Canadian Journal of Civil Engineering
Revue canadienne de génie civil

Novel Case-Control Observational Method for Assessing Effectiveness of Red-Light Cameras

Journal:	<i>Canadian Journal of Civil Engineering</i>
Manuscript ID	cjce-2016-0414.R1
Manuscript Type:	Article
Date Submitted by the Author:	10-Jan-2017
Complete List of Authors:	Abdulsalam, Amal; Carleton University, Rowlands, Dane; Carleton University Easa, Said; Ryerson University, Civil Engineering Abd El Halim, Abd El Halim; Carleton University
Keyword:	transp. & urban planning < MANUSCRIPT CLASSIFICATION, experimental research < type of paper to review, traffic control < Transportation, transportation < Computer Applications, Cross-sectional study, empirical Bayes, negative binomial, propensity score matching, red-light cameras, safety performance func

SCHOLARONE™
Manuscripts

Novel Case-Control Observational Method for Assessing Effectiveness of Red-Light Cameras

Amal Jasem Abdulsalam¹, Dane Rowlands², Said M. Easa³, and Abd El Halim O. Abd El Halim¹

¹ Department of Civil and Environmental Engineering, Carleton University, Ottawa, Ontario, Canada K1S 5B6

² The Norman Paterson School of International Affairs, Carleton University, Ottawa, Ontario, Canada K1S 5B6

³ Department of Civil Engineering, Ryerson University, Toronto, Ontario, Canada M5B 2K3

Corresponding author:

Amal Jasem Abdulsalam

1343 Meadowlands Dr. #1115

Ottawa, ON K2E7E8

Canada

amal.abdulsalam@carleton.ca

(613)-299-0706

ABSTRACT

Conventional methods for estimating the effect of an intervention suffer from selection bias, where the units that receive the treatment have different characteristics from those in the control group. This paper proposes a novel method that provides a specific criterion for selecting the control group. The method, called propensity score matching (PSM), was applied to the evaluation of red-light cameras (RLC) and its performance was compared with conventional cross-sectional and empirical Bayes methods. The application was performed using field data from the City of Ottawa involving 30 RLC intersections and 89 non-camera intersections observed for a period of 15 years. All three methods yielded fairly consistent results indicating an increase in property damage collisions and a decrease in injury and fatal collisions. Given the strong theoretical basis of the PSM method and its ability to produce a more stable and reliable estimator, the method is recommended as a viable alternative to the conventional methods.

Key Words: Cross-sectional study, empirical Bayes, negative binomial, observational study, propensity score matching, red-light cameras, regression, safety performance function.

1. Introduction

Traffic safety management is a process involving identification of sites with potential for safety improvement by using analytical methods and measures to protect road users from being killed or seriously injured (FHWA 2009). Completing scientifically rigorous traffic safety evaluations is a critical step in the process, because it not only provides a measure of effectiveness after implementing a treatment, but also enables agencies to select treatments that are expected to be effective after identifying sites with potential for safety improvement (FHWA 2011). Accurate estimation of treatment effectiveness is critical in transportation safety. Three commonly used observational methods to determine the effectiveness of traffic safety treatment presented in the literature are cross-sectional methods, empirical Bayes (EB) before-after methods, and case-control methods (Sasidharan and Donnell 2013). It is important to recognize the strengths and issues associated with the various methods to help identify a suitable method given the constraints of each specific evaluation. Table 1 summarizes the advantages and limitations of each of the methods used to analyze observational data in traffic safety.

The EB method is currently considered the most reliable in terms of arriving at estimates that minimize the likelihood of significant bias (Wood et al. 2015). There are, however, several practical limitations that may preclude the use of the cross sectional and EB methods, including (Gross 2013):

1. Confounding factors: several improvements may be implemented simultaneously, making it difficult to isolate the effect of a single treatment using a before-after study. Similarly, changes in traffic volume, driver population, vehicle mix, and other factors may occur over the analysis time period in a before-after study.

2. Sample size: it is sometimes difficult to find an adequate sample of sites where the treatment of interest has actually been implemented. Results from a limited sample will have a high level of statistical uncertainty. When there are few or no sites being treated with the treatment of interest, a before-after study is difficult to employ.
3. Study period: a EB study requires a time sequence, where it is necessary to implement a treatment and wait for sufficient data in the after period. While data collection can be time consuming for any safety evaluation, waiting several years after implementation is a practical concern in EB studies.
4. Selection bias: Traditional cross-sectional methods, such as the negative binomial (NB) method, aim to estimate the effect of an intervention by modeling the relationship between the covariates and the outcome. However, this method suffers from selection bias, where the units who receive the treatment may have different characteristics from those in the control group. These pre-existing differences between the groups must be controlled to obtain unbiased estimates of the treatment effects. It is crucial to try to separate out the causal effect of the treatment from the effect of these pre-existing differences between the treated and control groups.

Given the limitations associated with the EB and cross-sectional methods, alternative evaluation methods are needed to estimate the safety effectiveness of a treatment. The objective of this paper is to introduce and apply the propensity score matching (PSM) method in estimating the impact of red-light cameras (RLC) on traffic collisions. The PSM method has become a popular approach to estimate causal treatment effects by modeling the relationship between the covariates and the treatment assignment. The PSM method has been widely used in the multiple fields, such as medicine (Stone et al. 1995; Wang et al. 2005; Gum et al. 2001; Akematsu et al. 2012), communications (Kincaid 2009), economics (Lechner 2002), and policy evaluation (Harknett 2006).

PSM is a relatively new method that has had few applications in traffic safety. The first application of propensity scores in traffic safety research (Davis 2000) argued that the selection mechanism for the treatment should be included as a part of a traffic safety evaluation. It suggested that, when this does not occur, both EB and cross-sectional analysis methods become inconsistent with the true outcome. Other applications of propensity scores in traffic safety include estimating child safety restraint effectiveness (Elliott et al. 2009), estimation of design exception safety impacts (Wood and Porter 2013), safety effects of intersection lighting (Sasidharan and Donnell 2013), and assessing the impacts of speed cameras on safety (Li et al. 2013). Though widely implemented in diverse disciplines in many contexts, to the best of the authors' knowledge, it remains untried in red-light camera safety research by traffic management scholars. It should be thought of as an additional tool available to investigators when estimating treatment effects in studies where potential bias may exist.

The overall objective of the PSM method is to replicate a randomized experiment in two ways: first, by comparing treated and control units who look as if they could have been randomly assigned to treatment or control status; and second, by forming the comparison groups without the use of the outcome, thus preventing intentional or unintentional bias in selecting a particular sample to achieve a desired result.

