Accident Analysis and Prevention 42 (2010) 582–588

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Quasi-induced exposure: The choice of exposure metrics

Álvaro Gómez Méndez*, Francisco Aparicio Izquierdo

Automobile Research Institute (INSIA), Campus Sur de la UPM, Carretera de Valencia km. 7, 28031-Madrid, Spain

ARTICLE INFO

Article history:
Received 10 April 2009
Received in revised form 22 July 2009
Accepted 6 October 2009

Keywords:
Quasi-induced exposure
Accident risk

ABSTRACT

The quasi-induced exposure method is widely used to estimate exposure and risks of different groups of drivers and vehicles. Essentially, this method assumes that non-at-fault or passive parties in two-vehicle collisions represent a random sample of the populations on the road. Most previous works have used the whole sample of collisions to estimate exposure.

There has been some concern about possible biases in quasi-induced estimates. In this paper, we argue that (1) biases are mainly due to differences in accident avoidance abilities, speeds and injury risks, and (2) because the influence of these three factors on the probability of being non-at-fault is not the same for every crash type, differences may arise among non-at-fault populations, in which case some crash types would provide a more accurate estimate of exposure than others.

We explore the direction of biases due to speed, accident avoidance ability and injury risk in four accident types: accidents between vehicles travelling on different lanes in two-way, two-lane undivided roads; accidents between vehicles travelling on different lanes on multilane roads; intersection accidents; and accidents between vehicles travelling on the same lane. Our analysis shows that more research would be needed concerning the effect of speed on head-on crashes on undivided roads, and crashes on multilane roads.

1. Introduction

The lack of appropriate data on exposure is one of the greatest problems that road safety analyses have faced so far. It is widely accepted that the number of kilometers driven by a particular group of drivers/vehicles best represents exposure. However, the characteristics of the methods more frequently used to collect this indicator prevent its use in in-depth studies. This lack of data is common to most countries and spheres of study (Rumar et al., 1997). For this reason, indirect exposure measures, generally known as *induced exposure* measures, have been developed. The main feature of these estimates is that they are obtained directly from accident data bases, which, ideally, ensures complete homogeneity between accident and exposure data.

Most previous applications can be classified into two large variants. The first is the so-called *quasi-induced exposure*, based on identifying the at-fault and non-at-fault parties in two-vehicle crashes. Its fundamental hypothesis is that the samples of non-at-fault drivers/vehicles are representative of the populations on the road at the time of the accident. For a specific group, accident risk is given by the quotient between the number of at-fault and the number of non-at-fault drivers/vehicles. This method was first used at the beginning of the seventies and since then has seen a large number of applications. Some recent examples are: Cheuk Hing et al. (2003), Kirk and Stamatiadis (2001), Lyles et al. (1991), Rice et al. (2003), Rueda Domingo et al. (2004), Stamatiadis and Deacon (1995, 1997), Stamatiadis et al. (1999), Staplin and Lyles (1991), US GAO (1994), Yannis et al. (2005).

The second variant is based on identifying an accident type or types, called *non-relevant accidents*, where the attribute of interest has little or no influence. This variant has been particularly used in evaluating the impact of certain primary safety devices. For example, for studying the influence of anti-lock braking systems (ABS), Burton et al. (2004), Evans and Gerrish (1996) and Evans (1998), start out from the hypothesis that this system has little influence on the probability of being struck in the side of the car in side-impact collisions. The numbers of vehicles involved in this accident type therefore give the amounts of exposure of vehicles with and without ABS. In the case of the electronic stability programme (ESP), rear-end crash is usually taken as the non-relevant accident (Lie et al., 2005; Scully and Newstead, 2007; Tingvall et al., 2003). It should be noted that this method does not necessarily make use of the concept of *responsibility*.

The non-relevant accident identification-based method assumes that the attribute of interest (ABS, ESP) does not affect all types of collisions in the same way, from which we can obviously deduce that some types are closer to true exposure than others.

* Corresponding author. Tel.: +34 91 336 30 14; fax: +34 91 336 53 02.
E-mail addresses: alvaro.gomezmendez@yahoo.es (Á.G. Méndez), francisco.aparicio@upm.es (F.A. Izquierdo).

0001-4575/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2009.10.003

The applications of the quasi-induced exposure method have in most cases ignored this possibility. The fact that most works use the whole sample of two-vehicle crashes to determine exposure shows that the starting point is the following basic hypothesis: the populations of non-at-fault drivers/vehicles, and therefore the estimates of exposure, do not vary among different accident types. It is remarkable, though, that the test of this hypothesis has generally been omitted.

