Accident Analysis and Prevention 59 (2013) 337–345

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Do lower income areas have more pedestrian casualties?

Robert B. Noland*, Nicholas J. Klein, Nicholas K. Tulach

Alan M. Voorhees Transportation Center, Edward J. Bloustein School of Planning and Public Policy, Rutgers University, New Brunswick, NJ 08901, USA

ARTICLE INFO

Article history:
Received 6 February 2010
Received in revised form 31 May 2013
Accepted 5 June 2013

Keywords:
Road safety
Pedestrians
Negative binomial models
Bayesian analysis
Deprivation
Low income
Car ownership

ABSTRACT

Pedestrian and motor vehicle casualties are analyzed for the State of New Jersey with the objective of determining how the income of an area may be associated with casualties. We develop a maximum-likelihood negative binomial model to examine how various spatially defined variables, including road, income, and vehicle ownership, may be associated with casualties using census block-group level data. Due to suspected spatial correlation in the data we also employ a conditional autoregressive Bayesian model using Markov Chain Monte Carlo simulation, implemented with Crimestat software. Results suggest that spatial correlation is an issue as some variables are not statistically significant in the spatial model. We find that both pedestrian and motor vehicle casualties are greater in lower income block groups. Both are also associated with less household vehicle ownership, which is not surprising for pedestrian casualties, but is a surprising result for motor vehicle casualties. Controls for various road categories provide expected relationships. Individual level data is further examined to determine relationships between the location of a crash victim and their residence zip code, and this largely confirms a residual effect associated with both lower income individuals and lower income areas.

1. Introduction

Pedestrian fatalities and injuries in New Jersey constitute a large fraction of total vehicle crashes and disproportionately occur in lower income communities. About 20% of crash fatalities in New Jersey are pedestrians resulting in roughly 150 pedestrian fatalities and roughly ten times as many pedestrian injuries each year. About 29% of pedestrian fatalities and 41% of pedestrian injuries occur in the lowest income quartile Census block groups. By comparison, about 16% of motor vehicle fatalities and 20% of motor vehicle injuries occur in these block groups. Notably, while the proportion of crashes that involve a pedestrian are higher than in other states, the total casualty rate in New Jersey is one of the lowest of any state. One probable reason for this disparity is that New Jersey is more urbanized than other states. An unknown question is why pedestrian casualties are more likely in lower income neighborhoods; the analysis presented here seeks to examine why this is the case.

Recent research has established that spatial analysis techniques can help to explain associations between area-based factors and road crashes. This includes a spatial analysis of crashes in Honolulu, Hawaii, and an analysis for all of England (Levine et al., 1995; Noland and Quddus, 2004). These and other analyses find that total fatalities or injuries are associated with land use characteristics, road types, and area-based demographic factors. Commercial land uses frequently have more pedestrian-related crashes (Kim et al., 2006; Lightstone et al., 2001). Larger roads are associated with more crashes, possibly representing larger traffic flows. More deprived areas also tend to have more crashes, and in particular, those that injure pedestrians (Graham and Glaister, 2003; Loukaitou-Sideris et al., 2007).

One issue with area-based analysis of crash data is that there may be spatial correlation. In general, we would expect that a given spatial unit would be affected by characteristics of neighboring units. Omission of spatial correlation, if it exists, may lead to biased estimates. Because crash data is non-normally distributed (i.e., zero counts in some units), count data models are typically used and accounting for spatial correlation in these models requires a more complex estimation approach. We use the Crimestat v4.0 (Levine, 2010) software package to estimate these models.[2]

Our primary objective in this analysis is to examine the spatial factors associated with pedestrian casualties with a focus on understanding why lower income areas tend to suffer more crashes. As a comparative analysis we also examine motor-vehicle only casualties to determine whether there are distinct differences in associations. This is an ecological analysis of area-based factors and their association with casualties; we make no judgment on how these factors affect individual crashes. Disaggregate data is

* Corresponding author. Tel.: +1 848 932 2859.
E-mail address: rnoland@rutgers.edu (R.B. Noland).

[2] Crimestat is freely available at http://www.icpsr.umich.edu/CrimeStat/download.html. Version 4 is currently in beta testing and was kindly provided by Ned Levine.

0001-4575/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aap.2013.06.009

also analyzed based on the zip code of the victim's residence, which provides a proxy for the victim's socio-economic status. Various cross-tabulations of crash location with residence location are examined to disentangle whether low income areas or low income individuals are more likely to be associated with more pedestrian casualties.

2. Data

For this analysis, we combined data from four primary sources, Plan4Safety[3] crash data, 2000 US Census data, the Longitudinal Employer-Household Dynamics Data for 2008, and GIS layers obtained from the state of New Jersey.

