Accident Analysis and Prevention 91 (2016) 175–182

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Safety-in-numbers: Estimates based on a sample of pedestrian crossings in Norway

Rune Elvik

Institute of Transport Economics, Gaustadalleen 21, NO-0349 Oslo, Norway

ARTICLE INFO

Article history:
Received 29 December 2015
Received in revised form 14 February 2016
Accepted 9 March 2016
Available online 16 March 2016

Keywords:
Safety-in-numbers
Cyclists
Pedestrians
Pedestrian crossings
Accident prediction model

ABSTRACT

Safety-in-numbers denotes the tendency for the risk of accident for each road user to decline as the number of road users increases. Safety-in-numbers implies that a doubling of the number of road users will be associated with less than a doubling of the number of accidents. This paper investigates safety-in-numbers in 239 pedestrian crossings in Oslo and its suburbs. Accident prediction models were fitted by means of negative binomial regression. The models indicate a very strong safety-in-numbers effect. In the final model, the coefficients for traffic volume were 0.05 for motor vehicles, 0.07 for pedestrians and 0.12 for cyclists. The coefficient for motor vehicles implies that the number of accidents is almost independent of the number of motor vehicles. The safety-in-numbers effect found in this paper is stronger than reported in any other study dealing with safety-in-numbers. It should be noted that the model explained only 21% of the systematic variation in the number of accidents. It therefore cannot be ruled out that the results are influenced by omitted variable bias. Any such bias would, however, have to be very large to eliminate the safety-in-numbers effect.

1. Introduction

A transport policy objective in many countries is to curtail growth in the use of cars and promote walking or cycling. Important reasons for adopting this policy objective are to reduce global warming and improve public health. Pedestrians and cyclists have a higher injury rate per kilometre of travel than car occupants (Bjørnskau, 2015). If more people walk or cycle, one may expect the number of traffic injuries to increase. A counterargument is that the injury rate for pedestrians and cyclists is not constant, but subject to a "safety-in-numbers" effect, which means that the larger the number of pedestrians or cyclists, the lower the injury rate for each pedestrian or cyclist. If sufficiently strong, this protective effect may to a large extent counteract and perhaps eliminate the increase in the number of injuries that would otherwise be expected when there is more walking or cycling.

Unfortunately, there are many problems in estimating the safety-in-numbers effect and particularly in determining the causality of the effect. In the first place, data on the number of pedestrians or cyclists tend to come from short-term counts that may be associated with considerable uncertainty (Kröyer, 2015). In the second place, the reporting of accidents involving pedestrians or cyclists in official accident statistics is very low, in particular for cyclists (Lahrmann, 2015). In the third place, nearly all studies of safety-in-numbers rely on cross-sectional data, which make it difficult to establish causal relationships. In a recent review, Elvik and Bjørnskau (2016) concluded that no studies of safety-in-numbers have controlled adequately for all relevant confounding variables and that one cannot conclude that these studies have uncovered a causal relationship.

A fourth problem is that there are many ways of developing and fitting the accident prediction models by means of which the safety-in-numbers effect is estimated (Hauer, 2015) (note: the word "effect" is used as shorthand only and does not necessarily imply a claim of causality). Results may vary depending on, for example, which variables a model includes and how the statistical relationship between these variables and the number of accidents is modelled.

While it is difficult to establish causal relationships in a single study, replication of studies may, as the number of studies grows, reveal consistent patterns that may at least suggest causality. Consistency in the relationship between a cause and its effect (same cause, same effect) is one of the oldest criteria of causality. If a safety-in-numbers effect has been reproduced consistently in a range of countries and during a long period of time, that at least shows that it reflects a general tendency, which is robust with respect to the many differences between the individual studies.

E-mail address: re@toi.no

http://dx.doi.org/10.1016/j.aap.2016.03.005

The main objective of this paper is to estimate the safety-in-numbers effect in a sample of pedestrian crossings in Norway. Part of the sample was used in a previous study (Elvik et al., 2013), but it has now been enlarged as explained in the next section. The paper tries to implement the stepwise approach to regression modelling in road safety proposed by Hauer (2015).

2. Data and method

2.1. Sample of pedestrian crossings

A data set consisting of 389 marked pedestrian crossings in and close to Oslo has been created by merging four data sets. The four data sets are described in reports by Amundsen and Sætre (2009), Sætre et al. (2010), Sørensen et al. (2010) and Sørensen and Nævestad (2012). These pedestrian crossings were selected for detailed safety inspections for one or more of the following reasons:

1. Accident history: crossings with a history of accidents were selected.
2. Accident severity: crossings where accidents were severe, in particular where there had been fatal accidents, were selected.
3. Speed limit: crossings located on roads with a speed limit of 50 or 60 km/h were selected.
4. Complaints: crossings for which the public had made complaints were selected.

