Published in final edited form as:

J Transp Eng. 2011 July ; 137(7): 455–465. doi:10.1061/(ASCE)TE.1943-5436.0000225.

Event-Based Modeling of Driver Yielding Behavior at Unsignalized Crosswalks

Bastian J. Schroeder, Ph.D.^{1,*} and Nagui M. Rouphail, Ph.D.²

¹Senior Research Associate Institute for Transportation Research and Education (ITRE) North Carolina State University Centennial Campus, Box 8601 Raleigh, NC 27695-8601 USA
Bastian_Schroeder@ncsu.edu Tel.: +1-919-515-8565 Fax: +1-919-515-8898

²Director Institute for Transportation Research and Education (ITRE) North Carolina State University Centennial Campus, Box 8601 Raleigh, NC 27695-8601 USA Rouphail@ncsu.edu
Tel.: +1-919-515-1154 Fax: +1-919-515-8898

Abstract

This research explores factors associated with driver yielding behavior at unsignalized pedestrian crossings and develops predictive models for yielding using logistic regression. It considers the effect of variables describing driver attributes, pedestrian characteristics and concurrent conditions at the crosswalk on the yield response. Special consideration is given to ‘vehicle dynamics constraints’ that form a threshold for the potential to yield. Similarities are identified to driver reaction in response to the ‘amber’ indication at a signalized intersection. The logit models were developed from data collected at two unsignalized mid-block crosswalks in North Carolina. The data include ‘before’ and ‘after’ observations of two pedestrian safety treatments, an in-street pedestrian crossing sign and pedestrian-actuated in-roadway warning lights.

The analysis suggests that drivers are more likely to yield to assertive pedestrians who walk briskly in their approach to the crosswalk. In turn, the yield probability is reduced with higher speeds, deceleration rates and if vehicles are traveling in platoons. The treatment effects proved to be significant and increased the propensity of drivers to yield, but their effectiveness may be dependent on whether the pedestrian activates the treatment.

The results of this research provide new insights on the complex interaction of pedestrians and vehicles at unsignalized intersections and have implications for future work towards predictive models for driver yielding behavior. The developed logit models can provide the basis for representing driver yielding behavior in a microsimulation modeling environment.

INTRODUCTION

The safety of pedestrians on US roadways is a national public health issue. The National Highway Traffic Safety Administration (NHTSA, 2006) lists a total of 4,784 pedestrian fatalities in 2006 and 61,000 pedestrian injuries in traffic collisions, with 46% of fatalities occurring during times of evening peak traffic and poor lighting conditions between the hours of 3pm and midnight. The risk of pedestrian injuries is elevated for older pedestrians, with 31% of fatalities being in the age group of over 55 (NHTSA, 2006), despite a lower rate of exposure. The rate of pedestrian deaths per 100,000 people is 66% higher compared to the under 55 age group. The risk is also elevated for children. In an evaluation of pedestrian and bicycle injuries based on hospital data, 38.5% of motor vehicle roadway injuries were to children under the age of 14 (FHWA, 1999). In terms of contributing

*Corresponding Author .

factors, 22.2% of pedestrian fatalities were related to improper crossings and 14.2% due to failure to yield right of way at the crossing. In comparison, only 1.5% of fatalities were related to a failure to obey traffic signals (NHTSA, 2006). The low compliance rates of motorists in terms of yielding to pedestrians at unsignalized locations has been documented through national research efforts at roundabouts (Rodegerts et al., 2007) and mid-block crossing locations (Fitzpatrick et al., 2007).

This research explores factors that may predict the likelihood of yielding from microscopic behavioral characteristics. The focus is on unsignalized pedestrian crossings, where the interaction of drivers and pedestrians can be described in terms of pedestrian gap acceptance and driver yielding behavior. While behavior at signalized pedestrian crossings can be evaluated through signal timing parameters (Schroeder, Rouphail, and Lehan, 2009), behavior at unsignalized crossings is more difficult to quantify. Oftentimes, crossing effectiveness and safety of unsignalized crossings are described in terms of aggregated characteristics, such as site-average percentages of driver yields. This assumption oversimplifies the complexity of the interaction. Some of the research questions that arise: Are drivers more likely to yield at low speeds? Is a driver's propensity to yield reduced if traveling in a platoon of vehicles? Does assertive pedestrian behavior or the presence of multiple pedestrians affect driver behavior? What is the effect of pedestrian crossing treatments on driver behavior? With an increasing frequency of unsignalized pedestrian crossings in the US, both at mid-block locations and at modern roundabouts, these and similar questions need to be addressed.

In this research, the authors collected event-specific, microscopic behavioral data on the interaction of drivers and pedestrians at two unsignalized mid-block crosswalks. While questions of driver behavior and compliance are important issues at signalized crossings as well, the present effort was strictly focused on unsignalized locations. The collected data include continuous, discrete and binary variables characterizing the interaction of two agents; one driver and one pedestrian. Special attention is given to parameters describing the 'vehicle dynamic constraints' of drivers approaching the crosswalk. The research includes before and after evaluations of two unsignalized mid-block crossings with two different pedestrian safety treatments installed; an in-street pedestrian crossing sign and pedestrian-actuated in-roadway warning lights as described in sections 2B and 4L of the US Manual on Uniform Traffic Control Devices (MUTCD: FHWA, 2003).

This paper presents the data collection methodology, discusses selection and analysis of the behavioral parameters, and applies various statistical modeling techniques to the data. The paper concludes with a discussion of how the resulting logit models can be used to describe driver yielding behavior in a microsimulation environment.

Research Objective

The objective of this research is to explore factors affecting driver yielding behavior at unsignalized pedestrian crossings. The underlying research hypothesis is that drivers have a choice of whether or not to yield to a pedestrian and that this choice is a function of observable and quantifiable parameters, including vehicle dynamics, driver and pedestrian attributes and concurrent events at or near the crosswalk. The analyzed data are event-based and thus microscopic in nature. They do not take into account site-specific or other macroscopic characteristics of the interaction, except for their implicit effects on the pedestrian and driver populations.

Literature

A review of the literature demonstrated that the interaction of drivers and pedestrians at unsignalized crossings is not fully understood and that it further lacks computational procedures. The pedestrian chapter in the Highway Capacity Manual, HCM (TRB, 2000) states that in the case of zebra-striped crossing, the procedure for calculating pedestrian delay “does not apply, because pedestrians have the right-of-way”. The method assumes full compliance with legislation that requires drivers to yield to pedestrians in the crosswalk. However, experience and observational studies (Fitzpatrick et al, 2007 and Rodegerts et al, 2007) show that the compliance rate of drivers at marked but unsignalized crosswalks is likely to vary significantly across locations and for different crossing treatments. At the same time, there is likely to be some yielding behavior, even at locations without demarcation of the crosswalk, rendering the HCM2000 procedures inapplicable at most locations. In the ongoing development of the new version of the manual (HCM2010) this limitation is not sufficiently addressed, suggesting that this analysis gap will continue to exist.

