6-2022

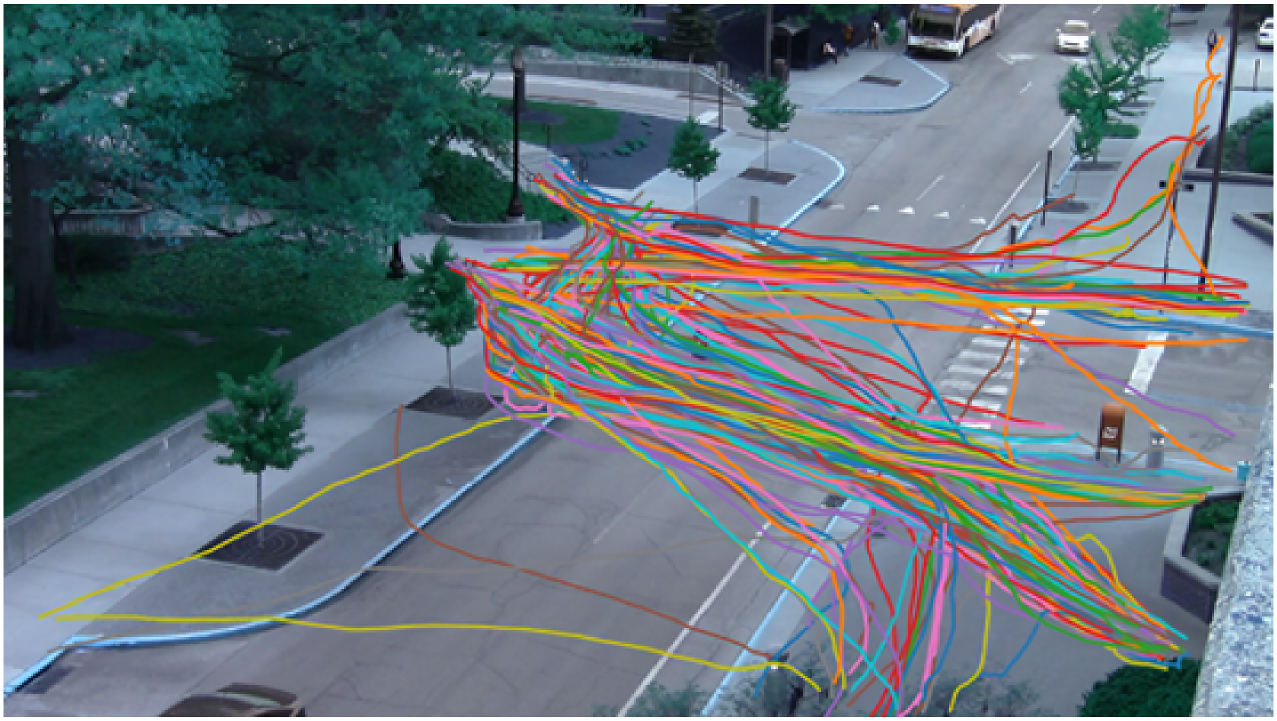
Smart Interaction - Pedestrians and vehicles in a CAV environment

Yunchang Zhang
zhan2854@purdue.edu

Jon D. Fricker

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Final Report #36
June 2022



Smart Interaction - Pedestrians and vehicles in a CAV environment

Yunchang Zhang
Jon D. Fricker

PURDUE UNIVERSITY®

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Report No. 36
Project Start Date: 01/01/2019
Project End Date: 12/31/2020

Report Date: June 2022

Smart Interaction – Pedestrians and Vehicles in a CAV Environment

Yunchang Zhang

Graduate Researcher

Jon. D. Fricker

Professor

Purdue University

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

ACKNOWLEDGEMENT AND DISCLAIMER

Funding for this research was provided by the Center for Connected and Automated Transportation under Grant No. 69A3551747105 of the U.S. Department of Transportation, Office of the Assistant Secretary for Research and Technology (OST-R), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Suggested APA Format Citation:

Zhang, Y., and Fricker, J.D. (2022). Smart Interaction – Pedestrians and Vehicles in a CAV Environment, Technical Report Nr. 36, Center for Connected and Automated Transportation, Purdue University, West Lafayette, IN.

Contact Information

Samuel Labi
3000 Kent Ave., West Lafayette, IN
Phone: 7654945926
Email: labi@purdue.edu

Jon. D. Fricker
550 Stadium Mall Dr.
W. Lafayette, IN
Phone: (765) 494-2205
Email: fricker@purdue.edu

CCAT
University of Michigan Transportation
Research Institute
2901 Baxter Road
Ann Arbor, MI 48152

uumtri-ccat@umich.edu
(734) 763-2498
www.ccat.umtri.umich.edu

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Technical Report Documentation Page

1. Report No.	**2. Government Accession No.**	**3. Recipient's Catalog No.**
CCAT Report #36	N/A	N/A

4. Title and Subtitle Smart Interaction – Pedestrians and Vehicles in a CAV Environment	**5. Report Date** June 2022
	6. Performing Organization Code N/A
7. Author(s) Yunchang Zhang, Jon. D. Fricker	**8. Performing Organization Report No.** N/A
9. Performing Organization Name and Address Center for Connected and Automated Transportation Purdue University, 550 Stadium Mall Drive, W. Lafayette, IN 47907; and University of Michigan Ann Arbor, 2901 Baxter Road, Ann Arbor, MI 48109	**10. Work Unit No.** N/A **11. Contract or Grant No.** Contract No. 69A3551747105
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE, Washington, DC 20590	**13. Type of Report and Period Covered** Final Report. 01/01/2019 - 12/31/2020 **14. Sponsoring Agency Code** OST-R

15. Supplementary Notes

Conducted under the U.S. DOT Office of the Assistant Secretary for Research and Technology's (OST-R) University Transportation Centers (UTC) program.

16. Abstract

At "semi-controlled" crosswalks with yield signs and markings, negotiations as to the right-of-way occur frequently between pedestrians and motorists, to determine who should proceed first. This kind of "negotiation" often leads to traffic delay and potential conflicts. To minimize misunderstandings between pedestrian and motorist that can have serious safety consequences, it is essential that we understand the decision-making process as they interact in real street-crossing situations. This project employs a game-theoretic approach to investigate the joint behaviors of pedestrians and motorists from the perspective of safety. Assuming bounded rationality for each player, the quantal response equilibrium is a special kind of game with incomplete information. Explanatory variables such as conflicting risks and time savings can be incorporated into the payoff functions of the "players" via expected utility functions. Finally, model parameters can be estimated using an expectation maximization algorithm.
The game-theoretic framework is applied to model pedestrian-motorist interactions at a semi-controlled crosswalk on a university campus. The estimation results indicate that the likelihood of pedestrian-vehicle conflict can be quantified. The results can lead to control measures that facilitate the negotiation between pedestrian and motorist and reduce the conflict risk at semi-controlled crosswalks and can help the design of an intelligent risk assessment system at crosswalks.

17. Key Words	**18. Distribution Statement**
Pedestrian crossings, Pedestrian-motorist interaction, Pedestrian wait behavior	No restrictions.

19. Security Classif. (of this report)	**20. Security Classif. (of this page)**	**21. No. of Pages**	**22. Price**
Unclassified	Unclassified	35 pages	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1. Background and Problem Statement

Unsignalized crosswalks also known as "semi-controlled" crosswalks, are a common feature in urban street networks. Instead of signals, the marked crosswalk is "controlled" by "yield to pedestrian" signs. See

Figure 1(a). As a result, there are frequent situations where pedestrians and motorists have to negotiate the right-of-way to determine who should go first. Consequently, a large number of pedestrian-motorist interactions (PMIs) occur at such crosswalks, which may cause conflicts. Faulty communication between pedestrians and motorists often occur. Drivers do not always yield to the pedestrians, and therefore pedestrians often need to stop walking to avoid collision. Failure of a driver to yield to a pedestrian has been documented as one of the most common causes of vehicle-pedestrian crashes at semi-controlled crosswalks (Schneider et al., 2017; Shaon et al., 2018). To address safety issues, it is essential that we understand as much as possible about the behaviors of pedestrians and motorists as they interact in real street-crossing situations.

Consequently, confusion between pedestrians and motorists leads to unnecessary delays for both pedestrians and vehicles and increase risks to pedestrian safety. An investigation of pedestrian crossing behavior and waiting behavior at such locations can be useful in developing policies and control strategies to enhance a pedestrian's perceived safety and improve the level of service (LOS) at unsignalized intersections.

The model is named zebra-crossing game that is to quantify the probability of confusion and probability of conflict between pedestrian and motorist. The quantified probability of conflict can be used as a real-time risk warning system for CAV deployment at crosswalks.

(a) (b) (c)

(a) Sign at Semi-Controlled Crosswalk. (MUTCD, 2009)
(b) One-way Grant Street, 2017 (Jon D. Fricker)
(c) One-way North University Street at Second Street, 2017. (Jon D. Fricker)

Figure 1 Semi-Controlled Crosswalk.

1.2. Existing Studies

1.2.1. Pedestrian-Motorist Interactions

There is significant literature on pedestrian-motorist interactions (PMIs) at semi-controlled crosswalks. In observational experiments, a PMI can be extracted as an event using the judgment of observer(s) (Fricker and Zhang, 2019).

For pedestrian crossing behavior, discrete choice models are often applied to estimate whether a pedestrian will accept (Y = 1) or reject (Y = 0) a gap, if a PMI occurs. Generally, older and female pedestrians are more likely to wait in the curb area, hesitate before crossing, and maintain a lower speed while crossing at mid-block crosswalks (Cloutier et al., 2017; Sucha et al., 2017; Alver and Onelcin, 2018). Moreover, researchers (Pawar et al., 2015; Pawar and Patil, 2016) have examined the impact of spatial or temporal gaps on pedestrian decision-making at semi-controlled crosswalks. Many of those studies conclude that vehicle dynamics (i.e., the distance from crosswalk to interacted vehicle and the vehicle approach speed) are variables that significantly influence pedestrian crossing behavior. Furthermore, traffic and environmental factors, such as the number of lanes (Zhang et al., 2019), traffic flow rate (Cassidy et al., 1995; Hamed, 2001; Chen and Wang, 2015), and crossing surface materials (Cloutier et al., 2017) have been proved to affect pedestrian crossing choices.

