



A framework for estimating the safety effects of roadway lighting at intersections

Eric T. Donnell^{a,*}, Richard J. Porter^{b,1}, Venkataraman N. Shankar^{a,2}

^a Department of Civil and Environmental Engineering, The Pennsylvania State University, University Park, PA 16802, United States

^b Department of Civil and Environmental Engineering, University of Utah, 122 So. Central Campus Drive, Suite 104, Salt Lake City, UT 84112, United States

ARTICLE INFO

Article history:

Received 23 January 2010

Received in revised form 8 May 2010

Accepted 22 June 2010

Keywords:

Night-to-day crash ratios

Lighting

Statistical analysis

Traffic safety

Intersections

ABSTRACT

National- and state-level guidance documents conclusively state that fixed lighting improves intersection safety. The sentiment is consistent with other design and safety manuals and is supported by a series of consistent safety findings; however, most published lighting-safety research is focused on rural, stop-controlled intersections and is limited by several methodological issues. The relationship between safety and intersection lighting at rural, signalized and urban locations is not as well documented. Methodological advancements in highway safety analysis justify new estimations of the safety effects of intersection lighting. This paper describes a proposed framework to estimate the safety effects of fixed lighting at a variety of intersection types and locations. Several key issues are explored including availability of relevant crash, lighting, and roadway inventory data; relevant data element structures; proposed analysis taxonomies to assess lighting-safety effects within and across different intersection classifications; specification and estimation of models to estimate expected crash frequencies during day and night; techniques to interpret model parameters, including variable elasticity; and tests of model transferability across states. A sample framework execution using Minnesota intersection data is provided. Results indicate a much lower overall safety benefit from lighting than published studies, but are consistent with estimates included in *Highway Safety Manual* research.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The American Association of State Highway and Transportation Officials' (AASHTO) Strategic Highway Safety Plan implementation guide for addressing unsignalized intersection collisions (Neuman et al., 2003) indicates that lighting is a proven safety countermeasure. Likewise, the AASHTO implementation guide to address signalized intersection collisions (Antonucci et al., 2004) suggests that roadway lighting is a tried safety countermeasure, but specific safety effects are not cited. The Minnesota Department of Transportation (Mn/DOT) reports that intersection illumination has the greatest benefit-to-cost ratio (B/C = 21.0) of all safety improvements included in their *Traffic Safety Fundamentals Handbook* (Mn/DOT, 2001).

These national- and state-level guidance documents conclusively state that fixed lighting improves intersection safety. The sentiment is consistent with other design and safety manuals and is supported by a series of consistent safety findings. Most findings are applicable to rural, stop-controlled intersections. Re-

sults of observational with–without comparisons show 25% lower nighttime crash rates (Preston and Schoenecker, 1999), 39% fewer nighttime crashes (Schwab et al., 1982), and 31% lower night-to-day crash ratios (Isebrands et al., 2004, 2006) at intersections with fixed lighting, compared to those without. Before–after studies show reductions in nighttime crash rates of 45–52% (Walker and Roberts, 1976; Lipinski and Wortman, 1978), reductions in nighttime crash frequency of 13–49% (Schwab et al., 1982; Isebrands et al., 2004, 2006; Walker and Roberts, 1976; Green et al., 2003), reductions in night-to-day crash ratios of 22–40% (Preston and Schoenecker, 1999; Isebrands et al., 2004, 2006; Lipinski and Wortman, 1978), and substantial reductions in fatal and injury crashes (Schwab et al., 1982; Lipinski and Wortman, 1978) following the installation of fixed intersection lighting.

A meta-analysis of 37 research studies indicates a 30% reduction in nighttime junction crashes (Elvik, 1995). Further analysis determined that the nighttime crash results found in the studies could not likely be attributed to regression to the mean or secular crash trends. The International Commission on Illumination summarized the results of 62 lighting and crash studies from 15 countries (CIE, 1992). The studies included only before–after analyses involving the installation or upgrading of lighting along roadways. Eighty-five percent of the results showed lighting to be beneficial with an overall average reduction of nighttime crashes of at least 30% due to lighting improvements. At rural intersections, the percent

* Corresponding author. Tel.: +1 814 863 7053.

E-mail addresses: edonnell@engr.psu.edu (E.T. Donnell), richard.jon.porter@utah.edu (R.J. Porter), shankarv@engr.psu.edu (V.N. Shankar).

¹ Tel.: +1 801 585 1290; fax: +1 801 585 5477.

² Tel.: +1 814 865 9434.

reduction in nighttime crashes ranged from 26% to 44%. The *Highway Safety Manual* reports an expected 38% reduction in nighttime fatal and injury intersection crashes following intersection lighting installation. The number provided is a general crash reduction, without specificity to setting (e.g., urban or rural) or intersection type (e.g., four-leg, three-leg, signalized, unsignalized).

The safety benefits of intersection lighting are extensively documented, but a new, comprehensive study is needed. Most previous work is limited to rural, stop-controlled intersections. The relationship between safety and intersection lighting at rural, signalized intersections as well as urban intersections is not as well documented. In addition, methodological advancements made in highway safety analysis justify new estimations of the direction and magnitude of the safety effect. Most studies that were included in the aforementioned meta-analysis were completed between 1948 and 1989, prior to a number of road safety modeling advancements. Three fundamentals worth mentioning in this context are:

- (1) The relationship between crash frequency and traffic is non-linear. Comparisons of crash rates, which assume linear crash-traffic relationships, may lead to incorrect safety conclusions (Hauer, 2005).
- (2) Observed crash frequencies vary randomly; a before–after or with–without comparison does not distinguish between random fluctuations and actual safety effects. Comparisons of *expected* crash frequencies are more meaningful.
- (3) Intersection safety is influenced by a number of characteristics in addition to lighting. Omitting these variables from analysis leads to biased findings, particularly if these features are correlated with lighting presence. A safety assessment of intersection lighting should include the expected crash frequencies conditioned on the values of these other safety influencing features.

