vehicle at the conflict point	20.01	2.00	67.10	6.43
Platoon Lead	Dummy	Vehicle was the first vehicle in a series of successive vehicles	0.76	0	1	0.43
Truck	Dummy	Vehicle was a truck, van, or SUV	0.33	0	1	0.47
Environment						
Red	Dummy	Conflict occurred immediately after a red light	0.16	0	1	0.37
Bike First	Dummy	The cyclist reached the conflict point first	0.59	0	1	0.49
Pedestrian	Dummy	Peds were simultaneously crossing intersection	0.46	0	1	0.50

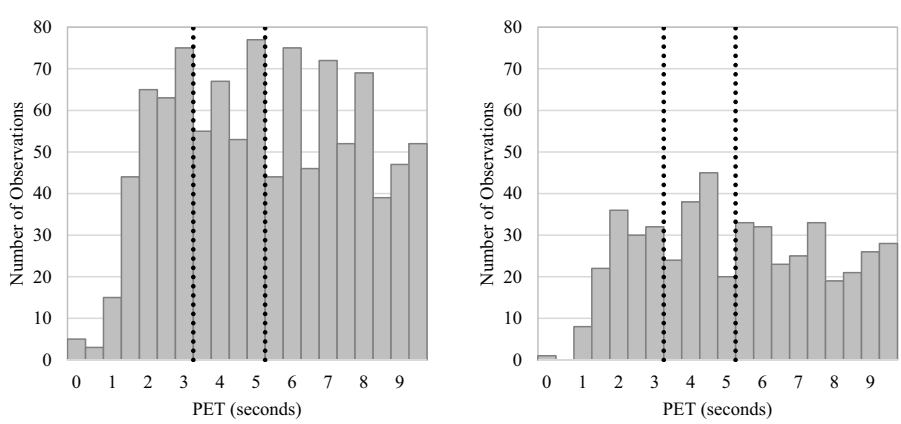

Fig. 4. Distribution of PET for male cyclists (a) and female cyclists (b).

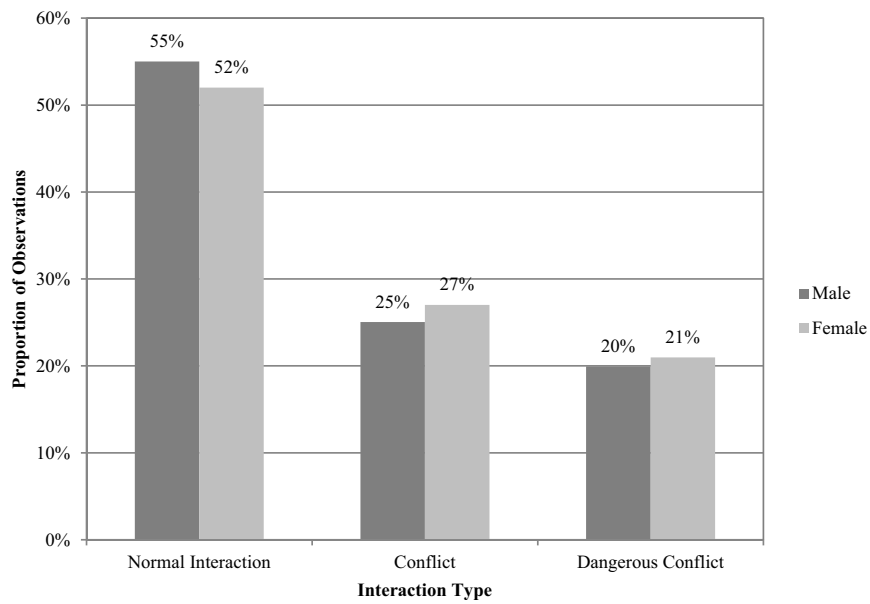

Fig. 5. Proportion of Interaction Type According to Gender.

Table 2
Model Results for Ordered Logit Models, OL1 and OL2.

