PalArch's Journal of Archaeology of Egypt / Egyptology

DRIVERS' PERCEPTION OF ACCIDENT RISK USING LATENT CLASS MODELS

Thomas Edison Guerrero Barbosa[1] , *Sir-Alexci Suarez Castrillon*[2]

[1,2] Faculty of Engineering. Francisco de Paula Santander University, Ocaña, Colombia.

Thomas Edison Guerrero Barbosa, Sir-Alexci Suarez Castrillon. Drivers' Perception Of Accident Risk Using Latent Class Models-- Palarch's Journal Of Archaeology Of Egypt/Egyptology 19(4), 323-330. ISSN 1567-214x

Keywords: Latent Classes, Risk Perception, Road Accidents, Driver Behaviour.

ABSTRACT

The human factor is considered one of the main causes of accidents. It is the driver of the vehicle who makes the decisions when driving. The task of driving a vehicle involves a set of sensory perceptions that must be in adequate conditions in order to minimise the risk of this task and avoid accidents. This work seeks to identify, through Latent Class models, how different types of drivers behave when faced with four situations/variables that constitute a risk of accidents while driving: driving speed, driving against the flow of traffic, overtaking a vehicle on a curve, and driving under the influence of alcohol or drugs. The results show that there are two classes of populations associated with risk perception in drivers, which can be categorised as Cautious and Risky. Individuals grouped in Class A would be women, with age groups between 18-35 years and over 46 years, and with a different occupation than the one asked in the survey or if unemployed. Class B would be men, aged 36 to 45, with occupations such as students and employees.

INTRODUCTION

Several factors are involved in the occurrence of crashes with different contributions. In each accident, the contribution of human, road (road and street) and vehicle factors is different (Mehryar et al., 2022). The human factor (behaviour) causes between 70% to 90% of road crashes; therefore, studying it is one of the keys to reducing road traffic crashes (Luna-Blanco, 2013). According to the WHO (2018) in its report on the global road safety situation, reports a total of 1.35 million deaths annually due to road crashes. While this is an alarming figure and a death (whatever it may be) is an assessable loss, there are other phenomena associated with road traffic injuries that are also of concern. For example, road traffic injuries are now the leading cause of death among 5-29 year-olds. The burden is disproportionately borne by pedestrians, cyclists and motorcyclists, particularly those living in developing countries. The report suggests that the price paid for mobility is too high (WHO, 2018).

Factors causing accidents can be grouped into two categories: internal and external factors. External factors - for example - are related to road, weather and vehicle conditions (Caliendo et al., 2007; Gomes, 2013; Montero-Salgado et al., 2022); on the other hand, the human condition is related to the road conditions . On the other hand, the human condition is considered to be the most influential of the internal factors and the major cause of road accidents (Machado et al., 2014). Regardless of the factor causing the accident, it is possible to find in the literature different modelling approaches that seek to address the problem, among which are known spatial analysis (Hamim & Ukkusuri, 2022), logit modelling (Gerald et al., 2022) among others. According to (Mayou et al., 1993) it is of little interest to study the psychological and psychiatric consequences of accidents since most road accidents are experienced as less threatening, both because the trauma can be very brief and because only a minority of accidents are life-threatening.

A number of research studies (Bamberg et al., 2003; Francis et al., 2004) have addressed the relationship between driving behaviour and perceived risk of road accidents. In the 1980s and 1990s, Homel (2012) was able to test the relationship between driver behaviour and knowledge of traffic rules.

Regarding studies focused on road safety, Latent Class (LC) models have been used for various purposes such as: identification of key factors affecting crash severity (De Ona et al., 2013; Kumar et al., 2017; Weiss et al., 2016), behaviour and perception of safety by cyclists (Rossetti et al., 2018), patterns that determine accidents in cyclists (Kaplan & Prato, 2013), among others.

This paper asks the following question: are there different classes or clusters in the population that have differences in the perception of risk in driving accidents? In order to answer this question, a latent class approach is used to define whether such a class or cluster effect exists within the population.

The content of this research paper is distributed as follows: The methodological approach on the use of Latent Class models is described in Chapter 2. Then, in Chapter 3, the details of the survey that allowed the collection of the survey, which includes a Stated Preference experiment with choice tasks under hypothetical scenarios, are described. The analysis of the main results of the research is carried out in Chapter 4. Finally, Chapter 5 presents the conclusions of the research.