Driver compliance behavior (willingness to yield) has also been investigated using discrete choice models. Schroeder and Rouphail (2010; 2011) conducted field observations related to driver yielding behavior at semi-controlled crosswalks in North Carolina. Descriptive variables, such as pedestrian characteristics, vehicle dynamics, and environmental factors, were collected and used in a choice modeling framework. Fricker and Zhang (2019) further applied a partial proportional odds model in a driver's willingness to yield study, considering different street operations. A detailed review of literature regarding PMIs can be found in Camara et al. (2019 and Amado et al. (2020).

The PMIs at an unsignalized crosswalk constitute a "zebra crossing" game involving two players – pedestrian and motorist. The decision of one player is a response to the decision of the other. Consequently, considering only the decision of one player is not sufficient to describe the communication between two players. The remedy is the game theoretic approach that can assign a utility function to each combination of actions.

1.2.2. Game-Theoretic Approach

Only a few studies focus on the game-theoretic approach to modeling PMIs at semi-controlled crosswalks. Elvik (2014) categorized ten classic games in road user behavior studies, but the game between pedestrian and motorist was not among them. Bjørnskau (2017) first developed a zebra crossing game with the perfect rationality assumption to explore the bicyclist-vehicle interaction at unsignalized crosswalks in Norway. Two kinds of Nash Equilibria (Cycle/Yield and Yield/Driver) were found based on ordinal responses. Field observations confirmed that Cycle/Yield was the perfect Nash Equilibrium (NE) and was the most frequent solution in the real world. However, the Walk/Yield solution (the bicyclist gets off the bicycle, negotiates with the driver, and walks over the crosswalk) is not a Nash Equilibrium solution. The Walk/Yield outcome implies that bicyclists may not know whether or not the interacted driver is aggressive. A recent study of pedestrians and motorists at unsignalized crosswalks (Fricker and Zhang, 2019) frequently observed the Walk/Yield solution. Therefore, some real behaviors of pedestrian and motorist are missing in the NE with complete information.

To capture uncertainties in the decision-making process, evolutionary game theory has been adopted, assuming players are not perfectly rational. Chen et al. (2016) integrated evolutionary game theory with cumulative prospect theory (CPT) in modeling PMIs at unsignalized crosswalks. The proposed evolutionary game framework has the potential to model the phenomenon of behavioral differences among pedestrians in a group. However, the evolutionary game framework is incorporated into a microsimulation platform that introduces a large number of parameters to be calibrated. Therefore, it is difficult to generalize the evolutionary game framework when we consider more sophisticated PMI models (Talebpour et al., 2015). To address these limitations, modified game-theoretic models such as Quantal Response Equilibrium (McKelvey and Palfrey, 1995) have been developed. In Quantal Response Equilibrium (QRE), players are assumed to make decisions with the lowest perceived costs that are subject to errors. A recent study (Arbis and Dixit, 2019) applied the QRE to model the lane-changing "game" between an on-ramp driver and a mainline driver and their results revealed that the QRE can accurately model the expected number and variance of driving strategies when the lane-changing game occurs.

Based on the literature review, the zebra crossing game has not been adequately investigated. There is little research that considers the joint behavior of driver and pedestrian at a semi-controlled crosswalk. In addition, drivers and pedestrians behave stochastically rather than rationally. Pure Nash Equilibrium with complete information may not be adequate to analyze the zebra crossing game. Consequently, the objective of this study is to propose a game-theoretic framework to model the joint behaviors of pedestrians and motorists at semi-controlled crosswalks. Three research questions are addressed in this study:

1. How relax the assumptions of perfect rational behavior and complete information in the game setting?
2. How analyze the game between pedestrian and driver with incentives of time savings and conflict avoidance?
3. How reduce the likelihood of conflict between pedestrians and drivers at unsignalized crosswalks?

2. METHODOLOGY

2.1. Data Collection

The data for this study were collected at two semi-controlled crossing locations on the Purdue University campus. Video recordings were made in Spring 2017 when Grant Street was a one-way northbound street. The street had two 12-ft. wide lanes with a speed limit of 25 mph. We observed 170 PMIs in a 40-minute period. See Figure 1**(b). Additional video recordings were made in Spring 2017, when University Street was a one-way northbound street. The street had two 10-ft wide lanes (plus a 4-foot bicycle lane) with a speed limit of 25mph. Video recordings were made at four different 40-minute periods (Zhang and Fricker, 2021). We observed a total of 1437 pedestrian-motorist interactions. See Figure 1**(c).

2.2. Definitions

A *pedestrian-motorist interaction* is defined as the behavior of either party when in the area of influence of the other (Fricker and Zhang, 2019):

- A pedestrian enters the curb area and intends to cross.
- The subject driver is aware of the pedestrian's intention and then responds to the pedestrian (yields or doesn't yield).

An interaction does not occur if:

- A pedestrian arrives at the curb area, but there is no vehicle close enough to the crosswalk to affect the pedestrian's crossing decision.
- A vehicle approaches the crosswalk, but there are no pedestrians in the curb area.

There are two types of pedestrian crossing decisions for each interaction:

- The subject pedestrian crosses immediately (Y = 1).
- The subject pedestrian waits and yields to the interacting motorist (Y = 0).

There are two types of driver's yielding behavior for each interaction:

- The driver yields by stopping or slowing down (Y = 1).
- The driver doesn't yield to the pedestrian (Y = 0).

Considering each combination of decisions, there are two special cases of interactions:

- *Confusion*: The driver yields, and pedestrian does not cross.

- *Conflict*: The driver doesn't yield, and pedestrian chooses to cross.

Confusion event involves delay (efficiency issue), and conflict event brings the safety issue. Therefore, confusion and conflict events are of special interests at semi-controlled crosswalks, and it is essential to model the "zebra-crossing game" to quantify the probability of confusion and probability of conflict for operational and safety assessments.

3. MODEL FORMULATION

The Quantal Response Equilibrium (QRE) builds on games with *incomplete information*. Such a game is formulated as: $G = \left\langle N,\ \{S_i\}_{i=1}^{N},\ \{A_i\}_{i=1}^{N},\ \{T_i\}_{i=1}^{N},\ P(t_1,\ldots,t_T),\ \{u_i\}_{i=1}^{N} \right\rangle$,

1. The set of players N = {Pedestrian =1; Driver =2}.
2. The set of states S_t as described by explanatory variables in Table 1.
3. The set of actions for each player: pedestrian A_1 = {Cross; Not Cross} and driver A_2 = {Yield; Not Yield}.
4. The set of types for each player: T_1 = {Aggressive; Cautious} given the state S_i.
5. A joint probability Distribution: $P_t = \{P_{aggressive}; P_{cautious}\}$ over types, $P_1 = \{ P_{aggressive} = P_{cross}; P_{cautious} = 1 - P_{cross} \}$ and $P_2 = \{ P_{aggressive} = 1 - P_{yield}; P_{cautious} = P_{yield} \}$.
6. The payoff function of each player: $\mu_i : T_1 \times A$. See Section 4.1 and Table 2.

Table 1 Explanatory Variables

Variable	**Description**
V_{ped}	The approach speed (ft/s) of pedestrians when a pedestrian enters the curb area. Using Google Maps, the distance covered by a pedestrian every 34 milliseconds (one video frame) is converted into a speed. (mean = 3.34 ft/s; sd = 2.47 ft/s)
d_{veh}	The distance of interacted vehicle to the conflict point when the interaction begins (in feet). If an interaction occurs when the subject pedestrian arrives at the curb, we paused the video and calculated the distance to vehicle using Google Maps. (mean = 72.02 ft; sd = 54.14 ft) d_{veh} stands for the distance buffer of the vehicle.
d_{ped}	The <u>direct</u> distance of the subject pedestrian to the interacted vehicle when interaction begins (in feet). Let the pedestrian distance to the conflict point as $d_{conflict}$. $d_{ped} = \sqrt{d_{conflict}^2 + d_{veh}^2}$ (mean = 76.30 ft; sd = 51.72 ft). d_{ped} stands for the pedestrian's distance buffer.
V_{veh}	The approach speed (ft/s) of interacted vehicles when a pedestrian enters the curb area. Using Google Maps, the distance covered by a vehicle every 34 milliseconds (one video frame) is converted into a speed. (mean = 12.50 ft/s; sd = 10.29 ft/s)
Sex	A binary variable taking the value 1 if the subject pedestrian is a male (44.0%), 0 for female pedestrian (56.0%).

Group	The number of pedestrians in the subject pedestrian's curb area, including the subject pedestrian. (mean = 2.30; sd = 2.02)
Cross	A binary variable taking the value 1 if the pedestrian crosses; 0 otherwise.
P_{cross}	The interacted driver's belief that the pedestrian will cross.
Yield	A binary variable taking the value 1 if the driver yields; 0 otherwise
P_{yield}	The subject pedestrian's belief that the interacted driver will yield.
Note: mean = average value; sd = standard deviation	

First, the state of nature of an interaction will determine the type of player that is measured by probability P_t. Then, the type of player determines the action that the subject player will choose in probability.

- P_{yield} and P_{cross} are the parameters of the joint probability distribution.
- P_{yield} represents the pedestrian's belief that the interacted driver will yield. In other words, P_{yield} represents the pedestrian's belief that the interacted driver will be cautious.
- P_{cross} denotes the driver's anticipation that the subject pedestrian will cross. In other words, P_{cross} denotes the driver's anticipation that the subject pedestrian will be aggressive.

The payoff function of each player is defined as $u_i : T_i \times A$. Payoff functions are interrelated by players' actions and beliefs (P_{yield} and P_{cross}), and PMIs can be quantified.