In addition, most previous work focused on the nighttime benefit of lighting, which appears substantial. However, highway safety countermeasures represent permanent applications or changes to the roadway environment and their net effect on total crashes (both day and night) should be assessed and compared to other alternative safety improvements. Such an approach will help policy-makers and highway engineers make more informed decisions concerning the overall net effect of a safety countermeasure and are consistent with the safety algorithms in the *Highway Safety Manual*.

Harwood et al. (2007) developed an accident modification factor (AMF) for at-grade intersections along urban and suburban arterials. The AMF is based on a meta-analysis of nighttime crash reductions following lighting installation on road segments by Elvik and Vaa (2004). Harwood et al. (2007) decided that the lighting effectiveness measures for road segments were the best available information to apply to intersections. The nighttime crash reductions reported by Elvik and Vaa (2004) were modified to consider total crashes (i.e., day and night) by using proportions of day and night crashes from Minnesota and North Carolina. When local jurisdictions can collect crash severity data, and stratify crashes by daytime and nighttime periods, Eq. (1) is recommended to compute a total crash AMF for lighting intersections along urban and suburban arterials:

$$AMF_{int} = 1 - [(1 - 0.36p_{fni} - 0.72p_{ini} - 0.83p_{pni})p_{ni}] \quad (1)$$

where p_{fni} is the proportion of total nighttime accidents for unlighted intersections that involve a fatality; p_{ini} , the proportion of total nighttime accidents for unlighted intersections that involve only a non-fatal injury; p_{pni} , the proportion of total nighttime acci-

dents for unlighted intersections that involve property-damage-only; and p_{ni} is the proportion of total accidents for unlighted intersections that occur at night.

Using the combined data from Minnesota and North Carolina, Harwood et al. (2007) illustrated that an appropriate intersection lighting AMF for total crashes is 0.96, or an expected 4% reduction in total crashes after the installation of roadway lighting at intersections on urban and suburban arterials. This analysis underscores the importance of considering all time periods when considering the effect of a safety countermeasure. While roadway lighting does indeed offer significant safety improvements, this effect is limited to nighttime exposure which represents only 20–25% of the total exposure. If approximately 75–80% of the total exposure occurs during the daytime period, the net effect of roadway lighting on total crashes will be significantly lower than nighttime safety effects alone. A more global analysis is needed to address the limitations of past research, including consideration of all potential safety-influencing intersection features in addition to lighting and the effect of roadway lighting on both daytime and nighttime crash frequency.

This paper describes a proposed framework to estimate the safety effects of fixed lighting at a variety of intersection types and locations. Several key issues are explored including:

- availability of relevant crash, lighting, and roadway inventory data;
- relevant data element structures;
- proposed analysis taxonomies to assess lighting effects within and across different intersection classifications;
- specification and estimation of models to estimate expected crash frequencies during day and night;
- techniques to interpret model parameters, including variable elasticity; and
- tests of model transferability across states.

The framework is defined using data and data structures available in California and Minnesota. Execution of one phase of the framework is illustrated using Minnesota intersection data followed by conclusions and recommendations.

2. Available data and data structures

Studies that explore empirical relationships between highway safety and roadway infrastructure elements or traffic engineering measures are commonly retrospective in nature, employing historical traffic, roadway, and crash data. The data may be housed in different divisions of one or more state government agencies, but usually those for transportation and law enforcement. The scope and applicability of empirical safety research are often limited by the availability of such data in a practically useable format, the accuracy and completeness of the coded data, and the ability to “link” the traffic, roadway and crash data.

The *Highway Safety Information System (HSIS)* is a multistate database³ managed by the University of North Carolina Highway Safety Research Center and Lendis Corporation, under contract with the Federal Highway Administration (FHWA). Participating HSIS states were selected based on data quality and the ability to merge electronically coded crash- and infrastructure-related files. The HSIS database is often a primary data alternative for highway safety research involving roadway infrastructure components, including research efforts associated with production of the *Highway Safety Manual* and *Safety Analyst*. It was selected for this study

³ Data are current for California, Illinois, Maine, Minnesota, North Carolina, Ohio and Washington. Older data are available from Michigan and Utah.

as well after a comprehensive review of state databases and the availability of major data elements needed to accomplish the study objectives.

Only California and Minnesota *HSIS* files contained the needed intersection-level geometric design, fixed lighting and traffic volume data. A summary of the data elements available and their respective structures (e.g., continuous, categorical, binary) is provided in the following sections.

The crash-related California and Minnesota files consist of three ‘sub-files’: crash, vehicle and occupant. The crash sub-files include basic data related to the crash event, such as the time and location of the crash, weather and lighting conditions at the time of the crash, crash severity, crash type, and other similar characteristics (Council and Mohamedshah, 2007a,b). The vehicle and occupant sub-files include information specific to each vehicle and each occupant within each vehicle involved in a crash. This includes vehicle type, factors contributing to the crash, occupant age, gender, and seating position. With this sub-file structure, one crash sub-file row may correspond to multiple vehicle rows (e.g., a crash involving three vehicles will have three row entries in the vehicle sub-file). Similarly, one vehicle row may correspond to multiple occupant rows (e.g., a vehicle with four occupants will have four row entries in the occupant sub-file). Crash identification numbers are the mechanisms by which information in all of the sub-files are merged and analyzed.

California and Minnesota intersection files contain data for nearly 25,000 and 26,600 intersections, respectively. Roadway inventory files include both a roadlog and intersection sub-file. The intersection files contain intersection-specific variables such as intersection type (i.e., number of intersection legs, type of traffic control), median presence, channelization presence and type, left-turn restrictions (e.g., permitted vs. restricted), number of through travel lanes, and the average annual daily traffic (AADT) on each crossing street. Additional roadway characteristics (e.g., lane width, shoulder width) of each crossing street are determined by merging the intersection files from California and Minnesota with respective ‘roadlog’ files, which contain geometric information for primary, “trunk” highways, other state-maintained municipal and county roads, and non-state-maintained municipal and county roads. A binary, lighting presence variable is provided in the California intersection file. No information related to the type of lighting or lighting design parameters (e.g., illuminance, luminance, uniformity, mounting height, spacing, fixture type) is provided. The Minnesota intersection file contains a five-level, categorical lighting variable: no lighting, partial lighting, full lighting, continuous lighting, and point lighting. Potential discrepancies in coding of this variable were discovered by the researchers and it was transformed to a binary lighting presence variable following a number of exploratory modeling runs. No data on lighting design were available in Minnesota.

