

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

Joshua Stipancic, Sohail Zangenehpour, Luis Miranda-Moreno, Nicolas Saunier, Marie-Axelle Granié

# ▶ To cite this version:

Joshua Stipancic, Sohail Zangenehpour, Luis Miranda-Moreno, Nicolas Saunier, Marie-Axelle Granié. Investigating the gender differences on bicycle-vehicle conflicts at urban intersections using an ordered logit methodology. Accident Analysis & Prevention, 2016, 97, pp.19-27. 10.1016/j.aap.2016.07.033 . hal-01356836

# HAL Id: hal-01356836 https://hal.science/hal-01356836

Submitted on 22 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# INVESTIGATING THE EFFECTS OF GENDER ON BICYCLE-VEHICLE CONFLICTS AT URBAN INTERSECTIONS USING AN ORDERED LOGIT METHODOLOGY

- 3
- 4 5
- 6 Joshua Stipancic, Corresponding Author, PhD Candidate
- 7 Department of Civil Engineering and Applied Mechanics, McGill University
- 8 Room 391, Macdonald Engineering Building, 817 Sherbrooke Street West
- 9 Montréal, Québec, Canada H3A 0C3
- 10 Email: joshua.stipancic@mail.mcgill.ca
- 11

# 12 Sohail Zangenehpour, PhD Candidate

- 13 Department of Civil Engineering and Applied Mechanics, McGill University
- 14 Room 391, Macdonald Engineering Building, 817 Sherbrooke Street West
- 15 Montréal, Québec, Canada H3A 0C3
- 16 Email: sohail.zangenehpour@mail.mcgill.ca
- 17
- 18 Luis Miranda-Moreno, Associate Professor
- 19 Department of Civil Engineering and Applied Mechanics, McGill University
- 20 Room 268, Macdonald Engineering Building, 817 Sherbrooke Street West
- 21 Montréal, Québec, Canada H3A 0C3
- 22 Phone: (514) 398-6589
- 23 Fax: (514) 398-7361
- 24 Email: luis.miranda-moreno@mcgill.ca
- 25
- 26 Nicolas Saunier, Associate Professor
- 27 Department of Civil, Geological and Mining Engineering
- 28 Polytechnique Montréal, C.P. 6079, succ. Centre-Ville
- 29 Montréal, Québec, Canada H3C 3A7
- 30 Phone: (514) 340-4711 x. 4962
- 31 Email: nicolas.saunier@polymtl.ca
- 32

# 33 Marie-Axelle Granié, Research Director

- 34 Institut français des sciences et technologies des transports, 304 Chemin de la Croix Blanche
- 35 13300 Salon de Provence, FR
- 36 Phone: 04 90 57 79 79
- 37 Email: marie-axelle.granie@ifsttar.fr
- 38
- Word count: 6670 words + 10 tables/figures x 250 words (each) = 9170 words
- 40
- 41
- 42
- 43
- 44
- 45
- 46 April 26<sup>th</sup>, 2016
- 47

## 1 ABSTRACT

2 In the literature, a crash-based modeling approach has long been used to evaluate the factors that contribute 3 to cyclist injury risk at intersections. However, this approach has been criticized as crashes are required to 4 occur before contributing factors can be identified and countermeasures can be implemented. Moreover, 5 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. 6 7 Despite recent developments, few methodologies and little empirical evidence exist on bicycle-vehicle 8 interactions at intersections using video-based data and statistical analyses. This study investigates bicycle-9 vehicle conflict severity and evaluates the impact of different factors, including gender, on cyclist risk at 10 urban intersections with cycle tracks. A segmented ordered logit model is used to evaluate postencroachment time between cyclists and vehicles. Video data was collected at seven intersections in 11 12 Montreal, Canada. Road user trajectories were automatically extracted, classified, and filtered using a 13 computer vision software to yield 1,514 interactions for analysis. The discrete choice variable was created by dividing post-encroachment time into normal interactions, conflicts, and dangerous conflicts. 14 15 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 16 conflicts. Furthermore, exogenous segmentation was beneficial in comparing different segments of the 17 18 population within a single model. Male cyclists, with all else being equal, were less likely than female 19 cyclists to be involved in conflicts and dangerous conflicts at the studied intersections. Bicycle and vehicle 20 speed, along with the time of the conflict relative to the red light phase, were other significant factors in 21 conflict severity. These results will contribute to and further the understanding of gender differences in 22 cycling within North America.

23

24

25

26 27

Keywords: discrete choice; cycling; surrogate safety; gender; urban; order logit

#### 1 INTRODUCTION

2 Road safety is a substantial concern for transportation professionals due to the high economic and social 3 cost of traffic crashes (Abdel-Aty, 2003). In 2012, traffic crashes resulted in 2077 fatalities and over 4 165,000 injuries in Canada, where cyclists and pedestrians account for approximately 18% of both fatalities 5 and injuries annually (Transport Canada, 2014). While the safety of motorists has commanded much 6 attention, the protection of vulnerable road users has become common only recently (Kockelman & Kweon, 7 2002). In North America, cyclists are twelve times more likely to be killed than motor vehicle drivers 8 (Moore, Schneider IV, Savolainen, & Farzanah, 2011; Strauss, Miranda-Moreno, & Morency, 2014), and 9 269 cyclist fatalities occurred in Canada between 2008 and 2012 (Transport Canada, 2014).