Engineering assessments of unsignalized pedestrian crossings typically report the percentage of drivers who yield to pedestrians. For example, Fitzpatrick et al, 2007 summarized previous evaluations of pedestrian-actuated in-roadway warning lights similar to one of the treatments investigated in this research. In a meta-analysis of five previous studies at 11 sites, the average driver yield rate was 66% with a range from 8% to 100%. For in-street pedestrian warning signs, the authors found a range of yielding from 44% to 97% from 13 different sites. This suggests a large variation across sites and implies ambiguity about treatment effectiveness. The pedestrian countermeasure selection system PEDSAFE (Harkey and Zegeer, 2004) includes links to similar case studies demonstrating the effect of different treatments on the percentage of driver yielding.

But what factors, other than, or in addition to the presence of a treatment, impacts driver yielding behavior? Conceptually, the behavior of drivers yielding to pedestrians is similar to a driver reaction to the onset of ‘amber’ at a signalized intersection, where it is recognized that the ability of a driver to come to a stop is a function of driver reaction time and an assumed deceleration rate (ITE, 1982). Consequently, the amber phase is timed long enough to allow vehicles unable to stop before the stop bar to proceed through the intersection. In this research, it is hypothesized that at an unsignalized pedestrian crossing, the event of a pedestrian stepping into the crosswalk influence area (CIA) corresponds to the event of a signal light status changing to amber. Accordingly, a yield is only feasible if a driver can reasonably react to the arrival of a pedestrian given the vehicle travel speed, distance from the conflict area, and maximum (comfortable) deceleration rate for the individual driver. In other words, yielding is unlikely (if not impossible) if the vehicle is subject to *vehicle dynamics constraints*. The effect of vehicle speed on the propensity of a driver to yield was established empirically for pedestrian crossings at roundabouts (Geruschat and Hassan, 2005) and is intuitively supported by pedestrian safety treatments emphasizing traffic calming and speed reduction (Harkey and Zegeer, 2004).

From prior research, it is further known that pedestrian attributes have an impact on yielding. Yielding has been empirically linked to pedestrian assertiveness and brightly-colored clothing (Harrel, 1993), the positioning of the pedestrian relative to the edge of curb (Geruschat and Hassan, 2005), and the presence of multiple pedestrians (Sun et al., 2002) at the crossing. It can further be argued that concurrent events at the crosswalk have a contributing effect. For example, the presence of queues or congestion beyond the crosswalk or the occurrence of a yield event in the opposing direction or adjacent travel lane may enhance the likelihood of yielding. For example, it has been established that drivers are more likely to yield at the entry of a modern roundabout than at the exit leg (Ashmead et al.,

2005, Geruschat and Hassan, 2005, and Rodegerts et al., 2007). Intuitively, yielding is also impacted by driver and traffic characteristics, including the type of vehicle, the vehicle position within a platoon of vehicles, and driver courtesy.

In an attempt to analyze yielding patterns more closely, Sun et al (2002) applied logit and probit models to observational data from an unsignalized pedestrian crossing site. The authors demonstrated the feasibility of using a discrete choice modeling approach through model validation results and found that drivers are more likely to yield to a group of pedestrians, that older drivers were more likely to yield than younger ones, that male drivers were more likely to yield than female ones, and that drivers of heavy vehicles were more likely to yield than drivers of passenger cars. The latter was attributed to a high percentage of public transit vehicles in the heavy vehicle class. Finally, the authors found that the yield probability increased if an adjacent yield was already underway. The authors did not consider elements of vehicle dynamics such as speed and position of the vehicle.

METHODOLOGY

The objective of this effort was to build upon findings of earlier research to develop an event-based predictive model of driver yielding. The analysis presumes that the yield decision is influenced by vehicle dynamics, driver and traffic characteristics, pedestrian attributes, and concurrent events at the crosswalk. Consequently, a data collection methodology had to be devised to capture these explanatory variables and the discrete outcome of the yield response from field observations. First, the interaction event and key variables are defined.

Definitions

A pedestrian-driver interaction event is defined as a pedestrian arriving in the crosswalk influence area (CIA) while a driver is on the approach of the crosswalk. At this point in time (t_1):

- The pedestrian has indicated that he or she is intending to cross at the facility (rather than continuing along the sidewalk).
- The driver is aware of the pedestrian's intention and must respond (make the decision to slow down or continue through the crosswalk).
- The observer is aware that an event sequence is about to take place (from video observation) and records the time and event parameters.

From the onset of a pedestrian-driver interaction event at time t_1 , there are two potential outcomes to driver yielding behavior:

- I. Yield [Y] – The approaching vehicle creates a crossing opportunity for the pedestrian by either coming to a full stop (hard yield, HY) or by decelerating to accommodate the pedestrian (soft yield, SY).
- II. Non Yield [NY] – The driver decides not to yield to the pedestrian; either because it is impossible to execute from its current position and speed, or because he/she is unwilling to yield.

This paper focuses only on the study of yield and non-yield events. Pedestrian gap acceptance data will be analyzed and reported in a future manuscript based on completed research (Schroeder, 2008).

Field Data Collection and Variables

In order to collect data that include vehicle speeds and relative positions to the crosswalk for each event, a laser speed gun (LIDAR) was time-synchronized with video recordings of the crosswalk. The LIDAR device was able to store each speed measurement and associated distance of the vehicle from the observer at a manufacturer-specified accuracy of ± 1 mph (1.6 km/h) (Laser Atlanta, 2008). An independent comparison of speed measurement devices (Gates, Schrock, and Bonneson, 2004) confirmed that at 35 mph approach speeds a LIDAR device was 98% accurate at ± 0.5 mph tolerance and 100% accurate for ± 1.5 mph tolerance. According to ITE (2000) the *cosine error* of measurement is less than 1 mph at a roadside angle less than 10 degrees, and is thus within human observer error if the measurements are made in a near head-on fashion. With the observer positioned at a known and constant distance from the crosswalk, it is possible to infer the distance between the vehicle and the crosswalk.

The time-synchronized video and laser speed observations were post-processed in the laboratory. For the speed measurements, it was assumed that the vehicle speed remained approximately constant between the time the pedestrian arrived in the CIA (t_1) and the time of the LIDAR measurement (due to the reaction time of the observer). Accordingly, the relative displacement of the vehicle during that time can be calculated to obtain an estimate at time t_1 . All additional variables were collected as defined in Table 1. The set of variables spans various elements of the interaction, including vehicle platooning (PLT), pedestrian assertiveness (AST), presence of a treatment (TRTMT) and treatment activation (FLASH). The latter two variables distinguish between 'before' and 'after' analysis periods (TRTMT), and whether the treatment was activated by the pedestrian (FLASH). For each event, all parameters were extracted manually by recording time stamps for pedestrian and vehicle arrivals at the crosswalk (continuous variables) and noting the state of all binary variables. An event is recorded if at the time a pedestrian has shown clear intention of crossing at the facility, a vehicle is approaching the crosswalk. All observations were made by the same observer to ensure consistency in reporting. The analysis was performed using various procedures in the SAS software package (SAS Institute Inc., 2001).

Site Description

The methodology was applied to two unsignalized mid-block crosswalks in North Carolina. The sites each have two-lane cross-sections and were selected partly because of heavy pedestrian activity and vehicle flows, and partly because of planned pedestrian crossing treatment installations at both sites.