Intersection crash counts were determined using county, route number, and milepost variables. Crashes occurring along each intersecting street, within a 250-foot radius of the center of the intersection, were counted for each intersection. Intersection crash counts were recorded in total (i.e., all crashes), by severity (i.e., fatal, injury, or property-damage-only) and by time-of-day (i.e., day or night). Historical records of sunrise and sunset times, by location and season, were used to classify crashes by time-of-day. The final California and Minnesota databases were structured as a panel. Each row represented one intersection observed for a 1-year period. Data from 1999 to 2002 were used for California (i.e., each California intersection was represented by up to four row entries depending on whether there were missing data and other coding issues); data from 2001 to 2004 were used for Minnesota. Modeling considerations associated with these panel structures are discussed in the following sections.

3. Safety effects of intersection lighting: evaluative aspects

The current study framework is designed to explore the relationship between intersection lighting and three safety performance measures:

- expected number of daytime crashes;
- expected number of nighttime crashes; and
- expected percent change in night-to-day crash ratio (see Eq. (2)).

$$\frac{\left(\frac{N}{D}\right)_w - \left(\frac{N}{D}\right)_{wo}}{\left(\frac{N}{D}\right)_{wo}} \quad (2)$$

where N is the expected number of nighttime crashes for all intersections; D , the expected number of daytime crashes for all intersections; w , the subscript to indicate intersection lighting presence; and wo is the subscript to indicate no intersection lighting.

The selected safety measures allow comparisons to similar, frequently-used performance measures in previous work (i.e., night crashes and night-to-day ratio) but also address the need to determine the overall safety effect of lighting on total crashes. As noted in the introduction (Section 1), an intersection lighting AMF has been offered by Harwood et al. (2007) for inclusion in the first edition of the *Highway Safety Manual*. The AMF was developed by modifying the results of a previous research effort (Elvik and Vaa, 2004), applicable to road segments but seen as the best available information to use for intersections. An alternative approach to develop an intersection lighting AMF is proposed in this study based on the framework described, which makes use of statistical models applied to a large sample of cross-sectional data.

An observational, before–after study has been offered as a superior alternative to a cross-sectional study for estimating the safety effectiveness of traffic engineering measures (Hauer, 2005). In the former, a treatment is applied at a location during the analysis period and the change in crashes after the treatment is applied is compared to an estimate of the expected number of crashes that would have occurred had the treatment not been applied. The intent of the observational before–after method is estimation of an unbiased safety effect. The empirical Bayes method has been advocated as a way to address selection bias (i.e., that traffic safety countermeasures are likely to be applied to sites experiencing a recent, unusually high number of crashes). The reader is referred to Hauer (1997) and Persaud and Lyon (2007) for a discussion of the methodology, and to Monsere and Fischer (2008) for an example related to reduced illumination levels of freeway lighting. The before–after study offers what appears to be considerable strength in separating the actual safety effect from safety influences that are not of interest to the researcher or practitioner. However, it may be difficult to identify a large sample of treated sites to produce a safety effect estimate with a high level of certainty in the case of roadway lighting. Further, it is rare that roadway lighting is the only safety countermeasure applied to a site. Because the warrants noted earlier contain both safety and operational considerations, lighting is often implemented as part of an overall intersection improvement (e.g., conversion of stop-controlled to signalized intersection). As such, the safety effects of roadway lighting may be confounded with other safety treatments applied at a site.

In cross-sectional studies, no treatment is applied to a location; rather, a site with a particular attribute is compared to a site without the same attribute over the same time period. The strengths of a before–after study noted earlier are limitations in a cross-sectional study; one cannot be completely confident in making a leap from statistical associations from cross-sectional model estimates

to causation. Another weakness of a cross-sectional study is that it is difficult to determine the reason why traffic measures of interest (e.g., roadway lighting) have been installed at one location and not at other, similar locations (i.e., selection bias may be an issue here as well). Advantages of cross-sectional studies are that a much larger sample of sites with and without roadway lighting can be identified. Furthermore, many variables can be included in a cross-sectional model which permits opportunities to develop safety-related hypotheses for future causative research.

Very few (less than 5) intersections were identified where lighting was installed at some point during the previously defined observation periods in California and Minnesota. Therefore, a framework built around the cross-sectional approach is the focus of the remainder of this paper. Multivariate regression models of expected day and night intersection crash frequency that include traffic volume, intersection design, traffic control features, and roadway lighting variables were estimated.

3.1. Statistical modeling alternatives of expected crash frequency

Several cross-sectional modeling approaches are considered in the proposed framework. In each method, the first step is to estimate the expected daytime and nighttime crash frequencies as a function of a set of explanatory variables. The negative binomial regression model is a logical choice to estimate the expected number of intersection crashes per year, and is a common approach to model intersection crash frequency (e.g., Poch and Mannering, 1996; Bauer and Harwood, 1996; Washington et al., 2005) because it accounts for the overdispersion common in crash data. The general functional form of the negative binomial regression model is:

$$\ln \lambda_i = \beta X_i + \varepsilon_i \quad (3)$$

where λ_i is the expected number of crashes at intersection i ; β , the vector of estimable regression parameters; X_i , the vector of geometric design, lighting presence, and traffic volume data; and ε_i is the gamma-distributed error term.

The mean–variance relationship for the negative binomial distribution is:

$$\text{Var}(y_i) = E(y_i)[1 + \alpha E(y_i)] \quad (4)$$

where $\text{Var}(y_i)$ is the variance of observed crashes y occurring at intersection i ; $E(y_i)$, the expected crash frequency at intersection i ; α is the overdispersion parameter.

