Accident Analysis and Prevention 82 (2015) 180–191

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Estimating the safety effects of lane widths on urban streets in Nebraska using the propensity scores-potential outcomes framework

Jonathan S. Wood*, Jeffrey P. Gooch, Eric T. Donnell

Department of Civil and Environmental Engineering, The Pennsylvania State University, 212 Sackett Building, University Park, PA 16802, USA

ARTICLE INFO

Article history:
Received 25 February 2015
Received in revised form 14 April 2015
Accepted 5 June 2015
Available online 16 June 2015

Keywords:
Lane width
Traffic safety
Urban streets
Propensity scores
Mixed effects Poisson and negative binomial regression
Generalized boosting
Highlights

ABSTRACT

A sufficient understanding of the safety impact of lane widths in urban areas is necessary to produce geometric designs that optimize safety performance for all users. The overarching trend found in the research literature is that as lane widths narrow, crash frequency increases. However, this trend is inconsistent and is the result of multiple cross-sectional studies that have issues related to lack of control for potential confounding variables, unobserved heterogeneity or omitted variable bias, or endogeneity among independent variables, among others. Using ten years of mid-block crash data on urban arterials and collectors from four cities in Nebraska, crash modification factors (CMFs) were estimated for various lane widths and crash types. These CMFs were developed using the propensity scores-potential outcomes methodology. This method reduces many of the issues associated with cross-sectional regression models when estimating the safety effects of infrastructure-related design features. Generalized boosting, a non-parametric modeling technique, was used to estimate the propensity scores. Matching was performed using both Nearest Neighbor and Mahalanobis matching techniques. CMF estimation was done using mixed-effects negative binomial or Poisson regression with the matched data. Lane widths included in the analysis included 9 ft, 10 ft, 11 ft, and 12 ft. Some of the estimated CMFs were point estimates while others were functions of traffic volume (i.e., the CMF changed depending on the traffic volume). Roadways with 10 ft travel lanes were found to experience the highest crash frequency relative to other lane widths. Meanwhile, roads with 9 ft travel lanes were found to experience the lowest relative crash frequency. While this may be due to increased driver caution when traveling on narrow lanes, it is possible that unobserved factors influenced this result. CMFs for target crash types (sideswipe same-direction and sideswipe opposite-direction) were consistent with the values currently used in the Highway Safety Manual (HSM).

1. Introduction

Lane width selection in the urban environment is challenging. While wider lanes may improve traffic flow, the availability of right-of-way may be limited due to high-density development, parking, allocation of bike lanes, etc. With the many factors that influence lane width decisions, designers have to consider the trade-offs associated with the optimal lane width in a given context. To assist in this decision making, the safety performance of lane width dimensions in urban areas should be quantified.

The existing literature on the safety effects of lane widths in urban areas provides highly variable findings. While the methodological approaches used to assess lane widths in the existing literature vary, most are focused on the application of cross-sectional statistical (regression) models. The objectives of this study are to overcome limitations of the cross-sectional regression method in estimating the safety effects of lane width in urban areas by using the propensity scores-potential outcomes framework. Data from four cities in Nebraska are used to develop lane width CMFs. The application of the propensity scores-potential outcomes framework has the following advantages over cross-sectional regression models:

- Selection bias is reduced via matching (Angrist and Pischke, 2014; Rosenbaum, 2002);
- The issue of common support is considered and accounted for (the effect of the treatment should only be estimated within the region of common support) (Guo and Fraser, 2010). Common support is defined as the range of values for each variable that are shared between the treated and comparison groups. When this is

* Corresponding author. Tel.: +1 435 760 2781.
E-mail addresses: jsw277@psu.edu (J.S. Wood), jpg5152@psu.edu (J.P. Gooch), edonnell@engr.psu.edu (E.T. Donnell).

http://dx.doi.org/10.1016/j.aap.2015.06.002

not done, the estimates are based on extrapolation and are often biased (Sasidharan and Donnell, 2013); and
- The results are robust to specification errors in the regression models (Kang and Schafer, 2007).

The paper is organized into five subsequent sections. The first describes previous literature covering the relationship between safety and lane width in urban areas, as well as propensity score applications in traffic safety. Based on this review, the methodology used for the development of the CMFs is subsequently described. Descriptive statistics of the data are provided in the third section of the paper. The matching and modeling techniques used to identify comparable treatment and comparison sites are then discussed, followed by the analysis results. Conclusions and recommendations for future research are outlined in the last section of the paper.

2. Literature review

2.1. Safety and lane width

Numerous studies have attempted to quantify the relationship between lane widths and crash frequency on urban roadways. One study estimated CMFs for 10-, 11- and 13-foot lane widths, relative to a 12-foot lane width (baseline condition), using three years of crash data (2007–2009) from Utah (Le and Porter, 2013). There were 1577 road segments, totaling approximately 380 miles, included in the dataset. The road segments were from Illinois and were collected using a combination of Highway Safety Information System (HSIS) files and Google Earth. Negative binomial regression was used to estimate the CMFs. A CMF of 0.800 and 0.868 were developed for 11- and 13-foot travel lanes, relative to the baseline condition of 12-foot travel lanes. The results, however, were not statistically significant. Ten foot lanes had higher crash frequencies than 12-foot lanes and the results were statistically significant at the 95 percent confidence level (CMF of 2.956).