	OL1			OL2		
	$y^* = \beta_1 \boldsymbol{X} + \varepsilon$			$y^* = \beta_1 \boldsymbol{X} + \beta_3 \boldsymbol{w} + \varepsilon$		
Explanatory variables	**Parameter**	**z stat**	**p value**	**Parameter**	**z stat**	**p value**
Bike Speed	**0.0193**	**2.07**	**0.039**	**0.0190**	**2.03**	**0.042**
Helmet	–	–	–	–	–	–
Vehicle Speed	**0.0187**	**2.31**	**0.021**	**0.0194**	**2.38**	**0.017**
Truck/Van	–	–	–	–	–	–
Platoon Leader	–	–	–	–	–	–
Red	**−0.7420**	**−4.78**	**0.000**	**−0.7460**	**−4.80**	**0.000**
Bike First	–	–	–	–	–	–
Pedestrian	–	–	–	–	–	–
Stanley	**−0.3477**	**−2.36**	**0.018**	**−0.3590**	**−2.43**	**0.015**
Peel	–	–	–	–	–	–
Mackay	−0.2505	−1.49	0.136	−0.2451	−1.46	0.145
Metcalfe	−0.2103	−1.54	0.123	−0.2123	−1.56	0.119
Denis	–	–	–	–	–	–
Union	*−0.4244*	*−1.80*	*0.072*	*−0.4313*	*−1.83*	*0.067*
Male	*N/A*	*N/A*	*N/A*	−0.1459	−1.38	0.169
Tau 1	0.5381			0.4482		
Tau 2	1.7778			1.6890		
Number of cases	1514			1514		
Log likelihood at convergence	−1494.23			−1493.29		
Log likelihood for constants- only model	−1522.09			−1522.09		
Pseudo R^2	0.0183			0.0189		

is negative, which indicates that males have a lower propensity for conflicts than females with all else being equal, the gender variable is not statistically significant (only 80% confidence), Model OL2 was inferior to OL1 by log-likelihood ratio test (LR of 1.88 compared to a chi-square value of 3.84). Based on this result, any evidence for variation in conflict severity between males and females is inconclusive.

Results for the SOL model are presented in Table 3. Non-significant β_1 coefficients indicate that a parameter is not significantly correlated with conflict severity for female cyclists. Non-significant β_2 coefficients indicate that parameter is homogeneous across genders. This formulation eliminates unnecessary variables and allows for immediate testing of differences between genders. Importantly, gender was found to be highly significant in this model. The SOL model was compared to OL1 using a log-likelihood ratio test. The null hypothesis was that safety is homogeneous across genders and segmentation is unnecessary. LR was calculated to be 11.08, compared to the critical value from the chi-square distribution of 5.99 (2° of freedom at 95% confidence). The null hypothesis was falsified and segmentation is justified. From the results of the OL models, gender appears to have no effect on safety. However, it is not that there is no effect; it is that variation exists both in propensity and in the value of the parameters. Allowing for this variation is necessary to reveal gender's true effect, representing an ideal scenario for an exogenously segmented model.

3.3. Model interpretation

Interpretation of the SOL model is facilitated by calculating the true value of each coefficient for both genders. For females, coefficients take the β_1 values from the model estimation. For males,

Table 3
Model Results for Segmented Ordered Logit Model, SOL.

	β_1			β_2		
	$y^* = \beta_1 \boldsymbol{X} + \beta_2 \boldsymbol{wX} + \beta_3 \mathbf{w} + \varepsilon$					
Explanatory variables	Parameter	z stat	p value	Parameter	z stat	p value
Bike Speed	–	–	–	**0.0272**	**2.31**	**0.021**
Helmet	–	–	–	–	–	–
Vehicle Speed	–	–	–	**0.0250**	**2.38**	**0.017**
Truck/Van	–	–	–	–	–	–
Platoon Leader	–	–	–	0.2395	1.63	0.104
Red	**−0.7713**	**−4.99**	**0.000**	–	–	–
Bike First	–	–	–	–	–	–
Pedestrian	–	–	–	–	–	–
Stanley	**−0.3774**	**−2.56**	**0.010**	–	–	–
Peel	–	–	–	–	–	–
Mackay	–	–	–	**−0.4946**	**−2.41**	**0.016**
Metcalfe	*−0.2384*	*−1.75*	*0.080*	–	–	–
Denis	–	–	–	–	–	–
Union	−0.8953	−2.21	**0.027**	0.6657	1.35	0.178
	β_3			β_3		
Male	−1.1703	−3.79	**0.000**	−1.1703	−3.79	**0.000**
Tau 1	−0.2007					
Tau 2	1.0455					
Number of cases	1514					
Log likelihood at convergence	−1488.69					
Log likelihood for constants- only model	−1522.09					
Pseudo R^2	0.0219					

Table 4
Coefficient Values for Male and Female Cyclists.