The appropriateness of the negative binomial (NB) regression model is based on the significance of the overdispersion parameter. When α is not significantly different from zero, the negative binomial model reduces to the Poisson model. The method of maximum likelihood is used to estimate the model parameters. The likelihood function for the NB model is shown in Eq. (5).

$$L(\lambda_i) = \prod_{i=1}^N \frac{\Gamma(\theta + y_i)}{\Gamma(\theta) y_i!} \left[\frac{\theta}{\theta + \lambda_i} \right]^\theta \left[\frac{\lambda_i}{\theta + \lambda_i} \right]^{y_i} \quad (5)$$

where N = total number of intersections in the sample; Γ = gamma function; and $\theta = 1/\alpha$.

If the database structure includes only one row per intersection (i.e., the crash counts for each intersection are summed over the entire analysis period) issues of temporal correlation among the crash counts are not an issue. However, if the data are structured as a panel (i.e., one intersection is entered as one row for each year), then temporal correlation may be an issue, likely resulting in underestimated standard errors of the model parameters (Greene, 2003). There are several count regression estimators that address panel data. These include the random effects negative binomial regression (RENB) model, generalized estimating equations (GEE), and the negative multinomial (NM) regression model.

The general functional form of a RENB model is as follows:

$$\ln \lambda_{it} = \beta X_{it} + u_i \quad (6)$$

where λ_{it} is the expected number of crashes at intersection i in year t ; β , the vector of estimable regression parameters; X_{it} , the vector of geometric design, lighting presence, and traffic volume data; and u_i is the random effect for i th intersection ($\exp(u_i)$ is gamma-distributed with mean 1 and variance α ; α is also the overdispersion parameter in the NB model).

Additional details concerning the RENB model and its application in transportation safety research can be found in Shankar et al. (1997) and Chin and Quddus (2003).

Applications of the GEE to model crash occurrence can be found in Lord and Persaud (2000) and Wang et al. (2006). The GEE approach was developed by Liang and Zeger (1986) and Zeger and Liang (1986) to produce more efficient and unbiased regression estimates for use in analyzing longitudinal or repeated measures data with non-normal response variables. Because segment-level crash data are correlated over time, the covariance structure of the data is modeled. The working correlation matrix may be independent, exchangeable, unstructured, or autoregressive. For a detailed overview of each working correlation matrix structure in the context of transportation safety, refer to Wang et al. (2006).

The NM regression model was proposed for use as a safety prediction method by Ulfarsson and Shankar (2003). The main difference between the NM and NB models relates to the random error term. In the NM model, this term is section-specific rather than observation-specific like in the NB model. The NM model reduces to the NB model when a roadway segment contains only a single observation. For a detailed derivation of the model, refer to Guo (1996).

Ulfarsson and Shankar (2003) compared the NB, RENB, and NM models for median crossover crash frequencies. The authors found that the NB model outperformed the RENB model when indicator variables capturing spatial and temporal effects were explicitly considered in the model estimation. The NM model outperformed both the NB and RENB models; however, the parameter estimates for many explanatory variables included in all three models were of similar sign and magnitude. The principal difference in the models was that the standard errors of the parameter estimates were generally underestimated in the NB model when compared to the NM model.

3.2. Endogeneity

As described previously, intersection lighting may be installed for a variety of reasons. Intersections may be selected for lighting installation based on an historical record of high crash frequencies. Ignoring this relationship may result in biased parameter estimates. An instrumental variable approach has been offered in the road safety modeling literature as a possible way to address a similar issue in the context of ice warning signs (Carson and Mannering, 2001) and left-turn lanes (Kim and Washington, 2006). Applying this method to lighting includes estimation of a binary logit lighting presence model. The computed probabilities for lighting presence are then used as an instrument for the endogenous binary lighting presence indicator. The authors of the current study have explored this approach as part of the proposed lighting safety framework, but several challenges exist:

- (1) The presence of lighting was difficult to predict with available exogenous variables. An instrument that is not highly correlated with the variable for which it is an instrument (i.e., a “weak” instrument) may result in parameter bias that is larger than that introduced by endogeneity.

- (2) The econometrics literature regarding the use of instrumental variables in count models is scarce. Although some methods have been proposed (e.g., Windmeijer and Santos Silva, 1997), repeated examples of their application does not exist.
- (3) The endogeneity argument can be made for a number of variables that often appear in intersection safety models (e.g., left-turn lanes, lighting presence, signalization, speed limit). Safety applications to date address only one potentially endogenous variable in the models, while treating other variables as exogenous variables.

Potential endogeneity of lighting presence is not addressed further in this paper. Additional research exploring the aforementioned complexities of endogeneity in the safety modeling context is recommended.

3.3. Analysis taxonomies and tests of model transferability

General analysis taxonomies for California and Minnesota intersections are shown in Figs. 1 and 2, respectively.

At the top of the taxonomies, the expected number of nighttime and daytime crashes is estimated in separate models for all intersections. The information contained in lower levels of the taxonomy is included in these “all intersection” models as indicator variables, and therefore represent fixed effects. The expected number of daytime and nighttime crashes can also be estimated at lower levels of the analysis taxonomy. Such an approach may provide

more informative safety assessments of roadway lighting for various intersection forms and traffic control types, provided that the sample size is sufficient to estimate the models.

If separate models are estimated using data from different states, an important consideration is to test for transferability of the models. To do this, the same specification (i.e., same dependent and independent variables with same functional form) should be used for each state model. A likelihood ratio test to determine if the regression parameters are transferable is illustrated below (Washington et al., 2003), with notation adapted to the current framework:

$$X^2 = -2[LL(\beta_p) - LL(\beta_{CA}) - LL(\beta_{MN})] \tag{7}$$

where $LL(\beta_p)$ is the log-likelihood at convergence of the pooled regression model (includes California and Minnesota samples); $LL(\beta_{CA})$, the log-likelihood at convergence of the California intersection model; and $LL(\beta_{MN})$ is the log-likelihood at convergence of the Minnesota intersection model.