10 Recent growth in bicycle activity and infrastructure improvement have increased awareness of 11 bicycle safety issues in North America. In this context, the complex nature of cyclist-vehicle interactions 12 must be understood by examining the factors that contribute to cyclist safety (Klop & Khattak, 2007; 13 Moore, Schneider IV, Savolainen, & Farzanah, 2011). One method for evaluating these factors is traffic 14 crash modelling. Crash models evaluate collision frequency using count regression models, or crash injury 15 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 16 17 recent summary of literature on crash frequency and severity modelling is provided by Mannering and Bhatt (2014). Additional summaries were compiled by Miranda-Moreno (2006), Lord and Mannering (2010), and 18 19 Savolainen et al. (2011). Traditional safety models calibrated with crash data are reactive, requiring crashes 20 to occur before causes can be identified and countermeasures can be implemented. Additionally, crash-21 based methods require long observation periods. Given the low rate of crash occurrence, many years of 22 accident data are required. Finally, the study of dangerous behaviors and other human factors can often not 23 be investigated.

24 As an alternative or complementary approach, surrogate safety techniques analyze interactions and 25 conflicts rather than crashes. Conflicts are events that are physically and predictably related to traffic crashes, and are placed immediately below collisions in Hyden's model (Hyden, 1987), presented in 26 27 FIGURE 1. As the pyramid model suggests, traffic interactions are inherently ordered through their 28 proximity to a potential collision, or severity, as measured by various indicators. Modelling interactions 29 rather than collisions provides several benefits. Interactions occur much more frequently than collisions 30 and statistically sufficient data can be collected in a shorter time period. The use of interactions and conflicts 31 is proactive, rather than reactive, and surrogate measures are insensitive to crash underreporting 32 (Kockelman & Kweon, 2002) and human factors can be incorporated in the analysis.

33



34 35

# 36 FIGURE 1 Safety pyramid based on Hyden's classic model (Hyden, 1987)

Despite the benefits offered by a proactive surrogate approach, dangerous bicycle behaviours in vehicle-bicycle interactions, and their associated factors, have rarely been studied using video analysis,

surrogate measures of safety, and regression models. A particular issue that remains to be investigated using 1 2 microscopic conflict data is the relationship between biological gender and cyclist safety. Important 3 behavioral differences have been identified for both pedestrians and motorists of different genders at 4 intersections (Holland & Hill, 2007; Santamarina-Rubio, Perez, Olabarria, & Novoa, 2014; Tom & Granie, 5 2011). However, very little is known about this in the cyclist context. Gender differences and their 6 associated risks may help to explain the disparity in ridership currently experienced in urban populations 7 as well as the implementation of the appropriate designs and facilities. In most countries, females cycle less 8 than males, particularly in English speaking countries that are less bicycle-friendly, including Canada, the 9 U.K., Australia, and the U.S. (Garrard, Handy, & Dill, 2012). Route conditions and vehicle interactions 10 greatly influence individuals' likelihood to cycle (Winters, Davidson, Kao, & Teschke, 2011) and female cyclists, much more than males, prefer to use routes with maximum separation from motorized traffic 11 12 (Garrard, Rose, & Lo Kai, 2008). Studies in the city of Montreal show that routes with separated cycle 13 tracks attract higher cyclist volumes than those without (Strauss & Miranda-Moreno, 2013) and that streets 14 with cycle tracks have a lower injury risk (Lusk, et al., 2011).

15 While behavioural differences may exist between genders, recent research has only shown minor differences in actual crash risk for males and females (Kaplan, Vavatsoulas, & Prato, 2014; Martinez-Ruiz, 16 et al., 2014). The purpose of this study is to estimate a segmented ordered logit model for bicycle-vehicle 17 18 interactions at urban intersections with cycle tracks. Intersections, being one of the most dangerous 19 locations in the urban network, deserve specific attention (Wang & Abdel-Aty, 2008). The objectives of 20 this research are to investigate bicycle-vehicle conflicts and contributory factors using an ordered-logit 21 modelling approach and to examine the influence of gender and other cyclist, vehicle, and environmental 22 factors on conflict severity.