Explanatory variables	Female	Male
	β_1	$\beta_1 + \beta_2$
Bike Speed	–	0.0272
Helmet	–	–
Vehicle Speed	–	0.0250
Truck/Van	–	–
Platoon Leader	–	0.2395
Red	−0.7713	−0.7713
Bike First	–	–
Pedestrian	–	–
Stanley	−0.3774	−0.3774
Peel	–	–
Mackay	–	−0.4946
Metcalfe	−0.2384	−0.2384
St Denis	–	–
Union	−0.8953	−0.2296
Male	−1.1703	−1.1703

Table 5
Sample probabilities of conflict severity for male and female cyclists.

	Crescent		Stanley		Union	
	Female	Male	Female	Male	Female	Male
P (Normal Interaction)	55%	63%	64%	71%	75%	68%
P (Conflict)	25%	22%	22%	18%	16%	20%
P (Dangerous Conflict)	20%	15%	14%	11%	9%	12%

the β_1 and β_2 parameters are summed to obtain the true value, shown in Table 4. Variables related to cyclist attributes led to several interesting results. The primary observation is the negative, highly significant coefficient for the gender variable. In general, this coefficient had the largest magnitude. This result suggests that, all else being equal, when interacting with vehicles, males are less likely than females to have conflicts and dangerous conflicts. This result implies that males are less likely to be involved in more severe interactions, as defined by the PET intervals, and it can be concluded that male cyclists who interact with vehicles are safer than female cyclists who interact with vehicles. A second important observation is that the coefficient for bike speed is positive, but only for male cyclists. For females, cyclist speed has no effect on their potential for conflict at the studied sites. For male cyclists, travelling at a faster speed increases the propensity for conflicts and dangerous conflicts.

With regards to vehicle attributes, increasing vehicle speed increased conflict propensity, again for males only. Additionally, males were more likely to engage in a conflict with the first vehicle in the platoon, although the strength of this result was relatively low. The only significant environmental factors were the red light variable and several site-specific constants. With regards to the site-specific variables, Crescent, Peel, and St Denis showed no significant variation in conflict occurrence. Both males and females had a lower propensity for conflicts at Stanley, Metcalfe, and Union (although the strength of the effect at Union varied by gender). Only males had a lower propensity for conflicts at Mackay, while for females the site was no different than the reference site at Crescent.

Sample probabilities of conflict severity were calculated for several sites, Crescent, Stanley, and Union to demonstrate the impact of the modelling results. The cases assumed cyclist and vehicle travelling at the respective mean speeds, with all other binary variables set to zero. Sample probabilities for normal interaction, conflict, and dangerous conflict were calculated for males and females and are provided in Table 5.

These results, which are a similar to observed percentages, show that males were safer at both Crescent and Stanley, as their probability of normal interaction was higher, and the probabilities of conflict and dangerous conflict were both lower. In fact, males were safer in many conditions, because any variables with positive coefficients were offset by the large negative coefficient on the male dummy variable. At Union, females were observed to be safer than males. However, this was the only one of the seven sites where this was true. The effect of the highly positive male variable can only be offset by high cyclist speeds, high vehicle speeds, or through conflicts with a platoon leader. In order to achieve parity between the genders, the cyclist or vehicle speeds must be increased by two standard deviations (or one standard deviation each).

4. Discussion

The segmented logit model identified several significant environmental factors (the red light variable and several site-specific constants). Conflicts and dangerous conflicts were less likely to occur immediately after a red light phase. The reason could be that during the green phase, motorists may be less aware of present cyclists, and may make manoeuvres with less caution. Drivers are familiar with yielding to opposing traffic following a red light, and so act more cautiously. After waiting at a red light, motorists are likely to be aware of cyclists waiting adjacent to them, and so yield to them with sufficiently safe spacing. Additionally, speeds of both cyclist and vehicle are lower after a red light phase, which may also decrease the severity of conflicts. With regards to vehicle attributes, increasing vehicle speed increased conflict propensity, for males only. This result is intuitive, as vehicles travelling at higher speeds have less time to take evasive action, and the interactions with road users are expected to have lower PETs. Why conflict occurrence is independent of vehicle speeds for females is unclear, though likely related to unobserved factors including physiology, experience, and behaviour, which are themselves gender-related. Models of gender and behaviour are not developed as part of this study, but could be included in future work.