The X^2 statistic is chi-squared (χ^2) distributed. The degrees of freedom for the test is equal to the sum of the number of regression parameters in the California and Minnesota models minus the number of regression parameters in the pooled model. Based on the data structures described previously, the vector of explanatory variables (X) for the transferability test would contain major and minor road approach traffic volumes, a lighting presence indicator, and the type of traffic control present at the intersection (signalized or unsignalized). The null hypothesis of the likelihood ratio

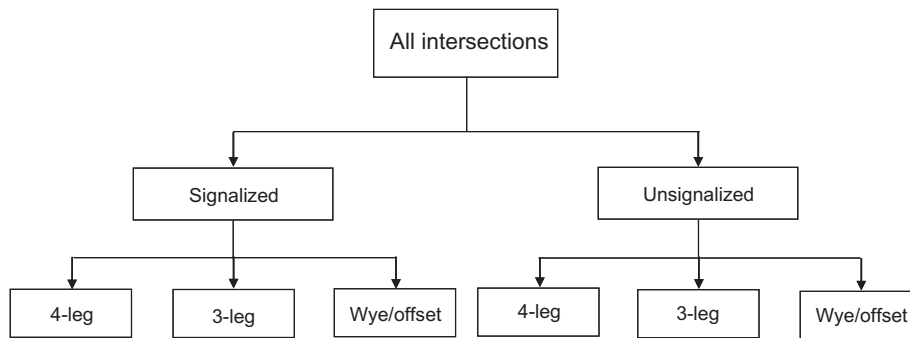


Fig. 1. California analysis taxonomy. (Wye intersections have three approach legs, none meeting at right-angles. An offset intersection includes two closely-spaced tee intersections. For the purposes of illustrating an analysis taxonomy using California HSIS data, these two intersection forms were grouped together; however, they could also be separated to test for differences in the expected crash frequency.)

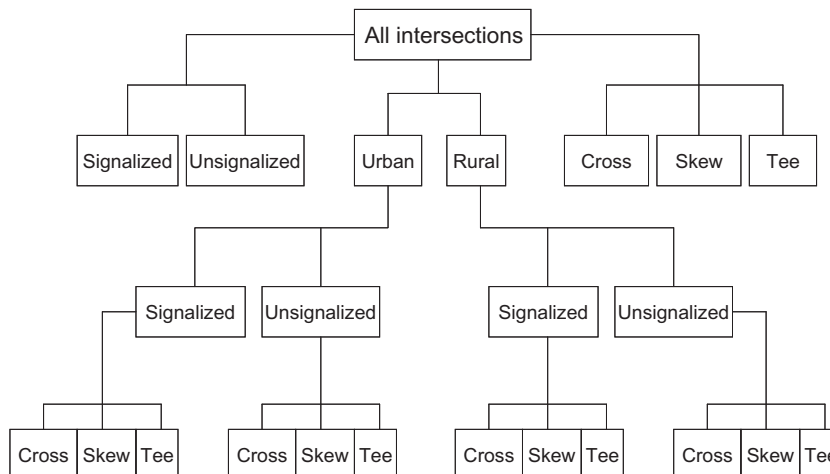


Fig. 2. Minnesota analysis taxonomy.

test for spatial transferability is that the regression parameters are equal between the two models. If the models are not transferable across states, full information models should be estimated separately for each state. These models would include all of the possible explanatory variables included in the analysis database.

4. Example modeling application

In this section, data from Minnesota intersections are used to illustrate execution of the proposed data and modeling framework. All intersection data contained in the Minnesota HSIS roadway inventory, intersection, and crash data files were used to develop the analysis database. Four years (1999–2002, inclusive) of crash and corresponding roadway inventory data were used in the analysis. A total of 25,832 observations (from 6464 intersections) were available for model estimation. Table 1 shows the definitions and descriptive statistics for relevant variables. A total of 888 of the 6464 intersections were signalized, the remaining 5576 operated under stop-controlled conditions.⁴ There were three intersection forms coded in the database: cross, tee, and skew. Approximately 49% of the intersections were four-leg, “cross” intersections. Nearly 40% were three-leg, “tee” intersections. The remaining 11% of the intersections were skew. Tee and cross intersection types were combined into a single category after preliminary estimation runs, resulting in a dummy variable for skew. There were 38,437 reported intersection crashes in the database.

Estimation results for the Minnesota, “all-intersection” analysis (top-level of taxonomy shown in Fig. 2) are shown in Table 2 using NB regression models.⁵ The explanatory variables showing positive associations with the expected daytime crash frequency were major and minor road traffic volumes, indicator for signal control, indicator for lighting presence, skewed intersection form indicator, depressed median indicator, and paved right-shoulder indicator. Explanatory variables showing negative associations with expected daytime crash frequency were percent heavy vehicles on the major road, posted speed limit indicator, no access control indicator, and the paved left-shoulder indicator. Using the parameter estimates from the NB regression models and mean values for all explanatory variables other than lighting, Eq. (2) estimates a difference in night-to-day crash ratio of –11.9%. This indicates that the expected night-to-day crash ratio at intersections in Minnesota with lighting is approximately 12% lower than the expected night-day crash ratio at locations without lighting. As a means of comparison, the percent change in the night-to-day crash ratio considering observed, as opposed to expected, crashes is –28.0%. This latter estimate is not based on statistical models and is similar to metrics used in some previous lighting-safety studies. Without controlling for othersafety-influencing intersection design features, and using only observed crash data, the night-day ratio appears to overestimate the safety benefits of lighting when compared to the night-day ratios estimated from models of expected crash frequencies.

Elasticity represents a measure of responsiveness of one variable to a change in another. In the intersection lighting context, it is interpreted as the percent change in expected crash frequency given a 1% change in a continuous explanatory variable. The elasticity of the expected crash frequency for continuous explanatory variable ‘k’ at intersection ‘i’ during time period ‘j’ is defined as:

⁴ Of the 5576 stop-controlled intersections identified in the database, only 44 were two-way stop-controlled (0.8%). For the purposes of this analysis, all stop-controlled intersections were grouped into one categorical variable to avoid sampling zeros in the analysis data.

⁵ Ongoing research to execute the remainder of the proposed modeling framework is underway. Given the aforementioned results from Ulfarsson and Shankar (2003), it is not expected that the safety effect estimates derived from the NB regression models will change significantly if other modeling approaches are used. Standard errors reported in Table 2 may be slightly underestimated.

