Analyzing the Effect of All-Red Intervals in Crash Reduction: A Case Study of Heckman Correction at Urban Signalized Intersection Crashes

Sabyasachee Mishra, Ph.D., P.E. (corresponding author)
Research Assistant Professor
National Center for Smart Growth Research and Education
1112J Preinkert Field House (Building 054)
University of Maryland
College Park, MD 20742
Phone: 301-405-9424
Fax: 301-314-5639
e-mail: mishra@umd.edu

Xiaoyu Zhu, PhD.
Faculty Research Associate
National Center for Smart Growth Research and Education
1226C Preinkert Field House (Building 054)
University of Maryland
College Park, MD 20742
Phone: 301-405-9386
Fax: 301-314-5639
e-mail: xyzhu@umd.edu

Total Word Count: Words (5438) + Number of Tables and Figures (6x250) =6,938

Date Submitted: August 1, 2012

Submitted for Peer Review and for Compendium of Papers CD-ROM at the 92nd Annual Meeting of the Transportation Research Board (TRB) in January 2013.

Abstract

All-Red (AR) interval is designed as a method of clearance interval to safely clear vehicles that enter the intersection dilemma zone. The provision of AR is generally expected to reduce the occurrence of crashes, though there are situations that AR is not proved to be effective because it is used at intersections with a higher potential for crashes. This controversial result however, does not indicate that the AR interval is a contributing cause of crashes. Therefore, the self-selection bias of signal designs needs to be corrected when estimating their effect in improving safety.

To address the selection-bias problem at signalized intersections, a Heckman two-stage approach is adapted. First, a probit model is developed to explain the interrelationship between the AR interval and highway geometry, traffic volume, and environmental variables. Second, the selection bias term (or Heckman correction) is included in the second stage to build two negative binomial models for locations with and without an AR interval. Further, average treatment effects (ATE) and effect of treatment on the treated (TT) are estimated to examine the effect of AR intervals on the whole sample and treated sample, respectively. Three-year crash data on urban signalized intersections in the Detroit metro area is used to validate the proposed models. The results show that a random intersection with an AR interval will reduce crashes by 36 percent when compared to a non-AR interval intersection. For treated intersections (with AR interval) there is a 51 percent reduction of total crashes compared to intersections without treatment (if not designed with AR interval). The AR interval is a meaningful advance in reducing crashes by 15 percent.

Keywords: Self-selection bias, Heckman Two-step correction model, All-Red Interval, Probit Model, Negative Binomial Model

1. Introduction

Intersections are a critical subject in traffic control and operation. According to the National Highway Traffic Safety Administration's (NHTSA) Traffic Safety Facts 2009 Report, there were more than 2.2 million reported intersection-related crashes, resulting in 7,043 fatalities (20.8% of total fatalities) in 2009 *(1)*. Among the intersection crashes, 1.06 million (around 50%) crashes took place at traffic-signalized intersections. To improve intersection safety, crash causation study is a popular approach to investigate the impact of traffic, geometry, and environmental factors. Traffic signal timing design should further be concerned with signalized intersections, because the majority of crashes are caused by red light running and signal timing issues at these locations.

All-Red (AR) clearance intervals, in which all movements receive a red indication, were implemented to reduce crashes by providing additional time for vehicles to clear the intersection before conflicting traffic movements are released. AR may also be useful in mitigating amber dilemma zone problems, particularly at high speed intersections. Without an AR clearance interval, the yellow interval is followed immediately by a green interval for the opposing movements. This allows conflicting movements to start directly after the yellow interval. Kim and Washington (2006) proved that crashes, particularly those related to signal violations at signalized intersections could be reduced by an AR clearance interval. While, inadequate AR interval may cause intersection crashes if there is a deviation from the recommended signal timing practices *(2)*. A study by Zador et al. (1985) found that AR intervals at signalized intersections are commonly ignored and there is a statistically significant relationship between AR intervals that are too short and an increasing number of crashes *(3)*. Moreover, some locations with AR intervals are not effective in reducing crashes *(4)*. A study by the Minnesota Department of Transportation shows that short-term improvement (up to one year before-and-after implementation of an AR clearance interval) is beneficial in reducing intersection crashes related to signal violations. On the other hand, long-term (more than two years before-and-after implementation of the AR clearance interval) research findings indicate that the short-term benefits are not sustained.

To investigate the effect of an AR interval in reducing the occurrence of intersection crashes, the issue of a self-selection bias[1] needs to be addressed first. For example, the crash occurrence rate at an intersection is expected to be less with AR than without AR intervals. On the other hand, the AR intervals are also applied at intersections with more complex traffic and geometry design or higher potential for conflicts. Then, crashes at some intersections with AR may be greater than similar intersections without AR. The situation does not mean that AR interval is one of the causes of crashes, due to the nature of the self-selection bias. The intersection in question may have already had a higher crash rate which necessitated the use of an of AR design.

