Estimating the Safety Effects of Automated Vehicle Technologies: Bayesian Inference of an Induced Exposure Model That Allows for Classification Uncertainty

Gary A. Davis, Raphael Stern, Christopher Cheong and Amirhossein Kiani

Dept. of Civil, Environmental, and Geo-Engineering

University of Minnesota

Traditionally, the effectiveness of safety-related modifications has been assessed using data from police-reported crashes. Certain automated vehicle (AV) technologies, mostly deployed at SAE levels 1 and 2, have now been on the road long enough that it is in principle possible to compare their crash experience, as reflected in crash databases, to the experience of similar, but unequipped, vehicles. Since crash frequency is determined both by crash risk and by exposure to crash opportunities, estimating crash rates requires measures of both crash frequency and exposure, the latter usually expressed as either vehicles-miles of travel (VMT) or total entering vehicles. Aggregated measures of exposure are often available from traffic counting programs but when one seeks to compare crash risk for vehicle or driver subgroups, such as impaired versus unimpaired drivers, passenger vehicles vs heavy trucks, or AV-equipped vehicles vs unequipped vehicles, exposure measures for the subgroups are needed. When a direct measure of exposure is unavailable, as is often the case, induced exposure methods attempt identify additional, control, events in the crash database that in some way reflect the desired exposure measure (Carr 1970; Thorpe 1967). That is, if we let λ_A denote the crash rate for driver or vehicle group A and λ_B denote the crash rate for group B then, if a reliable control group can also be identified, induced exposure methods can estimate the crash rate ratio, $\theta=\lambda_A/\lambda_B$, without having to obtain direct measures of the exposure for groups A and B. (Lyles et al. 1990; Davis and Gao 1993; Davis and Yang 2001)

Induced exposure methods have recently been applied to estimate the potential safety effects, for vehicles equipped with AV technologies, (e.g. Cicchino 2022; Leslie et al 2021; PARTS 2022). The studies used crash data aggregated from several different states in the USA, from a range of different road types, and from a restricted range of vehicle model-years. The applicability of the above to particular road types or to smaller geographic regions is an open question. A second constraint is that these studies all relied on access to proprietary data bases, provided by OEMs, listing which AV technologies were present on individual vehicles. In this paper we describe an induced exposure approach for estimating the safety-related effects of vehicle automation technologies for smaller geographic regions, using data readily available to state agencies.

Our study focuses on the effect of automated emergency braking (AEB) on rear-ending crashes occurring on interstate highways in Minnesota's Twin Cities metro region. Three sources provided our data: (1) From the Minnesota DOT we obtained computerized crash records for the reported rear-ending crashes on the Metro District interstates during the year 2020. (2) From the Minnesota Dept. of Public Safety we obtained computerized crash records for all vehicles involved in reported crashes during 2020, which included the vehicle identification number (VIN) for the involved vehicles. (3) From NHTSA's CISS database we obtained the AVOID and VPICDECODE datasets for the year 2020. The VPICDECODE contains vehicle information proved by decoding the vehicle's VIN while the AVOID dataset contains information regarding the presence/absence of several vehicle automation technologies as determined by the crash investigators.

To apply the induced exposure method it was necessary to classify each vehicle as being either a Target (rear-ending) or a Control (rear-ended) and as either having the AV technology present or absent. The first classification used the crash damage location code: vehicles with front end damage were coded as Targets and vehicles with rear-end damage were coded as Controls. The basic assumption is that the probability of a Target vehicle appearing in the crash database is proportional to its crash rate and its exposure while the probability a Control vehicle appears is proportional to its exposure. Using NHTSA's VINDECODER app it is possible to obtain information regarding presence/absence of automation technologies on a vehicle but this information is often incomplete. Measurement error issues are not infrequent in applied statistics, however, and when it is possible to model the error it can be possible to account for this in an analysis (Carroll et al 1995). To this end we cross- tabulated the CISS vehicles' VPICDECODE categories with their AVOID categories to construct a relationship between the two datasets, which then allowed us to treat the VINDECODER codes for the Minnesota crashes as partially informative regarding presence/absence of a range of AV technologies.