3.1. Expectation Utility Functions

3.1.1. Pedestrian Utility

Expectation utility functions for actions that the subject pedestrian may choose can be expressed as Equation 1a and Equation 2a:

$$EU_{Cross} = p_{yield}(b_1 v_{ped}^2 + b_2 sex + b_3 Group) \quad (1a)$$

$$EU_{Cross} = p_{yield} a_1 v_{ped}^2 \quad (1b)$$

$$EU_{DoNotCross} = b_4 d_{ped}^2 + b_5 d_{ped} + b_6 \quad (2a)$$

$$EU_{DoNotCross} = a_2 + a_3 d_{ped} \quad (2b)$$

Equations 1a and 2a become Equations 1b and 2b after eliminating variables that were found to be not statistically significant. If a pedestrian chooses to cross, the pedestrian will be motivated by reduced travel time, but will undertake the risk of getting hit by a vehicle. P_{yield} represents the pedestrian's belief that the interacted driver will *yield*.

1. If the driver chooses to yield, then P_{yield} = 1, and the expected utility of crossing will be equivalent to $EU_{Cross} = a_1 v_{ped}^2$, which denotes the benefits of saving time.

2. Otherwise, if the driver does not yield, then P_{yield} = 0, and the expected utility is equivalent to EU_{Cross} = 0, which represents the cost of a potential conflict.

3.1.1. Driver Utility

Expectation utility functions for the actions that the interacted driver may choose are expressed as Equation 3a and Equation 4a.

$$EU_{Yield} = a_4 d_{veh}^2 + a_5 d_{veh} + a_6 \tag{3a}$$

$$EU_{Yield} = a_4 d_{veh}^2 + a_5 d_{veh} + a_6 \tag{3b}$$

$$EU_{DoNotYield} = (1 - p_{cross})(b_{10} v_{veh}^2 + b_{11} v_{veh}) + b_{12} \tag{4a}$$

$$EU_{DoNotYield} = (1 - p_{cross}) a_7 v_{veh}^2 + a_8 \tag{4b}$$

Equations 3a and 4a become Equations 3b and 4b after eliminating variables that were found to be not statistically significant. If a driver chooses not to yield, the driver will be motivated by reducing travel time, but will assume the risk of hitting the pedestrian.

1. P_{cross} represents the driver's belief that the pedestrian will *cross*. If the pedestrian chooses to cross, then P_{cross} = 1, 1 – P_{cross} = 1 – 1 = 0, and the expected utility of not yielding will (by Equation 4b) be equivalent to $EU_{DoNotYield} = a_8$ which represents the cost of a potential conflict to the driver.

2. If the pedestrian does not cross, then P_{cross} = 0, 1 – P_{cross} = 1 – 0 = 1, and the expected utility (by Equation 4b) is equivalent to $EU_{DoNotYield} = a_7 v_{veh}^2 + a_8$, which denotes the benefit of saving time.

3.2. Logit Quantal Response Equilibrium

A logit quantal response function is a particular class of quantal response function that has been widely used in the study of choice behavior. The logit quantal response function assumes that the players' anticipations are accurate on average, but subject to some errors that follow an extreme value distribution (McKelvey and Palfrey, 1995). Equation 5 and Equation 6 take advantage of the expected utilities $\overline{EU}_{Cross}$, $\overline{EU}_{DoNotCross}$, $\overline{EU}_{Yield}$, and $\overline{EU}_{DoNotYield}$ to address the variability in the costs across a population (Watling, 2006).

$$p_{cross} = \frac{\exp[\overline{EU}_{Cross}(p_{yield})]}{\exp[\overline{EU}_{cross}(p_{yield})] + \exp[\overline{EU}_{DoNotCross}]} \tag{5}$$

$$p_{yield} = \frac{\exp[\overline{EU}_{Yield}(1-p_{cross})]}{\exp[\overline{EU}_{Yield}(1-p_{cross})]+\exp[\overline{EU}_{DoNotYield}]} \tag{6}$$

Additionally, the players' anticipations P_{yield} and P_{cross} are determined by the logit quantal response functions in Equations 5 and 6.

- P_{cross} is the probability of the subject pedestrian choosing the strategy $\{cross\}$ with the belief that the interacted driver will yield with the probability P_{yield}.
- P_{yield} is the probability of the interacted driver choosing the action $\{yield\}$ with the anticipation that the subject pedestrian will cross with probability P_{cross}.

The existence and uniqueness of logit QRE has been asserted by McKelvey and Palfrey (1995).

Mathematically computing P_{yield} and P_{cross} is a fixed-point problem. P_{yield} and P_{cross} are fixed points of functions $P_{yield} = F(P_{cross})$ and $P_{cross} = H(P_{yield})$ (Brouwer, 1911).

Model parameters are estimated by maximum likelihood estimation. Let ΔEU_{Cross} and ΔEU_{Yield} be latent indices for pedestrian decisions and motorist decisions:

$$\Delta EU_{Cross} = EU_{Cross} - EU_{DoNoCross} \tag{7}$$

$$\Delta EU_{Yield} = EU_{Yield} - EU_{DoNoYield} \tag{8}$$

The log-likelihood function of pedestrian decisions can be constructed:

$$LL_{ped}(a_1, a_2, a_3; y, X) = \sum_i \{\ln[\varphi(\Delta EU_{Cross})] * I\{y_i = 1\} + \ln[1-\varphi(\Delta EU_{Cross})] * I\{y_i = 0\}\} \tag{8}$$

where

- $y_i = 1$ represents that the subject pedestrian chooses the action $\{cross\}$.
- $y_i = 0$ represents that the subject pedestrian chooses the action $\{not\ cross\}$.
- $\varphi(.)$ is the cumulative distribution function of the logistic distribution.

Similarly, the log-likelihood function of driver decisions can be constructed:

$$LL_{veh}(a_4, a_5, a_6, a_7, a_8; y, X) = \sum_j \{\ln[\varphi(\Delta EU_{Yield})] * I\{y_j = 1\} + \ln[1-\varphi(\Delta EU_{Yield})] * I\{y_j = 0\}\} \tag{10}$$

where

- y_j = 1 represents that the interacted driver chooses the action {*yield*}.
- y_j = 0 represents that the interacted driver chooses the action {*not yield*}.

Therefore, let μ denote the vector of all model parameters, and the log-likelihood can be expressed as:

$$LL(\mu; y, X) = LL_{ped}(a_1, a_2, a_3; y, X) + LL_{veh}(a_4, a_5, a_6, a_7, a_8; y, X) \tag{11}$$

3.3. Solution Algorithm - Expectation Maximization

Expectation Maximization (EM) can be applied iteratively to generate solutions of logit QRE. P_{yield} and P_{cross} can be considered as latent variables. For a pair of initial probabilities$\{P_{cross,i}, P_{yield,i}\}$, μ_i is generated by maximizing the total log-likelihood function (Dixit and Denant-Boemont, 2014):

$$\begin{aligned}\max_{\mu_i} LL(\mu_i; y, X) = LL_{ped}(a_1, a_2, a_3; y, X, P_{yield,i}) + \\ LL_{veh}(a_4, a_5, a_6, a_7, a_8; y, X, P_{cross,i})\end{aligned} \tag{12}$$

For μ_i, a new pair of probabilities $\{P_{cross,i+1}, P_{yield,i+1}\}$is generated based on Equation 13 and Equation 14:

$$p_{cross,i+1} = \frac{\exp[\overline{EU}_{Cross}(p_{yield,i}, \mu_i)]}{\exp[\overline{EU}_{Cross}(p_{yield,i}, \mu_i)] + \exp[\overline{EU}_{DoNotCross}(\mu_i)]} \tag{13}$$

$$p_{yield,i+1} = \frac{\exp[\overline{EU}_{Yield}(1 - p_{cross,i}, \mu_i)]}{\exp[\overline{EU}_{Yield}(1 - p_{cross,i+1}, \mu_i)] + \exp[\overline{EU}_{DoNotYield}, \mu_i]} \tag{14}$$

Equations 12-14 can be applied iteratively until $\{P_{cross,i}, P_{yield,i}\}$ converges. The pseudocode of the proposed EM algorithm is shown in Figure 2.

The convergence of the EM algorithm has been shown by Wu (1983). The logit QRE usually converges within 150 iterations. We use traceplots (see Figure 3) of P_{cross} and P_{yield} with 150 iterations to show the convergence of the logit QRE. P_{cross} and P_{yield} converge to 0. 0.848 and 0.475, respectively.

Algorithm 1 Expectation Maximization

Initialization $\leftarrow \epsilon,\ v = 0,\ P_{0,cross},\ P_{0,yield}$.