In this paper, we seek a method to capture the effect of signal timing on crash exposure, using AR intervals at urban intersections as a case study. In the next section, the literature on crash occurrence and their associated variables is discussed. Then, the

[1] Self-selection bias appears where observations in the data select themselves into a group, causing a biased sample. It arises where the characteristics of the observations which cause them to select themselves in the group create abnormal or undesirable conditions in the group (Heckman 1979).

methodology is presented for applying a two-step procedure (introduced by Heckman (2001)) to negative binomial models, to correct self-selection bias *(5)*. The data for this case study is summarized in the following section and empirical analysis is conducted to examine the signal timing and roadway design effect in reducing the number of crashes. Finally, we conclude with a discussion on the importance of this approach and its useful implications for intersection safety.

2. Literature review

To examine the effect of an AR interval and other factors (such as traffic and geometry design) on crash occurrence, causal factors are widely studied by crash prediction models. In this section, various studies to mitigate crash occurrence are discussed based on objectives, methods, and explanatory variables.

2.1 Objectives for Crash Occurrence Study

Crash occurrence has long been studied for mid blocks and intersections. Rural intersections, such as Abdel-Aty and Nawathe (2006), and Kim et al. (2007) *(6-7)*, and urban intersections, such as Lord and Persuad (2003) and Greibe (2003) *(8-9)*, are both commonly studied. But, there are different focuses between rural and urban areas because of the nature of density of development in the vicinity. Driver behavior changes by the access points in terms of cluster of developments around the intersections. Other locations, such as rural two-lane highways (see Ma et al., 2007) are also studied *(10)*. In addition, there are various targets in past studies. For example, Songchitruksa and Tarko (2006) studied the right-angle crashes at signalized intersections *(11)*. Wang and Abdel-Aty (2008) modeled the detailed left-turn crash occurrence according to conflict patterns *(12)*. Lee and Abdel-Aty (2005) focused on vehicle-pedestrian crashes only in their study *(13)*.

2.2. Methods for Crash Occurrence Study

The most frequent and widely accepted approach in past studies is using Poisson and Binomial regression models *(10, 14)*. Recently, there have been some improvements based on this approach such as random effect negative binomial applied by Chin and Quddus (2003) *(15)*. Lord et al. (2005) compared the Poisson, Poisson-gamma and zero-inflated regression models in crash occurrence *(16)*. Other than this basic approach, researchers have applied several statistical models as well; such as generalized estimating equations (GEE) *(8)*, Generalized Linear Modeling (GLIM) *(17)*, and Analysis of Variance (ANOVA) *(18)*. In addition, Abdel-Aty and Nawathe (2006) proposed a neural networks method as a novel approach in crash prediction model *(6)*. Other more complicated models have been developed to fit multiple tasks in the research. For example, Kweon and Kockelman (2003) studied crash exposure and severity simultaneously by ordered probit model, conditioned on Poisson *(19)*. Huang and Abdel-aty (2010) developed a Bayesian analysis to fit multilevel data from macro analysis (geographic region) to micro analysis (occupant injury) *(20)*.

2.3 Factors Affecting Crash Occurrence

Generally, traffic, geometry and environmental factors are taken into account while developing crash causation studies. For instance, Annual Average Daily Traffic (AADT) has proven to be highly correlated with crash occurrence *(7, 21, 22)*. Left turn lanes are addressed to reduce number of crashes in several studies by Kim et al. (2006), Oh et al. (2003), and Harwood et al. (2000) *(21-23)*.

Guo et al., (2010) found that the size of the intersection, the traffic conditions by turning movement, and the coordination of signal phase have significant impacts on intersection safety (25). Wang and Abdel Aty (2006) found that having a large number of phases per cycle, i.e. indicated by the left-turn protection on the minor roadway, with high speed limits on the major roadway, and in high population areas are correlated with high rear-end crash frequencies (12). But the effect of signal timing on intersection safety is limited in the literature. An investigation of effects of signal timing information in the crash occurrence is discussed in this paper.

In conclusion, Poisson and negative binomial regression are still preferred methods to model crashes occurrence and prediction. The causal relationship between signal timing of urban intersections and crashes has received limited attention in the literature. In this research, we will use the two-step Heckman correction approach adjusted for a negative binomial model to study the crash occurrence. In addition to controlling for roadway design, the research will focus on signal timing (AR interval).