Table 1
Variable definitions and descriptive statistics of minnesota data.

Continuous variables	Min.	Max.	Mean	Standard deviation
Night crash frequency (per year)	0	28	0.366	0.969
Day crash frequency (per year)	0	55	1.121	2.457
Major road average daily traffic	40	77,430	8284	9381
Percent heavy vehicles on major road	0	61.11	8.8880	5.1092
Minor road average daily traffic	1	77,430	3164	5179
Categorical variables	Proportion in sample			
Area type indicator (1 = urban/suburban; 0 = rural)	1: 0.446 0: 0.554			
Traffic control indicator (1 = signal; 0 = stop-control)	1: 0.137 0: 0.863			
Lighting indicator (1 = present; 0 = not present)	1: 0.421 0: 0.579			
Intersection type indicator (1 = skew; 0 = cross or tee)	1: 0.110 0: 0.890			
Speed indicator* (1 = 50 mph or greater; 0 otherwise)	1: 0.673 0: 0.327			
No access control indicator* (1 = no access; 0 = partial access control)	1: 0.943 0: 0.057			
Depressed median indicator* (1 = depressed median; 0 = barrier or no median)	1: 0.116 0: 0.884			
Paved left-shoulder indicator* (1 = paved shoulder; 0 = unpaved or no shoulder)	1: 0.458 0: 0.542			
Paved right-shoulder indicator* (1 = paved shoulder; 0 = unpaved or no shoulder)	1: 0.510 0: 0.490			

* Indicates that data were used for the major intersecting roadway only.

$$E_{x_{ijk}}^{i,j} = \frac{\partial \lambda_{ij}}{\partial x_{ijk}} \times \frac{x_{ijk}}{\lambda_{ij}} \tag{8}$$

Eq. (8) reduces to the following expressions for the log–log (Eq. (9)) and log-linear (Eq. (10)) functional forms, respectively (the two types of functional forms between the expected crash frequency and the ADT and percent heavy vehicle variables):

$$E_{x_{ijk}}^{i,j} = \beta_k \tag{9}$$

$$E_{x_{ijk}}^{i,j} = \beta_k x_{ijk} \tag{10}$$

Elasticity for indicator variables, termed *pseudo-elasticity* by Lee and Mannering (2002), is the percent change in expected crash frequency given a change in the value of the indicator variable from zero to unity. The elasticity of the expected crash frequency for indicator variable ‘k’ at intersection ‘i’ during time period ‘j’ is defined as:

$$E_{x_{ijk}}^{i,j} = \exp(\beta_k) - 1 \tag{11}$$

The elasticities computed based on the daytime and nighttime crash frequency models shown in Table 2 are provided in Table 3.

The elasticities for major and minor road traffic volumes are similar across the nighttime and daytime crash frequency models. In both cases, however, the nighttime elasticity is lower than the daytime elasticity, and the minor road elasticity is lower than the major road elasticity. A 1% increase in major road average daily traffic is associated with an approximate 0.6% increase in crash frequency. A 1% increase in minor road traffic volume is associated with an increase in the daytime and nighttime crash frequency of less than 0.2%. This finding indicates that the major road traffic volume has a greater effect on intersection crashes than the minor road traffic volume. A 1% increase in the heavy vehicle percentage on the major road is associated with a decrease in the expected daytime and nighttime crash frequency.

Table 2
Negative binomial crash frequency models for Minnesota intersections.

Variable	Daytime model			Nighttime model		
	Parameter estimate	Standard error	t-Stat	Parameter estimate	Standard error	t-Stat
Constant	−6.5373	0.151	−43.37	−6.8986	0.206	−33.53
Log major road average daily traffic	0.6011	0.015	38.88	0.5737	0.021	26.85
Percent heavy vehicles on major road	−0.0092	0.003	−3.30	−0.0168	0.004	−4.37
Log minor road average daily traffic	0.1603	0.007	21.79	0.1262	0.010	12.29
Area type indicator (1 = urban/suburban; 0 = rural)	−0.0992	0.029	−3.39	−0.4212	0.040	−10.52
Traffic control indicator 1 = signal; 0 = stop-control)	0.6445	0.031	21.07	0.7120	0.041	17.20
Lighting indicator (1 = present; 0 = not present)	0.0477	0.031	1.56	−0.0791	0.042	−1.86
Intersection type indicator (1 = skew; 0 = cross or tee)	0.4862	0.031	15.86	0.4845	0.040	12.17
Speed indicator* (1 = 50 mph or greater; 0 otherwise)	−0.1601	0.022	−7.36	−0.1131	0.029	−3.88
No access control indicator* (1 = no access; 0 = partial access control)	−0.0416	0.038	−1.09	−0.0148	0.047	−0.31
Depressed median indicator* (1 = depressed median; 0 = barrier or no median)	0.0851	0.034	2.53	0.1712	0.043	3.99
Paved left-shoulder indicator* (1 = paved shoulder; 0 = unpaved or no shoulder)	−0.1163	0.042	−2.75	−0.2907	0.053	−5.44
Paved right-shoulder indicator* (1 = paved shoulder; 0 = unpaved or no shoulder)	0.0798	0.041	1.93	0.2153	0.052	4.11
Dispersion parameter (α)	0.9487	0.024	40.31	0.8948	0.041	21.99
	Number of observations = 22,058			Number of observations = 22,058		
	LL (constant only) = −28,954.27			LL (constant only) = −17,842.15		
	LL (full model) = −27,084.31			LL (full model) = −15,165.85		
	Pseudo R^2 = 0.1355			Pseudo R^2 = 0.1259		

* Indicates that data were used for major intersecting roadway only.

Table 3
Elasticities for daytime and nighttime crash frequencies.