Results of the safety effectiveness of RLC treatment derived from the PSM method will be compared to the results obtained from the EB method and the traditional cross-sectional method that employs an NB regression model. The data considered in this comparison are from the City of Ottawa, which included application of the RLC treatment on selected intersections, along with a reference group of untreated intersections in Ottawa.

2. Background of Traffic Safety Modeling Methodologies

Propensity Score Matching Method

One of the primary challenges of estimating the effect of a treatment using data from an observational study is the issue of confounding bias. A confounder is a variable that predicts both the treatment and the outcome, and therefore may impair the ability to make causal inferences about the effect (Lanza et al. 2013). A propensity score is a statistical tool that allows researchers to make more accurate causal inferences by balancing non-equivalent groups that may result from using a non-randomized design.

The basic idea of PSM method is to match an untreated group to the treated group such that the matched group is comparable to the treated group in all aspects of characteristics except the treatment. Before matching, the distribution of potential confounders is usually different between exposed and unexposed, while the matched groups will show only a slight residual imbalance. The PSM method enables the matching of covariates to be reduced to a single dimension, the propensity score. The propensity score is the probability that a unit is selected into the treatment group conditional on observed covariates. Similar groups can then be defined clearly as those with similar propensity scores, and by this avoid selection bias to ensure that the difference between the treatment and control groups can be attributed to the treatment (Austin 2011). Therefore, treated and untreated units with similar propensity scores can be compared to obtain the treatment effect.

The primary criteria for selecting untreated subjects whose propensity score is close to that of a treated subject is called nearest neighbor (NN) matching (Rosenbaum & Rubin 1985). Nearest neighbor matching selects for matching to a given treated subject that untreated subject whose propensity score is closest to that of the treated subject. The nearest neighbor matching technique generally selects k matched controls for each treated unit. The nearest neighbor matching method uses a algorithm, which cycles through each treated unit one at a time, selecting the available control unit with the smallest distance to the treated unit.

The observed outcome for the matched group can then be used as the counterfactual, and the treatment effect is estimated simply as the difference between the average outcomes of the two groups. The impact of a treatment for an individual i , noted τ_i , is defined as the difference between the potential outcome in case of treatment and the potential outcome in absence of treatment (Caliendo and Kopeinig 2005). That is,

$$[1] \qquad \tau_i = Y_{1i} - Y_{0i}$$

where Y_{1i} = the potential outcome where individual i receives treatment and Y_{0i} = the potential outcome where individual i does not receive treatment. The fundamental evaluation problem arises because only one of the potential outcomes is observed for each individual i . The unobserved outcome is called counter-factual outcome. Hence, estimating the individual treatment effects, τ_i , is not possible. Therefore, the first step is to estimate the propensity scores. This approach collapses multiple potentially confounding variables into a single score, the propensity score. In order to accomplish this, binary logit or probit regression can be used (Guo and Fraser 2010). The conditional probability of an entity receiving treatment is given by the covariates (X) and outcomes (Y). When the treatment assignment mechanism is unconfounded, the propensity score ($\hat{p}$) is represented as shown in the following equation (Sasidharan and Donnell 2013):

$$[2] \qquad \hat{p} = E(D = 1|X)$$

where D = treatment status ($D = 1$, treated and $D = 0$, untreated); X = covariates influencing treatment selection. The $\hat{p}$ is unknown in observational studies and must be estimated from the sample based on important covariates influencing the treatment selection. The $\hat{p}$ is commonly estimated using logistic regression. The propensity scores help to balance the covariates in the treated and untreated groups when the data are non-randomized. Given a population of units denoted by i , if the propensity score $\hat{p}_i$ is known, then the average treatment effect on the treated (ATET) group can be estimated as follows (Caliendo and Kopeinig 2005, Heinrich et al. 2010):

$$[3] \qquad \text{ATET} = E [E(Y_{1i}|D_i = 1, \hat{p}_i) - E(Y_{0i}|D_i = 1, \hat{p}_i)]$$

where the outer expectation is over the distribution of $(\hat{p}_i | D_i = 1)$ and Y_{1i} and Y_{0i} are the potential outcomes in the two counterfactual situations of (respectively) treatment and no treatment. The ATET is the effect of the treatment on the outcome among those who received the treatment. The syntax for estimating treatment effects in StataSE using NN matching, is as follows (SSCC 2015):

[4]   `teffects nnmatch (y x1 x2...xn) (t), atet`

where y = outcome, $x_1 x_2 \dots x_n$ = independent variables and t = treatment indicator. Variables need to be specified for the outcome model but not the treatment model. The ATET may be of great interest, as it increases understanding of how exposure to a certain risk ultimately affected those individuals who were exposed to the risk. An essential aim of the analysis phases of this method is to prevent, reduce, and assess bias, confounding, and chance, so that a causal unbiased association between exposure and outcome is estimated (Jespen et al. 2004). A summary of the PSM method is presented in Figure 1.

Cross-Sectional Method

A cross-sectional study is performed using generalized linear models (GLMs), such as Poisson and negative binomial (NB) regression models. The selection of the type of regression model is described in Fig. 2. In most count datasets, the conditional variance is greater than the conditional mean, a phenomenon known as over-dispersion (Lord and Mannering 2010). Therefore, the Poisson regression model often fits count data poorly, as indicated by a deviance or Pearson chi-square test. The NB regression model can be used to overcome problems associated with over-dispersed data. The negative binomial regression model takes the following general form (Donnell et al. 2009):

$$[5] \quad \lambda = \exp (\omega + X_i \beta_i + L_i \theta_i)$$

where λ = expected number of collisions on a roadway segment/intersection; ω = intercept; X_i = variables that influence λ (e.g., road geometry); β_i = coefficients that quantify the relationship between X and λ ; L_i = a binary (dummy) variable indicating the presence of the treatment; and θ_i =

coefficient that quantify the relationship between L and λ . The regression coefficients (β_i and θ_i) are estimated using the method of maximum likelihood.

Cross-sectional studies are also called prevalence studies, where exposure and outcome are measured simultaneously. Prevalence rates can be compared between groups, but the terms exposure and outcome are treacherous because the sequencing of the two is impossible to assess (Hulley et al. 2001). Therefore, a cross sectional study is generally used to provide the basis for a subsequent EB study.

Empirical Bayes Method

The modeling approach used for this method combines prior observed data to derive an estimate for the outcomes of interest. The EB approach uses a reference group to estimate a safety performance function (SPF). The SPF is used in combination with observed “before” collisions at the treated sites to predict the number of collisions that would have been expected in the after period for the treated entities had the treatment not been implemented. This prediction is then compared to the expected number of collisions in the after period with the treatment. The specific variables and process required to use the EB method can be found elsewhere (Hauer 1997; Persaud and Lyon 2007). An additional discussion of the methods used to estimate the SPF and weighting factors are described below. In the EB method, the weighting factors are used to jointly consider the observed number of collisions at the treated sites and the expected number of collisions at other similar sites to estimate the expected number of collisions at the treated sites. In the EB method described in the Highway Safety Manual, a negative binomial regression model is used to estimate the SPF. The following functional form is generally used for intersections (Hauer 1997):

[6] $E(\kappa) = \beta_0 V^{\beta_1}$

where $E(\kappa)$ = mean collision frequency per year for the prior distribution; V = annual average daily traffic (AADT); β_0, β_1 = model parameters.

