EVALUATING INDIVIDUAL HETEROGENEITY IN THE PROBABILITY OF CRASH INVOLVEMENT

M. Bobermin & S. Ferreira

U.PORTO

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Citta
FEUP UNIVERSITY OF PORTO
FCTUC UNIVERSITY OF COIMBRA

SUMMARY

INTRODUCTION

- Driver Behavior Questionnaire (DBQ) (Reason et al., 1990) one of the most applied tools to evaluate **driver behavior**

- Drivers are associated with a score representing the frequency of committing a series of behaviours, usually divided into **errors** and **violations**

Table 3. Main DBQ items.

Item
Attempt to drive away from traffic lights in third gear.
Check your speedometer and discover that you are unknowingly travelling faster than the legal limit.
Lock yourself out of your car with the keys still inside.
Become impatient with a slow driver in the outer lane and overtake on the inside.
Drive as fast along country roads at night on dipped lights as on full beam.
Attempt to drive away without first having switched on the ignition.
Drive especially close or 'flash' the car in front as a signal for that driver to go faster or get out of your way.
Forget where you left your car in a multi-level car park.
Distracted or preoccupied, realise belatedly that the vehicle ahead has slowed, and have to slam on the brakes to avoid a collision.
Intend to switch on the windscreen wipers, but switch on the lights instead, or *vice versa*.
Turn left on to a main road into the path of an oncoming vehicle that you hadn't seen, or whose speed you had misjudged.
Misjudge your gap in a car park and nearly (or actually) hit

J. Reason et al. 1990

INTRODUCTION

- Several studies have investigated the hypothesis that errors and violations have a distinct **effect on crash liability**

- Many of those studies applied zero-order correlations, assuming that parameters were fixed across observations, which does not account for the heterogeneity across individuals

INTRODUCTION

- An analysis considering **age**, **sex**, driving **exposure**, and **DBQ scores** using a **random-parameter logit model** to investigate their influence on crash involvement

- This allows understanding how the same behavior may distinctly affect crash occurrence due to unobserved variables

Methodology

Driver Behavior Questionnaire and Participants

- 1,321 participants
- DBQ version with 20 items and three dimensions (Er, OV, and AV).


51.6%


M: 32
R: 18-82


17%

Table 5: CATPCA dimension loadings values of the items from the 3D-20 version

Item	Dimensions 1 Errors	2 Ordinary Vio	3 Aggressive Vio
DBQE01	**0.50**	0.07	0.05
DBQE02	**0.59**	0.06	0.09
DBQE03	**0.46**	0.04	0.02
DBQE04	**0.58**	0.18	0.07
DBQE05	**0.44**	0.20	-0.09
DBQE07	**0.59**	-0.01	0.20
DBQE08	**0.55**	0.18	0.14
DBQV09	0.03	**0.79**	0.16
DBQV10	0.14	**0.78**	0.08
DBQV14	**0.51**	0.17	0.16
DBQV18	0.07	0.19	**0.75**
DBQV19	0.07	0.23	**0.67**
DBQV20	0.11	-0.06	**0.56**
EXTE01	**0.57**	0.14	0.05
EXTE02	**0.69**	0.11	0.03
EXTV04	0.05	0.08	**0.44**
EXTV07	0.18	**0.60**	0.07
EXTV08	0.20	**0.55**	0.10
EXTV11	0.22	**0.54**	0.19
EXTV12	0.08	0.12	**0.66**
%Variance Explained	16.186	12.382	10.702
Cronbach's alpha	0.77	0.73	0.67

Bobermin et al., 2021

Methodology

Random-parameter Binary Logit Model

Prob(Y=1|x) = F (x, β)
Prob(Y=0|x) = 1-F (x, β)

involved in a crash (Y=1) or not (Y=0)
vector x drivers' demographics and DBQ subscale scores

$$\text{Prob}(Y = 1|x) = \frac{e^{\beta x}}{1 + e^{\beta x}} = \Lambda(\beta x)$$

parameters β reflect the impact of changes in vector x on the probability, assumed as a logistic distribution