The first crosswalk is adjacent to the campus of Queens University on Selwyn Avenue in Charlotte, NC. The two-lane site has wide 20 foot lanes (6.1 meter) with a posted speed limit of 35 mph (56.3 km/h). The roadway has on-street parking starting approximately 100 feet (30.5m) to the north of the crosswalk. The crosswalk is used by students, faculty and residents in the communities adjacent to the university. The road functions primarily as a north-south arterial. Because of concerns for pedestrian crossing safety, the City of Charlotte decided to install pedestrian-actuated in-pavement flashing lights at the crosswalk. This site will be referred to as Site A in this paper.

The second crosswalk is on the campus of North Carolina State University (NCSU) on Sullivan Drive in Raleigh, NC. The facility has standard 12-foot (3.7m) lanes and a speed limit of 25 mph (40.2 kph). The crosswalk is used primarily by students walking to class from adjacent surface parking lots. The road functions as access to the parking facility, but is also used by through traffic avoiding heavier-traveled arterial roads to the north and south. The NCSU Department of Transportation decided to install an in-road pedestrian warning

sign at the crosswalk to increase driver yielding behavior. This site will be referred to as Site B. Figure 1 shows photos of the data collection sites and close-up views of the treatments. The authors recorded approximately six hours of video for each phase and each site, resulting in a total of 1074 yield interaction events.

DATA ANALYSIS

Data were collected describing driver characteristics, pedestrian characteristics, the dynamics of the closest approaching vehicle, and the circumstances surrounding the pedestrian-vehicle interaction event (Table 1). Three different response variables were evaluated. The first, YIELD, is a binary indicator describing whether the driver yielded or not. The categorical variable Y_ORDERED, further divides yields into ‘soft yields’ (SY) and ‘hard yields’ (HY). In this context, a HY describes an event where the vehicle comes to a complete stop near the crosswalk. Y_TYPE is a binary variable that only considers yield events and has two levels SY and HY.

Descriptive Statistics

The authors initially examined descriptive statistics to observe general trends and variability in the data. The results in Table 2 show the mean and standard deviation for all 604 data points at Site A, including before (361) and after (243) treatment conditions, as well as yield (105) and non-yield (499) events.

The response variable, YIELD, is a binary variable, so its mean is equivalent to the overall observed yielding rate. The overall yielding rate at Site A is 17.4%, with 15.0% yielding in the ‘before’ case and 20.9% yielding after treatment implementation ($p=0.065$). The ordered variable Y_ORDERED shows a similar trend, and Y_TYPE suggests that 66.7% of yields were classified as hard yields. The average observed speed (SPEED_FT) at the site is 40.34 ft/sec (27.5mph or 44.26 km/h) and is significantly higher for non-yield events ($p<0.0001$). The installation of the treatment resulted in a minor reduction in average speed of 1.4mph ($p=0.076$).

The analysis of vehicle data shows evidence of frequent vehicle platooning (PLT=67.1%), which indicates an effect from an upstream signal. Overall, only few events are associated with an adjacent yield (ADY=8.4%), vehicle queuing (QUE=0.7%), or heavy vehicles (HEV=1.5%). Comparing the data for yield and non-yield events, almost 3 times as many yields were associated with the actuated flasher (FLASH). Yield events also were associated with a higher likelihood of pedestrian assertiveness (AST), significantly lower speeds and longer distances from the crosswalk (both $p<0.0001$). Consequently, yield events are associated with higher time-to-collision (TTC, $p<0.0001$) and a lower necessary deceleration rates (DECEL, $p=0.0017$). Interestingly, the treatment was only activated in 32.4% of the events captured in the ‘after’ case.

Table 3 depicts the descriptive statistics for Site B with a total sample size of 470 data points (265 ‘before’ and 205 in the ‘after’ condition). Overall, 158 yields and 312 non-yields were observed, suggesting a yielding rate of 33.6%. The yielding percentage increased significantly after installation of the treatment from 26.8% to 42.4%, corresponding to a 58% increase in yielding ($p=0.0004$). The facility statistics show a mean speed of 36.6 ft/s (24.9mph or 40.1km/h) in the before condition and 33.3 ft/s (22.70mph or 36.5km/h) in the after period ($p=0.0001$). Most other variables appear to be consistent in the before and after data sets, with slight differences explained by random variation ($p>0.1$). Only the variable PREV had a significantly lower mean in the ‘after’ data set ($p=0.006$).

Several variables exhibit significant differences when comparing the ‘yield’ and ‘non-yield’ data sets. Yielding events appear to be positively correlated with ADY, AST, MUP, PXW, TRIG, TRTMT, DIST1, and TTC, and inversely correlated with DECEL and SPEED. Surprisingly, the effect of platoons (PLT) is not significant at this site ($p=0.187$). With an overall percentage of 38.3%, platoons are less common here than at Site A site. More importantly, the lower speed of vehicles (and thus platoons) may decrease a driver’s perceived risk of a rear-end collision when yielding. Consistent with the other site, driver yielding behavior seems to be associated mostly with vehicle dynamics and the behavior of the pedestrian.

Vehicle Dynamics Constraints

It was argued earlier that drivers may be subject to vehicle dynamics constraints (VDC) when making their decision to yield to a pedestrian. Figure 2 shows the relationship between vehicle speed and distance by yield outcome for both sites. The figure shows non-yield events as stars, soft yields as circles and hard yields as squares. A curve representing a constant deceleration rate of 10 ft/sec^2 (3.048 m/sec^2) has been superimposed, calculated from the relationship $\text{DECEL} = (\text{SPEED} \times \text{SPEED}) / (2 \times \text{DIST1})$. This deceleration rate has traditionally been used in US signal timing applications (ITE, 1982) when calculating the length of the amber phase for drivers and is largely consistent with today’s practice that uses a deceleration rate of 11.2 ft/sec^2 (3.4 m/sec^2) (FHWA, 2003). At signals, the amber phase is timed long enough, so that drivers that are too close and too fast to stop can proceed through the intersection without running the red light and others can decelerate at this ‘comfortable’ deceleration rate.

The relationship in Figure 2 supports the notion of VDC affecting the likelihood of drivers to yield. At Site A no yield events were observed above 10 ft/sec^2 and for Site B only 2 events were observed above the threshold. The figures further show that a large number of non-yields are observed well below the theoretical yield threshold. This is explained from the overall low propensity to yield at the sites, independent of VDC. It is also important to keep in mind that non-yielding vehicles that were far from the crosswalk may have been part of a fast-moving vehicle platoon, or simply were drivers that had no interest in yielding.

From the relationships shown, the authors hypothesized that for a ‘high-yield’ site, characterized by low speeds, narrow lanes, a courteous driver population, and assertive pedestrians, the VDC threshold could be a fairly accurate predictor for yield and non-yield events. This will be explored in future research at other sites.

MODEL DEVELOPMENT

Variable Selection

Given the strong indication of the impact of VDC, the authors decided to exclude all observations with a deceleration rate in excess of 10 ft/sec^2 from further analysis. The reasoning for excluding these data is that they would introduce a bias in classifying these as ‘non-yield events’ in the logistic model. These particular drivers may very well have yielded, had they only been a little further away from the crosswalk at the time of the pedestrian arrival. The goal of the following analysis then is to assess what other factors may explain the yielding behavior of drivers, who actually had a choice to yield. The remaining data set contained 540 and 430 observations for Sites A and B, respectively. Additionally, the authors decided to exclude parameters HEV and QUE because 2×2 interaction tables with the YIELD response showed several cells with too few observations, resulting in *scarce* data (Agresti, 2007). The remaining explanatory variables to be considered in the regression model are highlighted by checkmarks in table 1.