One of the few studies available that have addressed the differences among accident types is a recently published analysis by Keall and Newstead (2009). They compared the actual number of kilometers driven by drivers of different age and sex, and by cars of different types, with the number of drivers and vehicles involved in six accident types, that were considered as potential candidates for providing the estimates of exposure. These six candidates were: multivehicle crashes where drivers/vehicles were judged to be non-at-fault; rear-end crashes in which damage occurs to the rear of the vehicles considered; side crashes in which damage occurs to: the left side, the right side or either side of the vehicles considered; and all multivehicle crashes. The number of kilometers driven was estimated from odometer readings; in the case of drivers, it had to be assumed that the age and sex of the driver are equal to those of the owner. The authors found that the number of drivers/cars involved in rear-end crashes or as non-at-fault parties in multi-vehicle crashes were, within the context of their study, the most accurate estimates of exposure. They stressed (p. 27) that: "an important finding of this study is that even the crash type identified as best for measuring exposure will produce biased induced exposure risk estimates. This in itself may not be a serious drawback as long as the direction and extent of the bias can be recognized". The authors discuss some of the factors that may contribute to biases in induced exposure estimates, although they omitted some important variables, such as speed (Jiang and Lyles, 2007).

In this paper, we will try to present in a more systematic manner the factors that influence the probability of a driver being non-at-fault in a two-car crash. We expect that this discussion will be useful in determining the direction of bias in quasi-induced exposure analysis, even when information about kilometers driven is lacking. We will consider as potential candidates for estimating exposure non-at-fault populations in four accident types: accidents between vehicles travelling on different lanes in two-way, two-lane undivided roads; accidents between vehicles travelling on different lanes on multilane roads; intersection accidents; and accidents between vehicles travelling on the same lane. This classification is different from that used by Keall and Newstead (2009), who did not disaggregate the overall non-at-fault population, and who actually consider some crash types that may be highly correlated with the responsibility of a crash; for example, a significant proportion of drivers/vehicles hit in the side of the vehicle are at-fault parties in intersection accidents.

2. Problem formulation

We will start out with the following assumptions:

(1) The number of kilometers driven by a particular group of drivers/vehicles is the target measure of exposure.
(2) For a given set of road and environmental conditions, populations of non-at-fault parties may vary significantly among different accident types. This is so because, firstly, a number of factors influence the probability of being non-at-fault in a crash, for the same amount of exposure, and, secondly, these influences may not be the same for every type of crash.
(3) If, in a particular research, it is found that the populations of non-at-fault parties actually vary among different types of crashes, there is necessarily one type of crash (maybe two or more types with identical non-at-fault populations) whose population of non-at-fault parties comes the closest to the true distribution of kilometers driven. This type of crash should be used as the exposure metrics for that particular research.

The type of accident that should be selected to estimate exposure is the one for which, for any two groups of drivers/vehicles, the value of the relative probability of being non-at-fault is closer to one. To illustrate this, suppose two groups, *1* and *2*; let us call:

e_1,e_2 = actual exposure of groups *1* and *2*, as a proportion of one.
$\widehat{e}_1^i$, $\hat{e}_2^i$ = the estimates of exposure of groups *1* and *2*, given by a certain type *i* accident; it is equal to the quotient between the number of non-at-fault parties in type *i* accidents, and the total number of accidents of this type.
P_1^i, P_2^i = probabilities of groups *1* and *2* of being non-at-fault in type *i* accident; it is equal to the quotient between the number of non-at-fault parties in type *i* accidents, and exposure.

There is a simple relationship between actual and estimated relative exposure:

$$\frac{\widehat{e}_1^i}{\widehat{e}_2^2} = \frac{e_1 P_1^i}{e_2 P_2^i} \quad (1)$$

So, obviously, the type of accident should be selected in such a manner that (P_1^i/P_2^i) is as close to one as possible. Then, knowledge about the factors influencing the probability of being non-at-fault, in different accident types, is required. We argue that this probability is fundamentally a function of speed, accident avoidance ability and injury risk. The directions of these relationships are further investigated in Section 3.

Lastly, we think it is worth stressing our choice of kilometers driven as the target measure of exposure, because some previous studies have been somewhat ambiguous when dealing with this issue. Sometimes, the *classical definition* of the quasi-induced exposure measure is offered, without further specification. This classical definition can be given the following rough formulation: "a random sample of the drivers and vehicles on the road at the time of the accident". This definition may suggest density (number of vehicles per unit length of road) more than kilometers driven. However, most previous works of validation (Stamatiadis and Deacon, 1997; Kirk and Stamatiadis, 2001; Keall and Newstead, 2009) have focused on the comparison between quasi-induced measures and kilometers driven.

3. Factors of influence on the probability of being non-at-fault in a collision, in different accident types

We will discuss three fundamental factors of influence: speed, accident avoidance ability and risk of injury. This section is organised as follows. In Section 3.1, we will present the effects whose influences have been documented in previous research, or can be determined on the basis of generally accepted traffic safety principles. These effects will be referred to as 'known effects'. On the other side, we will discuss in Section 3.2 the 'uncertain effects', for which detailed theoretical and empirical research is lacking. In Section 3.3 we will briefly mention some effects that we regard as less important.

We will use a classification of accidents according to the positions and directions of the vehicles previous to the accident. We expect that our analysis will demonstrate that traditional classifications, which are mainly oriented to the description of the configuration of the collision and the points of impact of the vehicles involved (e.g. *frontal*, *side*, *sideswipe* and *rear-end*) are less valid. We propose four main groups: accidents between vehicles travelling on different lanes, opposite directions (two-way, two-