3. Methodology

The approach used in this study is adapted from a standard Heckman two-step procedure, introduced by Heckman (1979 and 2001) *(5, 26)* and applied by Zhou and Kockelman (2007) *(27)*. Instead of the Ordinary Least Square (OLS) model coupled with the deduction effect found in previous research, the negative binomial model is used in the second procedure and the ratio treatment effect is calculated in this study. This method provides consistent estimates of explanatory variables by correcting self-selection bias. Simple closed-form expressions for treatment parameters: the average treatment effect (ATE) and the effect of treatment on the treated (TT) are derived to quantify the effects of individuals' self-selection.

3.1 Model Specification

Negative Binomial (NB) Regression is a common approach in crash occurrence/exposure studies. The negative binomial model is employed to relax the restriction on the variance of y in the Poisson model. The formulation is

$$y_i \sim NB(\mu_i, \kappa) \tag{1}$$

$$log\mu_i = X_i\beta \tag{2}$$

The treatment effects of AR design need to be estimated by comparing the number of crashes at intersections with and without AR intervals.

An AR interval is assigned for a particular intersection (receives treatment) if the utility D_i^* of doing so is positive and not assigned if the utility is negative.

$$D_i = \begin{cases} 1 & \text{if } D_i^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

$$D_i^* = Z_i\theta + \varepsilon_i^D \qquad (4)$$

where D_i denotes the observed treatment decision of the intersection i ($D_i = 1$ for intersections with AR and $D_i = 0$ for intersections without AR); Z_i is a row vector of explanatory variables for intersection i; ε_i^D is the error term for unobserved variables. Crash occurrence at AR and non- non-AR intersections are assigned to follow two NB distributions. The model is formulated as follows.

$$y_i^1 \sim NB(\mu_i^1, \kappa^1) \qquad log\mu_i^1 = X_i\beta^1 + \varepsilon_i^1 \qquad (5)$$

$$y_i^0 \sim NB(\mu_i^1, \kappa^0) \qquad log\mu_i^0 = X_i\beta^0 + \varepsilon_i^0 \qquad (6)$$

y_i^1 and y_i^0 are the number of crashes at AR and non-AR intersections following the negative binomial distributions with a mean of μ_i^1 and μ_i^0, respectively; X_i is a row vector of the explanatory variables for intersection i; ε_i^1 and ε_i^0 are error terms for unobserved variables.

Under the assumption of joint normal distribution across the three error terms, the two step estimation procedure applied to a negative binomial model is summarized as follows based on the application of an Ordinary Least Square (OLS) model by Heckman(2001) and Zhou and Kockelman (2007) *(5, 27)*:

(i) Obtain $\hat{\theta}$ from a probit model on the decision to take the treatment from equation (4)

(ii) Use $\hat{\theta}$ to compute the selection-correction terms

$$E(\varepsilon^D | z_i\hat{\theta}, D_i = 1) = \frac{\phi(z_i\hat{\theta})}{\Phi(z_i\hat{\theta})} \qquad (7)$$

$$E(\varepsilon^D | z_i\hat{\theta}, D_i = 0) = -\frac{\phi(z_i\hat{\theta})}{1 - \Phi(z_i\hat{\theta})} \qquad (8)$$

(iii) Instead of OLS, estimate two negative binomial models for groups with treatment and without treatment, respectively, with the inclusion of the appropriate selection-correction terms obtained from (ii).

$$log\mu_i^1 = X_i\beta^1 + \rho_1\sigma_1\frac{\phi(z_i\hat{\theta})}{\Phi(z_i\hat{\theta})} + \varepsilon_i^1 \qquad (9)$$

$$log\mu_i^0 = X_i\beta^0 + \rho_0\sigma_0(-\frac{\phi(z_i\hat{\theta})}{1-\Phi(z_i\hat{\theta})}) + \varepsilon_i^0 \qquad (10)$$

(iv) Use the estimation result of $\hat{\beta}^1, \hat{\beta}^0, \widehat{\rho_1\sigma_1}$, and $\widehat{\rho_0\sigma_0}$ obtained from step (iii) and $\hat{\theta}$ from step (i) to obtain point estimates of ratio treatment parameters ATE and TT.

$$ATE(x) = \frac{\hat{\mu}_i^1}{\hat{\mu}_i^0} = \frac{\exp(x\hat{\beta}^1)}{\exp(x\hat{\beta}^0)} = \exp[x(\hat{\beta}^1 - \hat{\beta}^0)] \qquad (11)$$

$$TT(x, z, D[z] = 1) = \exp[x(\hat{\beta}^1 - \hat{\beta}^0) + (\widehat{\rho_1\sigma_1} - \widehat{\rho_0\sigma_0})\frac{\phi(z_i\hat{\theta})}{\Phi(z_i\hat{\theta})}] \qquad (12)$$

ATE is the average treatment effect evaluated for a random sample. It represents the ratio change in the number of crashes when turning a random intersection from non-AR design to AR design. ATE >1 means intersections with AR have more crashes than without AR, and vice versa. TT is the expected ratio change from the treatment for those selected to be treated. Similarly to ATE, if TT>1, a treated intersection (with AR) has more crashes than if this intersection is without AR, and vice versa. A comparison of ATE and TT values can estimate the treatment effect. Implications of ATE and TT are discussed in the result section.