The EB analysis requires AADT and collision data for the treatment site for both before and after the treatment was implemented to calculate a site's expected collision frequency, N_{expected} . The SPF is incorporated into the EB analysis, $N_{\text{predicted}}$, to determine the average collision frequency at similar sites. The sites expected collision frequency can be calculated as follows (HSM 2010):

$$[7] \quad N_{\text{expected}} = w N_{\text{predicted}} + (1 - w) N_{\text{observed}}$$

$$[8] \quad w = \frac{1}{1 + k \sum_n N_{\text{predicted}}}$$

where N_{expected} = estimate of expected average collisions frequency for the study period, $N_{\text{predicted}}$ = predictive model estimate of the average collision frequency predicted for the study period under the given conditions, N_{observed} = observed collision frequency at the site over the study period, w = weighted adjustment to be placed on the predictive model estimate, n = number of years, and k = over-dispersion parameter of the associated SPF used to estimate $N_{\text{predicted}}$. A summary of the procedures of the EB method is presented in Fig. 3.

4. Application to RLC Network in Ottawa

4.1 Data Collection and Preparation

The data used for this paper were obtained from the City of Ottawa which included fifteen years (January 1, 1999 to December 31, 2013) of detailed traffic collision data. The collision data involved 34 RLC intersections and 88 comparison intersections, which consisted of 15,425 collisions. There was a total of 31,399 unit records in the database due to the multiple units involved in some collisions. Of the 15,425 collision cases, 0.09% were classified as fatal, 25.5% were classified as serious injury and the rest property-damage only (PDO). The collision database included information regarding the date, time, type, classification, road surface condition, vehicle type, vehicle maneuver, apparent driver action, pedestrian involvement, road alignment and initial direction of travel date of each incident. From the total of 34 RLC equipped intersections, 30 intersections were eventually included in this statistical analysis based on data availability.

Detailed drawings of the signalized intersections were not available at the City of Ottawa's traffic engineering department to record the road characteristics. Therefore, a geometry database was created using Google Earth images to record road characteristics of the RLC intersections including the number of through lanes, number of left turning lanes, presence of median, and right turn channelization. Additional traffic and signal timings, such as posted speed limits on the major and minor roads and the amber signal timing were obtained from the Traffic Operations department at the City of Ottawa.

The average annual daily traffic (AADT) volumes were used to attempt to adjust for exposure. AADT does not necessarily follow a linear distribution so a better approach was to introduce the natural log of the AADT as the independent variable. The AADT data were available for each RLC and comparison group intersection. Table 2 presents the descriptive statistics of the explanatory variables.

4.2 Evaluation Results

The statistical analysis methods were conducted to assess the association of the RLC program with changes in road collisions at intersections. The intersection collisions were aggregated into total, rear-end, angle, turning movement, fatal, non-fatal injury, and property damage only (PDO) vehicle collisions. StataSE version 14.0 was used for all statistical analyses. The results of the PSM method, cross-sectional method, and the EB method are presented next.

Propensity Score Matching Method

Propensity scores were estimated using the method described previously, with the propensity score being the probability that the intersection received the RLC treatment. The correlation of variables in the propensity score model are not considered to be a problem since the goal of the model is to balance the covariates, not to find statistically significant relationships between the covariates and the treatment (Conniffe et al. 2000).

The propensity score method selected for this study was the NN matching without replacement method, described by Sasidharan and Donnell (2013), to estimate the treatment effect by specifying variables for the outcome model rather than the treatment model. The individual from the comparison group is chosen as a matching partner for a treated individual that is closest in terms of propensity score. Several variants of NN matching are proposed, e.g. NN matching 'with replacement' and 'without replacement'. In the former case, an untreated individual can be used more than once as a match, whereas in the latter case it is considered only once. (Caliendo and Kopeinig 2005). The advantages of matching include transparency in that once matched, the groups to be compared can be explicitly described so that any reviewer can see whether the characteristics of the compared groups might be subject to confounding.

The model estimates separate propensity scores for each year of data. A binary logit model with the dependent variable indicating whether or not segments have the RLC treatment was used for propensity score estimation. The binary logit model estimation results were used to estimate the propensity scores for each of the road segments in the database. The selected models took the form of Equation 4 as shown in Table 3.

The quality of the propensity scores estimated in a sample are evaluated using two types of comparisons: comparing the distributions of the propensity scores across the two groups and comparing the distributions of each covariate across the two groups (Arbogast and Seeger 2012). The comparison of propensity score distributions is typically referred to as the evaluation of common support. Ideally, propensity score distributions would overlap entirely indicating that observations from both groups are available across the range of the propensity score. Methods for evaluating common support include graphical displays of the density distribution of the propensity score for both the treatment and control groups which can help to provide clear knowledge of the extent to which there is overlap in the propensity score between the two groups. Areas with only limited (or no) overlap between the two groups is more common in the tails of the density distribution, suggesting that units most (or least) likely to participate in the intervention are very different from the large majority of cases.

The relative frequency distributions of the estimated propensity scores for total collisions for both the treated and untreated (i.e., reference group) intersections are shown in Fig. 4. The `-pstest-` command in StataSE software was used to plot these graphs intended for checking covariate imbalance; to see if the matching made the treatment and control groups more alike. The amount of non-overlap of these two curves on the middle of the distribution identifies intersections who have a very low probability of treatment and are never treated. Conversely, the area of overlap will identify patients who have comparable propensity scores among treated and untreated subjects, and are therefore better candidates for inclusion in a comparative analysis.

It should be noted that the real purpose of matching is to balance the characteristics between the treatment and control groups. The comparison of the distribution of each covariate across the two groups is referred to as balancing. Unlike the evaluation of common support, assessment of balance takes place after conditioning on the propensity score. A balancing test is proposed by Rosenbaum and Rubin (1983) and applied by Dehejia and Wahba (2002). The units are first divided into blocks with similar propensity scores. Within each block, the t-test is used to test whether the distribution of covariates X is the same in both groups. If a block has unbalanced covariates, then that block is divided into smaller blocks and the evaluation is repeated. If differences still remain, then the specification for propensity score is revisited and higher order or interaction terms are included. Then the process is repeated again from the beginning. Such a stratification test was done using the `-psmatch2-` command in StataSE.