Variable	Elasticity* or pseudo-elasticity (%)	
	Daytime crashes	Nighttime crashes
Log major road average daily traffic	0.601*	0.574*
Percent heavy vehicles on major road	−0.082*	−0.149*
Log minor road average daily traffic	0.160*	0.126*
Area type indicator (1 = urban/suburban; 0 = rural)	−9.4	−34.4
Traffic control indicator (1 = signal; 0 = stop-control)	90.5	103.8
Lighting indicator (1 = present; 0 = not present)	4.9	−7.6
Intersection type indicator (1 = skew; 0 = cross or tee)	62.6	62.3
Speed indicator (1 = 50 mph or greater; 0 otherwise)	−14.8	−10.7
No access control indicator (1 = no access; 0 = partial access control)	−4.1	−1.5
Depressed median indicator (1 = depressed median; 0 = barrier or no median)	8.9	18.7
Paved left-shoulder indicator (1 = paved shoulder; 0 = unpaved or no shoulder)	−11.0	−25.2
Paved right-shoulder indicator (1 = paved shoulder; 0 = unpaved or no shoulder)	8.3	24.0

For the indicator variables shown in Table 3, urban/suburban locations are associated with approximately 9.4% fewer expected daytime crashes when compared to rural locations, all else being equal. The result may be due to lower speeds and lower crash

severities in urban/suburban areas, leading to a lower likelihood that a crash is reported. At night, the elasticity for the urban/suburban indicator is −34.4%. Urban/suburban areas are more likely to have more traffic and more ambient light than rural locations, thereby improving driver attentiveness and visibility, which results in fewer nighttime crashes when compared to rural locations. The expected crash frequency at intersections with a traffic signal was more than 90% higher than for stop-controlled intersections. The result is difficult to interpret without analyzing more disaggregated crash breakdowns by crash type or without traffic signal phasing data.

As expected, the lighting elasticity changes sign when comparing the expected daytime and nighttime crash frequencies. This suggests that the presence of lighting is associated with a near 5% increase in daytime crashes, but is associated with a near 8% reduction in nighttime crashes. As expected, a skewed intersection is associated with approximately 62% higher daytime and nighttime crash frequencies when compared to intersections meeting at right-angles. Higher posted speed limits are associated with fewer daytime and nighttime crashes when compared to lower posted speed limits on the major road. This may be the result of higher geometric design standards on high-speed roads, resulting in improved driver expectancy and speed consistency.

No access control on the major road is associated with a nominal decrease in daytime and nighttime crashes when compared to major roads with partial access control. A depressed median is associated with an increase in daytime and nighttime crash frequency when compared to a barrier or no median on the major road. When a depressed median is present on an intersection approach, drivers may find it difficult to determine turning opportunities resulting in erratic maneuvers that contribute to an increase

in crash occurrence. The paved left- and right-shoulder elasticities indicate opposite effects. A paved left-shoulder on the major road is associated with fewer daytime and nighttime crashes when compared to unpaved or no paved shoulders. The opposite is true of paved right-shoulders.

5. Conclusions and recommendations

This paper described a proposed framework to estimate the safety effects of fixed lighting at a variety of intersection types and locations based on available data sources. The framework represents a new, comprehensive way to assess the safety benefits of intersection lighting, a well documented topic with studies dating before a number of methodological advancements made during the last 15–20 years in the highway safety analysis field. In addition, the framework includes analysis of both day and night crash frequencies to promote cost-effective comparisons to other safety countermeasures that do not have a specific time-of-day effect. A number of data collection and modeling issues were discussed. A method to test for transferability of the proposed models across states was outlined.

Minnesota intersections are used to illustrate execution of one component of the data and modeling framework. Findings indicated that intersections with fixed roadway lighting have fewer expected nighttime crashes and more expected daytime crashes than intersections without lighting when accounting for traffic volume and other safety-related intersection design features. The differences for night crash frequency (–7.6%) and night-to-day crash ratio (–12%) are in the same direction, but much smaller than those in older, published literature which tend to be approximately 30%. The difference in total crashes, estimated using the combined results of the Minnesota day and night models is a 3% reduction in crashes at intersections with lighting when compared to those without lighting, very close to the 4% reduction proposed by Harwood et al. (2007). When using observed Minnesota crash counts alone, without controlling for other safety influencing features, the reduction in the night–day ratio is 28%, which is similar to past research. This suggests that safety effects of intersection lighting based on direct comparisons of reported crash counts alone may be overestimated, a possible result of omitted variable bias. Ongoing and future research will apply the proposed framework to the entire analysis taxonomies to determine specific safety effects of lighting by intersection type. Model estimation issues presented here, including the panel data structure, selection bias, and model transferability will be addressed in greater detail.

In addition to the aforementioned modeling issues, several data-related future research needs were identified through the proposed framework. First, most current roadway inventory and crash record databases that contain lighting data only indicate its presence. As with many observational cross-sectional studies in traffic safety, electronic coding of data can present methodological challenges (e.g., data quality, underreporting, etc.). A complete roadway lighting management system, which is linkable to electronic roadway inventory and crash records, would be a next logical step to develop a complete understanding of safety effects of fixed roadway illumination systems. The following are sample lighting data elements that may be included in this inventory: installation date, fixture type, mounting height, spacing, uniformity, and luminance. When appended to other electronic databases, such a system would permit consideration or comparison of alternative analysis methods, such as observational before–after, cross-sectional, epidemiological, or causal inference. Second, a complete safety assessment of roadway lighting should be based not only crash frequency, but also on crash severity. Combining estimates of frequency with severity outcomes will afford designers with the opportunity to compare the cost-effectiveness of a

contemplated fixed illumination system to a similar scenario without lighting. Future research should therefore consider estimating models of severity given that a crash has occurred at locations with and without lighting.