23

# 24 LITERATURE REVIEW

25 Existing efforts in injury severity modelling primarily concern motor vehicle occupants in single or multiple 26 vehicle crashes. Multiple modeling techniques have been proposed to investigate the relationship between 27 injury severity levels and associated factors. Mannering and Bhatt (2014) provide a comprehensive 28 literature review on the methods used in severity analysis in general, and Eluru et al. (2008) provides a 29 summary of literature concerning analysis of cyclist injuries. Much empirical evidence has been reported 30 in the literature identifying the key factors influencing injury severity, including characteristics of the 31 roadway, environment, and road user, for passengers, drivers and pedestrians. Moore et al. (2011) utilized 32 a multinomial logit and mixed logit to model cyclist injury severity at intersections and non-intersection 33 locations using seven years of crash data from Ohio. The study stated "the injury mechanisms are 34 substantially different [...] at intersection and non-intersection locations". Klop and Khattak (2007) studied 35 bicycle crash severity on rural roads in North Carolina using an ordered probit model and four years of 36 crash data. Visibility and weather conditions were found to most significantly increase collision severity. 37 Eluru et al. (2008) utilized data from the 2004 US national database to estimate a mixed generalized ordered 38 response logit (MGORL) model of cyclist injury severity. The MGORL generalizes more standard ordered 39 models, providing additional flexibility across observations. Cyclist age and vehicle speed were correlated 40 with injury severity. In the existing literature, gender is often a contributory factor to injury severity.

41 Cyclist gender studies have focussed largely on variation in preference or behaviour. Bernhoft and 42 Carstensen (2008) used a survey approach, finding that male cyclists in Denmark tended to act less 43 cautiously, but also felt safer than their female counterparts. However, many of the differences between genders became statistically insignificant when controlling for other factors, such as travel patterns, which 44 45 may themselves be gender specific. Similarly, a French study by Felonneau et al. (2013) noted that male cyclists tended to overestimate their ability more so than females whereas females tend to overestimate 46 47 their carefulness more than males and proneness to take risk as cyclist seems to be gender specific even at 48 an early age (Granié, 2011). Johnson, Newstead, Charlton, and Oxley (2011) studied red-light compliance 49 and found that male cyclists were more frequently non-compliant than females. While gender differences 50 are evident in terms of behaviour, behavioural differences do not necessarily create disparity in risk. Kaplan 51 et al. (2014) evaluated injury severity using a generalized ordered logit model on Danish crash data

1 collected over a 5-year period, and found no correlation between gender and injury risk. Martinez-Ruiz et 2 al. (2014) used 17 years of crash data from Spain to calculate the crash rate ratio by gender for different 3 ages. Without adjusting for cycling exposure, males were more likely to be involved in a collision than 4 females. When controlling for exposure, crash rates were approximately equal for males and females with 5 age being a stronger determinant of crash risk. As female cyclists seem to favour sites with cycle tracks 6 (Garrard, Rose, & Lo Kai, 2008) a possible explanation for this lack of difference in the risk of accidents 7 can be found in the behavior of male and female cyclists on road with cycle tracks.

8 While research on severity modelling includes consideration for cyclists and advanced 9 formulations, few studies have attempted to apply modeling techniques to surrogate measures for bicycle 10 safety (Zangenehpour, Miranda-Moreno, & Saunier, 2013; Zangenehpour, Strauss, Miranda-Moreno, & 11 Saunier, 2015). Analysis of differences between male and female cyclist risk by surrogate safety techniques 12 has also been limited. Rather than using independent models for males and females, a segmented modelling 13 approach is preferred in order to increase the number of observations available for estimation and to make 14 differences between genders instantly comparable. The reliance on crash data has meant dependence on 15 crash databases, which are subject to underreporting, errors, and omissions. This study attempts to address

16 these shortcomings through the estimation of a segmented ordered logit model of interaction severity.

#### 17 18 **МЕТНО**В

# 18 METHODOLOGY

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

22

## 23 Video Data Collection and Processing

24 The methodology for data collection and processing uses a similar approach to the one implemented in 25 several past studies (Jackson, Miranda-Moreno, St-Aubin, & Saunier, 2013; Saunier, Sayed, & Ismail, 2010; Sayed, Zaki, & Autey, 2013). Video data was collected using an inexpensive and commercially 26 27 available video camera which stores video and is powered internally. The camera is mounted to a using a 28 telescoping fibreglass mast to ensure a clear view above the intersection and approaches. The camera system 29 is introduced in detail by Zangenehpour, Miranda-Moreno, & Saunier (2015). Once collected, the video 30 data was processed using an open-source road user tracking software, Traffic Intelligence (Jackson, 31 Miranda-Moreno, St-Aubin, & Saunier, 2013; Saunier, Saved, & Ismail, 2010; Saunier, 2015). Road user 32 trajectories (position and speed at each frame) were automatically extracted from the video footage. Road 33 users were then classified as pedestrians, cyclists, or vehicles using a technique developed by 34 Zangenehpour, Miranda-Moreno, and Saunier (2014), and the trajectories were filtered to isolate the desired 35 interactions. Any erroneous interactions were removed manually from the data set.