The NN matching models, presented earlier in this section, estimated the average treatment effect on the treated (ATET) coefficients. The ATET coefficients, presented in Table 4, are the actual average expected decline in collisions per year at a treated intersection, i.e. for total collisions, each year there would be 1.7969 fewer total collisions per treated intersection compared to what is expected to have occurred without treatment. It is important to note that all of these estimated treatment effects are statistically significant (p -value < 0.05). These coefficients were

used to estimate the reduction in collisions due to the RLC treatment based on the before and after observations and the results are outlined in Table 5.

Cross-Sectional Method

A wide range of external factors that affect collisions at signalized intersections were considered. A key focus was to attempt to isolate and identify some of the factors which contributed to the occurrence of collisions at RLC intersections. The dependent variable (number of collisions) is estimated using a NB equation. To determine whether the zero-inflated negative binomial (ZINB) models provide an improvement over the traditional NB model, ZINB regression was carried out for the different collision types modeled earlier. The likelihood ratio test of $\alpha = 0$ was significantly different from zero for all models. This suggests that the data is over-dispersed and that a ZINB model is more appropriate than a zero-inflated Poisson model. The Vuong test suggested that the ZINB is not a significant improvement over a standard NB model for all collision types.

The independent variables were selected based on theoretical importance in previous studies, availability of data and statistical significance (p -value < 0.05). The key independent variable was the presence or absence of a RLC, *rlcinstalled*, the treatment variable. Statistics for Akaike Information Criterion (AIC) were checked to select the best-fitting model; as the criterion decreases the model becomes more desirable. The explanatory power of the model is measured by the maximized log likelihood of the predictor coefficients and error variance. The resulting best fit models are outlined below, where these models took the form of Equation 5 as follows:

$$[9] \text{ Total} = \exp(-10.3940 - 0.0853 \text{ rlcinstalled} + 1.2301 \text{ lnaadt} + 0.0185 \text{ thrulanes})$$

$$[10] \text{ Rear end} = \exp(15.1396 + 0.0733 \text{ rlcinstalled} + 1.5706 \text{ lnaadt} + 0.3967 \text{ fplt} + 0.3529 \text{ channelization})$$

$$[11] \text{ Angle} = \exp(-5.7207 - 0.2727 \text{ rlcinstalled} + 0.5776 \text{ lnaadt} + 0.1890 \text{ thrulanes} + 0.1915 \text{ leftturn})$$

[12] Turning

$$mvmt = \exp (-6.0292 - 0.3134rlcinstalled + 0.74271lnaad + 0.7147pplt - 0.0186speedl)$$

[13] Fatal = $\exp (-16.1071 - 1.5813rlcinstalled + 0.4193lnaad + 2.5163thrulanes)$

[14] Injury = $\exp (-6.9640 - 0.0383rlcinstalled + 0.7515lnaad - 0.2046median)$

[15] PDO = $\exp (15.1396 + 0.0733rlcinstalled + 1.5706lnaad + 0.3967fplt + 0.3529channelization)$

The dependent variable is the number of collisions per year, which is a rate or a number of events per time. Hence, the regression coefficients can be interpreted as the log of the rate ratio. The rate at which events occur is called the incidence rate; thus it is possible to interpret the coefficients in terms of incidence rate ratios (IRR) (UCLA Statistical Consulting Group 2007). Based on the IRR, if an intersection was to have a RLC treatment, while holding all other variables in the model constant, the rate for the different collisions types and classifications is presented in Table 6.

. The p -value of the RLC treatment variable was 0.000 for the angle collisions model and 0.001 for the turning movement collisions model. These results suggest that a strong degree of statistical significance exists and that the results are highly unlikely to have occurred by chance. Therefore, the null hypothesis that there is no difference between the treated and non-treatment cases is strongly rejected. The coefficients of RLC are negative for total collisions, angle collisions, and turning-movement collisions indicating their effect in reducing intersection collisions.

However, with a value of 0.063 the p -values of the RLC treatment variable for the best fit total collisions model is above the generally accepted maximum (0.05) needed to reject the null hypothesis of no significant effect. This result held for the other models and variables tested. Similarly, the best fit model for the rear-end collisions yielded a p -value of 0.219 for the treatment, which is far from the threshold for statistical significance of 0.05. This result suggests that the observed difference of the RLC treatment in the total and rear-end collisions is plausibly a chance

finding and that the idea that the hypothesis that there is no difference between the treatment and non-treatment intersections cannot be rejected. Failing to reject the null hypothesis according to the standard criteria of statistical significance, however, does not definitively mean that there is no treatment effect (Altman and Bland 1995). It is the size of the effect that determines the importance, not the presence of statistical significance (Crombie and Davis 2009). The model coefficients show that the presence of RLC is associated with an increased probability of rear-end collisions. The changes in the number of collisions due to the RLC treatment based on before-after observations are outlined in Table 7.

Empirical Bayes Method

The collision data provided by the City of Ottawa were used to develop SPF for different collision types and classifications. The negative binomial models were used to estimate the SPF (Table 8), based on Eq. 5. These models provide the expected average safety performance (collision frequency) for a group of typical signalized intersections representing a range of traffic volumes. Combining the "before" years of the annual collision experience at RLC intersections with the derived SPF, the number of total rear-end, angle, fatal, injury, and PDO collisions that would have been expected to occur if the treatments had not been implemented at each intersection were estimated. The expected number of collisions from the EB estimation is then compared to the actual (observed) number of collisions that occurred at each of the three years after the RLC were implemented. This comparison provides the basis for the safety effectiveness of the treatment to be assessed at each of the RLC sites.

Based on the SPF and the over-dispersion parameters derived from the data, the number of expected collisions had the RLC treatment not been installed was estimated. Table 9 shows the percentage difference between the expected and observed "after" (post-treatment) number of fatal, injury, and PDO collisions averaged over the 30 RLC sites. The models based on collision-impact types show a reduction in angle and turning movement collisions, but show an increase in rear-end collisions. It is speculated that the introduction of RLC tends to increase rear-end collisions due to abrupt braking by a driver as he/she approaches an intersection and realizes it has a RLC, causing the driver behind to collide into the back

of the front vehicle when following too closely. Consequently, the net result is a minimal increase in the total number of collisions. The increase in total collisions is due to the fact that rear-end collisions account for 37% of the total collisions, which skews the results.

The models based on collision classifications have shown that the RLC has had an encouraging safety effect as the treatment has reduced the number of non-fatal injury collisions, thereby saving lives and reducing the number of individuals injured at intersections. However, these treatments have also been shown to have increased the number PDO collisions.