4. Data

Data is collected on number of signalized intersection variables for the research including, (a) crash data, (b) traffic data, (c) highway geometry and (d) traffic signal data.

Crash data for this study is collected from the South Eastern Michigan Council of Governments (SEMCOG) for the years 2002 to 2004. The crash data collected contains information on crashes involving fatalities, injuries, and property damage for the Wayne county area, in Detroit, Michigan. The crash database consists of information such as the location, number and types of vehicles involved in the crash, speed limit, weather conditions under which the crash occurred, time of day of the crash, type of crash (angle, rear-end, side-swipe), and severity of crash. Exposure data is collected from the Wayne County 2008 traffic volume database and state of Michigan traffic volume database[2]. The county database consists of traffic volume for major and minor arterials, and the state database consists of traffic volume for major arterials, interstates, and expressways. The traffic volume data is prepared for study all locations in conjunction with the county and the state databases. Both the county and state prepare and maintain annual traffic volume data. Figure 1 shows the locations of all intersections used in the analysis.

[2] The county database did not consist of traffic volume on all major arterials, so the state traffic volume database is used in such cases.

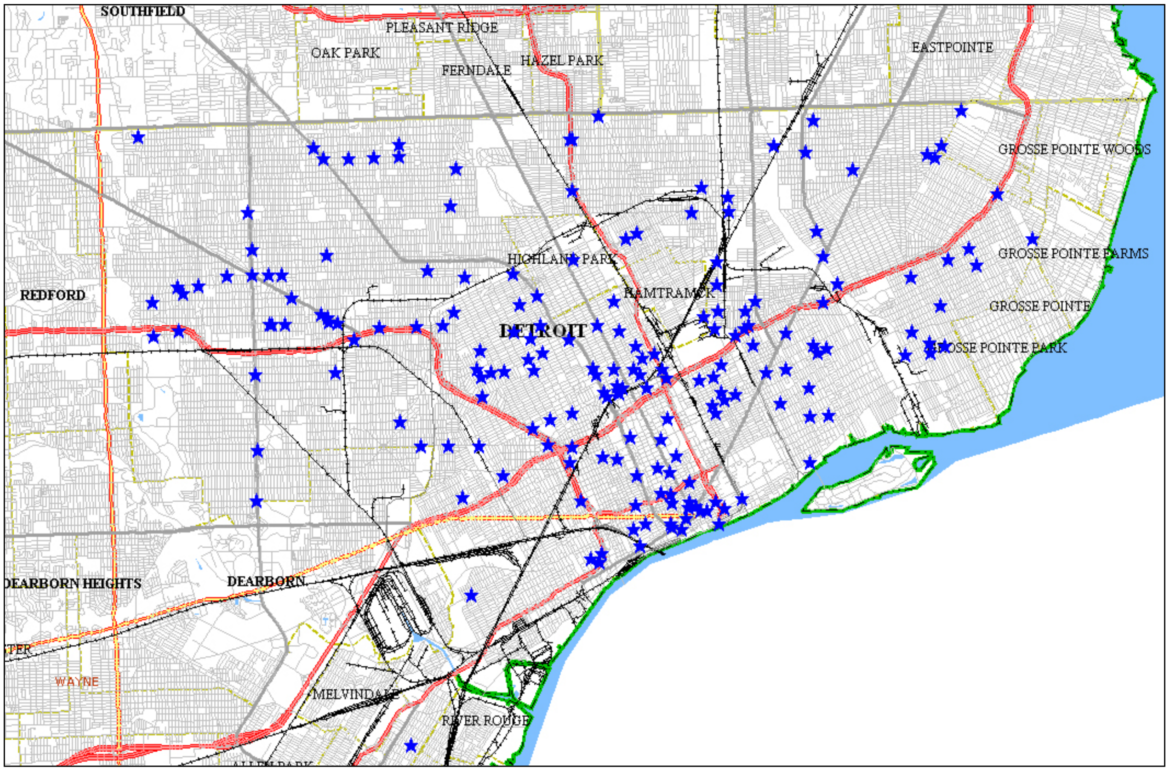

Figure 1 Signalized Intersections with Crash data in Wayne County, Michigan (2002-2004)

Highway geometry data is collected from the county asset management database and from satellite visuals in the absence of available data. The asset management database captures all improvements performed in the countywide highway network so the highway geometry data corresponds with crash data from a specific year. Highway geometry data consists of characteristics for major and minor streets including the total number of lanes; number of through, right, and left turn lanes, exclusive left and right turn lanes; presence of a median, including median type and width; parking availability; shoulder width; speed limit; and sight distance. Hazard ratings have been added in the database, information that was not originally found in the collected data. The methodology for hazard rating developed by Zegeer et al. (1994) is used in this research *(28)*. The hazard rating ranges from one (best) to seven (worst). The highway geometry data is collected within 250 feet of the intersection.