METHODOLOGY

Latent Class models are quite useful in modelling since the classes are able to capture unobserved heterogeneity (Magidson et al., 2020). The basic concept was introduced by Paul Lazarsfeld (1950) to construct typologies (or groups) from dichotomous variables as part of his more general latent structure analysis. Classes can represent different characteristics such as: choice sets, decision protocols, tastes, model structures, among others (Magidson et al., 2003).

For explanatory purposes we will consider that the estimate of the probability "P" that an individual "n" chooses alternative "i" from a CL model can be estimated from two components: a class-specific choice model and a class membership model. In the first component, the probability of choosing i conditional on class membership s is estimated, while in the second component, the probability of belonging to class s is calculated. The mathematical expression describing the estimation of the probability of choice using CL is shown as follows:

$$P_n(i) = \sum_{s=1}^{S} P_n(i|s)\, P_n(s)$$

Where the term $P_n(i|s)$ y $P_n(s)$ are the first and second components described above, respectively. Estimation from CL has been used extensively in the marketing area, however, it has become more widespread (Grover & Srinivasan, 1987). However, its use has been extended to other research fields such as economics, transport and geography (Boxall & Adamowicz, 2002; Greene & Hensher, 2003).

The data and survey design

Data were collected in Ocaña, a small city in Colombia and the second largest city in the department of Norte de Santander, approximately 400 km northeast of Bogotá (Colombia). It has an urban area of 7 km^2 and a population of 93,650 inhabitants (National Planning Department, 2011).

To elicit drivers' risk perceptions and the influence of their driving behaviour, a Stated Preference (SP) instrument was designed in which respondents were confronted with two hypothetical driving scenarios and asked to choose one of them.

The first part of the questionnaire sought to obtain socio-economic information from the respondent such as: gender, age, main occupation, level of education, marital status, type of disabilities, possession of driving license, age of driving license, personal monthly income, whether their current job involved driving a vehicle and whether they had been involved in a traffic accident. For the application of the CL models, the qualification of perception indicators is necessary. These indicators can be attitudinal when they represent perceptions of the individuals based on their life experience (Bahamonde-Birke et al., 2017) or perceptual when these perceptions are directly related to the alternatives (Raveau et al., 2010). On a scale of four responses (always, often, sometimes and never) each respondent was asked about their perception of each of the questions below (Barbosa et al., 2017).

1) Do you do other activities while driving (e.g. check the fuel level, check the speedometer level, adjust the stereo)?
2) Do you use your mobile phone while driving (answer and/or chat)?
3) Are you driving under sub-optimal conditions (e.g. under the influence of drugs, injury, sleep, stress)?
4) Do you wear a seat belt (or helmet in the case of motorbikes) when driving?

This was followed by a PD questionnaire, the design of which considered eight hypothetical driving scenarios with two alternatives each. Each alternative, in turn, was characterised by four attributes: i) driving speed, ii) driving against the flow of traffic, iii) overtaking a vehicle on a curve, and iv) driving under the influence of alcohol and drugs. The respondent was asked - for each hypothetical scenario presented - to choose the alternative that he/she considered to be the most accident-prone.

RESULTS

The Table 1 CL model is presented. This model characterises two classes within the population, which have been denoted by the letters A and B. Users were more likely to belong to Class A if they were female, with age groups between 18-35 years and over 46 years and with a different occupation than the one asked in the survey or if they were unemployed. On the other hand, Class B was characterised by males, aged between 36 and 45, with occupations such as students and employees.

Regarding the attributes that characterised the PD experiment, it could be observed that - for both classes - they are significant at a 95% confidence level and show a positive sign. This last aspect means that - independently of the class - individuals feel an increase in the perception of risk (or danger) as speed increases, a fact that is consistent with other findings in the literature such as those presented by Perdomo et al. (2014) and Machado et al. (2014). The same effect is found for the variables θ_{ccv} , θ_{avc} and θ_{cad} , suggesting that when an individual drives on the wrong side of the road, overtakes a vehicle on a curve or drives under the influence of alcohol and/or drugs, his or her perception of risk increases.

The parameter associated with class (θ_{CLASE}) yielded significance at a 95% confidence level, suggesting that there are indeed clusters within the phenomenon studied. It is interesting to note that, although for both classes the variables Driving speed and Driving under the influence of alcohol and/or drugs are significant and common, it is evident that Class A punishes it much more than Class B. Note that for Class A, the effect of Driving speed on risk perception is 2.6 times higher than for Class B. On the other hand, for Class A, the effect of Driving under the influence of alcohol and/or drugs on risk perception is 3 times higher than for Class B. Under this analysis, it can be considered that those individuals who are part of Class A are a "Cautious" group while those in Class B could be called more "Risky" than those in Class A.