The pedestrian crossings are not representative of all pedestrian crossings in Oslo and its suburbs. In particular, the mean number of accidents per crossing is likely to be considerably higher than for a typical pedestrian crossing in Oslo and its suburbs. The variables recorded for each pedestrian crossing are listed in Table 1.

There are three groups of variables. The first group consists of dependent variables, i.e. variables whose values are influenced by the other groups of variables listed in Table 1. These variables include the count of injury accidents and the counts of injured road users according to injury severity.

The second group of variables describe traffic volume. Traffic volume is indicated both by means of estimates of the annual average daily number of cars, pedestrians or cyclists and by means of counts made when the pedestrian crossings were inspected. AADT is almost always estimated on the basis of short-term counts. As shown by Mensah and Hauer (1998), using an average value for traffic volume, rather than the volume prevailing at the time of each accident, may lead to bias when estimating the relationship between traffic volume and the number of accidents. AADT is, however, very often the only available data on traffic volume.

In most of the pedestrian crossings, traffic counts were made when the crossings were inspected. These counts were typically made during daytime on weekdays and for a period of six hours. Based on the counts, the number of pedestrians and cyclists crossing the road during the maximum hour was estimated. One could argue that these estimates might be more strongly related to the number of accidents than AADT, since most accidents involving motor vehicles and either pedestrians or cyclists happen in daytime when hourly traffic volume is higher than at night.

The third group of variables listed in Table 1 are various characteristics of the road layout and traffic control at the pedestrian crossings. This includes the number of directions from which vehicles may approach a pedestrian crossing (arms: an indicator of the number of traffic movements a pedestrian or cyclist must attend to when crossing the road), the number of lanes, the presence of a refuge, the presence of traffic signal control, speed limit and the 85th percentile speed of approaching motor vehicles.

2.2. Analytic choices

Hauer (2015) emphasizes the importance of making all analytic choices when developing an accident prediction model explicitly and stating the reasons for the choices that were made. Unless models are developed this way, one cannot know whether the final model was the best possible model, given the available data and the intended use of the model, or whether it was inferior. In this paper, the main analytic choices are:

1. Are the independent variables so highly correlated that there could be co-linearity problems in developing a model?
2. Which set of variables describing traffic volume is most closely related to the count of accidents?
3. Which of the other independent variables should be included in a model?

2.3. Correlations among variables

To help answer the first question, a correlation matrix (Pearson correlation coefficients) was estimated. It is shown in Table 2. Most of the correlations are minor or moderate. Only three correlations are quite strong. These are the correlations between the two

Table 1
Variables recorded for each pedestrian crossing.

Variable	Definition and explanation	Mean	Minimum	Maximum
Group 1: Dependent variables				
Accidents.	Count of police reported injury accidents during 5 years	1.296	0	11
Slightly injured road users	Count of slightly injured road users during 5 years	1.527	0	11
Seriously injured road users	Count of seriously injured road users during 5 years	0.082	0	3
Fatally injured road users	Count of fatally injured road users during 5 years	0.008	0	1
Group 2: Traffic volume variables				
Motor vehicle volume	Annual average daily number of motor vehicles (AADT)	8181	145	28200
Pedestrian volume	Estimated annual average daily number of pedestrians crossing the road	233	0	5000
Cyclist volume	Estimated annual average daily number of cyclists crossing the road	35	0	589
Count of cars	Count of cars made during daytime when data were collected about each pedestrian crossing	7339	33	25863
Pedestrians in maximum hour	Count in pedestrians in the hour with the largest number (short-term count)	62	0	1571
Cyclists in maximum hour	Count of cyclists in the hour with the maximum number (short-term count)	8	0	252
Group 3: Other independent variables				
Arms to be observed	The number of approaches (directions) from which vehicles enter a pedestrian crossing	2.666	2	4
Number of lanes	The number of lanes for vehicles	2.224	1	6
Presence of refuge	If there is a refuge for pedestrians or not (dichotomous variable; 1 if refuge, 0 otherwise)	0.550	0	1
Signal control	Presence of traffic signals (1 if present, 0 otherwise)	0.111	0	1
Speed limit	Speed limit in kilometres per hour (30, 40, 50 or 60)	52.237	30	60
85-percentile speed	The speed below which 85% of motor vehicles travel (km/h)	44.759	16	80