Logit Model Form

The analysis of categorical data requires different statistical methodologies from interval or continuous data (Washington et al. 2003). Driver yielding behavior is measured as a binary variable with the two values 1=yield, 0=non-yield.

There are several options for analyzing the data. The most intuitive approach is to utilize the YIELD response variable in a *binary logistic regression*, to predict the likelihood of drivers yielding. Other approaches that utilize an ordered dependent variable (NY, SY, and HY) include a multinomial logit form and a cumulative logit model for ordered responses. Both of these model forms were explored, but did not result in conclusive results. The binary logit model is further preferred in a practical context, because of ease of interpretation and more importantly the ability to implement it in a microsimulation environment. The distinction between SY and HY events can then be modeled as a two-stage binary logit, that first predicts the likelihood of yielding, and then the likelihood of HY given that a yield was predicted in the first level. In the initial model, all data are evaluated to predict the probability of yielding, $P(\text{Yield})$. In the second stage only those events that resulted in a yield are evaluated to predict the likelihood of a hard yield, $p(\text{HY}|\text{Yield})$.

Following the discussion in Agresti (2007) the logistic regression model describing the log odds of drivers yielding, Y is given by:

$$\text{logit} [P(Y=1)] = \log \left(\frac{P(Y=1)}{1 - P(Y=1)} \right) = \alpha + \sum_{i=1}^m \beta_i x_i \quad \text{Equation 1}$$

with intercept α and parameters β_i describing the effects of m explanatory variables x_i on the yield response. Keeping all other effects fixed, a one-unit increase in the variable x_i has a multiplicative effect of e^{β_i} on the odds of Y . This is referred to as the *odds ratio* of the effect. The model parameters can be obtained through maximum likelihood estimation in SAS statistical software. The probability estimates for the yield response can be obtained using the exponential function as follows:

$$P(Y|\vec{x}) = \frac{e^{\alpha + \sum_{i=1}^m \beta_i x_i}}{1 + e^{\alpha + \sum_{i=1}^m \beta_i x_i}} \quad \text{Equation 2}$$

The yield probabilities calculated from equation 2 can be used to quantify the impact of a change in a binary variable on the likelihood of yielding, or to plot the change in $P(Y|\vec{x})$ by varying a continuous variable. By multiplying probabilities $P(\text{Yield})$ and $P(\text{HY}|\text{Yield})$ the likelihood of HY and SY outcomes can be predicted from the overall data set.

Additional insight on the absolute fit of a model to the data is given by the max-rescaled R-Square (max-R^2) statistic (SAS, 2007). It describes the amount of overall variability in the data explained by the model, similar to the R^2 statistic commonly used in linear regression. Generally, a higher max-R^2 indicates that more of the variability in the response variable is explained by the model.

Modeling Approach

Assuming a binary logit form, two different modeling approaches are applied. First, an *unrestricted model* is evaluated, which uses all eligible parameters. It is motivated by the desire to understand which variables suggest an impact on yielding from the empirical data. This initial modeling approach utilizes all feasible variables and lets a model selection

algorithm determine which variables significantly improve model fit. The resulting model parameter set is then compared to research hypotheses. The parameters are also scrutinized with respect to feasibility of implementation in a microsimulation context. The authors used the forward selection algorithm (with entry significance level 0.05) in SAS PROC LOGISTIC for this unrestricted model. The algorithm adds variables to the model based on the Chi-Square statistic, until the significance levels of all remaining variables exceed the desired confidence level.

After exploring the feasibility of the unrestricted parameters, a second logistic model was limited to variables that can be implemented in simulation. According to Agresti (2007) it is important to consider the practical significance of a logit model in addition to statistical fit. Consequently, the *restricted model* only considers variables that are available in a simulation context to predict yielding behavior. In particular, simulation models use algorithms to predict changes in behavior at certain times in the simulation. In this context, it is assumed that the driver makes the yield decision at the instant the pedestrian arrives at the crosswalk. Therefore, the variables under consideration for the restricted model are by definition limited to those that are observable and quantifiable at this instant. Consequently, the restricted model is strictly *predictive* using data available at the onset of the interaction event, rather than *explanatory*, which would include variables observed after the yield decision has already been made. The resulting model was developed from only two sites and should therefore not be applied to other sites. However, the motivation behind this exercise is to guide future research towards developing comprehensive predictive models for driver yielding behavior.

The authors presume that the following explanatory variables can be quantified at the time of pedestrian arrival and are considered as predictors in the restricted model: AST, FLASH, HEV, MUP, NEAR, PLT, TRTMT, SPEED_FT, DECEL. These variables include descriptors of the crosswalk geometry (TRTMT), pedestrian behavior (AST, FLASH, MUP, NEAR), and the approaching driver (DECEL, HEV, PLT, SPEED_FT). The remaining variables are not observable at the onset of an event and therefore have been excluded from consideration in the restricted model. It is important to highlight the difference between variables TRTMT and FLASH, both of which are associated with treatment installation. The variable TRTMT strictly distinguishes between data collected before and after treatment installation as a binary indicator variable. It is applicable to both test sites. The variable FLASH only applies to Site A, where this binary indicator distinguishes whether or not a pedestrian in fact activated the in-pavement flashing lights treatment. It is hypothesized that FLASH is a better indicator of treatment effectiveness at Site A than TRTMT.

RESULTS

Yield Models for Site A

Table 4 shows the logistic models for Site A. The unrestricted model suggests that drivers are more likely to yield if an adjacent yield is present (ADY), if the pedestrian is waiting at the nearside of the road (NEAR), if the pedestrian triggers a yield by stepping into the roadway (TRIG), if a previous pedestrian is still in the crosswalk (PXW) and if the in-pavement pedestrian flasher is activated (FLASH). Similarly, the likelihood of yielding is decreased if the driver is traveling in a platoon (PLT), if the previous vehicle failed to yield (PREV) and if yielding would require greater deceleration rates (DECEL). This last effect is particularly interesting, considering that vehicles subject to vehicle dynamics constraints have been removed from the analysis. Note that the DECEL parameter is essentially an interaction term of DIST1 and SPEED_FT, neither of which are significant at the selected confidence level. The use of the interaction term without the individual effects is justified,

because the term has scientific meaning (the deceleration rate) and is therefore not a conventional interaction term.

When all other parameters are kept constant, the odds ratio e^{β_i} represents the multiplicative effects of each variable x_i on the odds of yielding (Agresti, 2007). For example, the estimated odds of a yield occurring if there is an adjacent yield is 2.63 times the odds for a non-ADY event. If the vehicle travels in a platoon, the odds for a yield are reduced 0.47 times and are also reduced 0.36 times if the previous vehicle did not yield. The strongest effect is evident from TRIG, followed by PXW, FLASH and ADY. The flasher actuation increased the odds of yielding 2.79 times compared to a non-actuation. The odds in yielding are further reduced 0.65 times for each additional 1ft/sec² in necessary deceleration rate.