The Wayne county signal database is used to extract signal timings of crash locations. The database consists of cycle length, green interval, yellow interval (amber), red interval, all red interval, phase diagrams, actuated or pre-timed signal control, size of the signal head, pedestrian walk time, and signal phases including exclusive turn (right or left) phase timing. The signal timing data is collected for one location at a time. The database is restricted to signalized intersections and only to the Wayne county area. To eliminate bias in the database, locations with no crashes for the three-year study period are also added. Definitions and descriptive statistics for the variables used in the analysis are presented in Table 1(a) and Table 1(b). In Table 1(a) descriptive statistics of continuous variables are shown and in Table 1(b) descriptive statistics of categorical variables is presented. Further, the descriptive statistics of AR interval and non-AR interval are shown separately in each Table.

Table 1(a) Descriptive Statistics for Continuous Variables

Continuous Variable	**Street**	**AR Intersections (137)**			**Non-AR Intersections (48)**		
		Min.	**Mean**	**Max.**	**Min.**	**Mean**	**Max.**
Total Crashes	Total	0.00	18.79	54.00	3.00	20.56	51.00
	Fatality	0.00	0.04	3.00	0.00	0.03	4.00
	Injury	0.00	5.15	21.00	0.00	4.50	21.00
	PDO	0.00	13.60	45.00	0.00	16.06	45.00
log_e (AADT)	Major	5.42	9.32	10.71	7.68	9.53	10.90
	Minor	5.34	8.33	10.11	6.47	8.44	9.71
# of Driveways	Major	0.00	2.13	10.00	0.00	2.00	9.00
	Minor	0.00	1.97	12.00	0.00	2.04	8.00
# of Lanes	Major	2.00	3.66	9.00	0.00	3.81	7.00
	Minor	2.00	2.58	8.00	2.00	2.83	8.00
# of Left Turn Lanes	Major	0.00	0.27	1.00	0.00	0.21	1.00
	Minor	0.00	0.32	1.00	0.00	0.40	1.00
# of Right Turn Lanes	Major	0.00	0.32	1.00	0.00	0.40	1.00
	Minor	0.00	0.11	1.00	0.00	0.10	1.00
Median Width (feet)	Major	0.00	4.74	37.00	0.00	6.48	46.00
	Minor	0.00	2.93	46.00	0.00	5.81	50.00
Shoulder Width (ft)	Major	0.00	3.95	11.00	0.00	4.42	11.00
	Minor	0.00	3.69	12.00	0.00	3.60	11.00
Posted Speed Limit (mph)	Major	15.00	31.02	55.00	25.00	31.98	35.00
Cycle Length (sec)	Intersection	50.00	59.34	90.00	50.00	59.58	80.00
All Red Phase (sec)	Major	1.00	1.17	3.00	0.00	0.04	1.00
	Minor	1.00	1.18	2.00	0.00	0.13	1.00
Amber Phase (sec)	Major	4.00	4.18	6.00	4.00	4.25	5.00
	Minor	4.00	4.23	5.00	4.00	4.38	5.00

A total of 185 observations are used to generate descriptive statistics, and further used in the analysis procedure. Total crashes for a three-year period are examined, excluding pedestrian crashes. Table 1(a) shows the crash data for AR and non-AR signalized intersections. AR intersections consist with AR intervals on both approaches. The remaining intersections are categorized as “Non-AR Intersections”. Total crashes (the dependent variable) is a sum of the fatality, injury (three types), and property damage only (PDO). It is evident that the average number of PDO crashes at locations with AR is less than non-AR locations. There are very minor differences in the number of Injury and Fatal crashes. When compared to the total number of crashes, AR locations have fewer crashes than non-AR locations. Averages for exposure variables such as Annual Average Daily Traffic (AADT), highway geometry such as number of right and left turning lanes, protected left and right turns, median width, median type, number of driveways, hazard rating, acceleration leg, parking lanes, shoulder width are considered.