Concerning gender differences in conflicts, results suggest that, all else being equal, when interacting with vehicles, males are less likely than females to have conflicts and dangerous conflicts. This result implies that male cyclists who interact with vehicles are safer than female cyclists who interact with vehicles. The results demonstrate that, although there are some specific situations where females were safer than males, in general, there are many more cases where males are as safe, if not safer, than female cyclists. Even if this type of cycling infrastructure, with a bidirectional cycle track located to the left of other travel lanes, is rather common in English speaking countries (Johnson et al., 2011), these kinds of interactions with vehicles could seem complex to manage for less experienced cyclists, or those with little confidence on their own cycling abilities, such as a segment of the female population (Handy, 2014). The existence of an on-road path dedicated to cyclists may encourage less experienced cyclists to transfer those skills learned as drivers and thus to focus their attention, when the light is green for them, on traffic coming across and forgetting to take into account the vehicles traveling in the same direction as them. However, in this configuration of intersection, experience as a pedestrian is more appropriate to handle the situation as a cyclist. This points to some broad policy implications for transportation professionals. Namely, building cycling facilities that appear safe to novice cyclists, but provide challenging interactions with vehicles (such as intersections with cycle tracks) may contribute both to an overall reduction in safety, but also to a disparity in safety across segments of the population (gender, experience level, etc.). In general, transportation facilities should match perceived and actual risk to elicit appropriately safe behaviour from all road users.

The result that, all else being equal, female cyclists were more prone to conflicts with vehicles should be confirmed in future studies with other intersection configurations and other cities. This may help to explain why fewer females choose to cycle and therefore contribute to the disparity in ridership by gender. It is known that females prefer routes with fewer vehicular interactions (Garrard et al., 2008), and that they feel less safe than their male counterparts (Bernhoft and Carstensen, 2008). However, as stated above, this does not necessarily mean that males are safer overall. Interestingly for males, results show that a higher cyclist speed resulted in an increased conflict severity, while conflict severity was independent of speed for females. Therefore, there is a compensating effect between gender and speed. Although being male reduces the chances of conflict, cycling faster increases the chances. Males travelling at high speeds (25–30 km/h) are no safer than females. This could indicate that males who travel at faster speeds are also less risk averse, and appears to support previous behavioural research that, even though male cyclists are generally safer, they are less cautious (Bernhoft and Carstensen, 2008) and over-estimate their own competency while female cyclists tend to underestimate their cycling skills (Felonneau et al., 2013), which can have negative consequences on interaction management for both genders. Additionally, as the data analyzed focussed on interactions between cyclists and left turning vehicles, cyclists crossing against the red light were excluded from the analysis. However, as Johnson et al. (2011) showed, males are more likely to run red lights. Although those that run red lights may avoid the studied conflict type, this is a dangerous behaviour that may lead to more severe, though less frequent, conflicts and more severe, though less frequent, collisions. These cyclists are not present in this dataset because they are not interacting with vehicles turning left (the only conflict type considered) and are instead conflicting with vehicles travelling through the intersection in the transverse direction. Additional study is required to determine the effect of red light compliance on cyclist safety. Lastly, males were more likely to engage in a conflict with the first vehicle in the platoon. This may indicate that males tend to have a more aggressive cycling behaviour, or that the behavioural approach for managing different types of interactions is different for males and females, but a more detailed analysis would be beneficial.

5. Conclusions

This paper proposes a methodology based on video analysis and an ordered-based modeling approach to investigate the relationship between cyclist gender and conflict occurrence at urban intersections with cycle tracks. The combination of video-based surrogate safety measures and ordered regression models show promise for future use in conflict analysis. Even basic model formulations were able to estimate parameters with significant correlation to conflict severity. Although the distribution of conflict severity showed no difference between genders, the segmented model identified a distinct difference in conflict severity for different genders, with males having a lower propensity of conflicts and dangerous conflicts than females, all else being equal. Therefore, the segmentation technique was successful in observing variation across segments of the population. Furthermore, the effect of other variables could be quantified.