36

#### 37 Computation of Surrogate Safety Measures

38 Popular surrogate measures of safety include time-to-collision (TTC) and post-encroachment time (PET). 39 TTC is "the time required for two vehicles to collide if they continue at their present speed and on the same 40 path" (van der Horst, de Goede, de Hair-Buijssen, & Methorst, 2014) or more generally if their movements 41 remain unchanged, which can include variations in speed and direction. PET is the difference in time 42 between two road users occupying the same location in space, or the potential conflict point (Peesapati, 43 Hunter, & Rodgers, 2013). TTC is measured continuously and, depending on the choice or motion 44 prediction method, will yield several measurements over time when there is a collision course (when some 45 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 46 47 and are complementary in the analysis of bicycle conflicts (van der Horst, de Goede, de Hair-Buijssen, & 48 Methorst, 2014). Surrogate safety data has been successfully collected using video-based detection systems 49 and extracted using computer vision techniques (Jackson, Miranda-Moreno, St-Aubin, & Saunier, 2013; 50 Sayed, Zaki, & Autey, 2013). PET was selected as the surrogate measure of safety in this research as all 51 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 using Traffic Intelligence, 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 seconds;
- Alternative 1: Conflict with PET between 3 and 5 seconds, and;
- Alternative 2: Dangerous conflict with PET less than 3 seconds.

These thresholds have been used successfully in previous work, such as Zangenehpour, Strauss, Miranda-Moreno, & Saunier (2015), 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 & Kweon, 2002).

17

26 27

8

9

10

11

# 19 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 & 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,  $y^*$ , is defined by

$$y^* = \boldsymbol{\beta} \boldsymbol{X} + \boldsymbol{\varepsilon} \tag{1}$$

where *X* is the vector of explanatory variables,  $\beta$  is the vector of unknown parameters, and  $\varepsilon$  is the logistic distributed random error term. This propensity is bound by unknown thresholds,  $\tau_i$ , which delineate alternatives. For the case presented herein, the probability of each choice alternative is

$$P(y=0) = CDF(\tau_1 - \beta X) \quad \text{(normal interaction)}$$
(2)  

$$P(y=1) = CDF(\tau_2 - \beta X) - CDF(\tau_1 - \beta X) \quad \text{(conflict)}$$
  

$$P(y=2) = 1 - CDF(\tau_2 - \beta X) \quad \text{(serious conflict)}$$

36 where *CDF* is the cumulative distribution function of the logistic distribution defining  $\varepsilon$ .

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

44 45

46

$$y^* = \boldsymbol{\beta}_1 \boldsymbol{X} + \boldsymbol{\beta}_2 \boldsymbol{w} \boldsymbol{X} + \boldsymbol{\beta}_3 \boldsymbol{w} + \boldsymbol{\varepsilon}$$

47 where *X* are the basic variables and  $\beta_1$  are their coefficients, *wX* and  $\beta_2$  are the segmented variables and 48 their coefficients, respectively, and *w* and  $\beta_3$  are the segmentation variable (gender) and its coefficient, 49 respectively.

50 The conflict severity model was developed in three steps, with each step incorporating additional 51 gender heterogeneity. An ordered logit model (OL1) was estimated assuming homogeneity between

(3)

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.

#### 9 DATA DESCRIPTION

#### 10

8

## 11 Site Selection

12 Seven test sites were selected along Maisonneuve Boulevard, a one-way urban arterial in downtown 13 Montreal, Canada. The sites, shown in FIGURE 3, featured two travel lanes and a fully separated bidirectional cycling facility, or cycle track, along Maisonneuve. The sites along Maisonneuve were 14 15 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 16 cycling corridors on the island of Montreal, and the specific intersections were chosen because of their 17 18 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. 19 20 Approximately 4 hours 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. 21

22 Similar to Wang and Abdel-Aty (2008), a single traffic interaction was analyzed. Conflicts were 23 limited to those between cyclists traveling straight through the intersection and left-turning vehicles. An 24 example of a dangerous conflict is provided in FIGURE 2. Given the site geometry and design of the cycling 25 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 26 27 interactions were removed manually from the data set. Several interactions were also missed by the tracking 28 algorithm, which could not be included as part of this study. Once filtered, 1514 useable observations 29 remained for analysis.

30



31

32 FIGURE 2 Example of a dangerous bicycle-vehicle conflict



5 6

7 FIGURE 3 Study sites along Maisonneuve at Crescent (a), Stanley (b), Peel (c), Mackay (d),

(g)

8 Metcalfe (e), St Denis (f), and Union (g)

#### 9 **Data Exploration**

10 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 11 12 explanatory variables, including cyclist gender helmet use, whether the vehicle was a truck, van, or SUV, 13 and whether the vehicle was the first in a platoon, were extracted manually from the video. As in previous 14 studies (Johnson, Newstead, Charlton, & Oxley, 2011), gender was classified based on physical appearance. 15 The environmental attributes included variables representing if the interaction occurred immediately after 16 a red light phase (both cyclist and vehicle were simultaneously waiting at a red light before the interaction 17 occurred) and site-specific constants for each site, with Crescent used as the reference. Dummy variables 18 were used to indicate whether the cyclist was the first object to reach the conflict point and whether 19 pedestrians (or contra-flow cyclists) were present in the intersection at the time of the conflict, capturing 20 the behavioural freedom of the road users. A summary is provided in TABLE 1. Site specific data has been 21 omitted for brevity.