Table 1(b) Descriptive Statistics for Categorical Variables

Categorical Variable	**Street**	**AR Intersections (%)**		**Non-AR Intersections (%)**	
		No	**Yes**	**No**	**Yes**
Painted Median	Major	26.27	73.73	97.91	2.09
	Minor	0.03	97.97	97.91	2.09
Curbed Median	Major	83.21	16.79	91.19	8.81
	Minor	92.71	7.29	89.58	10.42
No Median	Major	90.51	9.49	25.00	75.00
	Minor	9.00	91.00	16.67	83.33
Speed Limit Less than 30mph	Intersection	35.76	64.24	52.08	47.92
Non Freeway Section	Intersection	0.06	0.94	10.41	89.59
Left Turn Protection on Major Street	Intersection	82.48	17.52	62.50	37.50
Number of Lanes 3 or Less on Major Street	Intersection	70.80	29.20	77.08	21.92
Physical Divider on Major Street	Intersection	59.12	39.88	56.25	43.75
Physical Divider Without Barrier	Intersection	74.45	25.55	77.08	22.92
Cycle Length Over 60 sec	Intersection	76.64	23.36	79.16	20.84
Acceleration Leg	Major	95.62	4.38	0.00	0.00
	Minor	91.97	8.03	87.50	12.50
Parking Lane	Major	19.70	80.30	27.08	72.92
	Minor	80.29	19.71	72.91	27.09
Sight Distance	Major	63.50	36.50	70.83	29.17
	Minor	68.61	31.39	68.75	31.25
Hazard Rating >3	Major	12.40	87.60	12.50	87.50
	Minor	36.49	63.51	37.50	62.50
Exclusive Left Turn Phase	Major	92.70	7.30	95.83	4.17
	Minor	91.97	8.03	97.91	2.09

Among the environmental variables available, sight distance and roadway type are considered. A number of signal timing variables such as cycle length, amber time, pedestrian crossing time, presence and absence of protected and permitted left turns are considered in the analysis.

5. Results

In this section, we first present the results of the probit model along with a description of its corresponding significant variables. Second, the negative binomial results are discussed and explained. The treatment effect of an AR interval is then compared and explained.

The first step analyzes the choice of AR design on both major and minor approaches using a binary probit model. The sign of the coefficients in Table 2 are as expected and reasonable. The results show that the presence of protected left turn signal on the major street of a signalized intersection reduces the probability of having an AR interval. This refers to the case of a protected left turn where the vehicles are only permitted to make a left then within a designated time, signaled by a green arrow. In contrast, for permissive left turn locations, vehicles can make a left even when there is no green arrow for the left turn, but the signal is green for the opposite through movement. A painted median warrants an AR interval because with no physical barrier separating two opposite movements there might be a tendency on the part of the driver to cross the intersection dilemma zone at a yellow interval. The tendency to cross the intersection is the result of a reduced intersection island distance, in the case of a painted median on the major street. A major street speed limit, for instance less than 30 miles per hour, increases the probability of a required AR interval. Note that the intersections analyzed in this study are located in an urban Central Business District area where traffic volume is very high and intersections are separated by smaller distances. A number of other variables were tested but were found to be not significant.

Table 2 Result for Binary Probit Model on All-Red Design Location Choice

Variable	**Coefficient**	**t-stat**
Constant	0.485 **	2.679
Left Turn Protection on Major Street	-0.538*	-2.300
Painted Median on Major Street	1.074*	2.033
Major Street Speed Limit Less than 31mph	0.420*	1.972
R^2	0.07	
Sample Size	185	

Note: ***=significant at 99%; **=significant at 95%; *=significant at 90%

Table 3 shows negative binomial model results for (1) AR Model, and (2) Non-AR Model. A set of common variables are used for both models. For the AR intersections the result is shown in column 2 and 3. Ln(AADT) for major and minor streets is positively correlated with the total number of crashes. It is observed that if AADT on the major or minor street increases, then the total number of crashes in an urban signalized intersection increases. Non-freeway sections are the intersections not close to a freeway ramp, transition area, weigh station or rest area. These locations are likely to have more crashes at AR intersections. The probability of crash occurrence decreases if the intersection approaches have fewer lanes. Intuitively, lower volume intersections have fewer crashes than higher volume intersections. The selection-correction term is not significant for the AR Model.