4.3 Comparison of Methodologies

Table 10 presents a summary of the collision prediction estimates of the different methods examined in this paper. The observed and the counterfactual number of collision estimates are very similar qualitatively, though there are some differences in the magnitude of the estimated effects. This finding is important as it is not always possible or practical to conduct an EB study to assess the safety effectiveness of a treatment. The PSM methodology appears to be a viable alternative to the EB methodology. Using multiple techniques that address different deficiencies in each specific method also adds to the robustness and confidence that the user can have in the evaluation of different traffic-related treatments.

It should be noted that the EB method currently used in traffic safety research (AASHTO 2010) does not have a formal, quantitative mechanism built-in to ensure that treated and untreated sites are similar with regards to site-specific features. If the untreated sites are not similar, the SPF developed for use in the EB method may yield biased results (Davis 2000). Thus, alternative evaluation methods are needed to overcome the challenges associated with the EB method, and they must be able to address selection biases that are common to observational studies in general.

7. Conclusions

This paper has introduced the PSM method for the evaluation of RLC and compared its performance to that of the traditional cross-sectional and EB methods. The PSM method can be viewed as an additional tool that is available to investigators as they try to estimate the effects of treatments in traffic studies where potential bias may exist. This method has not been previously applied in assessing the treatment effect of RLC on different collision types and classifications. Based on this study, the following comments are offered:

1. The results of this study show that the cross-sectional and EB methods yielded similar collision predictions compared to the PSM method. Therefore, the PSM method appears to be a viable alternative that adds to the robustness and confidence in the evaluation of different traffic-related treatments.
2. The greater analytical flexibility, nuance and confidence that can be acquired through multiple assessment methods can also help to identify combinations of treatments that can enhance the safety benefits of different traffic management techniques. By examining more specifically the conditions in which RLC are associated with certain types of collisions, policy-makers can combine treatments to create the best conditions for reducing accident rates, injuries, and fatalities.
3. Besides demonstrating the value of using multiple evaluation methods, this study has shown the effects and potential value of using RLC to affect collision profiles for key intersections. All evaluation methods indicate that while these cameras reduce overall collision rates modestly, the aggregate effect is the result of a substantial decrease in angle and turning-movement collisions that is partly offset by an increase in rear-end collisions. More importantly, the results indicate that RLC lead to important reductions in fatal and non-fatal injury collisions, though PDO collisions increase. From a safety perspective, therefore, red light cameras may be viewed as a valuable instrument for managing traffic safety problems in urban areas.

4. When estimating the treatment effects using observational data, the PSM method is preferred to regression-based methods for several practical reasons. First, unlike regression-based methods, where the researcher is usually tempted to continually modify the regression model to achieve the desired accuracy, in the PSM method the propensity score is estimated and a matched sample is constructed without any reference to the outcome (Rubin 2001). Second, in the PSM method, the outcomes are explicitly compared between treated and untreated subjects that have similar distributions of observed baseline covariates. In the regression-based methods, the degree of overlap between the distributions of baseline covariates for the two groups may be difficult to assess. Third, it is simpler to determine whether the propensity score model has been adequately specified conditional on the propensity score, but much more difficult to determine whether the regression model has been correctly specified (Austin 2011).
5. Future enhancements to the EB method should consider integrating propensity scores into the process for identification and selection of a suitable reference group. Future work should also consider how to deal with the hidden bias (bias from unmeasured covariates) in the EB method. When propensity score methods are used with cross-sectional data, the results are less likely to be biased. Thus, the results of the propensity score analysis can likely be trusted more than traditional cross-sectional analysis methods in many cases. Future applications of the PSM method should test the sensitivity of the results by considering additional independent variables such as horizontal alignment and roadside elements in the PSM method. Finally, a larger sample of treatment and reference group sites may be considered in future research since statistically insignificant finding is often attributed to small sample sizes pertaining to sites and the collision frequency at each site.

Acknowledgements

The authors are grateful to two anonymous reviewers for their thorough and most helpful comments. The authors also wish to thank the staff at the Customer and Traffic Information Services at the City of Ottawa and Mr. Stuart Edison and his team at the Traffic Management Centre, City of

Ottawa for providing the data used in this paper. Also, the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Carleton University is gratefully acknowledged.

References

- Akematsu, Y., Tsuji, M. 2012. Measuring the effect of telecare on medical expenditures without bias using the propensity score matching method, *Telemedicine and e-Health*, 2012, Vol. 18, No. 10, pp. 743-747.
- Altman D., Bland J. 1995. Statistics notes: Absence of evidence is not evidence of absence. *BMJ*. 311: 485. Available from <http://www.ncbi.nlm.nih.gov/pubmed/7647644> [accessed 01 March 2017].
- American Association of State Transportation Officials (AASHTO). 2010. Highway Safety Manual. AASHTO, Washington, D.C.
- Arbogast, P. and Seeger, J. 2012. Summary Variables in Observational Research: Propensity Scores and Disease Risk Scores. Agency for Healthcare Research and Quality. Effective Health Care Program Research Report Number 3. Available from <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=1084> [accessed 01 March 2017].
- Austin, P.C. 2011. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*. 46(3): 399–424.
- Caliendo, M., and S. Kopeinig. 2005. Some practical guidance for the implementation of propensity score matching. [online] IZA Discussion Paper No. 1588. Bonn, Germany. Available from <http://ftp.iza.org/dp1588.pdf> [accessed 01 March 2017].
- Cochran WG. 1968. The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*. 24(2): 295-313.
- Conniffe D., Gash V., O'Connell P.J. 2000. Evaluating state programmes: "natural experiments" and propensity scores. *The Economic and Social Review*, Vol. 31, No. 4, pp. 283-308. Available from <https://core.ac.uk/download/files/153/7045479.pdf> [accessed 01 March 2017].
- Crombie, I. and Davis, H. 2009. What are confidence intervals and p-values? 2nd edition. Hayward Medial Communications. [online] Available from https://grunigen.lib.uci.edu/sites/all/docs/gml/what_are_conf_inter.pdf