	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 first in a platoon	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	.365
Bike First	Dummy	The cyclist reached the conflict point first	0.59	0	1	.493
Pedestrian	Dummy	Peds were simultaneously crossing intersection	0.46	0	1	.499

#### 1 TABLE 1 Variables and Statistics for All Captured Interactions

A preliminary exploration of the raw variable data was conducted with descriptive statistics. Of the 1,514 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 FIGURE 4.







13 The proportion of conflicts, dangerous conflicts, and normal interactions were determined for males 14 and females and are provided in FIGURE 5. As the distributions seem to be drawn from the same underlying 1 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.

5



#### 6 7

7 FIGURE 5 Proportion of Interaction Type According to Gender

# 8 RESULTS

#### 9 10 **Model Estimation**

The models were estimated using all available parameters and systematically eliminating statistically 11 12 insignificant variables. In the results that follow, variables retained were significant at 80% confidence or 13 greater. Variables significant at 90% confidence are italicized, and variables significant at 95% confidence 14 are bolded. Results for the ordered logit models are presented in TABLE 2. In Model OL1, which assumed 15 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 16 17 constants-only model. The null hypothesis was that none of the variables help to explain conflict severity. 18 LR for Model OL1 was 55.72, compared to the critical value from the chi-square distribution of 14.07 (7 degrees of freedom at 95% confidence). The null hypothesis was rejected, and OL1 was superior. Adding 19 20 the gender variable in Model OL2 produced only minor changes in the model results. Although the male 21 dummy variable is negative, which indicates that males have a lower propensity for conflicts than females 22 with all else being equal, and although the gender variable was significant at 80% confidence, Model OL2 23 was inferior to OL1 by log-likelihood ratio test (LR of 1.88 compared to a chi-square value of 3.84). Based 24 on this result, any evidence for variation in conflict severity between males and females is inconclusive.

		OL1	OL2				
	<i>y</i> *	$= \beta_I X + \varepsilon$		$y^* = \boldsymbol{\beta}_1 \boldsymbol{X} + \boldsymbol{\beta}_3 \boldsymbol{w} + \boldsymbol{\varepsilon}$			
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 <sup>2</sup>		0.0183			0.0189		

#### 1 TABLE 2 Model Results for Ordered Logit Models, OL1 and OL2

3

4 Results for the SOL model are presented in TABLE 3. Non-significant  $\beta_1$  coefficients indicate that 5 a parameter is not significantly correlated with conflict severity for female cyclists. Non-significant  $\beta_2$ 6 coefficients indicate that parameter is homogeneous across genders. This formulation eliminates 7 unnecessary variables and allows for immediate testing of differences between genders. Importantly, gender 8 was found to be highly significant in this model. The SOL model was compared to OL1 using a log-9 likelihood ratio test. The null hypothesis was that safety is homogeneous across genders and segmentation 10 is unnecessary. LR was calculated to be 11.08, compared to the critical value from the chi-square distribution of 5.99 (2 degrees of freedom at 95% confidence). The null hypothesis was falsified and 11 segmentation is justified. From the results of the OL models, gender appears to have no effect on safety. 12 13 However, it is not that there is no effect; it is that variation exists both in propensity and in the value of the 14 parameters. Allowing for this variation is necessary to reveal gender's true effect, representing an ideal 15 scenario for an exogenously segmented model.

		β1			$\beta_2$		
		J	$w^* = \beta_1 X + \beta_2 w$	$\lambda X + \beta_3 w + \varepsilon$			
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	
		ßз			βз		
Male	-1.1703	-3.79	0.000	-1.1703	-3.79	0.000	
Tau 1			-0.20	07			
Tau 2			1.04	55			
Number of cases			151	4			
Log likelihood at convergence			-1488	.69			
Log likelihood for constants- only model			-1522	.09			
Pseudo R <sup>2</sup>			0.02	19			