$\mu = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$.

while $P_{v,cross} - H(P_{v,yield}) > \epsilon$ Or $P_{v,yield} - F(P_{v,cross}) > \epsilon$ **do**

$LL(\mu; y, X, P_{v,cross}, P_{v,yield}) = LL_{ped}(a_1, a_2, a_3; y, X, P_{v,yield}) + L_{veh}(a_4, a_5, a_6, a_7, a_8, P_{v,cross})$

$\mu_{v+1} \in \arg\min LL(\mu; y, X, P_{v,cross}, P_{v,yield})$

$P_{v+1,cross} = H(P_{v,yield}; \mu_{v+1})$

$P_{v+1,yield} = F(P_{v,cross}; \mu_{v+1})$

$v = v + 1$

end

Figure 2 Expectation Maximization Algorithm

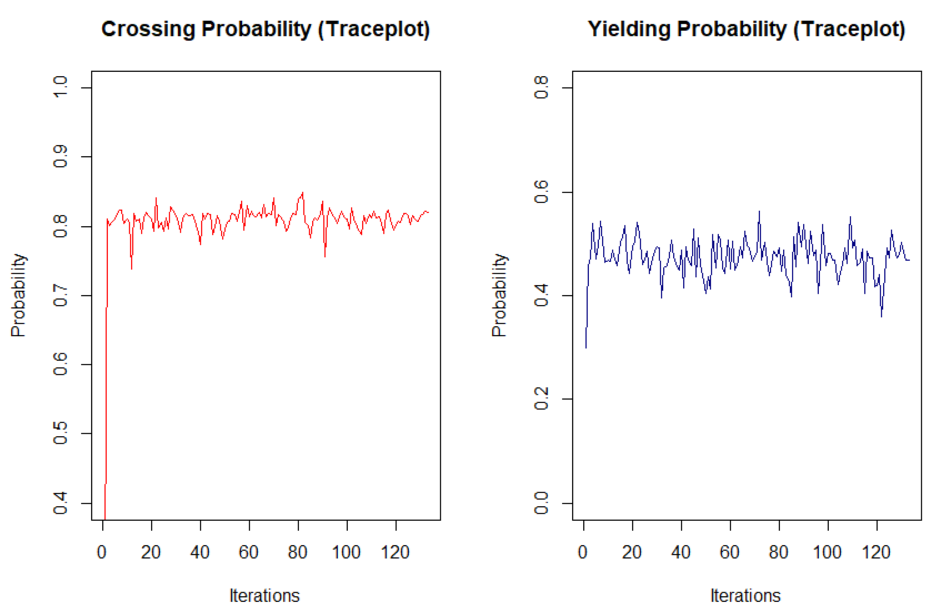


Figure 3 Trace plots of Fixed-Point Iterations

4. RESULTS

4.1. Estimation Results

The estimation results are shown in **Table 2**. Parameters and variables used in Equations 1b, 2b, 3b and 4b are shown in the first column of **Table 2**. "Bootstrap standard deviation" is a surrogate estimate of standard error (Train, 2009). The z-score is an indication of how far from zero an estimated coefficient is. If the absolute value of a z-score is greater than 1.96, we can conclude that the estimated parameter is "significant" at the 95 percent confidence level. Note that a6 represents the intercept in the Equation 3b. Including the intercept will make the model more flexible even if the intercept is not significant.

Table 2 Estimation Results for Model Parameters

Parameter and Variable	Coefficient	Bootstrap Standard Deviation	z_score	$P > \|z\|$
a_1 V_{ped}^2	0.245	0.043	5.654	0.000
a_2	1.920	0.513	3.746	0.000
a_3 d_{ped}	-0.024	0.005	-4.682	0.000
a_4 d_{veh}	0.054	0.023	2.389	0.017
a_5 d_{ped}^2	-0.00030	0.00013	-2.382	0.017
a_6	-0.464	0.593	-0.782	0.434
a_7 V_{veh}^2	0.057	0.010	5.484	0.000
a_8	-1.072	0.506	-2.119	0.034

4.1.1. Pedestrian Dynamics

The estimation of parameter a_1 is associated with the approach speed (ft/s) of a pedestrian when he/she enters the curb area. A positive a_1 represents the increase in the likelihood of a pedestrian crossing if the pedestrian approaches the curb with a higher velocity. The pedestrian approach speed has a strong correlation with the aggressiveness of the subject pedestrian.

4.1.2. Vehicle Dynamics

Similarly, the estimation for parameter a_6 is associated with the approach speed (ft/s) of a driver when a PMI occurs. A positive a_6 represents the increase in the likelihood of a driver not yielding if the vehicle approaches the crosswalk with a higher speed.

4.1.3. Pedestrian Distance to Conflict Point

The estimation parameters a_2 = 1.92 and a_3 = -0.024 are related to the pedestrian's action – {*not cross*}. A negative a_3 represents the decrease in the likelihood of a pedestrian not crossing if the direct distance of the subject pedestrian to the interacted vehicle is higher.

4.1.4. Vehicle Remaining Distance

d_{veh} also reveals a non-linear effect on a driver's decision. If the distance to the conflict point is less than 89.49 ft, the expected utility function for {*yield*} increases with the increase in distance to the conflict point. Therefore, the probability of {*yield*} increases.

When d_{veh} > 89.49 ft, the expected utility function for {*yield*} decreases as d_{veh} increases. There are two possible explanations.

1. A driver may observe the pedestrian's behavior first and then respond to it because there's an adequate buffer for the driver to "think about" the best action.
2. There may be no need for driver to yield because, if the pedestrian leaves the curb area quickly and crosses quickly, the driver will not have to Yield (as defined in **Table 1**).

4.2. Payoff Matrix

The payoff matrix can be derived by applying the estimated coefficients in **Table 2**. Each entry in Table **3** represents the payoff from each action. Each player will choose an action to maximize the payoff.

Table 3 Payoff Matrix

		Driver	
		Yield	Do Not Yield
Pedestrian	Cross	$[0.245v_{ped}^2,\ 0.054d_{veh} - 0.0003d_{veh}^2 - 0.464]$	[0, -1.072]
	Do Not Cross	$[1.92 - 0.024d_{ped},\ 0.054d_{veh} - 0.0003d_{veh}^2 - 0.464]$	$[1.92 - 0.024d_{ped},\ 0.057v_{veh}^2 - 1.072]$
Pedestrian (mean case)	Cross	[2.74, 1.86]	[0, -1.072]
	Do Not Cross	[1.49, 1.86]	[1.49, 7.84]

If the mean values of explanatory variables are used in the payoff functions, the entries in the last two rows of **Table 3**. These entries represent the expected utilities or payoff of each action. In game theory, a *strictly dominant* strategy is defined as one that always provides greater utility than another strategy for one player. When a vehicle approaches the crosswalk with the speed of 5 ft/s, the payoff for {not yield} is $0.057 * 5^2 - 1.072 = 0.353$, and the payoff for {yield} is no less than 0.353 if d_{veh} lies in the range [16.7, 163.3] by solving the equation. The dominant strategy will be {yield} for the driver. Dominant strategy is a key point in perfect Nash equilibrium game, assuming perfect rationality for every player. In the Quantal Response Equilibrium, perfect rationality is relaxed to the bounded rationality, and either dominant or dominated strategy can be chosen in a probability.

It is clear in the "mean value" case that there is no strictly dominant strategy for both pedestrian and driver. For example, if the pedestrian chooses the action {*cross*}, the potential outcome is 2.74 when the interacted motorist chooses to *yield* or is 0 when the interacted motorist chooses *not to yield*. If the pedestrian chooses the action {*not cross*}, the potential outcome is always 1.49. The outcomes for the action {*cross*} are not always greater than the outcomes for the action {*not cross*}. This demonstrates that the best pedestrian strategy depends on the driver's behavior (the probability of driver yielding as perceived by the pedestrian). Neither pedestrian behavior is dominant. In the absence of a dominant strategy, a player must decide on an action based on the expected action of the other player. The uncertainties in this zebra crossing game raise issues of safety and efficiency. Of special interest are the probabilities of the two special cases of interactions:

- *Confusion*: The driver yields, and pedestrian does not cross.
- *Conflict*: The driver does not yield, and pedestrian chooses to cross.

5. CONFLICT AND CONFUSION ANALYSIS

We define the conflict probability and the confusion probability as:

$$p_{conflict} = p_{cross} * (1 - p_{Yield}) \tag{15}$$

$$p_{confusion} = (1 - p_{Cross}) * p_{yield} \tag{16}$$

$P_{conflict}$ and $P_{confusion}$ are two important performance measures. $P_{conflict}$ is the probability that a conflict occurs, representing the conflict risk. $P_{confusion}$ is the probability that a misunderstanding between pedestrian and motorist happens, which is an efficiency measure at the semi-controlled crosswalk.

5.1. Conflict and Confusion Prediction

Sample calculations of the probability of conflict and the probability of confusion are provided for a pedestrian-motorist interaction (PMI) with covariate values d_{ped} = 51.80 ft, d_{veh} = 38.97 ft, v_{ped} = 3.69 ft/s, and v_{veh} = 28.24 ft/s.

1. Set initial parameters for the first step in the algorithm (Figure 2). Here, P_{cross} = 0.7, P_{yield} = 0.4, and ε = 0.001 are chosen.

2. Calculate the expected utilities. Equations 17-20 below are Equations 1b-4b for the sample calculations:

$$EU_{Cross} = 0.4 * 0.245 * 3.69^2 = 1.334 \tag{17}$$

$$EU_{DoNotCross} = 1.92 - 0.024 * 51.80 = 0.677 \tag{18}$$

$$EU_{Yield} = 0.054 * 38.97 - 0.0003 * 38.97^2 - 0.464 = 1.185 \tag{19}$$

$$EU_{DoNotYield} = (1 - 0.7) * 28.24^2 - 1.072 = 12.565 \tag{20}$$

3. Recalculate the probabilities using Equations 5 and 6: P'_{cross} = 0.659 and P'_{yield} = $1.142*10^{-5}$.

4. Stop if the convergence requirement is met, viz., $P'_{cross} - P_{cross} < \varepsilon$ and $P'_{yield} - P'_{yield} < \varepsilon$. Else, $P_{cross} = P'_{cross}$, $P_{yield} = P'_{yield}$, and repeat Steps 1-3.

5. At convergence, P_{cross} = 0.309 and P_{yield} = $2.224*10^{-13}$.

6. Calculate the probability of conflict as $P_{conflict}$ = 0.309 * (1 – $2.224*10^{-13}$) = 0.309 and the probability of confusion $P_{conflict}$ = (1 – 0.309) * $2.224*10^{-13}$ = $1.538*10^{-13}$.

The high value of $P_{conflict}$ reflects the low distance (38.97 ft) of the vehicle to the conflict point, a high vehicle approach speed (28.24 ft/s) and a moderate pedestrian approach speed (3.69 ft/s).