- Davis, G., 2000. Accident reduction factors and causal inference in traffic safety studies: a review. *Accident Analysis & Prevention*, 32: 95-109.
- Dehejia, R., Wahba, S. 2002. Propensity score matching methods for non-experimental causal studies. *Review of Economics and Statistics*.84:151–161.
- Donnell, E., Shankar, V., and Richard, P.(2009). Analysis of Safety Effects for the Presence of Roadway Lighting. [online] University Park, PA: The Pennsylvania State University. Available from http://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP05-19_SafetyReport.pdf [accessed 01 March 2017].
- Elliott, M. R., Kallan, M. J., & Durbin, D. R. (2009). Child restraint systems for young children during motor vehicle collisions. *American Journal of Public Health* , 99(9), 1540–1541. (DOI:10.2105/AJPH.2009.166421)
- Erke, A. 2009. Red-Light Cameras - A Meta-analysis of the Effects of Red-light Cameras on Collisions. *Accident Analysis and Prevention*. 41 (5): 897-905.
- Federal Highway Administration (FHWA). 2009. Safety Analyst Brochure. [online] Available from <http://www.fhwa.dot.gov/publications/research/safety/09023> [accessed 01 March 2017].
- Federal Highway Administration (FHWA). 2011. Highway Safety Improvement Program Manual. [online] Available from <http://safety.fhwa.dot.gov/hsip/resources/fhwasa09029/sec6.cfm> [accessed 01 March 2017].
- Garber, N.J., Miller, J.S., Abel, R.E., Eslambolchi, S. and Korukonda, S.K. 2007. The impact of red-light cameras (photo-red enforcement) on collisions in Virginia. Report no. VTRC 07-R2. Charlottesville, VA: Virginia Transportation Research Council.
- Gross, F. 2013. Case-control analysis in highway safety: Accounting for sites with multiple crashes. *Accident Analysis and Prevention*. 61: 87-96.
- Gum, P. A., Thamilarasan, M., Watanabe, J., Blackstone, E. H. , Lauer, M. S. 2001. Aspirin use and all-cause mortality among patients being evaluated for known or suspected coronary artery disease: a propensity analysis, *Journal of the American Medical Association*, Vol. 286, pp.1187-1194.
- Guo, S., & Fraser, M. 2010. Propensity score analysis. *Statistical Methods and Applications*. Thousand Oaks: Sage.
- Hallmark, S., Orellana, M., Fitzsimmons, E., McDonald, T., Matulac, D. 2010. Evaluating the Effectiveness of an Automated Red Light Running Enforcement Program in Iowa Using a Bayesian Analysis. Paper No.10-0489, TRB, Washington, D.C.

- Harknett K. 2006. Does receiving an earnings supplement affect union formation? Estimating effects for program participants using propensity score matching. *Evaluation Review*. 30: 741-778. (DOI:10.1177/0193841X06293411).
- Harwood, D.W., Council, F.M., Hauer, E., Hughes, W.E., and Vogt, A. 2000. Prediction of the Expected Safety Performance of Two-lane Rural Highways. Report No. FHWA-RD-99-207, Federal Highway Administration, McLean, VA.
- Hauer, E. 1997. *Observational Before-After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety*. Pergamon Press, Elsevier Science Ltd., Oxford, United Kingdom.
- Heinrich, C., Maffioli, A. and Vázquez, G. 2010. A Primer for Applying Propensity-Score Matching. Inter-American Development Bank Impact-Evaluation Guidelines Technical Notes No. IDB-TN-161. Available from <http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=35320229> [accessed 01 March 2017].
- Hulley SB, Cummings SR, Browner WS, et al. 2001. *Designing clinical research*. 2nd ed. Philadelphia: Lippincott Williams & Wilkins.
- Jepsen, P., Johnsen, S. P., Gillman, M. W., & Sørensen, H. T. 2004. Interpretation of observational studies. *Heart*, 90(8): 956–960. Available from <http://doi.org/10.1136/hrt.2003.017269> [accessed 01 March 2017].
- Kennedy, P. 2003. *A guide to econometrics*. 5th ed. Cambridge, MA: MIT Press.
- Kincaid, D.L. 2009. *Data Analysis Using Logistic Regression and Structural Equation Modelling (SEM)*. Hanoi, Vietnam: Centre for Community Health Research and Development (CCRD).
- Lanza, S. T., Moore, J. E., and Butera, N. M. 2013. Drawing causal inferences using propensity scores: a practical guide for community psychologists. *Am. J. Community Psychol.* 52, 380–392.
- Lechner, M. 2002. Program heterogeneity and propensity score matching: An application to the evaluation of active labor market policies. *Review of Economics and Statistics*. 84: 205-220. (DOI:10.1162/003465302317411488).

Li, H., D.J.Graham, and A. Majumdar. 2013. The impacts of speed cameras on road accidents: An application of propensity score matching methods. *Accident Analysis and Prevention*. 60: 148-157.

Lord, D., Mannering, F. 2010. The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives. *Transportation Research - Part A*; 44(5): 291–305.

Persaud, B. and Lyon, C. 2007. Empirical Bayes before-after safety studies: lessons learned from two decades of experience and future directions. *Accident Analysis and Prevention* 39: 546-55.

Reeve, B. B., Smith, A. W., Arora, N. K., & Hays, R. D. 2008. Reducing Bias in Cancer Research: Application of Propensity Score Matching. *Health Care Financing Review*, 29(4): 69–80.

Rubin, D. 2001. Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services & Outcomes Research Methodology* 2:169–188

Rosenbaum, P.R., Rubin, D.B. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*; 70:41–55.

Sasidharan, L. and Donnell, E. 2013. Application of propensity scores and potential outcomes to estimate effectiveness of traffic safety countermeasures: Exploratory analysis using intersection lighting data, *Accident Analysis & Prevention*, 50: 539-553.

Social Science Computing Cooperative (SSCC). 2015. Propensity Score Matching in Stata using teffects. [online] University of Wisconsin - Madison. Available from https://www.ssc.wisc.edu/sscc/pubs/stata_psmatch.htm [accessed 01 March 2017].

Stone R.A., Obrosky D.S., Singer D.E., Kapoor W.N., Fine M.J. 1995. Propensity score adjustment for pretreatment differences between hospitalized and ambulatory patients with community-acquired pneumonia. *Journal of Medical Care*.33: 56-66.

UCLA Statistical Consulting Group. 2007. Introduction to SAS. [online] Available from <http://www.ats.ucla.edu/stat/sas/notes2/> [accessed 01 March 2017].

Wang, P. S., Schneeweisse, S., Avorn, J., Fischer, M. A., Mogun, H, Solomon, D. H., Brookhart, M. A. 2005. Risk of death in elderly users of conventional vs. atypical antipsychotic medications. *New England Journal of Medicine*, Vol. 353, pp. 2335-2341.

- Wood, J.S. and Porter, R.J. 2013. Safety Impacts of Design Exceptions on Non-Freeway Segments. Session 679 of the 92nd Annual Meeting of the Transportation Research Board, Washington, D.C.
- Wood, J., Donnell, E., Porter, R. 2015. Comparison of safety effect estimates obtained from empirical Bayes before-after study, propensity scores-potential outcomes framework, and regression model with cross-sectional data. *Accident; Analysis and Prevention*. 75: 144-54.