As was done in earlier studies we fit a logit model of the form

$$P(Y = 1|x) = \frac{exp(\beta_0+\beta_1 x)}{1+exp(\beta_0+\beta_1 x)} \tag{1}$$

Where Y=1 for Target vehicles, Y=0 for Control vehicles, x=1 if AEB is present, x=0 if AEB is absent. The crash rate ratio is then found as $\theta=\exp(\beta_1)$, while the market share for AEB-equipped vehicles in our data can also be found via

$$\hat{r} = \frac{1}{\left(1+\frac{1+exp(\hat{\beta}_0+\hat{\beta}_1)}{1=exp(\hat{\beta}_0)}\right)\left(\frac{1-\hat{q}}{\hat{q}}\right)} \tag{2}$$

Here q = P(x=1) is the fraction of AEB-equipped vehicles in our sample. Bayes estimates of the desired quantities were computed using the Markov Chain Monte Carlo software WinBUGS (Lunn et al. 2013). Our model can be seen as an adaptation of a WinBUGS example that relates the incidence of cervical cancer to exposure to the herpes simplex virus (Spiegelhalter et al. 2007). Our results are summarized in Table 1.

The estimated market share for AEB-equipped vehicles on Twin Cities freeway during 2020 was about r=0.161, or 16%. Our estimated crash rate ratio, θ=0.196, indicates that the crash rate of AEB-equipped vehicles was only about 20% of that for unequipped vehicles. Interestingly, the estimated reductions reported in other studies (Cicchino 2022; Leslie et al 2021; PARTS 2022) range between about 34% and 49%, so our approximately 80% reduction is substantially greater than the other reports. We are currently working on replicating this result using data from additional years.

Table 1. Estimation Summary

Parameter	Posterior Mean	2.5%	97.5%
β_0	0.144	0.056	0.231
β_1	-1.66	-2.14	-1.19
θ	0.196	0.118	0.304
r	0.161	0.131	0.194
q	0.098	0.082	0.117

References

Carr, R. (1970) A statistical analysis of rural Ontario traffic accidents using induced exposure data, *Proceedings of the Symposium on the Use of Statistical Methods in the Analysis of Road Accidents,* OECD, Paris, 86-72.

Carroll, R., Ruppert, D., and Sefanski, L. (1995) *Measurement Error in Nonlinear Models*, Chapman &Hall.

Cicchino, J., (2022) Effects of forward collision waring and automatic emergency braking on rear-end collisions involving pickup trucks, researchpapers@iihs.org, Insurance Institute of Highway Safety, Arlington, VA.

Davis, G. and Gao, Y. (1993) Statistical methods to support induced exposure analyses of traffic accident data, *Transportation Research Record*, 1401, 43-49.

Davis G. and Yang, S. (2001) Bayesian identification of high-risk intersections for older drivers via Gibbs sampling, *Transportation Research Record*, 1746, 84-89.

Leslie, A., Kiefer, R., Meitzner, M., and Flannagan, C. (2021) Field effectiveness of General Motors advanced driver assistance and headlighting systems, *Accident Analysis and Prevention*, 159, 106275.

Lyles, R., Stamatriadis, P., Blower, D., and Campbell, K. (1990) *The Michigan Heavy Truck Study*, Final Report, College of Engineering, Michigan State University, Lansing, MI.

Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. (2013) *The BUGS Book: A Practical Introduction to Bayesian Analysis,* CRC Press

PARTS, (2022) *Real-world Effectiveness of Model Year 2015-2020 Advanced Driver Assistance Systems,* Partnership for Analytics Research in Traffic Safety, MITRE Corporation.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2003) *WinBUGS User Manual*, Version 1.4. http://www.mrc-bsu.cam.ac.uk/bugs.

Thorpe, J. (1967) Calculating relative involvement rates in accidents without determining exposure, *Traffic Safety Research Review*, 11, 3-8.