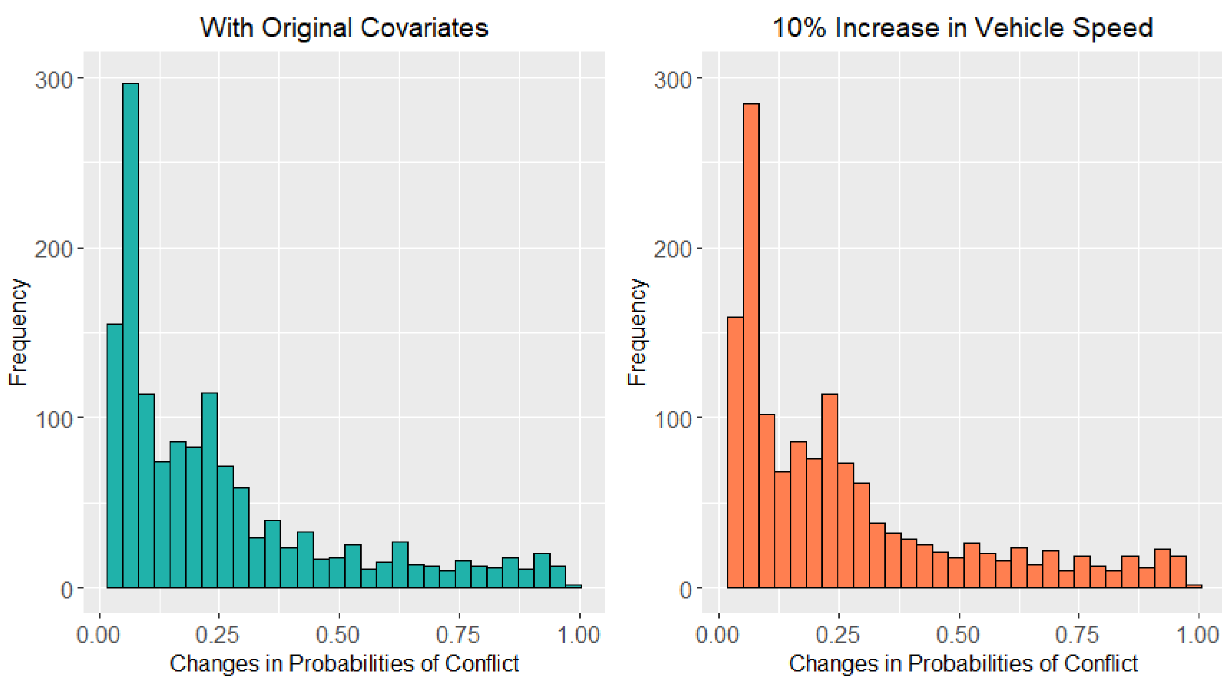
5.2. Relationship between Explanatory Variables and Conflict

An essential research question will be: How to reduce the likelihood of conflict between pedestrians and drivers at semi-controlled crosswalks? Based on Equations 15 and 16, the relationship between the explanatory variables and the likelihood of conflict $P_{conflict}$ is indirect. For example, the coefficient a_6 is associated with the approach speed (ft/s) of a driver when a PMI occurs. A change in vehicle approach speed will directly influence the probability of vehicle yielding P_{yield} as perceived by the pedestrian. A change in P_{yield} will result in a change in the probability of pedestrian crossing P_{cross}. The probability of conflict is derived as $P_{\text{conflict}} = P_{\text{cross}} * (1 - P_{\text{yield}})$, which is Equation 15. According to Equations 1b and 4b, the change of vehicle approach speed will result in changes in both P_{yield} and P_{cross}. Hence, the relationship between the explanatory variables and the likelihood of conflict is not direct.

For this reason, a sensitivity analysis was conducted by predicting changes in conflict probability if there is a 10% change in any one of the variables listed in Table 1, holding constant the values of all other variables. The sensitivity analysis is based on the following five steps:

1. For each pedestrian-motorist interaction (PMI), calculate the likelihood of conflict $P_{conflict}$ and confusion $P_{confusion}$ based on the six-step conflict prediction in Section 5.1, given the observed values of d_{ped}, d_{veh}, v_{ped} and v_{veh} for the PMI. The green bars in Figure 4(a) comprise the frequency histogram after $P_{conflict}$ has been calculated for all PMIs.

2.



(a) (b)

Figure 4 Distributions of Conflict Probability Before-and-After

3. Change by 10% the value of a given predictor variable for each PMI and calculate the new likelihood of conflict $P'_{conflict}$ and likelihood of confusion $P'_{confusion}$, based on Equations 15 and 16. The red bars in Figure 4 Distributions of Conflict Probability Before-and-After(b) represent the frequency distribution of the likelihood of conflict $P'_{conflict}$ for all PMIs, given a 10% increase in the original value of V_{veh} for each PMI.

4. Calculate $P'_{conflict} - P_{conflict}$ for the current PMI and add the result to the frequency distribution of changes in conflict probability for V_{veh}.

5. After all PMIs have been examined, draw the histogram for the frequency distribution of changes in conflict probability for V_{veh}. See Figure 5(a).

6. Repeat Steps 1-4 for the other explanatory variables and build histograms, as shown in Figure 5(b), Figure 5(c) and Figure 5(d).

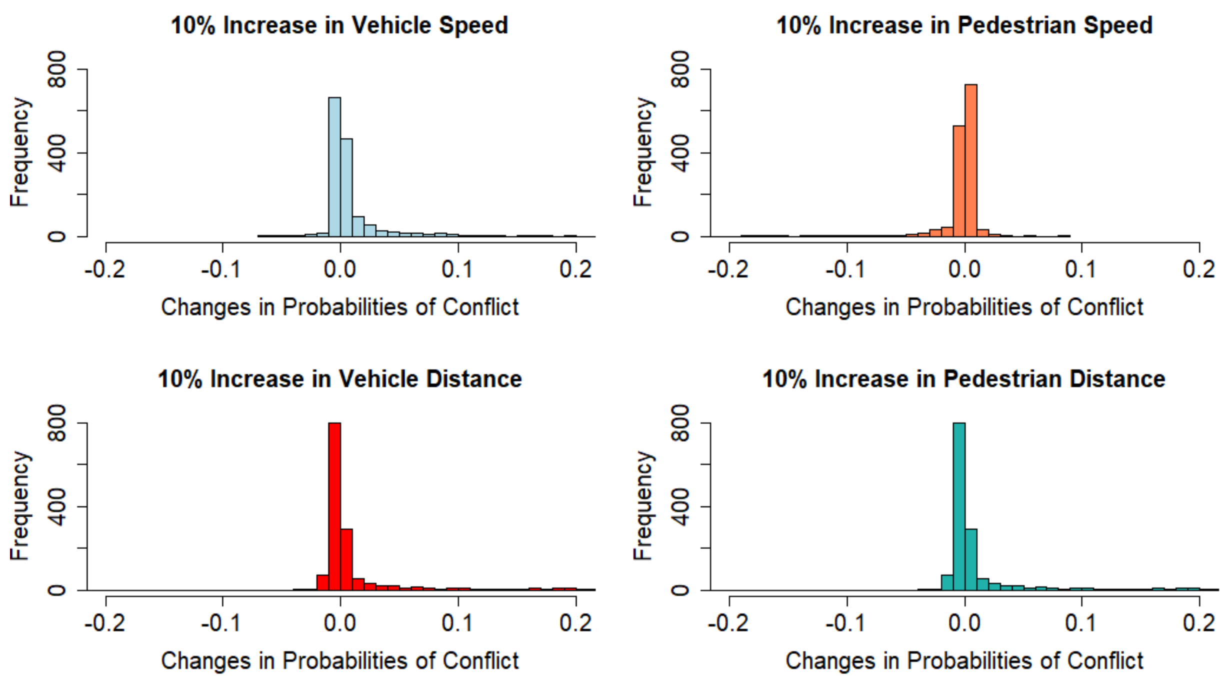


Figure 5 Sensitivity Analysis

The histograms in Figure 5 indicate that an increase in an explanatory variable will not necessarily lead to an expected change in the likelihood of conflict. For example,

- An increase in the vehicle approach speed V_{veh} will result in a lower probability of yield P_{yield}, based on Equations 4b and 6.

- A decrease of P_{yield} will also result in a reduction of P_{cross}, based on Equations 1b and 5.

- However, an increase in $(1 - P_{\text{yield}})$ and the decrease of P_{cross} will not necessarily lead to a reduction in the probability of conflict, as calculated by Equation 15.

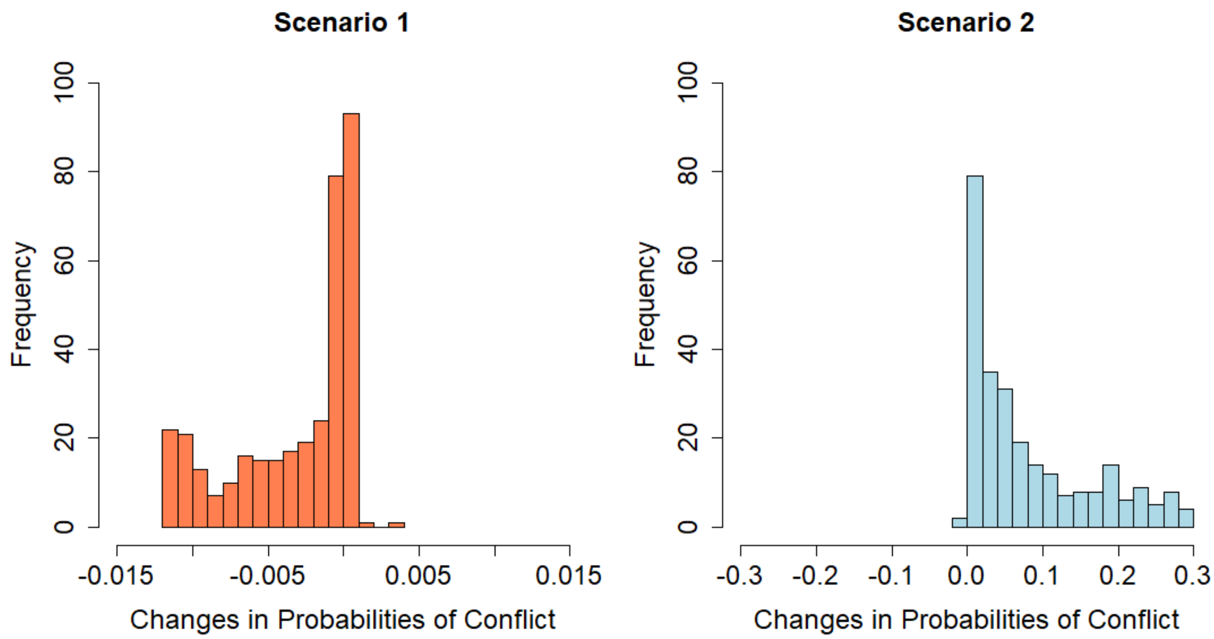
When we compare the top two histograms in Figure 5 to the bottom two histograms, it is apparent that Probabilities of Conflict are more sensitive to changes in distances than to changes in speeds. If the vehicle distance to the conflict point is increased by 10%, P_{conflict} is usually reduced, by as much as 0.01 in some cases. Perhaps surprisingly, the P_{conflict} response to a 10% increase in vehicle approach speed is very small, usually between -0.1 and +0.2. In response to changes in pedestrian approach speed and pedestrian distance to the conflict point, changes in P_{conflict} will be larger for distance, but without positive or negative tendencies in either case.