References

- Antonucci, N.D., et al., 2004. Guidance for the implementation of the AASHTO strategic highway safety plan: a guide for reducing collisions at signalized intersections, vol. 12. NCHRP Report 500, Transportation Research Board, Washington, DC.
- Bauer, K.M., Harwood, D.W., 1996. Statistical models of at-grade intersection accidents. Report No. FHWA-RD-96-125, Federal Highway Administration, McLean, VA.
- Carson, J., Mannering, F., 2001. The effect of ice warning signs on accident frequencies and severities. *Accident Analysis and Prevention* 33 (1), 99–109.
- Chin, H.C., Quddus, M.A., 2003. Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. *Accident Analysis and Prevention* 35 (2), 253–259.
- Commission Internationale de l'Éclairage, 1992. Road Lighting as an Accident Countermeasure, CIE No. 93. Vienna, Austria.
- Council, F.M., Mohamedshah, Y., 2007a. The Highway Safety Information System guidebook for the California state data files. Report No. FHWA-RD-00-137, Federal Highway Administration, Washington, DC.
- Council, F.M., Mohamedshah, Y., 2007b. The Highway Safety Information System guidebook for the Minnesota state data files. Report No. FHWA-RD-01-058, Federal Highway Administration, Washington, DC.
- Elvik, R., 1995. Meta-analysis of evaluations of public lighting as accident countermeasure. *Transportation Research Record* 1485, 112–123.
- Elvik, R., Vaa, T., 2004. *The Handbook of Road Safety Measures*. Elsevier Science.
- Green, E.R., Agent, K.R., Barrett, M.L., Pigman, J.G., 2003. Roadway lighting and driver safety. Report No. KTC-03-12/SPR247-02-1F. University of Kentucky, Lexington, KY.
- Greene, W.H., 2003. *Econometric Analysis*, fifth ed. Prentice Hall, Upper Saddle River, NJ.
- Guo, G., 1996. Negative multinomial regression model for clustered event counts. *Sociological Methodology* 26, 113–132.
- Harwood, D.W., Bauer, K.M., Richard, K.R., Gilmore, D.K., Graham, J.L., Potts, I.B., Torbic, D.J., Hauer, E., 2007. Methodology to predict the safety performance of urban and suburban arterials. NCHRP Web-only Document 129, TRB, Washington, DC.
- Hauer, E., 2005. Cause and effect in observational cross-sectional studies in road safety. In: Draft, Proceedings of Transportation Research Board Annual Meeting, Washington, DC.
- Hauer, E., 1997. *Observational before–after Studies in Road Safety*. Pergamon, Elsevier Science Ltd, Oxford, UK.
- Isebrands, H., Hallmark, S., Hans, Z., McDonald, T., Preston, H., Storm, R., 2004. Safety impacts of street lighting at isolated rural intersections: Part II, Year 1 report. Center for Transportation Research and Education, Iowa State University, Ames, IA.
- Isebrands, H., Hallmark, S., Hans, Z., McDonald, T., Preston, H., Storm, R., 2006. Safety impacts of street lighting at isolated rural intersections: Part II, Final report. Report No. MN/RC-2006-35, Minnesota Department of Transportation, St. Paul, MN.
- Kim, D.G., Washington, S., 2006. The significance of endogeneity problems in crash models: an examination of left-turn lanes in intersection crash models. *Accident Analysis and Prevention* 38 (6), 1094–1100.
- Lee, J., Mannering, F., 2002. Impact of roadside features on the frequency and severity of run-off-road accidents: an empirical analysis. *Accident Analysis and Prevention* 34 (2), 149–161.
- Liang, K.Y., Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73, 13–22.
- Lipinski, M.E., Wortman, R.H., 1978. Effect of illumination on rural at-grade intersection accidents. *Transportation Research Record* 611, 25–27.
- Lord, D., Persaud, B.N., 2000. Accident prediction models with and without trend: application of the generalized estimating equations (GEE) procedure. *Transportation Research Record: Journal of the Transportation Research Board* 1717, 102–108.
- Minnesota Department of Transportation, 2001. *Traffic Safety Fundamentals Handbook*. Office of Traffic Safety, St. Paul, MN.
- Monsere, C.M., Fischer, R.L., 2008. Safety effects of reducing freeway illumination for energy conservation. *Accident Analysis and Prevention* 40 (4), 1773–1780.
- Neuman, T.R., et al., 2003. Guidance for the implementation of the AASHTO strategic highway safety plan, vol. 5: a guide for addressing unsignalized intersection collisions. NCHRP Report 500, Transportation Research Board, Washington, DC.
- Persaud, B., Lyon, C., 2007. Empirical Bayes before–after studies: lessons learned from two decades of experience and future directions. *Accident Analysis and Prevention* 39 (3), 546–555.
- Poch, M., Mannering, F., 1996. Negative binomial analysis of intersection accident frequencies. *Journal of Transportation Engineering* 122 (2), 105–113 (ASCE).
- Preston, H., Schoenecker, T., 1999. Safety impacts of street lighting at rural intersections. Report No. 1999-17. Minnesota Department of Transportation, St. Paul, MN.
- Schwab, R.N., Walton, N.E., Mounce, J.M., Rosenbaum, M.J., 1982. Synthesis of safety research related to traffic control and roadway elements: highway lighting, vol. 2. Report No. FHWA-TS-82-233. Federal Highway Administration, Washington, DC (Chapter 12).

- Shankar, V.N., Albin, R.B., Milton, J.C., Mannering, F.L., 1997. Evaluating median cross-over likelihoods with clustered accident counts: an empirical inquiry using the random effects negative binomial model. *Transportation Research Record* 1635, 44–48.
- Ulfarsson, G.F., Shankar, V.N., 2003. Accident count model based on multiyear cross-sectional roadway data with serial correlation. *Transportation Research Record* 1840, 193–197.
- Walker, F.W., Roberts, S.E., 1976. Influence of lighting on accident frequency at highway intersections. *Transportation Research Record* 562, 73–78.
- Wang, X., Abdel-Aty, M., Brady, P.A., 2006. Crash estimation at signalized intersections: significant factors and temporal effect. *Transportation Research Record: Journal of the Transportation Research Board* 1953, 10–20.
- Washington, S., Karlaftis, M., Mannering, F., 2003. *Statistical and Econometric Methods for Transportation Data Analysis*. Chapman and Hall/CRC, Boca Raton, FL.
- Washington, S., Persaud, B., Lyon, C., Oh, J., 2005. Validation of accident models for intersections. Report No. FHWA-RD-03-037, Federal Highway Administration, McLean, VA.
- Windmeijer, F.A.G., Santos Silva, J.M.C., 1997. Endogeneity in count data models: an application to demand for health care. *Journal of Applied Econometrics* 12, 281–294.
- Zeger, S.L., Liang, K.Y., 1986. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 42, 121–130.