One limitation of this study was related to the technique for data extraction. Determining gender from video footage is a subjective exercise (Johnson et al., 2011), although less subjective techniques are more invasive. Furthermore, relatively few observations were made of female cyclists. Future work should consider increasing the scope of data collection to ensure more observations of female cyclists are available for analysis. A data set with more observations of female cyclists may help to show the true effect of several variables, enable more significant parameters to be estimated, and may in fact reveal more significant variables related to conflict occurrence for females. However, as current methods require manual extraction of the data, this may be practically difficult. With this proof of concept established, future study should incorporate additional variables, including traffic and cyclist volumes, time of day, and cyclist age. Importantly, if the results obtained herein can be confirmed in other studies, the relationship between gender and risk can be quantified and used to explain the disparity in ridership currently observed in urban populations within North America.

Acknowledgements

Funding for this project was provided in part by the Natural Science and Engineering Research Council. The authors would like to thank Marina Jorge Amaral and Nawfal Kammah for their help with data preparation.

References

Abdel-Aty, M., 2003. Analysis of driver injury severity levels at multiple locations using. J. Safety Res. 34, 597–603.
Bernhoft, I.M., Carstensen, G., 2008. Preferences and behaviour of pedestrians and cyclists by age and gender. Transp. Res. Part F 11, 83–95.
Bhat, C.R., 1997. An endogenous segmentation mode choice model with an application to intercity travel. Transp. Sci. 31 (1), 34–48.
Eluru, N., Bhat, C.R., Hensher, D.A., 2008. A mixed generalized ordered response model for examining pedestrian. Accid. Anal. Prev. 40, 1033–1054.
Felonneau, M.-L., Causse, E., Constant, A., Contrand, B., Messiah, A., Lagarde, E., 2013. Gender stereotypes and superior conformity of the self in a sample of cyclists. Accid. Anal. Prev. 50, 336–340.
Garrard, J., Rose, G., Lo Kai, S., 2008. Promoting transportation cycling for women: the role of bicycle infrastructure. Prevenetative Med. 46 (1), 55–59.
Garrard, J., Handy, S., Dill, J., 2012. Women and cycling. In: City Cycling. MIT Press, Cambridge, MA, pp. 211–234.
Granié, M.-A., 2011. Différences de sexe et rôle de l'internalisation des règles sur la propension des enfants à prendre des risques à vélo. Rech. Transp. Sécurité, 34–41.
Handy, S., 2014. The contribution of comfort, convenience, and liking of bicycling to the bicycling gender gap: evidence from Davis, California. In: 5th International Conference on Women's Issues in Transportation, Paris, pp. 487–496.
Holland, C., Hill, R., 2007. The effect of age, gender and driver status on pedestrians' intentions to cross the road in risky situations. Accid. Anal. Prev. 39 (2), 224–237.
Hydén, C. (1987). The development of a method for traffic safety evaluation: the Swedish traffic conlficts technique. PhD Thesis, Lund, Sweden.
Jackson, S., Miranda-Moreno, L.F., St-Aubin, P., Saunier, N., 2013. Flexible, mobile video camera system and open source video analysis software for road safety and behavioral analysis. Transp. Res. Rec.: J. Transp. Res. Board 2365, 90–98.
Johnson, M., Newstead, S., Charlton, J., Oxley, J., 2011. Riding through red lights: the rate, characteristics and risk factors of non-compliant urban commuter cyclists. Accid. Anal. Prev. 43 (1), 323–328.
Kaplan, S., Vavatsoulas, K., Prato, C.G., 2014. Aggravating andmitigating factors associated with cyclist injury severity in Denmark. J. Safety Res. 50, 75–82.
Klop, J.R., Khattak, A.J., 2007. Factors influencing bicycle crash severity on two-lane: undivided roadways in North Carolina. Transp. Res. Rec.: J. Transp. Res. Board 1674, 78–85.
Kockelman, K.M., Kweon, Y.-J., 2002. Driver injury severity: an application of ordered probit models. Accid. Anal. Prev. 34, 313–321.
Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transp. Res. Part A 44 (5), 291–305.