Draft

List of Figure Captions

Figure 1 – Procedures of the PSM method

Figure 2 – Selection of regression model for the cross-sectional method

Figure 3 – Procedures of the empirical Bayes method

Figure 4 – Matching of estimated propensity scores

Draft

Table 1 – Methods for analysis of observational data in traffic studies

Type of observational study	Method Description	Advantages	Limitations
Cross-section method	Treatment and collisions are assessed simultaneous in each intersection at a given point or snapshot in time. Such studies are useful for establishing associations rather than causality and for determining prevalence, rather than incidence.	- Provides good estimate of prevalence (number of cases of a specific outcome that exist in a defined population at a specified point in time.) - Simple and inexpensive - Useful when the before period data are not available; data exhibits under or over-dispersion; or, the treatment installation date is unknown	- Not useful for rare events because in a one-time snapshot (rare outcomes or rare exposures may not be captured) - Does not permit distinction between cause and effect - Cannot infer temporal sequence between exposure and outcome if exposure is a characteristic
EB before-after method	Uses a reference group to estimate . An SPF is used in combination with observed “before” collisions at the treated sites to predict the number of collisions that would have been expected in the after period for the treated entities had the treatment not been implemented. This prediction is then compared to the expected number of collisions in the after period with the treatment.	- Mitigates regression-to-the-mean (RTM) bias - Can appropriately evaluate sites with zero collision counts during the analysis period - Accounts for differences in traffic volume and collision trends between the before and after periods at the treatment sites	- Large sample sizes and long follow-up (years of data) is necessary - Exposure may be linked to a hidden confounder. - No guidelines exist concerning how to most effectively select the group of reference sites for developing SPF. - The statistical significance of observed changes cannot be evaluated.
Case-control method	Match an untreated group to the treated group such that the matched group is comparable to the treated group in all characteristics except the treatment. The observed outcome for the matched group can then be used as the counterfactual, and the treatment effect is estimated simply as the difference between the average outcomes of the two groups.	- Feasible method for studying very rare disorders or those with long lag between exposure and outcome - Can use smaller sample sizes than cross-sectional studies - Cost/time effective when using previously collected exposures	- Relies on precision of previous records to determine exposure status (retrospective) - Cannot calculate the relative risk (incidence) because the cases already had it when the study begins - Difficult to infer temporal relationship between exposure and disease

Table 2 – Descriptive statistics of the explanatory variables

Variable	Number of Observations	Mean	Standard Deviation	Min	Max
Presence of RLC (<i>rlcinstalled</i>)	1,830	0.47611	0.49957	0	1
Natural log of AADT (<i>lnaad</i>)	1,830	10.1304	0.4728	8.15	11.19
Presence of a median (<i>median</i>)	1,830	0.57377	0.49466	0	1
Presence of a right turn channelization (<i>channelization</i>)	1,830	0.26230	0.44000	0	1
Number of thru lanes (<i>thrulanes</i>)	1,830	2.20492	0.59987	1	4
Number of left turn lanes (<i>leftturn</i>)	1,830	0.59016	0.58345	0	2
Fully-protected left turn signal (<i>fplt</i>)	1,830	0.26230	0.44000	0	1
Permitted-protected left turn signal (<i>pplt</i>)	1,830	0.20492	0.40375	0	1
Permitted left turn signal (<i>plt</i>)	1,830	0.68033	0.46648	0	1
Posted speed limit on major road (<i>speed1</i>)	1,830	52.9508	7.6510	40	70
Posted speed limit on minor road (<i>speed2</i>)	1,830	53.1148	6.7868	40	80

Table 3 – Syntax of selected models for the PSM method

Model Type	StataSE Syntax
Total	teffects nnmatch (<i>totalcollisions lnAADT median channelization fplt pplt speed1</i>) (<i>rlcinstalled</i>), atet
Rear end	teffects nnmatch (<i>rearend lnAADT median channelization leftturn thrulanes fplt</i>) (<i>rlcinstalled</i>), atet
Angle	teffects nnmatch (<i>angle lnAADT median channelization leftturn thrulanes plt fplt pplt speed2</i>) (<i>rlcinstalled</i>) , atet
Turning movement	teffects nnmatch (<i>turningmovement lnAADT median channelization leftturn thrulanes plt fplt speed1 speed2</i>) (<i>rlcinstalled</i>), atet
Fatal	teffects nnmatch (<i>fatal lnAADT median thrulanes</i>) (<i>rlcinstalled</i>), atet
Non-fatal Injury	teffects nnmatch (<i>injury lnAADT median channelization leftturn thrulanes plt speed1</i>) (<i>rlcinstalled</i>), atet
PDO	teffects nnmatch (<i>pdo lnAADT channelization leftturn thrulanes pplt speed2</i>) (<i>rlcinstalled</i>), atet

Table 4 – StataSE output of the PSM method

Model Number	Collision type & classification	ATET Coef.	Std. Err.	z	p > z	95% Conf. Interval	
1	Total collisions	-1.7969	0.6539	-2.75	0.0060	-3.0785	-0.5153
2	Rear end	0.8970	0.3629	2.47	0.0130	0.1858	1.6082
3	Angle	-0.4848	0.2140	-2.26	0.0240	-0.9043	-0.0653
4	Turning movement	-0.6338	0.2449	-2.59	0.0100	-1.1139	-0.1538
5	Fatal	-0.0378	0.0183	-2.06	0.0390	-0.0738	-0.0018
6	Non-fatal injury	0.3969	0.2014	1.97	0.0490	0.02058	0.7918
7	PDO	1.9543	0.5356	3.65	<0.000	0.9045	3.0040

Table 5 – PSM results of numbers of collisions avoided

Collision Types & Classification	AAOC ^a (Before RLC)	AAOC ^a (After RLC)	Coef. ^b	Effect without RLC (Col 1 – Col 3)	Percent Change (Col 2-Col 4) / Col 4
	(1)	(2)	(3)	(4)	(5)
Total Collisions	14.86	14.00	-1.7969	16.65	-16
Rear end	5.78	7.25	0.8970	4.88	49
Angle	3.43	2.38	-0.4848	3.92	-39
Turning movement	3.07	2.35	-0.6338	3.70	-37
Fatal	0.02	0.02	-0.0378	0.05	-72
Non-fatal injury	3.86	3.20	0.3969	3.46	-8
PDO	10.98	10.79	1.9543	9.02	20

^a AAOC = Average annual observed collisions per intersection.

^b Coefficients of PSC model from Table 2.