In the development of the restricted model, the authors went through several iterations in which variables were manually added and removed from the model, while comparing parameter significance and overall model fit statistics. In this paper, only the final iteration of this exercise is presented. The final model predicts that the likelihood of yielding is increased with pedestrian assertiveness (AST), which increases the odds of yielding 5.59 times. The activation of the flashing lights (FLASH) increased the odds of yielding by a factor 3.28 while vehicle platooning (PLT) decreased the odds by an odds ratio of 0.39. The deceleration rate also remains a significant factor, decreasing the odds of yielding by a factor of 0.68 for each 1 ft/sec² (0.305 m/sec²) increase in DECEL.

Following the estimation of P(Yield), a second-level restricted model was developed through a similar process to predict the probability of P(HY|Yield). The model predicts a decrease in the probability of a hard yield with indicators AST and FLASH equal one, and an increase with greater deceleration rate (DECEL). The continuous variable DECEL showed a p-value of 0.2284 and is therefore not statistically significant at the given sample size. The decision to retain this variable in the restricted model was motivated for reason of consistency with the first-stage yield prediction, and to assure some sensitivity of the model over the range of observed vehicle dynamics (speed and distance from the crosswalk). The effect of DECEL is intuitive, since a vehicle with a high DECEL value tends to not have enough time for a soft yield given its speed and proximity to the crosswalk. The model further suggests that a driver who has decided to yield is more likely to perform a soft yield for pedestrians who are assertive in their approach to the crosswalk and to those that activate the treatment. Conceptually, both assertive behavior and a flashing beacon activation can be interpreted as advanced notifications to the drivers. They therefore tend to elicit an earlier deceleration to a soft yield.

Yield Models for Site B

Applying the forward selection algorithm with a $p=0.05$ threshold, SAS produces an eight-variable unrestricted model for Site B (Table 5). The model includes parameters ADY, AST, NEAR, PREV, PXW, TRIG, TRTMT, and DECEL. The largest odds ratios are present for AST and TRIG, which increase the log odds of a yield occurring by factors 6.1 and 5.3, respectively. Effects of ADY, NEAR, and PXW are also high with odds ratios greater than 2. The treatment effect still increases the yield probability 1.7 times. There is further evidence that the necessary deceleration rate (DECEL) has a strong effect. Despite the fact that events with DECEL greater than 10 ft/sec² were excluded from model development, an increase of 1 ft/sec² in necessary deceleration rate still decreases the log odds of yielding 0.69 times.

Similar to Site A, the restricted model is limited to variables that can be implemented in microsimulation. The resulting five-variable restricted model predicts an increasing likelihood of yielding from AST, NEAR and TRTMT. The variables PLT and DECEL

decrease the likelihood of yielding. The strongest effect is evident from pedestrian assertiveness, which in the absence of the TRIG variable now has an odds ratio of 12.02. Keeping everything else constant, the assertiveness of a pedestrian therefore has a major impact on yielding behavior at this site, which is consistent with anecdotal field observations.

When predicting $P(HY|Yield)$ for Site B the resulting restricted model predicts that the likelihood of a hard yield is decreased by pedestrian assertiveness and increased with a greater deceleration rate. This relationship is consistent with the findings for Site A, except that the treatment effect was not significant here. So, while the static treatment at Site B affects the overall yield propensity, it fails to elicit the same effect on the type of yield.

Model Comparison

In the comparison of all developed models, statistical tests for the overall model (likelihood ratio, Wald, and Score tests) are significant at $p=0.0001$ for all models. However, the max- R^2 values for the models are fairly low and generally fall in the range of 0.10 to 0.40. This is expected, because of the complexity of the interaction of the two modes and because the models were strictly based on observable data. For example, the field data collection has no way of assessing “driver courtesy” or “sense of urgency”, where a particular driver is rushed to get to a destination on a particular day. Combined with a lack of understanding of yielding laws, it is generally expected that unobserved sources of variation are high when it comes to yielding behavior.

A comparison of the max- R^2 statistics for the unrestricted and restricted models suggests that some model fit is sacrificed when excluding variables that are not observable at the onset of the interaction event. For the MB-CTL, site the statistic is reduced from 0.365 to 0.267 and at Site B from 0.320 to 0.244 when moving from the unrestricted to the restricted models. However, a comparison of overall model fit using the $-2\log L$ statistic fails to reject the null hypothesis for both sites. The Chi-Square statistic for the difference of $-2\log L$ of 0.432 at 4 df MB-CTL has a p-value of 0.980 and the $-2\log L$ difference of 2.032 at 3 df results in a p-value of 0.566 at Site B. While no statistical test is possible, the AIC statistics support the notion that the overall fit of the unrestricted model is not significantly better than that of the restricted model for the $P(Yield)$ probabilities.

As a practical interpretation of the outcome of the restricted logit models, a graphical representation of the probability curves is presented. Figure 3 plots the quantities $P(Yield)$ and $P(HY|Yield)$ for changes in the explanatory variables AST, FLASH and DECEL for Site A. In particular, the graph for $P(Yield)$ shows a decreasing yield probability with increasing DECEL over the calibrated model range from 0 to 10ft/sec². Additionally, the graph shows that the probability curve is shifted upward if the base levels of the binary two variables (AST=0 and FLASH=0) are increased to 1. For example, at a constant deceleration rate of 5 ft/sec², the base yield probability of 9.2% is increased to 45.4% for assertive pedestrians (AST=1) and to 32.8% if the pedestrian activates the in-pavement flashing lights (FLASH=1). If both variables are true, the yield probability is predicted to be 65.2%. This graphical way of visualizing the logit model, clearly conveys the effect of the variables on the yield response.

Figure 3 further shows a set of probability curves for $P(HY|Yield)$ for Site A. The probability of a hard yield increases as a function of DECEL as discussed above. Similarly, changes in the binary levels of AST and FLASH from 0 to 1 shift the curves downward indicating a lower likelihood of a hard yield. Clearly, the shape of the figure would be reversed, if the logit model had predicted the quantity $P(SY|Yield)$ from the same data set.

Figure 4 shows the corresponding probability curves for Site B. Again $P(\text{Yield})$ decreases with DECEL, while a change of binary variables AST and TRTMT from 0 to 1 shifts the curves upward, representing a greater yield probability for the same deceleration rate. The curves for $P(\text{HY}|\text{Yield})$ for Site B show a greater likelihood of a hard yield with increasing DECEL and a downward shift resulting from changing the level of AST from 0 to 1. The figure further makes evident the stronger relative effect of the assertiveness variable when compared to the treatment effect. For example, at a deceleration rate of 5 ft/sec^2 , the installation of the treatment is predicted to increase $P(\text{Yield})$ from 13.7% to 23.2% for non-assertive pedestrians (65.5% to 78.4% for assertive pedestrians). However, assertive pedestrian behavior increased the predicted yield propensity from 13.7% to 65.5% without a treatment, and from 23.2% to 78.4% after treatment installation.

Implementation in Microsimulation

The yield prediction models developed in this research are given in a form that can be implemented in microsimulation, where they can enhance the existing representation of the interaction of pedestrian and drivers at unsignalized crossings. In most simulation models, the interaction of pedestrians and vehicles is not a main focus, but may be approximated through existing gap acceptance algorithms for vehicle operations (Schroeder and Roupail, 2007). In this previous research, the authors demonstrated how yield probabilities could be assigned globally to different driver types. However, the assumption of *global* parameters is insufficient to reflect the developed behavioral algorithms developed in this research.