These results may provide clues to control measures at the crosswalk. Controlling vehicle approach speed is possible, but it will not necessarily reduce conflict probabilities, based on the results of the game theory analysis. The findings with respect to vehicle distance to the conflict point may inform the design of traffic control at the crosswalk, or whether new controls at the semi-controlled crosswalk should be implemented at all. In this section, we mainly discuss the effectiveness of controlling the vehicle distance to the conflict point.

5.3. Relationship between Vehicle Distance to the Conflict Point and Conflict

The bottom left histogram in Figure 5 indicates that an increase in vehicle distance to the conflict point (d_{veh}) will result in a lower likelihood of conflict. For a zebra crossing game, the three scenarios identified in a previous paper (Zhang et al., 2020) are used:

1. If a vehicle is too close to yield to the subject pedestrian ($d_{\text{veh}} \leq 30$ ft), the normal pedestrian choice is to "let the vehicle go first", and it will cause little delay to the pedestrian.
2. If a vehicle is too far away ($d_{\text{veh}} > 120$ ft), the normal pedestrian choice is to cross without any hesitation, because the pedestrian will feel safe.
3. If the vehicle is neither too far from the crosswalk nor too close (esp. 40 ft to 50 ft, as demonstrated in (Zhang et al., 2020), the zebra crossing game will be more complicated and more instructive. We set a range for the parameter (30 ft $< d_{\text{veh}} \leq$ 80 ft) in the Scenario 3.
4. The remaining range for the parameter (80 ft $< d_{\text{veh}} \leq$ 120 ft) is set in the Scenario 4.



(a) (b)

Figure 6 Distributions of Changes in Conflict Probabilities under Scenario 1 and Scenario 2

The first and second scenarios will not involve dangerous situations. If a vehicle is too close to yield to the subject pedestrian, a 10% increase in distance is a small change (e.g., 20 feet versus 22 feet). Sensitivity analysis for only the Scenario 1 observations in the database (d_{veh} < 30 ft) indicates minor changes in the likelihood of conflict (-0.015 to 0.005) that are based on Equation 3b. See Figure 6(a).

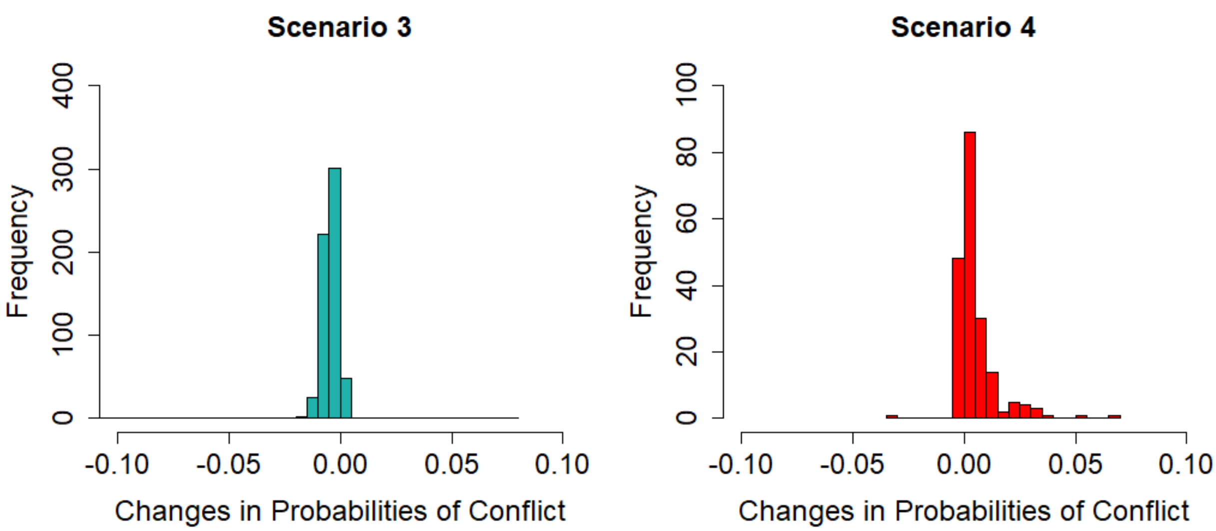
If a vehicle is "too far away", a 10% increase in distance will result in larger changes in d_{veh} (e.g., 200 feet versus 220 feet). Sensitivity analysis for only the Scenario 2 observations in the database (d_{veh} >120 ft) indicates major changes in the likelihood of conflict (0 to 0.3). If a vehicle is too far away (d_{veh} >120 ft), the normal pedestrian choice is to cross without any hesitation, because the pedestrian will feel safe. Therefore, a 10% increase in the vehicle distance to the conflict point will have little effect.

However, for the third scenario (30 ft < d_{veh} ≤80 ft), a large proportion of pedestrians must quickly consider the costs of being hit and the costs of delay. If at least one player has incorrect expectations concerning the behavior of the others, such inefficient communication can lead to unsafe situations. The results of the sensitivity analysis are shown in Figure 7(a). An increased vehicle distance to the conflict point usually reduces the likelihood of conflict in the Scenario 3.

An increased vehicle distance to the conflict point usually *increases* the likelihood of conflict in the Scenario 4 (80 ft < d_{veh} ≤ 120 ft). As we discussed in the Section 4.1.4, when d_{veh} > 89.49 ft, the expected utility function for {*yield*} decreases as d_{veh} increases. There are two possible explanations. (1) A driver may observe the pedestrian's behavior first and then respond to it because there's an adequate buffer for the driver to "think about" the best action. (2) There may be no need for driver to yield because, if the pedestrian leaves the curb area quickly and crosses quickly, the driver will not have to Yield.

This results for Scenario 3 and Scenario 4 make sense, but how does this finding help inform the design of control measures (if any) at the crosswalk? Because Figure 5 shows us that $P_{conflict}$ is sensitive only to d_{veh}, increasing d_{veh} from the current value for each PMI covered by Scenario 3 will do the most to reduce $P_{conflict}$ and improve safety. Two strategies suggest themselves:

1. First determine if a need for a control measure exists. The value of $P_{conflict}$ generated by the methods in this proejct can serve as a guide in that respect. For example, if too many PMIs fall above a certain value of $P_{conflict}$ (say, 0.15) in Figure 4(a), control measures may be justified.
2. Although active controls might be considered, it may be sufficient (and even preferable) to install passive controls such as speed humps at a distance from the crosswalk that reduces $P_{conflict}$ to an acceptable level. The findings of this study are consistent with a distance of 40 ft to 50 ft that was demonstrated in Zhang et al. (2020).



(a) (b)

Figure 7 Distributions of Changes in Conflict Probabilities under Scenario 3 and Scenario 4

5.4. Relationship between Vehicle Distance to the Conflict Point and Conflict

We collected data from three different semi-controlled crossing locations in eight different time periods, during which there are temporal variations in traffic volumes and pedestrian flows. These three datasets are separated into eight subsets based on the time period when we collected the data. The validity of probability of conflict is based on the following four steps:

1. Driver yielding rates for the **eight** datasets are calculated as:
   a. Derive the number of PMIs with Yield = 1 (from **Error! Reference source not found.**) at dataset *i* as $N_{yield,\ i}$.
   b. Derive the total number of PMIs at dataset *i* as $N_{total,\ i}$.

c. Derive the driver yielding rate for dataset i as $p_{yield} = \frac{N_{yield,\ i}}{N_{total,\ i}}$.

2. Conflict probabilities of conflicts in mean for the **eight** datasets are derived via the proposed six-step method (y-axis in Figure 8).

3. To delineate the relationship between the observed driver's yielding rates and calculated conflict probabilities, a linear regression ($R^2 = 0.71$) has been developed in Figure 8.

4. A negative relationship ($y = 0.499 – 0.431x$) between calculated conflict probabilities and observed driver yielding rates has been found, which indicates that an increase of driver yielding rate will result in a lower calculated conflict probability.