Table 6 – NB regression results

Independent variables	Dependent Variable and Selected Model													
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	Total		Rear end		Angle		Turning Movement		Fatal		Injury		PDO	
	Coef. (IRR)	p-value	Coef. (IRR)	p-value	Coef. (IRR)	p-value	Coef. (IRR)	p-value	Coef. (IRR)	p-value	Coef. (IRR)	p-value	Coef. (IRR)	p-value
constant	-10.3940	<0.000	15.1396	<0.000	-5.7207	<0.000	-6.0292	<0.000	-16.1071	0.2150	-6.6940	<0.000	-3.6673	-
<i>rlcinstalled</i>	-0.0853 (0.9182)	0.0630	0.0733 (1.0761)	0.2190	-0.2727 (0.7726)	<0.000	-0.3134 (0.7309)	0.0010	-1.5813 (0.2057)	0.1750	-0.0383 (0.9624)	0.5930	0.1294 (1.1381)	0.0350
<i>lnaad</i>	1.2301	<0.000	1.5706	<0.000	0.5776	<0.000	0.7427	<0.000	0.4193	0.7530	0.7515	<0.000	0.5628	<0.000
<i>fplt</i>	-	-	0.3967	<0.000	-	-	-	-	-	-	-	-	-	-
<i>pplt</i>	-	-	-	-	-	-	0.7147	<0.000	-	-	-	-	-	-
<i>plt</i>	-	-	-	-	-	-	-	-	-	-	-	-	-0.2792	<0.000
<i>speed1</i>	-	-	-	-	-	-	-0.0186	0.0030	-	-	-	-	0.0112	0.0080
<i>speed2</i>	-	-	-	-	-	-	-	-	-	-	-	-	-0.0143	0.0030
<i>median</i>	-	-	-	-	-	-	-	-	-	-	-0.2046	0.0070	-0.6755	<0.000
<i>thrulanes</i>	0.0185	0.6010	-	-	0.1890	0.0020	-	-	2.5163	0.0030	-	-	0.1241	0.0250
<i>leftturn</i>	-	-	-	-	0.1915	0.0030	-	-	-	-	-	-	-	-
<i>channelization</i>	-	-	0.3529	<0.000	-	-	-	-	-	-	-	-	0.6722	<0.000
	Goodness of fit statistics													
Alpha	0.2544		0.3076		0.5075		0.8553		0.5018		4.70E-15		0.4354	
Log likelihood	-2272.28		-1625.04		-1354.71		-1313.61		-1510.85		-17.16		-2044.85	
AIC	4554.56		3262.09		2721.43		2639.21		44.33		3031.71		4109.70	

Table 7 – NB regression estimates of expected collision

Collision Type/Classification	AAOC ^a (Before RLC)	AAOC ^a (After RLC)	IRR	Percent change ^b based on NB models (IRR)	Expected collisions if RLC were not installed based on IRR %
Total Collisions	14.86	14.00	0.9182	-8	16.07
Rear end	5.78	7.25	1.0761	8	5.34
Angle	3.43	2.38	0.7726	-23	4.21
Turning movement	3.07	2.35	0.7309	-27	3.89
Fatal	0.02	0.02	0.2057	-79	0.03
Non-fatal injury	3.86	3.20	0.9624	-4	4.01
PDO	10.98	10.79	1.1381	14	9.46

^a AAOC = Average annual observed collisions per intersection.

^b Percent change between AAOC after RLC and effect without RLC.

Table 8 – Statistical parameters of the estimated safety performance functions of various models

Model	Variable	Coefficient	Std. Error	Over-dispersion parameter	SPF Model Form
Total	<i>AADT</i>	-7.73E-06	6.70E-06	0.4586	$2.2390 * AADT^{-0.00000773}$
	Intercept	2.2390	0.2116		
Rear end	<i>AADT</i>	-1.23E-05	8.79E-06	0.6671	$1.5127 * AADT^{-0.0000123}$
	Intercept	1.5127	0.2743		
Angle	<i>AADT</i>	3.96E-06	8.14E-06	0.1773	$0.2561 * AADT^{0.00000396}$
	Intercept	0.2561	0.2625		
Turning Movement	<i>AADT</i>	-1.17E-05	9.03E-06	0.2825	$0.6960 * AADT^{-0.0000117}$
	Intercept	0.6960	0.2758		
Fatal	<i>AADT</i>	-1.89E-06	0.0001	3.64E-09	$-4.8638 * AADT^{-0.00000189}$
	Intercept	-4.8638	3.2912		
Injury	<i>AADT</i>	<0.0000	7.19E-06	0.1236	$0.8106 * AADT^{-0.00000529}$
	Intercept	0.8106	0.2238		
PDO	<i>AADT</i>	-8.65E-06	6.79E-06	0.4271	$1.9667 * AADT^{-0.00000865}$
	Intercept	1.9667	0.2137		

Table 9 – Statistical parameters of the estimated safety performance functions of various models

Collision Type/Classification	AAOC ^a (Before RLC)	AAOC ^a (After RLC)	Expected collisions after without RLC (EB model calculation)	Percent change ^b
Total Collisions	14.86	14.00	13.62	3
Rear end	5.78	7.25	5.29	37
Angle	3.43	2.38	3.32	-28
Turning movement	3.07	2.35	2.99	-21
Fatal	0.02	0.02	0.02	-25
Non-fatal injury	3.86	3.20	3.75	-15
PDO	10.98	10.79	10.11	7

^a AAOC = Average annual observed collisions per intersection.

^b Percent change between AAOC after RLC and effect without RLC

Table 10 – Percent change of observed and expected number of collisions based on alternative methods

Collision Impact Type/Classification	Percent change of observed and expected number of collisions		
	NB Method	EB Method	PSM Method
Total Collisions	-8	3	-16
Rear end	8	37	49
Angle	-23	-28	-39
Turning movement	-27	-21	-37
Fatal	-79	-25	-72
Non-fatal injury	-4	-15	-8
PDO	14	7	20

List of Symbols

λ = expected number of collisions on a roadway segment/intersection

ω = intercept

β_0, β_1 = model parameters

β_i = coefficients quantifying the relationship between X and λ

θ_i = coefficient quantifying the relationship between L and λ

τ_i = the impact of a treatment for an individual i

D = treatment status ($D = 1$, treated and $D = 0$, untreated)

$E(\kappa)$ = mean annual expected collision frequency for the prior distribution

k = over-dispersion parameter of the associated SPF used to estimate $N_{\text{predicted}}$

L_i = a binary (dummy) variable indicating the presence of the treatment

n = number of years

N_{expected} = estimate of expected average collisions frequency for the study period

N_{observed} = observed collision frequency at the site over the study period

$N_{\text{predicted}}$ = predictive model estimate of the average collision frequency predicted for the study period under the given conditions

$\hat{p}$ = propensity score

V = annual average daily traffic (AADT)

t = treatment indicator

w = weighted adjustment to be placed on the predictive model estimate,

X_i = variables influencing λ (e.g., traffic, road geometry, etc.)

X = covariates influencing treatment selection

x_n = independent variables

Y_{0i} = the potential outcome where individual i does not receive treatment.

Y_{1i} = the potential outcome where individual i receives treatment

y = outcome variable

Formatted: Font: 10 pt, Font color: Text 1

Draft

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1