Table 1. Estimation of CL parameters.

Variable	**Description**	**CL**	
		Value	**t-test**
Explanatory variables			
θ_{vel_A}	Driving speed - Class A	34.8198	13.624
θ_{ccv_A}	Driving against the line - Class A	6.8045	9.589

θ_{cad_A}	Driving under the influence of alcohol and/or drugs - Class A	8.3233	13.162
$\theta ASC1\_A$	Specific Constant of Alternative 1 - Class A	-10.81	-10.762
θ_{vel_B}	Driving speed - Class B	13.4422	3.805
θ_{avc_B}	Cornering - Class B	7.0678	4.597
θ_{cad_B}	Driving under the influence of alcohol and/or drugs - Class B	2.7707	3.839
$\theta ASC1\_B$	Specific Constant of Alternative 1 - Class B	0.726	0.718
θ_{CLASE}	Class membership parameter	2.1294	2.945
θ_{gen}	Gender (male)	-1.1686	-2.884
θ_{edad1}	Age between 18 and 25 years	1.091	2.019
θ_{edad2}	Age between 26 and 35 years	0.6331	1.642
θ_{edad3}	Age between 36 and 45 years	-0.5577	-1.449
θ_{ocup1}	Student	-1.7339	-2.582
θ_{ocup2}	Employee	-0.691	-1.522
θ_{ocup3}	Unemployed	0.5791	0.483
General report			
N	Number of observations		
$L(\theta)$	Log-likelihood	-594.03	
ρ^2	Rho squared index	0.446	

CONCLUSION

It was to estimate a model under the Latent Class approach. The estimated model allows the existence of two classes or clusters of population that perceive risk differently while facing driving manoeuvres to be evidenced. Likewise, this model allows identifying that the four attributes considered within the choice experiment (driving speed, driving against the line, overtaking a vehicle on a curve and driving under the influence of alcohol and/or drugs) had an impact on the perception of the risk of accidents affecting drivers. In light of the results, the CL model appears to show that the attributes Driving speed and Driving under the influence of alcohol and/or drugs are strongly penalised more by one Class than the other. In general, the variables included in the model estimation showed the expected signs and their behaviour was in line with theory.

The CL model clearly shows the definition of two Classes (A and B) within the experiment. On the one hand, there are individuals with a "Cautious" category, while there is another class that tends to be more "Risky". These two categories are framed by the phenomenon studied in this research: drivers' perception of accident risk. Individuals grouped in Class A would be women, with age groups between 18-35 years and over 46 years, and with a different occupation than the one asked in the survey or if unemployed. Class B would be men, aged 36 to 45, with occupations such as students and employees.

Future research could focus on trying to perform similar experiments on simulators and compare the results to accept or refute the findings of this study. It would also be interesting to consider the effect of other types of vehicles such as motorbikes, light vehicles, cargo vehicles, among others. The hypothesis in this case is that drivers' perception of accident risk is different depending on the type of vehicle.

REFERENCES

Bahamonde-Birke, F. J., Kunert, U., Link, H., & Ortúzar, J. de D. (2017). About attitudes and perceptions: Finding the proper way to consider latent variables in discrete choice models. Transportation, 44(3), 475-493.

Bamberg, S., Ajzen, I., & Schmidt, P. (2003). Choice of travel mode in the theory of planned behavior: The roles of past behavior, habit, and reasoned action. Basic and Applied Social Psychology, 25(3), 175-187.

Barbosa, T. E. G., Raveau, S., & Ortúzar, J. de D. (2017). Understanding drivers' perception of crash risk: A latent variable approach. Transport Studies, 21(2), Article 2. https://estudiosdetransporte.org/sochitran/article/view/207

Boxall, P. C., & Adamowicz, W. L. (2002). Understanding heterogeneous preferences in random utility models: A latent class approach. Environmental and Resource Economics, 23(4), 421-446.

Caliendo, C., Guida, M., & Parisi, A. (2007). A crash-prediction model for multilane roads. Accident Analysis & Prevention, 39(4), 657-670.

De Ona, J., López, G., Mujalli, R., & Calvo, F. J. (2013). Analysis of traffic accidents on rural highways using Latent Class Clustering and Bayesian Networks. Accident Analysis & Prevention, 51, 1-10.