Table 3 Results of Negative Binomial Models for Number of Crashes at All-Red and Non-All-Red Intersections

	All Red Model		**Non-All Red Model**	
Variable	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Constant	-0.1532	-0.207	2.5570	1.887
ln (AADT Major)	0.1428 **	1.987	-0.0327	-0.264
ln (AADT Minor)	0.1930***	3.207	0.1335	1.534
Non Freeway Section	0.4580*	1.709	0.3108	1.176
Number of Lanes 3 or less	-0.4052 **	-3.163	-0.6610***	-3.402
Selection-correction term	-0.1959	-0.568	0.4214	1.445
R^2	0.19			0.30
Sample Size	137			48

Note: ***=significant at 99%; **=significant at 95%; *=significant at 90%; others not significant

All of the variables used in the AR model are also used for the Non-All Red Model. But only the indicator of three or less lanes is significant. This suggests that when the number of lanes is three or less on the Non-All Red intersections, the probability of occurrence of crashes decreases. The selection-correction term is also not significant for the Non-All Red Model. Other variables tested but found to be not significant in the NB model include sight distance on the major street (both left and right turn), divider without barrier, hazard rating more than three on the major street, and parking on the major street.

The magnitude of treatment effects is shown in Table 4. ATE and TT parameters are estimated using equation (9) and (10). The ATE value of signalized intersections with an AR interval will have an average crash rate of 64% compared to intersections with no AR. Alternatively, AR intersections will, on average reduce the crash rate by 35.65 percent (i.e. 1-0.6435) compared to non-AR intersections. The TT value shows that for intersections already designed with AR the number of crashes experienced is 49.58 percent compared to non-AR intersections. The effect of AR on the treated intersections shows savings of 50.42 (i.e. 1-0.4958) percent of crashes. AR effect accounts for a 14.77 percent reduction in crashes (0.5042-0.3565).

Table 4 Treatment Effect

Parameter	**Value**	**Effect =(1-Value)**
ATE	0.6435	0.3565
TT	0.4958	0.5042
Difference (=ATE-TT)		0.1477

6. Conclusions

In this paper, we propose a two-stage approach to examine the effect of AR signal design in reducing the number of crashes. In the first stage, a probit model demonstrates the interdependence between AR and other geometry variables. In the second stage, negative binomial models analyze the crash causal factors including a self-selection bias correction term. Additionally, two treatment parameters (ATE and TT) are estimated to examine the effect on AR for all intersections, and treatment intersections. Three years of crash data is used for analysis of selected signalized intersections in the Detroit Metro area consisting of 185 samples. Within this dataset there are 137 treated samples and 48 untreated samples.

The probit model results show AR interval design is negatively correlated with protected left turns on major streets, and positively correlated on painted medians, and a lower speed limit on major streets. In the negative binomial model for the treated sample, AADT (both on the major and minor street) and roadway locations away from freeways are positively correlated with total intersection crashes. In the same sample, major streets with three or fewer lanes are negatively correlated with total intersection crashes. For untreated variables only major streets with three or fewer lanes is significantly related to total crashes. The AR effect accounts for a 14.77 percent reduction in the total number of crashes while controlling for traffic and geometry variables.

In addition to correcting the self-selection bias of an AR interval, this approach provides methodological implications to analyze other variables with selection bias. For some of Before-After studies in traffic safety, researchers need to collect crash data for a long duration before and after an event to evaluate a treatment. It is not always appropriate to do the before-after experiment considering the risk of crash caused by the experiment. With method developed in this paper, researchers do not need to wait for the outcome with treatment and risk to apply the treatment. Based on the current sample, with and without treatment, the treatment effect could be calculated with correction of selection bias and control of other variables. Furthermore, it should be noted that selection bias estimation and correction rely on the assumptions of a joint normal distribution, and may not be robust for departures from this distributional assumption (Heckman et al. 2001). More robust approaches tend to be nonparametric in nature. Future research should incorporate nonparametric distribution in the estimation of selection bias terms.

References

1. National Highway Traffic Safety Administration's (NHTSA), 2009. Traffic Safety Facts 2009 Report. http://www-nrd.nhtsa.dot.gov/Pubs/811402.pdf.