Lusk, A.C., Furth, P.G., Morency, P., Miranda-Moreno, L.F., Willet, W.C., Dennerlein, J.T., 2011. Risk of injury for bicycling on cycle tracks versus in the street. Inj. Prev., 131–135.
Mannering, F.L., Bhat, C.R., 2014. Analytic methods in accident research: methodological frontier and future directions. Anal. Methods Accid. Res., 1–22.
Martinez-Ruiz, V., Jiminez-Mejias, E., de Dios Luna-del-Castillo, J., Garcia-Martin, M., Jimenez-Moleon, J.J., Lardelli-Claret, P., 2014. Association of cyclists' age and sex with risk of involvement in a crash before and after adjustment for cycling exposure. Accid. Anal. Prev. 62, 259–267.
Miranda-Moreno, L. (2006). Statistical models and methods for identifying hazardous locations for safety improvements. PhD Thesis.
Moore, D.N., Schneider IV, W.H., Savolainen, P.T., Farzanah, M., 2011. Mixed logit analysis of bicyclist injury severity resulting from motor vehicle crashes at intersection and non-intersection locations. Accid. Anal. Prev. 43, 621–630.O'Donnell, C.J., Connor, D.H., 1996. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. Accid. Anal. Prev. 6, 739–753.
Peesapati, L.N., Hunter, M.P., Rodgers, M.O., 2013. Evaluation of postencroachment time as surrogate for opposing left-turn crashes. Transp. Res. Rec.: J. Transp. Res. Board 2386, 42–51.
Santamarina-Rubio, E., Perez, K., Olabarria, M., Novoa, A.M., 2014. Gender differences in road traffic injury rate using time travelled as a measure of exposure. Accid. Anal. Prev. 65, 1–7.
Saunier, N., Sayed, T., Ismail, K., 2010. Large-scale automated analysis of vehicle interactions and collisions. Transp. Res. Rec. 2147, 42–50.
Saunier, N., 2015. Traffic Intelligence, Retrieved from https://bitbucket.org/Nicolas/trafficintelligence/.
Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accid. Anal. Prev. 43 (3), 1666–1676.
Sayed, T., Zaki, M.H., Autey, J., 2013. Automated safety diagnosis of vehicle–bicycle interactions using computer vision analysis. Saf. Sci. 59, 163–172.
Strauss, J., Miranda-Moreno, L.F., 2013. Spatial modeling of bicycle activity at signalized intersections. J. Transp. Land Use 6 (2), 47–58.
Strauss, J., Miranda-Moreno, L.F., Morency, P., 2014. Multimodal injury risk analysis of road users at signalized and. Accid. Anal. Prev. 71, 201–209.
Tom, A., Granie, M.-A., 2011. Gender differences in pedestrian rule compliance and visual search at signalized and unsignalized crossroads. Accid. Anal. Prev. 43 (5), 1794–1801.
Transport Canada, 2014. Canadian Motor Vehicle Traffic Collision Statistics, Retrieved 3.12.14 from Transport Canada: http://www.tc.gc.ca/media/documents/roadsafety/cmvtcs2012_eng.pdf.
van der Horst, A.A., de Goede, M., de Hair-Buijssen, S., Methorst, R., 2014. Traffic conflicts on bicycle paths: a systematic observation ofbehaviour from video. Accid. Anal. Prev. 62, 358–368.
Vélo Québec, 2010. Bicycling in Québec in 2010. Vélo Québec.
Wang, X., Abdel-Aty, M., 2008. Analysis of left-turn crash injury severity by conflicting pattern using partial. Accid. Anal. Prev. 40, 1674–1682.
Winters, M., Davidson, G., Kao, D., Teschke, K., 2011. Motivators and deterrents of bicycling: comparing influencers on decisions to ride. Transportation 38 (1), 153–168.
Zangenehpour, S., Miranda-Moreno, L., Saunier, N., 2013. Impact of bicycle boxes on safety of cyclists: a case study in montreal. In: TRB Annual Meeting Compendium of Papers, Washington, DC.
Zangenehpour, S., Miranda-Moreno, L., Saunier, N., 2014. Automated classification in traffic video at intersections with heavy pedestrian and bicycle traffic. Transportation Research Board Compendium of Papers.
Zangenehpour, S., Miranda-Moreno, L.F., Saunier, N., 2015a. Video-based automatic counting for short-term bicycle data collection in a variety of environments. In: Transportation Research Board 94th Annual Meeting, Washington D.C.
Zangenehpour, S., Strauss, J., Miranda-Moreno, L., Saunier, N., 2015b. Are intersections with cycle tracks safer? A control-case study based on automated surrogate safety analysis using video data. In: TRB Annual Meeting Compendium of Papers, Washington, DC.