A similar analysis was performed using data from Florida (Abdel-aty and Radwan, 2000). Three years of crash data (1992–1994) for 566 segments of highway (urban, suburban, and rural, for a total of 1606 crashes) were used in the analysis. Negative binomial regression was used to estimate the safety effects of geometric features, including lane width. Lane width was included as a continuous variable in the statistical model. The findings indicated that segments with wider lanes experienced lower crash frequencies than segments with narrower lanes. The estimated effect was $\exp(-0.364 \times \text{Lane Width})$ for total crashes and was statistically significant at the 95 percent confidence level. The range of lane widths included in the analysis was not provided by the authors.

Other researchers used HSIS data from Illinois on urban, suburban, and rural roads to assess the association between infrastructure-related characteristics and crash frequency (Noland and Oh, 2004). Approximately 12.5 percent of all roads in Illinois was included in the analysis. Crash data from 1987 to 1990 were used. Fixed-effects negative binomial regression was used to analyze the data. The log of lane width was included in the model as a continuous variable. The findings indicated that crash frequency increases as lane width increases, but the result was not statistically significant. The range of lane widths included in the analysis was not provided by the authors.

In an analysis of 7500 miles of roadway data from Oregon, the effects of lane width on urban roadways were estimated (Strathman et al., 2001). Two years of crash data were used (1997–1998). Negative binomial regression was used for the analysis. Lane width was included in the models as a continuous variable. The findings indicated that on non-freeway urban roads, total crash frequency increased as lane width increased, but the findings were not statistically significant. For urban freeways, the findings indicated that wider lanes were associated with higher total crash frequency, and the results were statistically significant at the 99.9 percent confidence level. The range of lane widths included in the analysis was not provided by the authors.

The urban section of the American Association of State Highway and Transportation Officials' (AASHTO) Highway Safety Manual states that CMFs for lane and shoulder width in urban areas is a function of traffic volume, but does not provide any more information about lane width CMFs for urban streets. Although not explicit, the urban street lane width-safety discussion implies that the lane width CMF in other section of the HSM can be used to assess the safety effects of lane width decisions on urban streets (AASHTO, 2010). The lane width CMF contained in the rural two-lane highways section of the HSM varies based on traffic volume and applies to single-vehicle run-off-road (SVROR) and multi-vehicle same and opposite direction sideswipe crashes. There is a procedure to convert these specific crash types into a CMF for total crashes, which considers the proportion of SVROR and multi-vehicle sideswipe crashes among total crashes. Based on the proportion of these crash types in the dataset used for this paper (0.214), the HSM CMFs for total crashes with a baseline width of 12 ft are 1.01 for 11 ft, 1.05 for 10 ft, and 1.08 for 9 ft lane widths.

Other studies were found that considered the safety effects of lane widths, but did not explicitly consider urban areas (Hauer et al., 2004) or did not provide the modeling results (Hadi et al., 1995; Qin et al., 2005). Another study used a full Bayes before–after methodology to estimate the safety impacts of increasing lane width on an urban highway in New Jersey, but the study was limited to a single treated road, 16 reference road segments, and the initial lane width, final lane width, and change in lane width for the treated highway were not provided (Yanmaz-Tuzel and Ozbay, 2010). The results indicated that crashes were reduced by 28.1 percent due to the lane width increase.

The findings of these studies indicate conflicting safety relationships between lane width and safety on urban roadways. This is likely the result of many confounding issues present within the evaluation methodology, such as study design limitations (cross-sectional models), specifying the lane width as a continuous variable in cross-sectional statistical models, and small sample size.

2.2. Methods used to estimate crash modification factors

One focus of highway safety research is the quantification of safety for crash prediction. For roadway characteristics (such as lane width), this is done by developing CMFs. Ideally, CMFs would be determined using randomized experiments. However, factors such as ethics and cost (monetary and/or societal) prevent highway safety researchers from performing randomized experiments. As a result, other study designs using observational data have been utilized, attempting to isolate the safety effect for single variables (such as lane width). When developing CMFs, the preferred method in today's practice is the empirical Bayes (EB) before–after observational study, which combines observed and predicted crash data to reduce regression to the mean bias (Hauer, 1997). While this is the preferred analysis method, the safety effects of infrastructure characteristics, such as lane width or horizontal curve parameters (e.g., radius or superelevation), often cannot be analyzed by this process since they rarely undergo changes without significant reconstruction to the infrastructure. As a result, observational before–after methods are not feasible, thus cross-sectional methods must be employed.

As discussed previously, the propensity scores-potential outcomes framework has several advantages over cross-sectional