We include pedestrian and motor vehicle crash data from 2003 to 2007 extracted from the Plan4Safety database which is a comprehensive database of crashes that have taken place in New Jersey, based on police reports. We extracted records for casualties to pedestrian and motor vehicle occupants (driver and passenger) where the person involved was "killed", "incapacitated", or sustained a "moderate injury".[4] We also included records that listed pedestrians with "complaint of pain", or "null" values where there was another pedestrian involved that had a "moderate injury" or worse. The resulting crash database includes 8144 pedestrian casualties and 82,477 motor vehicle casualties. There may be some recorded injuries that were very minor or non-existent and would represent poor reporting by the police, especially for crashes with multiple pedestrians, thus for completeness we include all recorded pedestrians as injured. After extracting the data from the Plan4Safety database, we geo-coded the records using geolocation information from the Plan4Safety database and aggregated the counts of casualties to census block groups. The data contained the residence zip code of the victims, and we also geo-coded the crashes to the zip code area in which the crash occurred.

We obtained basic demographic data for block groups in New Jersey from the 2000 US Census Summary File 3. We excluded any block groups with zero resident population (43 in total) which reduces the total pedestrian casualties by 27 (0.3%) to 8117 casualties and motor vehicle casualties by 692 (0.8%) to 81,785 casualties.[5] The final database includes 6460 block groups. Employment data at place of work was derived from the Census Longitudinal Employer-Household Dynamics data for 2008.[6] This is distinct from other employment measures that measure the employed population within a given spatial unit. Measuring employment where people actually work is important as these are often areas where there is more pedestrian activity.

We obtained GIS road layers from the New Jersey Department of Transportation (NJDOT). For each block group, we calculated road density per square mile by functional class. We included the following functional classes: freeways and turnpikes, US highways (generally the largest arterial roads), New Jersey state highways (also relatively large arterials), two categories of county roads (500 and 600 designations, with 500 being higher category roads), and ramps and jughandles.

Jughandles appear to be a road engineering innovation originally unique to New Jersey. These are designed to avoid cross-traffic turns (left turns) and involve an off-ramp that loops around so that vehicles cross the main arterial flow at a right angle with a traffic signal. The rationale is to minimize vehicle conflicts by reducing cross-traffic turning crashes, while also improving operations via the elimination of dedicated cross-traffic (left) turn signals (Jagannathan, 2006).[7]

3. Hypotheses

Previous spatial analyses have generally found an association between the income level of an area and both pedestrian and vehicle casualties (Graham et al., 2005; Graham and Glaister, 2003; LaScala et al., 2000; Loukaitou-Sideris et al., 2007; Noland and Quddus, 2004). Thus we can hypothesize that lower income areas will be associated with more pedestrian casualties. One of our objectives is to understand why this is the case, therefore we also hypothesize that lower rates of vehicle ownership are associated with more pedestrian casualties, and that this will fully capture the effect of area-based income. Furthermore, we hypothesize that the effect of vehicle ownwership and income will have different associations with motor-vehicle passenger casualties.

Additional evaluated hypotheses include the impact of population and employment density, both proxies for land use and urban form. Previous research has suggested that population density is associated with reductions in pedestrian casualties while employment density is associated with increases (Ewing and Dumbaugh, 2009; Noland and Quddus, 2004). Road network density variables are also examined and it is hypothesized that more roads with higher speed traffic (i.e., those of a higher functional classification) will be associated with more pedestrian casualties.

The dataset also includes the zip code residence of many of the victims. Using this information, we compare the income level of each victim's residential neighborhood with the income level of the area where the crash occurred. We also include how many victims were in a crash in their home zip code area. Our hypothesis is that area-based income is associated with more casualties and that individual income is not. Individual income can vary substantially within a zip code and this is a limitation of this analysis. However, if one accepts area-based income as a proxy for individual income then this hypothesis can be tested and determine whether lower income areas are inherently more risky or whether lower income individuals take more risks.

4. Modeling approach and estimation methods

Our analysis uses a negative binomial model because crash data, and in particular pedestrian casualties, are rare events that are typically Poisson distributed. Estimation methods that assume normality cannot be used since crash outcomes are non-negative discrete counts with some dependent variables being equal to zero. In our dataset, 48% of the block groups have zero pedestrian casualties and 10% have zero counts of motor vehicle casualties. We tested our dependent variables and residuals and could not reject the hypothesis that these were not normally distributed. In addition, further tests on per capita rates were found to be non-normal necessitating use of a count estimation method.

Due to over-dispersion in the data, negative binomial models are used in place of Poisson regressions. The Poisson model assumes that the mean is equal to the standard deviation; this does not typically hold in empirical settings. Tests of over-dispersion for our models found that negative binomial regression was required.

[3] The Plan4Safety database is maintained by Rutgers University Center for Advanced Infrastructure and Transportation. More detailed information on the Plan4Safety database is available online at http://plan4safety.rutgers.edu/plan4safety/login.aspx.

[4] The data coded bicyclists as pedestrians within the pedestrian tab. These were removed from the dataset.

[5] Population was used as an off-set variable in our models. This represents an exposure measure and the zero population block-groups were elimated to avoid estimation problems.

[6] Available at http://lehd.ces.census.gov/.

[7] Further information on jughandle design is available at http://www.state.nj.us/transportation/eng/documents/RDM/sec6.shtm (accessed 01.08.12).