The logit models contain variables describing vehicle dynamics, which in turn are a function of car-following algorithms, speed distributions, and arrival patterns in simulation. Because of the nature of the simulation software, these values are updated every simulation time step (typically 0.1 seconds). Correspondingly, the yield probability functions are subject to change from one time step to the next. The vehicle dynamics element therefore requires the logit implementation to be in the form of *core model algorithms*, a term coined by the Federal Highway Administration's *Next Generation Microsimulation (NGSIM)* consortium (FHWA, 2004). As opposed to *globally* defined parameters, this allows for *time-sensitive* variables that are dynamically updated as the simulation progresses. While a portion of the logit variables would remain in the form of global parameters that are fixed throughout the simulation, other variables found in the logistic regression models become implicit functions of other algorithms in the simulation.

In particular, variables AST, FLASH, and TRTMT would become global attributes of the pedestrian (assertiveness and activation of the treatment) and of the system (treatment presence). The remaining variables (DECEL and PLT) are time-varying functions of the simulation algorithms, and thus not defined explicitly as pedestrian or driver attributes. These algorithms include the driver yielding and pedestrian crossing algorithms themselves. The actual logit functions are implemented in the form of core algorithms that read-in the state of the variables every simulation time step and re-calculate the decision probability. The decision outcome then becomes a random variable, and the yield responses depend on the outcome of a random experiment following the logit probability function.

In the implementation, some assumptions need to be made about behavioral consistency. In a true randomized experiment it is feasible that an entity would change its decision outcome repeatedly, which is undesirable. It is therefore recommended that once a response is "true" the decision is fixed. Once a driver decides to yield the action is followed through. More research would be necessary to predict that drivers change their behavior after the initial decision. The potential for implementing these behavioral models in microsimulation is an important contribution to the field. The NGSIM effort by the Federal Highway Administration recognizes the need to enhance pedestrian models and ranks this area 7th in a

top 10 list of modeling stakeholder requirements (FHWA, 2004). The six higher ranked problem statements are lane selection on arterials, oversaturated freeway flow, freeway lane distribution, weaving sections, two-way left-turn lanes, and response to variable message signs. However, while NGSIM is sponsoring the development of new core algorithms for simulation it will take time before limited resources become available to address aspects of pedestrian-vehicle interaction.

CONCLUSION

The findings presented in this document shed light on the complex interaction of pedestrians and vehicles at unsignalized crosswalks. The data showed clear evidence that it is important to account for vehicle dynamics constraints on the evaluation of crosswalks. The yield threshold in this study was identified to occur at a deceleration rate of 10ft/sec^2 (3.05m/sec^2), which interestingly coincides with signal timing practice for the amber phase at signalized intersections. Additionally, variables describing vehicle dynamics (specifically the necessary deceleration rate to come to a stop) proved important even for yielding situations below the vehicle dynamics threshold.

Vehicle yields also showed a strong correlation with the behavior of the pedestrian. An assertive pedestrian, who walks briskly in the approach to the crosswalk, appears to frequently elicit a response from the driver by increasing the likelihood of yielding. Furthermore, the data gave evidence of a strong effect of vehicle platoons in reducing the likelihood of yielding. In fact, the frequent occurrence of platoons may explain the low overall yielding rate at one of the studied sites. Finally, the analysis of two pedestrian treatments demonstrated an increase in the yield probability. In the case of a pedestrian-actuated treatment, the yield probability was sensitive to the activation of the treatment.

The research demonstrated that the data collection approach and analysis methodology can be combined with logistic regression techniques to develop predictive models for driver yielding behavior. The implications of such predictive models for driver yielding behavior is especially important in combination with modern microsimulation tools, which may eventually be used to predict performance of a future site. While the developed yield probability models are based on only two sites, their mathematical and logical form (restricted model) are suitable for implementation in the form of a new core algorithm in simulation.

Acknowledgments

The research leading up to this document was partially sponsored by the NCHRP 3-78 project, 'Crossing Solutions at Roundabouts and Channelized Turn Lanes for Pedestrians with Visual Disabilities' and NIH Bioengineering Research Partnership NIH Grant R01 EY12894-03, 'Pedestrian Access to Complex Intersections'. The authors would like to acknowledge the National Academies for their support and the members of the project team, who have provided continuous feedback to the research efforts. The authors would also like to thank the City of Charlotte, NC for facilitating the data collection efforts and the Federal Highway Administration Dwight D. Eisenhower Transportation Fellowship Program for graduate student financial support. This research was part of the doctoral dissertation by Dr. Bastian Schroeder that is available through North Carolina State University.

The project described was partially supported by Grant Number R01EY12894 from the National Eye Institute. This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Eye Institute or the National Institutes of Health.

REFERENCES

- Agresti, Alan. An Introduction to Categorical Data Analysis. Second Edition. John Wiley & Sons Inc.; Hoboken, New Jersey: 2007.