2. Kim, D. J., and Washington, S. 2006. The significance of endogeneity problems in crash models: An examination of left-turn lanes in intersection crash models. *Accident Analysis & Prevention* 38(6), 1094-1100.
3. Zador, P., Stein, H., Tarnoff, P., 1985. Effect of clearance interval timing on traffic flow and crashes at signalized intersections. *Institute of Transportation Engineers (ITE) Journal* 55(11), = 36-39.
4. Souleyrette, R. R., O'Brien, M. M., McDonald, T., 2004. Effectiveness of All-Red clearance interval on intersection crashes. Minnesota Department of Transportation, Report # 2004-26. http://www.lrrb.org/pdf/200426.pdf, accessed May 27, 2012.
5. Heckman, J.J., Tobias, J.L., Vytlaci, E.J., 2001. Four parameters of interest in the evaluation of social programs. *Southern Economic Journal* 68(2), 210-223.
6. Abdel-Aty, M., Nawathe, P., 2006. A novel approach for signalized intersection crash Classification and prediction. *Journal of Advances in Transportation Studies information* 9, 67-80.
7. Kim, D.J., Washington, S., Choi, K., 2007. Modeling crash outcome probabilities at rural intersections: application of hierarchical binomial logistic models. *Accident Analysis and Prevention* 39(1), 125-234.
8. Lord, D., Persuad, B. N., 2003. Accident prediction models with and without trend. *Transportation Research Record: Journal of the Transportation Research Board, No. 1717*, TRB, National Research Council, Washington, D.C., 102-108.
9. Greibe, P., 2003. Accident prediction models for urban roads. *Accident Analysis and Prevention 35* (1), 273-285.
10. Ma, J., Kockelman, K.M., Damien, P., 2007. Bayesian multivariate poisson-lognormal regression for crash prediction on rural two-lane highways. CD-ROM. Transportation Research Board of the National Academics, Washington D.C.
11. Songchitruksa, P., Tarko, A.P., 2006. Practical method for estimating frequency of right-angle collisions at traffic signals. *Transportation Research Record: Journal of the Transportation Research Board, No. 1953*, TRB, National Research Council, Washington, D.C., 89-97.
12. Wang, X., Abdel-Aty, M., 2008. Modeling left-turn crash occurrence at signalized intersections by conflicting patterns. *Accident Analysis and Prevention* 40 (1), 76-88.
13. Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. *Accident Analysis and Prevention* 37 (4), 775–786.
14. Shankar, V., Milton, J., Mannering, F.L., 1997. Modeling accident frequencies as zero-altered probability processes: An empirical inquiry. *Accident Analysis and Prevention* 29 (6), 829-837.
15. Chin, H.C., Quddus, M.A., 2003. Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. *Accident Analysis & Prevention* 35, 253–259.
16. Lord, D., Washington, S. P., John, N.I., 2005. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accident Analysis and Prevention* 37(1), 35-46.

17. Hadayeghi, A., Shalaby, A.S., Persaud, B. N., 2004. Macrolevel Accident prediction models for evaluating safety of urban transportation systems. *Transportation Research Record: Journal of the Transportation Research Board, No. 1840*, TRB, National Research Council, Washington, D.C., 87-95.
18. Abbas, K.A., 2004. Traffic safety assessment and development of predictive models for accidents on rural roads in Egypt. *Accident Analysis and Prevention* 36 (1), 149-163.
19. Kweon, Y., Kockelman, K., 2003. Overall injury risk to different drivers: combining exposure, frequency, and severity models. *Accident Analysis and Prevention* 35(4), 441–450.
20. Huang, H., Abdel-Aty, M., 2010. Multilevel data and Bayesian analysis in traffic safety *Accident Analysis & Prevention* 42(6), 1556-1565.
21. Persaud, B, N., Retting, R. A., Lyon, C. A., 2004. Crash reduction following installation of centerline rumble strips on rural two-lane roads. *Accident Analysis and Prevention* 36(6), 1073-1079.
22. Kim, D.J., Washington, S., Oh, Jutaek., 2006 . Modeling crash types: new insights into the effects of covariates on crashes at rural intersections. *Journal of Transportation Engineering, American Society of Civil Engineers* 32 (4), 282-292.
23. Oh, J., Lyon, C., Washington, S., Persaud, B., Bared, J., 2003. Validation of FHWA crash models for rural intersections: Lessons learned. *Transportation Research Record: Journal of the Transportation Research Board, No. 1840*, TRB, National Research Council, 41-49.
24. Harwood, D. W., Council, F. M., Hauer, E., and Vogt, A., 2000. Prediction of the expected safety performance of rural two-lane highways, *Report prepared by Midwest Research Institute*, Kansas City, MO, USA.
25. Guo, F., Wang, X., Abdel-Aty, M.A., 2010. Modeling signalized intersection safety with corridor-level spatial correlations. *Accident Analysis & Prevention,* 42, 84–92.
26. Heckman, J.J., 1979. Sample selection bias as a specification error. *Econometrica* 47(1), 153-162.
27. Zhou, B., Kocckelman, K.M., 2007. Self-selection in home choice: Use of treatment effects in evaluating the relationship between the built environment and travel behavior. *Transportation Research Record*: *Journal of Transportation Research Board*, *No. 2077,* TRB, National Research Council, Washington, D.C., 54-61.
28. Zegeer, C.V., Stewart, R., Council, F., Neuman, T. R., 1994. Accident relationships of roadway width on low-volume road, *Transportation Research Record*: *Journal of Transportation Research Board*, *No. 1445*, TRB, National Research Council, Washington, D.C., 1-10.