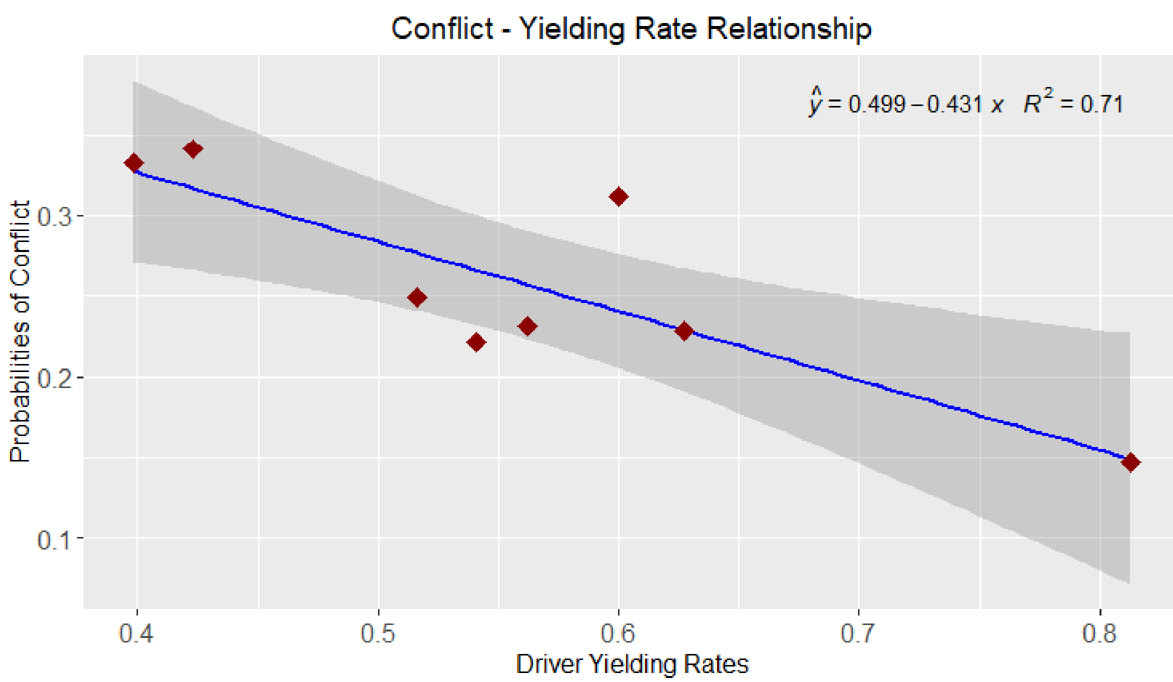


Figure 8 Relationship Between Probability of Conflict and Driver Yielding Rate

6. CONCLUSION

This project applies Quantal Response Equilibrium (QRE) to study the joint behavior of pedestrians and motorists from the perspective of risk at semi-controlled crosswalks. The essential research question is: How to reduce the likelihood of conflict between pedestrians and drivers?

Quantal Response Equilibrium involves players who have incomplete information about the other players. In the zebra crossing game, a player must decide on an action based on the *belief* that the other player will be aggressive or cautious. The results of these decisions can put a pedestrian and a motorist in conflict. Our analysis, in which QRE is applied to data for a semi-controlled crosswalk, produced the following results:

1. The probability of conflict $P_{conflict}$ when a pedestrian-motorist interaction (PMI) occurs can be quantified. This makes it possible to use *probability of conflict* as a pedestrian safety measure.
2. Because the relationship between the explanatory variables in the database and the likelihood of conflict $P_{conflict}$ is indirect, a sensitivity analysis was conducted to predict changes in $P_{conflict}$ if there is a 10% change in any one of the variables.
3. It was found that *probabilities of conflict* are not very sensitive to changes in vehicle or pedestrian approach speeds, but are sensitive to changes in vehicle distance to the crosswalk.

These results can provide a basis for deciding whether control measures should be installed at a semi-controlled crosswalk and, if needed, how they should be installed. An example using speed humps is presented. The information acquired from the QRE analysis can also be used to employ more active control measures.

This report is based on a published paper: Zhang, Y., & Fricker, J. D. (2021). Incorporating conflict risks in pedestrian-motorist interactions: A game theoretical approach. Accident Analysis & Prevention, 159, 106254.

7. STUDY SCOPE AND LIMITATIONS

This study investigated the "zebra crossing" game involving two players – pedestrian and motorist at semi-controlled crosswalks. 1437 pedestrian-motorist interactions were observed, and the joint behavior of driver and pedestrian as they interact in real street-crossing situations is explored. As more data gathered in our dataset, the model can be further updated, and the model performance is expected to be better. In addition, there are several limitations in the model-based game:

1. Limit information in data: the data used in the model was collected by a single analyst. The event-based framework only includes information of a PMI at one time stamp when the interaction occurs. Currently, we have developed pedestrian detection and tracking framework using deep-learning techniques. The pedestrian and vehicle trajectory data containing information over multiple time steps will be utilized in the future.
2. Restricted model: expected utility functions are fixed in the pre-specified model. Agents are unlikely to behave exactly in the way that the model-based game defines. Advanced computation models and model-free imitation learning techniques will be explored.

8. OUTPUTS, OUTCOMES, AND IMPACTS

8.1. Research Outputs

8.1.1. Synopsis of Project

The project exceeded expectations, in that the results went far beyond a basic "inventory" and categorization of interactions between pedestrians and motorists. Appropriate statistical analysis has revealed factors and relationships that are described in three papers listed in the following section.

8.1.2. List of Publications

Zhang, Y., Qiao, Y., & Fricker, J. D. (2020). Investigating Pedestrian Waiting Time at Semi-Controlled Crossing Locations: Application of Multi-State Models for Recurrent Events Analysis. *Accident Analysis & Prevention*, 137, 105437.

Zhang, Y., & Fricker, J. D. (2020). Multi-State Semi-Markov Modeling of Recurrent Events: Estimating Driver Waiting Time at Semi-Controlled Crosswalks. *Analytic Methods in Accident Research*, 100131.

Zhang, Y., & Fricker, J. D. (2021). Investigating temporal variations in pedestrian crossing behavior at semi-controlled crosswalks: a Bayesian multilevel modeling approach. Transportation research part F: traffic psychology and behaviour, 76, 92-108.

8.1.3. List of Presentations

Yunchang Zhang, Jon, D. Fricker (2020). "Multi-State Semi-Markov Models: An Application to Drivers' Gap Acceptance in front of Approaching Pedestrians at Unsignalized Crosswalks". *Accepted by Transportation Research Board 99th Annual Meeting*, January 2020.

Zhang, Y., & Fricker, J. (2021, June). Investigating Smart Traffic Signal Controllers at Signalized Crosswalks: A Reinforcement Learning Approach. In 2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS) (pp. 1-6). IEEE.

8.1.4. List of Outcomes and Highlights

This section will emphasize list of outcomes and highlights:

- "Semi-controlled" crosswalks are unsignalized but marked with "yield to pedestrian" signs.
- Road user trajectory dataset wad extracted from video recordings. 1437 pedestrian-motorist interactions were observed.

- Performance measures were developed and tested to evaluate the safety and efficiency of crossing locations under a full range of control measures.
- Simulation software was applied to replicate the base case in the data base and use the associated performance measures as the benchmark.
- At the other extreme, a scenario was created with fully autonomous vehicles and pedestrians who are fully compliant with control measures.

8.1.5. List of Impacts

This study improves the operation and safety of semi-controlled crosswalks by developing a database and identifying factors that affect pedestrian and motorist behavior.

1. This information will be used to test the impact of new technologies on crosswalk safety and performance.
2. A coupling project with INDOT is a perfect complement to this study, in that it offers opportunities to apply a variety of designs and control methods to other types of crossing locations.

LIST OF REFERENCES

Akaike, H. Factor analysis and AIC. Psychometrika. 1987. 52(3), 317-332.

Andersen, P. K., & Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. The annals of statistics, 1100-1120.

Bjørnskau, T. (2017). The Zebra Crossing Game–Using game theory to explain a discrepancy between road user behaviour and traffic rules. Safety science, 92, 298-301.

Brüde, Ulf, and Jörgen Larsson. Models for Predicting Accidents at Junctions where Pedestrians and Cyclists Are Involved. How Well Do They Fit? Accident Analysis & Prevention, 1993. 25: 499-509.

Camara, F., Romano, R., Markkula, G., Madigan, R., Merat, N., & Fox, C. (2018, March). Empirical game theory of pedestrian interaction for autonomous vehicles. In Proceedings of Measuring Behavior 2018. Manchester Metropolitan University.

Chung, Y. (2010). Development of an accident duration prediction model on the Korean Freeway Systems. Accident Analysis & Prevention, 42(1), 282-289.

Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2), 187-202.

Cloutier MS, Lachapelle U, d'Amours-Ouellet AA, Bergeron J, Lord S, Torres J. "Outta my way!" Individual and Environmental Correlates of Interactions Between Pedestrians and Vehicles during Street Crossings. Accident Analysis & Prevention, 2017. 104: 36-45.

Cross Validated, Test the Significant Differences Between the Coefficients Obtained by Logistic Regression on Two Data Sets, https://stats.stackexchange.com/questions/20330/test-the-significant-differences-between-the-coefficients-obtained-by-logistic-r, Accessed December. 17, 2018.

De Dominicis, R., & Manca, R. (1984). An algorithmic approach to non-homogeneous semi-Markov processes. Communications in Statistics-Simulation and Computation, 13(6), 823-838.

Fricker, J. D., & Zhang, Y. (2019). Modeling Pedestrian and Motorist Interaction at Semi-Controlled Crosswalks: The Effects of a Change from One-Way to Two-Way Street Operation. Transportation Research Record, 0361198119850142.

Harrell, F. E. Regression Modeling Strategies. Springer-Verlag, New York, 2015.

Himanen V, Kulmala R. An Application of Logit Models in Analysing the Behaviour of Pedestrians and Car Drivers on Pedestrian Crossings. Accident Analysis & Prevention, 1988. 20:187-97.

Guan, J., Wang, K., & Chen, F. (2016). A cellular automaton model for evacuation flow using game theory. Physica A: Statistical Mechanics and its Applications, 461, 655-661.

Guo, H.W., Gao, Z.Y., Yang, X.B., Jiang, X.B. Modeling pedestrian violation behavior at signalized crosswalks in China: a hazards-based duration approach. Traffic Inj. Prev. 2011. 12 (1), 96–103.

Guo, H.W., Wang, W.H., Guo, W.W., Jiang, X.B., Bubb, H. Reliability analysis of pedestrian safety crossing in urban traffic environment. Saf. Sci. 2012. 50 (4), 968–973.