- Ashmead, Daniel H.; Guth, David; Wall, Robert S.; Long, Richard G.; Ponchilla, Paul E. ASCE Journal of Transportation Engineering. Vol. Volume 131. ASCE; November. 2005 Street Crossing by Sighted and Blind Pedestrians at a Modern Roundabout; p. 812-821.
- FHWA. Injuries to Pedestrians and Bicyclists: An Analysis Based on Hospital Emergency Department Data. Federal Highway Administration (FHWA); McLean, VA: 1999. FHWA-RD-99-0781999
- FHWA. Manual on Uniform Traffic Control Devices for Streets and Highways (MUTCD). US Department of Transportation, Federal Highway Administration (FHWA); Washington, DC: 2003. Available online at <http://mutcd.fhwa.dot.gov2003>
- FHWA. NGSIM Task E.1: Core Algorithm Assessment. Department of Transportation. Federal Highway Administration; Washington, DC: 2004. Publication No. #FHWA-HOP-06-00092004
- Fitzpatrick, Kay; Turner, Shawn; Brewer, Marcus; Carlson, Paul; Ullman, Brooke; Trout, Nada; Park, Eun Sug; Whitacre, Jeff; Lalani, Nazir; Lord, Dominique. TCRP Report 112/NCHRP Report 562. Transportation Research Board; Washington, D.C.: 2007. Improving Pedestrian Safety at Unsignalized Intersections.
- Gates, Tim J.; Schrock, Steven D.; Bonneson, James A. Transportation Research Record: Journal of the Transportation Research Board. Vol. Volume 1870. Washington, DC: 2004. Comparison of Portable Speed Measurement Devices; p. 139-146.
- Geruschat, Duane R.; Hassan, Shirin E. Driver Behavior in Yielding to Sighted and Blind Pedestrians at Roundabouts. Journal of Visual Impairment and Blindness. May.2005 Volume 99(Number 5)
- Guth, David; Ashmead, Dan; Long, Richard; Wall, Robert; Ponchillia, Paul. Blind and Sighted Pedestrians' Judgment of Gaps in Traffic at Roundabouts. Human Factors. 2005 Summer;Vol. 47(No. 2):314–331. [PubMed: 16170941]
- Harkey, David L.; Zegeer, Charles V. PEDSAFE: Pedestrian Safety Guide and Countermeasure Selection System. Federal Highway Administration; Washington, D.C.: September. 2004 FHWA-SA-04-003
- Harrel WA. The Impact of Pedestrian Visibility and Assertiveness on Motorist Yielding. Journal of Social Psychology. 1993; Vol. 133(3):353–360.
- HCM. Highway Capacity Manual. Transportation Research Board; Washington, DC: 2000. 2000
- IIHS. Fatality Facts 2007. Insurance Institute for Highway Safety; 2007. URL: http://www.iihs.org/research/fatality_facts_2007/default.html
- Institute of Transportation Engineers. Transportation and Traffic Engineering Handbook. Prentice-Hall; Englewood Cliffs, NJ, USA: 1982. (ITE 1982)
- NCHRP 3-78 project web page. National Cooperative Highway Research Program. Transportation Research Board; Washington, D.C.: <http://www4.trb.org/trb/crp.nsf/All+Projects/NCHRP+3-78>
- NHTSA. Traffic Safety Facts: 2006 Data. National Highway Traffic Safety Administration (NHTSA), US Department of Transportation; Washington, DC: 2006. Publication Number DOT HS 810 8182006
- Rodegerdts, Lee; Blogg, Miranda; Wemple, Elizabeth; Myers, Edward; List, George; Flannery, Aimee; Troutbeck, Rod; Brilon, Werner; Wu, Ning; Persaud, Bhagwant; Lyon, Craig; Harkey, David; Carter, Daniel. NCHRP Report 572. Transportation Research Board; Washington, D.C.: 2007. Roundabouts in the United States. <http://onlinepubs.trb.org2007>
- SAS Institute Inc.. The SAS System for Windows. Release 9.02. Copyright 1999-2001 by SAS Institute Inc.; Cary, NC, USA: 2007.
- Schroeder, Bastian. Doctoral Dissertation in Civil Engineering. North Carolina State University; May. 2008 A Behavior-Based Methodology for Evaluating Pedestrian-Vehicle Interaction at Crosswalk. 2008
- Schroeder, Bastian; Roupail, Nagui; Lehan, Brendan. Observational Study of Pedestrian Behavior along a Signalized Urban Corridor. Presented at the 88th Annual Meeting of the Transportation Research Board; Washington, D.C.. 2009; 2009.
- Dazhi, Sun; Ukkusuri, Satish; Benekahal, Rahim; Waller, S. Travis Modeling of Motorist-Pedestrian Interaction at Uncontrolled Mid-Block Crosswalks. Presented at 82nd Annual Meeting of the Transportation Research Record; Washington, DC. 2003; 2003.
- Washington, SP.; Karlaftis, MG.; Mannering, FL. Statistical and Econometric Methods for Transportation Data Analysis. Chapman & Hall/CRC; Boca Raton, Florida: 2003. 2003

Figure 1.
Photos of Data Collection Sites

Figure 2.
Speed-Distance Relationship Graph

Figure 3.
Model Probability Plots for Restricted Models Site A

Figure 4.
Model Probability Plots for Restricted Models Site B

Table 1

Variable Definitions

Code	Variable Definition	Units	Model*
Response Variable			
YIELD	Binary Yield Outcome (Yield =1, Non-Yield=0)	Bin	✓
Y_ORDERED	Ordered Yield Event Outcome (3=HY=hard yield, 2=SY=soft yield, 1=NY=non-yield)	Ordered	
Y_TYPE	The Type of yielding outcome given that there is a yield (HY=1, SY=0)	Bin	
Binary Explanatory Variables			
ADY	Presence of an adjacent yield in the opposite direction; ADY=1 if a vehicle in the opposite lane has already yielded for the pedestrian at the crosswalk	Bin	✓
AST	The pedestrian is 'assertive'; AST=1 if the pedestrian exhibits assertive behavior in the approach of the crosswalk , indicated for example through fast walking pace	Bin	✓
FLASH	Indication whether the flashing beacon was actuated by the pedestrian; FLASH=1 if beacon was flashing during the yield event	Bin	✓
FOLL	Approaching vehicle has close follower; FOLL=1 if the vehicle has a follower at a short headway of approximately 2-4 seconds	Bin	✓
HEV	Approaching vehicle is a heavy vehicle; HEV=1 if the vehicle is anything larger than the equivalent of a 15-passenger van (dump truck, TTST, bus)	Bin	
MUP	There are multiple pedestrians present in the CIA; MUP=1 if the number of pedestrians waiting at the curb is greater than 1	Bin	
NEAR	Pedestrian is waiting on the near-side of the approaching vehicle; NEAR=1 if the pedestrian waits on the same side of the road that the vehicle is traveling on	Bin	✓
PLT	Approaching vehicle is part of a platoon of vehicles; PLT=1 if the headway to the following OR the previous vehicle was short (approximately 2-4 seconds)	Bin	✓
PREV	The previous vehicle passed without yielding; PREV=1 if the previous vehicle failed to yield to the same pedestrian waiting at the crosswalk	Bin	✓
PXW	A pedestrian from a previous event is still present in the crosswalk; PXW=1 if the driver has to account for a pedestrian who is still in the roadway from a previous event	Bin	✓
QUE	Vehicle is part of a queue of vehicles; QUE=1 if the queue or platoon that the vehicle is a part of is moving slowly due to some downstream congestion or incident	Bin	
TRIG	The pedestrian triggered the yield by stepping into roadway; TRIG=1 if the pedestrian actively seized the roadway before the driver action indicated a yield	Bin	✓
TRTMT	Presence of the 'in-pavement flasher' crossing treatment; TRTMT=1 if the treatment was installed and so is equivalent to the 'after' case	Bin	✓
Continuous Explanatory Variables			
DECEL	Deceleration rate necessary to come to a full stop prior to crosswalk; DECEL is calculated from measured speed and distance; $DECEL = (SPEED_FT * SPEED_FT) / (2 * DIST1)$	feet/sec ²	✓
DIST1	Vehicle position at the time of pedestrian arrival in crosswalk influence area measured in feet using a LIDAR speed measurement device	Feet	✓
SPEED_FT	Vehicle speed at the time of pedestrian arrival in crosswalk influence area measured in ft/sec using a LIDAR speed measurement device	Feet/sec	✓
TTC	Time until vehicle would theoretically arrive at the crosswalk; TTC is calculated from the measured speed and distance at the time pedestrian arrives in the crosswalk influence area; $TTC = DIST1 / SPEED_FT$	Sec	✓