#### 1 TABLE 3 Model Results for Segmented Ordered Logit Model, SOL

# 2

#### 3 Model Interpretation

4 Interpretation of the SOL model is facilitated by calculating the true value of each coefficient for both 5 genders. For females, coefficients take the  $\beta_I$  values from the model estimation. For males, the  $\beta_I$  and  $\beta_2$ 6 parameters are summed to obtain the true value, shown in TABLE 4. Variables related to cyclist attributes 7 led to several interesting results. The primary observation is the negative, highly significant coefficient for 8 the gender variable. In general, this coefficient had the largest magnitude. This result suggests that, all else 9 being equal, when interacting with vehicles, males are less likely than females to have conflicts and 10 dangerous conflicts. This result implies that males are less likely to be involved in more severe interactions, 11 as defined by the PET intervals, and it can be concluded that male cyclists who interact with vehicles are 12 safer than female cyclists who interact with vehicles. A second important observation is that the coefficient 13 for bike speed is positive, but only for male cyclists. For females, cyclist speed has no effect on their 14 potential for conflict at the studied sites. For male cyclists, travelling at a faster speed increases the propensity for conflicts and dangerous conflicts. Therefore, there is a compensating effect between gender 15 16 and speed. Although being male reduces the chances of conflict, cycling faster increases the chances. Males 17 travelling at high speeds (25 to 30 km/h) are no safer than females. This could indicate that males who 18 travel at faster speeds are also less risk averse, and appears to support previous behavioural research that, 19 even though male cyclists are generally safer, they are less cautious (Bernhoft & Carstensen, 2008) and

20 over-estimate their own competency (Felonneau, et al., 2013).

	Female	Male
Explanatory variables	βı	$\beta_1 + \beta_2$
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

#### 1 TABLE 4 Coefficient Values for Male and Female Cyclists

2

3 With regards to vehicle attributes, increasing vehicle speed increased conflict propensity, again for 4 males only. This result is intuitive, as vehicles travelling at higher speeds have less time to take evasive 5 action, and the interactions with road users are expected to have lower PETs. However, why conflict 6 occurrence is independent of vehicle speeds for females is unclear. Additionally, males were more likely to 7 engage in a conflict with the first vehicle in the platoon, although the strength of this result was relatively 8 low. This may indicate that males tend to have a more aggressive cycling behaviour, or that the behavioural 9 approach for managing different types of interactions is different for males and females, but a more detailed 10 analysis would be beneficial. A data set with more observations of female cyclists may help to show the 11 true effect of these variables and enable more significant parameters to be estimated.

12 The only significant environmental factors were the red light variable and several site-specific 13 constants. Conflicts and dangerous conflicts were less likely to occur immediately after a red light phase. During the green phase, motorists may be less aware of present cyclists, and may make manoeuvres with 14 15 less caution. After waiting at a red light, motorists are likely to be aware of cyclists waiting adjacent to 16 them, and so yield to them with sufficiently safe spacing. Drivers are familiar with yielding to opposing 17 traffic after a red light phase, and so act more cautiously. Additionally, speeds of both cyclist and vehicle 18 are lower after a red light phase, which may also decrease the severity of conflicts. With regards to the site-19 specific variables, Crescent, Peel, and St Denis showed no significant variation in conflict occurrence. Both 20 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, 21 22 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.

	Crescent		Stanl	ey	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%

#### **1 TABLE 5** Sample probabilities of conflict severity for male and female cyclists

2

3 These results show that males were safer at both Crescent and Stanley, as their probability of normal 4 interaction was higher, and the probabilities of conflict and dangerous conflict were both lower. In fact, 5 males were safer in many conditions, because any variables with positive coefficients were offset by the 6 large negative coefficient on the male dummy variable. At Union, females were observed to be safer than 7 males. However, this was the only one of the seven sites where this was true. The effect of the highly 8 positive male variable can only be offset by high cyclist speeds, high vehicle speeds, or through conflicts 9 with a platoon leader. In order to achieve parity between the genders, the cyclist or vehicle speeds must be 10 increased by two standard deviations (or one standard deviation each). The results demonstrate that, although there are some specific situations where females were safer than males, in general, there are many 11 more cases where males are as safe, if not safer, than female cyclists. Even if this type of cycling 12 13 infrastructure, with a bidirectional cycle track located to the left of other travel lanes, is rather common in 14 English speaking countries (Johnson, Newstead, Charlton, & Oxley, 2011), these kinds of interactions with 15 vehicles could seem complex to manage to less experienced cyclists, or those with little confidence on their own cycling abilities, such as females (Handy, 2014). The existence of an on-road path dedicated to cyclists 16 may encourage less experienced cyclists to transfer those skills learned as drivers and thus to focus their 17 18 attention, when the light is green for them, on traffic coming across and forgetting to take into account the 19 vehicles traveling in the same direction as them. However, in this configuration of intersection, experience 20 as a pedestrian is more appropriate to handle the situation from the cyclist's point of view. This result that 21 female cyclists were more prone to conflicts with vehicles should be confirmed on future studies with other 22 intersection configurations. It may help to explain why fewer females choose to cycle and therefore 23 contribute to the disparity in ridership by gender. It is known that females prefer routes with less vehicular 24 interactions (Garrard, Rose, & Lo Kai, 2008), and that they feel less safe than their male counterparts 25 (Bernhoft & Carstensen, 2008). These results suggest that the higher perception of risk for female cyclists 26 is linked to a higher actual risk as well. 27