Francis, J., Eccles, M. P., Johnston, M., Walker, A. E., Grimshaw, J. M., Foy, R., Kaner, E. F., Smith, L., & Bonetti, D. (2004). Constructing questionnaires based on the theory of planned behaviour: A manual for health services researchers.

Gerald, A., Mbue, I. N., Merlin, A. Z., & Achiri, A. T. (2022). A Systematic Review of Excessive Speed Monitoring and Control System for Accident Prevention on Cameroon Highways. World Journal of Engineering and Technology, 10(1), 28-39.

Gomes, S. V. (2013). The influence of the infrastructure characteristics in urban road accidents occurrence. Accident Analysis & Prevention, 60, 289-297.

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. Transportation Research Part B: Methodological, 37(8), 681-698.

Grover, R., & Srinivasan, V. (1987). A simultaneous approach to market segmentation and market structuring. Journal of Marketing Research, 24(2), 139-153.

Hamim, O. F., & Ukkusuri, S. V. (2022). Determining prominent factors across system hierarchies to improve road safety in LMICs: A case study of Bangladesh. Safety Science, 150, 105709.

Homel, R. (2012). Policing and punishing the drinking driver: A study of general and specific deterrence. Springer Science & Business Media.

Kaplan, S., & Prato, C. G. (2013). Cyclist-Motorist Crash Patterns in Denmark: A Latent Class Clustering Approach. Traffic Injury Prevention, 14(7), 725-733. https://doi.org/10.1080/15389588.2012.759654

Kumar, S., Toshniwal, D., & Parida, M. (2017). A comparative analysis of heterogeneity in road accident data using data mining techniques. Evolving Systems, 8(2), 147-155.

Lazarsfeld, P. F. (1950). The logical and mathematical foundation of latent structure analysis. Studies in Social Psychology in World War II Vol. IV: Measurement and Prediction, 362-412.

Luna-Blanco, R. (2013). Perception of risk and safety when driving vehicles. Carreteras: Revista técnica de la Asociación Española de la Carretera, 189, 48-56.

Machado, J. L., de Oña, J., de Oña, R., Eboli, L., & Mazzulla, G. (2014). A stated preference experiment for understanding drivers' risk perception. Procedia-Social and Behavioral Sciences, 162, 263-272.

Magidson, J., Eagle, T., & Vermunt, J. K. (2003). New developments in latent class choice models. Sawtooth Software Conference Proceedings, 89-112.

Magidson, J., Vermunt, J., & Madura, J. (2020). SAGE Research Methods Foundations. SAGE Publications Ltd. https://doi.org/10.4135/9781526421036

Mayou, R., Bryant, B., & Duthie, R. (1993). Psychiatric consequences of road traffic accidents. British Medical Journal, 307(6905), 647-651.

Mehryar, H. R., Alizadeh, N., Rostampour, F., Paryab, S., Bab, S., & Garkaz, O. (2022). Causes of road accidents in northwestern iran in the period 2010-2018. International Archives of Health Sciences, 9(1), 14. https://doi.org/10.4103/iahs.iahs_175_21

Montero-Salgado, J. P., Muñoz-Sanz, J., Arenas-Ramírez, B., & Alén-Cordero, C. (2022). Identification of the mechanical failure factors with potential influencing road accidents in Ecuador. International Journal of Environmental Research and Public Health, 19(13), 7787.

National Planning Department (2011). National Development Plan 2010-2014 Colombia. Bogotá DC. Imprenta Nacional de Colombia.

Perdomo, M., Rezaei, A., Patterson, Z., Saunier, N., & Miranda-Moreno, L. F. (2014). Pedestrian preferences with respect to roundabouts-A video-based stated preference survey. Accident Analysis & Prevention, 70, 84.

Raveau, S., Álvarez-Daziano, R., Yáñez, M. F., Bolduc, D., & de Dios Ortúzar, J. (2010). Sequential and simultaneous estimation of hybrid discrete choice models: Some new findings. Transportation Research Record, 2156(1), 131-139.

Rossetti, T., Guevara, C. A., Galilea, P., & Hurtubia, R. (2018). Modeling safety as a perceptual latent variable to assess cycling infrastructure. Transportation Research Part A: Policy and Practice, 111, 252-265.

Weiss, H. B., Kaplan, S., & Prato, C. G. (2016). Fatal and serious road crashes involving young New Zealand drivers: A latent class clustering approach. International Journal of Injury Control and Safety Promotion, 23(4), 427-443.

WHO (2018). Global status report on road safety 2018. https://www.who.int/publications-detail-redirect/9789241565684