Table 2

Descriptive Statistics for Site A

Variable	ALL DATA		BEFORE*		AFTER*		YIELDS		NON-YIELDS	
	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
Sample Size	604		361		243		105		499	
Response Variables										
YIELD	0.174	0.379	0.150	0.357	0.209	0.407	1.000	0.000	0.000	0.000
Y_ORDERED	1.289	0.662	1.266	0.655	1.324	0.671	2.667	0.474	1.000	0.000
Y_TYPE**	0.667	0.474	0.778	0.420	0.549	0.503	0.667	0.474		
Binary Factors										
ADY	0.084	0.278	0.086	0.281	0.082	0.275	0.229	0.422	0.054	0.226
AST	0.142	0.349	0.141	0.349	0.143	0.351	0.371	0.486	0.094	0.292
FLASH	0.131	0.337	0.000	0.000	0.324	0.469	0.276	0.449	0.100	0.300
FOLL	0.489	0.500	0.476	0.500	0.508	0.501	0.429	0.497	0.502	0.500
HEV	0.015	0.121	0.011	0.105	0.021	0.142	0.019	0.137	0.014	0.118
MUP	0.321	0.467	0.302	0.460	0.348	0.477	0.410	0.494	0.302	0.460
NEAR	0.511	0.500	0.501	0.501	0.525	0.500	0.448	0.500	0.524	0.500
PLT	0.671	0.468	0.681	0.467	0.656	0.476	0.505	0.502	0.706	0.456
PREV	0.678	0.468	0.701	0.459	0.643	0.480	0.648	0.480	0.684	0.465
PXW	0.132	0.339	0.150	0.357	0.107	0.309	0.390	0.490	0.078	0.268
QUE	0.007	0.081	0.011	0.105	0.000	0.000	0.019	0.137	0.004	0.063
TRIG	0.079	0.270	0.044	0.206	0.131	0.338	0.305	0.463	0.032	0.176
TRTMT	0.403	0.491	0.000	0.000	1.000	0.000	0.486	0.502	0.386	0.487
Continuous Factors										
DECEL	4.787	5.928	3.959	4.540	6.013	7.368	2.266	1.630	5.316	6.353
DIST1	335.433	256.526	381.892	273.254	266.697	212.314	406.698	254.921	320.467	254.588
SPEED_FT	40.341	9.419	3.959	4.540	6.013	9.274	36.754	10.738	41.094	8.949
TTC	8.376	6.002	9.468	6.369	6.753	4.999	11.027	5.934	7.818	5.871

* Refers to Presence or Absence of Crossing Treatment

** The Sample Size for Y_TYPE is only 105 observations, because it only considers HY and SY events

Table 3

Descriptive Statistics for Site B

Variable	ALL DATA		BEFORE*		AFTER*		YIELDS		NON-YIELDS	
	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
Sample Size	470		265		205		158			
Response Variables										
YIELD	0.336	0.473	0.268	.0444	0.424	0.495	1.000	0.000	0.000	0.000
Y_ORDERED	1.513	0.777	1.411	0.729	1.644	0.820	2.525	0.501	1.000	0.000
Y_TYPE**	0.525	0.501	0.535	0.502	0.517	0.503	0.525	0.501		
Binary Factors										
ADY	0.066	0.248	0.060	0.239	0.073	0.261	0.120	0.326	0.038	0.193
AST	0.091	0.289	0.072	0.258	0.117	0.322	0.209	0.408	0.032	0.176
FOLL	0.245	0.430	0.223	0.417	0.273	0.447	0.253	0.436	0.240	0.428
HEV	0.021	0.144	0.026	0.161	0.015	0.120	0.019	0.137	0.022	0.148
MUP	0.351	0.487	0.347	0.577	0.356	0.480	0.443	0.498	0.304	0.461
NEAR	0.626	0.485	0.619	0.487	0.634	0.483	0.665	0.474	0.606	0.489
PLT	0.383	0.487	0.366	0.483	0.405	0.492	0.342	0.476	0.404	0.491
PREV	0.294	0.487	0.343	0.476	0.229	0.421	0.246	0.433	0.317	0.466
PXW	0.213	0.410	0.189	0.392	0.244	0.430	0.348	0.478	0.144	0.352
QUE	0.051	0.220	0.042	0.200	0.063	0.244	0.139	0.347	0.038	0.193
TRIG	0.051	0.220	0.042	0.200	0.063	0.244	0.139	0.347	0.006	0.080
TRTMT	0.436	0.496	0.000	0.000	1.000	0.000	0.551	0.499	0.378	0.486
Continuous Factors										
DECEL	4.601	4.145	4.890	4.554	4.227	3.523	3.138	1.929	5.341	4.732
DIST1	202.624	124.729	216.215	148.238	185.055	112.888	224.699	126.678	191.445	137.476
LDECEL	1.267	0.699	1.311	0.731	1.210	0.654	0.987	0.577	1.407	0.714
LDIST1	5.085	0.723	5.135	0.752	5.019	0.681	5.231	0.679	5.011	0.735
SPEED_FT	35.147	8.998	36.584	8.586	33.289	9.197	33.418	10.011	36.023	8.320

Variable	ALL DATA		BEFORE*		AFTER*		YIELDS		NON-YIELDS	
	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
TTC	5.783	3.577	5.963	3.951	5.553	3.020	6.727	3.412	5.308	3.569

* Refers to Presence or Absence of Crossing Treatment

** The Sample Size for Y_TYPE is only 105 observations, because it only considers HY and SY events

Table 4

Results of Logistic Regression for Site A

	P(Yield) - Unrestricted		P(Yield) - Restricted		P(HY Yield) - Restricted	
	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio
Intercept	-0.12	.	-0.38	.	1.11 [*]	.
ADY	0.97 [*]	2.63	.	.	.	.
AST	.	.	1.72 ^{***}	5.59	-1.22 ^{**}	0.72
FLASH	1.03 ^{**}	2.79	1.19 ^{***}	3.28	-1.12 [*]	0.85
NEAR	0.59 [*]	1.80	.	.	.	.
PLT	-0.75 ^{**}	0.47	-0.96 ^{***}	0.39	.	.
PREV	-1.02 ^{**}	0.36	.	.	.	.
PXW	1.15 ^{**}	3.16	.	.	.	.
TRIG	1.80 ^{***}	6.04	.	.	.	.
TRTMT	.	.	.	.	.	.
DECEL	-0.43 ^{***}	0.65	-0.38 ^{***}	0.68	0.20	1.69
AIC ⁺	430.4		443.4		128.3	
-2 Log L ⁺	391.9		433.4		120.3	
R-Square	0.2286		0.1673		0.1193	
Max-rescaled R-2	0.3648		0.2671		0.1656	

*** p<0.001,

** p<0.01,

* p<0.05

⁺ Model Fit Statistics Akaike Information Criterion (AIC), and Log Likelihood Criterion (-2 Log L) are shown for Intercept and Covariates

Table 5

Results of Logistic Regression for Site B

	P(Yield) - Unrestricted		P(Yield) - Restricted		P(HY Yield) - Restricted	
	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio
Intercept	-0.32	.	-0.12	.	0.10	.
ADY	1.10 [*]	2.99	.	.	.	.
AST	1.81 ^{***}	6.10	2.49 ^{***}	12.02	-1.82 ^{***}	0.16
FLASH	.	.	.	.	.	.
NEAR	0.70 ^{**}	2.02	0.62 [*]	1.85	.	.
PLT	.	.	-0.49 [*]	0.61	.	.
PREV	-0.58 [*]	0.56	.	.	.	.
PXW	0.93 ^{**}	2.52	.	.	.	.
TRIG	1.66	5.28	.	.	.	.
TRTMT	0.54 [*]	1.72	0.65 ^{**}	1.91	.	.
DECEL	-0.38 ^{***}	0.69	-0.34 ^{***}	0.71	0.11	1.12
AIC ⁺	465.0		491.6		204.3	
-2 Log L ⁺	447.0		479.6		198.3	
R-Square	0.2338		0.1769		0.1073	
Max-rescaled R-2	0.3204		0.2442		0.1431	

p<0.001,**
p<0.01,*
p<0.05⁺ Model Fit Statistics Akaike Information Criterion (AIC) and Log Likelihood Criterion (-2 Log L) are shown for Intercept and Covariates