27

#### 28 CONCLUSIONS

29 This paper proposes a methodology based on video analysis and an ordered-based modeling approach to 30 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 31 32 promise for future use in conflict analysis. Even basic model formulations were able to estimate parameters 33 with significant correlation to conflict severity. Although the distribution of conflict severity showed no 34 difference between genders, the segmented model identified a distinct difference in conflict severity for 35 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 36 37 segments of the population. Furthermore, the effect of other variables could be quantified. Interestingly for 38 males, a higher cyclist speed resulted in an increased conflict severity, while conflict severity was 39 independent of speed for females. This appears to support previous research that states that male cyclists 40 tend to be less cautious and overestimate their own competency, while female cyclists tend to underestimate their cycling skills, which can have negative consequences on interaction management for both genders. 41 42 However, as stated above, this does not necessarily mean that males are safer overall. As Johnson, et al. (2011) showed, males are more likely to run red lights. Although those that run red lights may avoid 43 conflicts, this is a dangerous behaviour that may lead to more severe, though less frequent, conflicts and 44

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. The absence of these least risk-averse males from the data set may inflate the observed differences between genders. Additional study is required to determine the effect of red light compliance on cyclist safety.

6 Vehicle speed was observed to increase conflict severity for males only, and males were more likely 7 to have a severe conflict with the first vehicle in a platoon. This may indicate that males tend to have a more 8 aggressive cycling behaviour, or that the behavioural approach for managing different types of interactions 9 is different for males and females, but a more detailed analysis would be beneficial. Conflicts and dangerous 10 conflicts were less likely to occur immediately after a red light phase. Drivers are familiar with yielding to 11 opposing traffic after a red light phase, and so act more cautiously. Additionally, speeds of both cyclist and 12 vehicle are lower after a red light phase, which may also decrease the severity of conflicts.

13 While the effects of many variables could be explained, the parameters of other variables require 14 additional consideration. For example, the effect of the platoon leader variable was heterogeneous with 15 respect to gender, although this variable was statistically significant at only 80% confidence. Explanation of these results requires more detailed analysis to describe behavioural differences in specific traffic 16 17 interactions for males and females. Further limitations of this study were related to the technique for data 18 extraction. Determining gender from video footage is a subjective exercise (Johnson, Newstead, Charlton, 19 & Oxley, 2011), although less subjective techniques are more invasive. Future work should consider 20 increasing the scope of data collection to ensure more observations of female cyclists are available for 21 analysis. More observations would allow a better estimation of the parameters, and may in fact reveal more 22 significant variables related to conflict occurrence for females. However, as current methods require manual 23 extraction of the data, this may be practically difficult. With this proof of concept established, future study 24 should incorporate additional variables, including traffic and cyclist volumes, time of day, and cyclist age. 25 Importantly, if the results obtained herein can be generalized, the relationship between gender and risk can be quantified and used to explain the disparity in ridership currently observed in urban populations within 26 27 North America.

28

# 29 ACKNOWLEDMENT

30 Funding for this project was provided in part by the Natural Science and Engineering Research Council.

31 The authors would like to thank Marina Jorge Amaral and Nawfal Kammah for their help with data

- 32 preparation.
- 33

#### 1 **REFERENCES**

- Abdel-Aty, M. (2003). Analysis of driver injury severity levels at multiple locations using. *Journal of Safety Research*, 34, 597-603.
- Bernhoft, I. M., & Carstensen, G. (2008). Preferences and behaviour of pedestrians and cyclists by age
   and gender. *Transportation Research Part F*(11), 83-95.
- Bhat, C. R. (1997). An Endogenous Segmentation Mode Choice Model with an Application to Intercity
   Travel. *Transportation Science*, *31*(1), 34-48.
- 8 Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for
   9 examining pedestrian. Accident Analysis and Prevention, 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. *Accident Analysis and Prevention*(50), 336-340.
- Garrard, J., Handy, S., & Dill, J. (2012). Women and Cycling. In *City Cycling* (pp. 211-234). Cambridge,
   MA: MIT Press.
- Garrard, J., Rose, G., & Lo Kai, S. (2008). Promoting transportation cycling for women: The role of
   bicycle infrastructure. *Prevenetative Medecine*, 46(1), 55-59.
- 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. *Recherche Transports Sécurité*, 34-41.
- Handy, S. (2014). The contribution of comfort, convenience, and liking of bicycling to the bicycling
   gender gap: Evidence from Davis, California. 5th International Conference on Women's Issues in
   Transportation, (pp. 487-496). Paris.
- Holland, C., & Hill, R. (2007). The effect of age, gender and driver status on pedestrians' intentions to
   cross the road in risky situations. *Accident Analysis and Prevention*, 39(2), 224-237.
- Hyden, 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.
   *Transportation Research Record: Journal of the Transportation Research 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. *Accident Analysis and Prevention*, 43(1), 323-328.
- Kaplan, S., Vavatsoulas, K., & Prato, C. G. (2014). Aggravating andmitigating factors associated with
   cyclist injury severity in Denmark. *Journal of Safety Research*(50), 75-82.
- Klop, J. R., & Khattak, A. J. (2007). Factors Influencing Bicycle Crash Severity on Two-Lane, Undivided
   Roadways in North Carolina. *Transportation Research Record: Journal of the Transportation Research Board*, 1674, 78-85.
- Kockelman, K. M., & Kweon, Y.-J. (2002). Driver injury severity: an application of ordered probit
   models. *Accident Analysis and Prevention*, *34*, 313-321.
- Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and
   assessment of methodological alternatives. *Transportation Research 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. *Injury Prevention*, 131-135.
- Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier
   and future directions. *Analytic Methods in Accident Research*, 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. *Accident Analysis and Prevention*, *62*, 259-267.