Kadali BR, Rathi N, Perumal V. Evaluation of Pedestrian Mid-block Road Crossing Behaviour Using Artificial Neural Betwork. Journal of Traffic and Transportation Engineering (English Edition), 2014. 1:111-9.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American statistical association, 53(282), 457-481.

Leaf WA, Preusser DF. Literature Review on Vehicle Travel Speeds and Pedestrian Injuries. US Department of Transportation, National Highway Traffic Safety Administration, 1999.Williams R. Generalized Ordered Logit/Partial Proportional Odds Models for Ordinal Dependent Variables. Stata Journal, 2006. 6: 58.

Lee, R. (2000). The Lee-Carter method for forecasting mortality, with various extensions and applications. North American actuarial journal, 4(1), 80-91.

Li, B.B. A model of pedestrians’ intended waiting times for street crossings at signalized intersections. Transport. Res. Part B. 2013. 51 (1), 17–28.

Lord S, Cloutier MS, Garnier B, Christoforou Z. Crossing Road Intersections in Old Age—With or Without Risks? Perceptions of Risk and Crossing Behaviours among The Elderly. Transportation Research Part F: Traffic Psychology and Behaviour, 2018. 55: 282-96.

Michalaki P, Quddus MA, Pitfield D, Huetson A. Exploring the Factors Affecting Motorway Accident Severity in England Using the Generalised Ordered Logistic Regression Model. Journal of Safety Research, 2015. 55: 89-97.

Nam, D., Mannering, F., 2000. An exploratory hazard-based analysis of highway incident. Transp. Res. Part A 34 (2), 85–102.

Ouhbi, B., & Limnios, N. (2003). Nonparametric reliability estimation of semi-Markov processes. Journal of Statistical Planning and Inference, 109(1-2), 155-165.

Papadimitriou E. Theory and Models of Pedestrian Crossing Behaviour along Urban Trips. Transportation Research Part F: Traffic Psychology and Behaviour, 2012. 15: 75-94.

Papadimitriou E, Lassarre S, Yannis G. Introducing Human Factors in Pedestrian Crossing Behaviour Models. Transportation Research Part F: Traffic Psychology and Behaviour, 2016. 36:69-82.

Paselk, T., Mannering, F., 1993. Use of duration models for predicting vehicular delay at a US/Canadian border crossing. Transportation 21, 249±270.

Prentice, R. L., Williams, B. J., & Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. Biometrika, 68(2), 373-379.

Sal Amati K, Schroeder B, Geruschat D, Rouphail N. Event-Based Modeling of Driver Yielding Behavior to Pedestrians at Two-Lane Roundabout Approaches. Transportation Research Record: Journal of the Transportation Research Board, 2013. 2389: 1-1.

Schroeder BJ, Rouphail NM. Event-Based Modeling of Driver Yielding Behavior at Unsignalized Crosswalks. Journal of Transportation Engineering, 2010. 137: 455-65.

Sucha M, Dostal D, Risser R. Pedestrian-Driver Communication and Decision Strategies at Marked Crossings. Accident Analysis & Prevention, 2017. 102: 41-50.

Sun, D., Ukkusuri, S. V. S. K., Benekohal, R. F., and Waller, S. T. Modeling of Motorist-Pedestrian Interaction at Uncontrolled Mid-block Crosswalks. 82nd Annual Meeting of the Transportation Research Record, Washington, D.C., 2003.

Tiwari, G., Bangdiwala, S., Saraswat, A., & Gaurav, S. (2007). Survival analysis: Pedestrian risk 21 exposure at signalized intersections. Transportation research part F: traffic psychology and 22 behaviour, 10(2), 77-89.

Wang X, Abdel-Aty M. Analysis of Left-turn Crash Injury Severity by Conflicting Pattern Using Partial Proportional Odds Models. Accident Analysis & Prevention, 2008. 40: 1674-82.

Washington, S. P., Karlaftis, M. G., & Mannering, F. (2010). Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC.

Wei, L. J., Lin, D. Y., & Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American statistical association, 84(408), 1065-1073.

Williams R. GOLOGIT2: Stata Module to Estimate Generalized Logistic Regression Models for Ordinal Dependent Variables. Stata Software Components. Boston College Department of Economics; 2016.

Yang, X. , Huan, M. , Abdel-Aty, M. , Peng, Y. , & Gao, Z. . (2015). A hazard-based duration model for analyzing crossing behavior of cyclists and electric bike riders at signalized intersections. Accident Analysis & Prevention, 74, 33-41.

Yannis G, Papadimitriou E, Theofilatos A. Pedestrian Gap Acceptance for Mid-block Street Crossing. Transportation Planning and Technology, 2013. 36: 450-62.

Ye, F., Lord, D., 2014. Comparing three commonly used crash severity models on sample size requirements: multinomial logit, ordered probit and mixed logit models. Anal. Methods Accid. Res. 1, 72–85.

Zhang, Yunchang (2019): Pedestrian – Vehicle Interactions at Semi-Controlled Crosswalks: Explanatory Metrics and Models. figshare. Thesis.

Zhuang X, Wu C. Pedestrians' Crossing Behaviors and Safety at Unmarked Roadway in China. Accident Analysis & Prevention, 2011. 43:1927-36.

APPENDIX: JOURNAL PAPERS PUBLISHED FROM THIS WORK

CCAT Project Title: Smart Interaction – Pedestrians and Vehicles in a CAV Environment

Paper 1

Zhang, Y., Qiao, Y., & Fricker, J. D. (2020). Investigating Pedestrian Waiting Time at Semi-Controlled Crossing Locations: Application of Multi-State Models for Recurrent Events Analysis. *Accident Analysis & Prevention, 137*, 105437.

Abstract:

"Semi-controlled" crosswalks are unsignalized, but clearly marked with "yield to pedestrian within crosswalk" signs. Ideally, pedestrians can cross the street immediately after they arrive at the curb. However, real world observations show that pedestrians and vehicles are often involved in non-verbal "negotiations" to decide who should proceed first. This kind of "negotiation" often causes delays for both parties and may lead to unsafe situations. The study in this paper was based on video recordings of the waiting behaviors of 2059 pedestrians interacting with 1003 motorists at selected semi-controlled crosswalks. One such location experienced a conversion from one-way operation to two-way operation, which provided a rare opportunity for a before-and-after study at that location. Multi-state Markov models were introduced as a novel approach to correlate the dynamic process between recurrent events. Time-varying covariates related to pedestrian characteristics, traffic condition, and vehicle dynamics (distance and speed) turned out to be significant.

The analytical method developed in this study provides a tool to dynamically model pedestrian waiting decisions with uncertainties. Model results reveal that, after the conversion from one-way to two-way operation, the probability of a pedestrian accepting a lag decreases from 69.7% to just below 60% on the same street. In addition, pedestrians are more hesitant to cross a two-way street than a one-way street. Countermeasures that increase motorist yielding rate or reduce pedestrian confusion will enhance safety such crossing locations.

Paper 2

Zhang, Y., & Fricker, J. D. (2020). Multi-State Semi-Markov Modeling of Recurrent Events: Estimating Driver Waiting Time at Semi-Controlled Crosswalks. *Analytic Methods in Accident Research*, 100131.

Abstract:

At "semi-controlled" crosswalks, signs and markings are present, but delay to pedestrians and motorists is largely the result of the "negotiation" between the two parties to determine who yields. This paper proposes a novel approach using multi-state semi-Markov models to investigate motorists' delay and their interactions with pedestrians. Motorist waiting behavior can be divided into a series of gap acceptance decisions as part of a Markov Chain. Each gap acceptance decision is modeled as a specific transition between two states in the Markov Chain.

To demonstrate the reliability of the proposed models, multi-state semi-Markov models are estimated for the waiting behavior of more than 1,000 drivers in the presence of pedestrians at semi-controlled crosswalks. The multi-state semi-Markov models are capable of dealing with specific challenges related to (i) the need to account for recurrent events and (ii) a generalized framework for vehicle delay estimation and simulation at semi-controlled crosswalks. The extent to which motorists behave more aggressively and impatiently as their delay increases is demonstrated. Differences in behavior for operators of buses and trucks were also identified. The semi-Markov method is also able to deal effectively with the "pulsing" arrival patterns of pedestrians at crosswalks as university classes begin and end nearby and handle temporal heterogeneity. Finally, to address aggressive driver behavior, several safety implications are discussed.

Paper 3

Zhang, Y., & Fricker, J. D. (2021). Investigating temporal variations in pedestrian crossing behavior at semi-controlled crosswalks: a Bayesian multilevel modeling approach. *Transportation Research Part F: Traffic Psychology & Behavior*, 76, 92-108.

Abstract:

"Semi-controlled" crosswalks are unsignalized, but have clear pavement markings and "yield to pedestrian" signs. At these locations, pedestrians and motorists frequently interact to determine who should proceed first. When interacting with drivers, pedestrian crossing decisions are complex events that involve a variety of human responses, as well as vehicle dynamics, traffic characteristics, and environmental conditions. In addition, these complexities can be subject to temporal effects. Without considering temporal variations in pedestrian-motorist interaction, statistical methods could lead to biased coefficient estimates and inaccurate conclusions.

The study developed a Bayesian multilevel logistic regression (BMLR) model to capture heterogeneities in pedestrian interaction behavior during four different time periods. The proposed method incorporates time-specific effects that vary randomly between time-periods based on a weakly informative prior. The results indicate significant factors, some of which confirm previous research and some that are new ways to explain pedestrian behavior at the individual level. The identification of variables such as FlowOn and FlowWait sheds light on the interactions between pedestrians – providing more information than the single GroupSize measure.

Some consequent safety implications are discussed from the perspectives of vehicle dynamics, vehicle flow rate and pedestrian volume. The more detailed metrics developed in this paper will provide a valuable starting point. for the design of crosswalk controls that will foster a higher degree of compliance and less delay.