$\frac{1}{2}$	Miranda-Moreno, L. (2006). Statistical models and methods for identifying hazardous locations for safety improvements PhD Thesis
3	Moore D N Schneider IV W H Savolainen P T & Farzanah M (2011) Mixed logit analysis of
4 5	bicyclist injury severity resulting from motor vehicle crashes at intersection and non-intersection locations. Accident Analysis and Prevention 43, 621-630
6	O'Donnell C. J. & Connor, D. H. (1006). Predicting the severity of motor vehicle accident injuries using
7	models of ordered multiple choice Accident Analysis and Prevention 6, 739-753
8	Passanati I. N. Hunter, M. P. & Podgars, M. O. (2013). Evaluation of Postanaroachmant Time as
0	Surrogate for Opposing L eft-Turn Crashes, Transportation Research Record: Journal of the
10	Transportation Research Board(2386), 42-51.
11	Santamarina-Rubio, E., Perez, K., Olabarria, M., & Novoa, A. M. (2014). Gender differences in road
12	traffic injury rate using time travelled as a measure of exposure. Accident Analysis & Prevention,
13	65, 1-7.
14	Saunier, N. (2015). <i>Traffic Intelligence</i> . Retrieved from https://bitbucket.org/Nicolas/trafficintelligence/
15	Saunier, N., & Sayed, T. (2007). Automated Analysis of Road Safety. Transportation Research Record:
16	Journal of the Transportation Research Board (2019), 57-64.
17	Saunier, N., Sayed, T., & Ismail, K. (2010). Large-Scale Automated Analysis of Vehicle Interactions and
18	Collisions. Transportation Research Record(2147), 42-50.
19	Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of
20	highway crash-injury severities: A review and assessment of methodological alternatives.
21	Accident Analysis and Prevention, 43(3), 1666-1676.
22	Sayed, T., Zaki, M. H., & Autey, J. (2013). Automated safety diagnosis of vehicle-bicycle interactions
23	using computer vision analysis. Safety Science, 59, 163-172.
24	Strauss, J., & Miranda-Moreno, L. F. (2013). Spatial modeling of bicycle activity at signalized
25	intersections. The Journal of Transport and Land Use, 6(2), 47-58.
26	Strauss, J., Miranda-Moreno, L. F., & Morency, P. (2014). Multimodal injury risk analysis of road users
27	at signalized and. Accident Analysis and Prevention, 71, 201-209.
28	Tom, A., & Granie, MA. (2011). Gender differences in pedestrian rule compliance and visual search at
29	signalized and unsignalized crossroads. Accident Analysis & Prevention, 43(5), 1794-1801.
30	Transport Canada. (2014). Canadian Motor Vehicle Traffic Collision Statistics. Retrieved December 3,
31	2014, from Transport Canada:
32	http://www.tc.gc.ca/media/documents/roadsafety/cmvtcs2012_eng.pdf
33	van der Horst, A. A., de Goede, M., de Hair-Buijssen, S., & Methorst, R. (2014). Traffic conflicts on
34	bicycle paths: A systematic observation of behaviour from video. Accident Analysis and
35	Prevention, 62, 358-368.
36	Vélo Québec. (2010). Bicycling in Québec in 2010. Vélo Québec.
37	Wang, X., & Abdel-Aty, M. (2008). Analysis of left-turn crash injury severity by conflicting pattern using
38	partial. Accident Analysis and Prevention, 40, 1674-1682.
39	Winters, M., Davidson, G., Kao, D., & Teschke, K. (2011). Motivators and deterrents of bicycling:
40	comparing influencers on decisions to ride. Transportation, 38(1), 153-168.
41	Zangenehpour, S., Miranda-Moreno, L. F., & Saunier, N. (2015). Video-based automatic counting for
42	short-term bicycle data collection in a variety of environments. Transportation Research Board
43	94th Annual Meeting. Washington D.C.
44	Zangenehpour, S., Miranda-Moreno, L., & Saunier, N. (2013). Impact of Bicycle Boxes on Safety of
45	Cyclists: a Case Study in Montreal. TRB Annual Meeting Compendium of Papers. Washington,
46	DC.
47	Zangenehpour, S., Miranda-Moreno, L., & Saunier, N. (2014). Automated Classification in Traffic Video
48	at Intersections with Heavy Pedestrian and Bicycle Traffic. Transportation Research Board
49	Compendium of Papers.