$$\text{Prob}(Y = 1|x) = \int \Lambda(\beta x) f(\beta|\varphi) d\beta$$

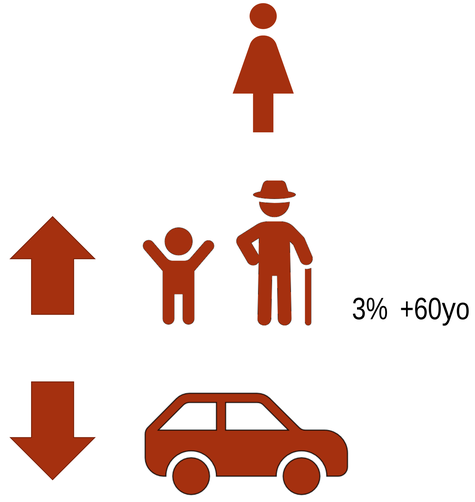
a continuous density function of β is introduced to allow parameters to vary, f ($\beta|\phi$), with ϕ referring to a vector of parameters of that density function (mean and variance)

Results and Analysis

Variable	Estimated value	Standard Error	P[Z>z]	Marginal Effects	Confidence Intervals
Nonrandom parameters					
Age	-.023***	.006	.000	-.003***	(-.034; -.012)
Sex	-.362***	.125	.004	-.052***	(-.607; -.118)
Exposure 02 - two to three times a week	.269	.216	.214	.038	(-.155; .692)
Exposure 03 - four to five times a week	.357*	.213	.093	.051*	(-.060; .774)
Exposure 04 - six or more times a week	.394**	.197	.046	.056**	(.008; .779)
Means for random parameters					
Constant	-1.068***	.263	.000	--	(-1.585; -.552)
Score Errors	.065	.151	.669	.009	(-.232; .361)
Score Ordinary Violations	.263***	.087	.003	.037***	(.093; .433)
Score Aggressive Violations	-.415***	.130	.002	-.059***	(-.670; -.159)
Scale parameters for dists. of random parameters					
Constant	.983***	.098	.000	--	(.0791; 1.175)
Score Errors	.861***	.121	.000	--	(.0624; 1.097)
Score Ordinary Violations	.274***	.055	.000	--	(.167; .381)
Score Aggressive Violations	1.109***	.133	.000	--	(.848; 1.370)

***, **, * Significance at 1%, 5%, 10% level, respectively.

- Age, sex, and the exposure do not have a random effect, fixed parameters

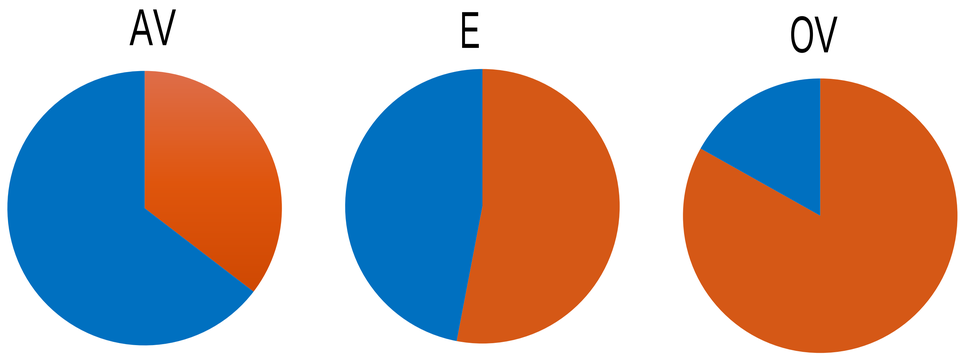


Results and Analysis

Variable	Estimated value	Standard Error	P[Z>z]	Marginal Effects	Confidence Intervals
Nonrandom parameters					
Age	-.023***	.006	.000	-.003***	(-.034; -.012)
Sex	-.362***	.125	.004	-.052***	(-.607; -.118)
Exposure 02 - two to three times a week	.269	.216	.214	.038	(-.155; .692)
Exposure 03 - four to five times a week	.357*	.213	.093	.051*	(-.060; .774)
Exposure 04 - six or more times a week	.394**	.197	.046	.056**	(.008; .779)
Means for random parameters					
Constant	-1.068***	.263	.000	--	(-1.585; -.552)
Score Errors	.065	.151	.669	.009	(-.232; .361)
Score Ordinary Violations	.263***	.087	.003	.037***	(.093; .433)
Score Aggressive Violations	-.415***	.130	.002	-.059***	(-.670; -.159)
Scale parameters for dists. of random parameters					
Constant	.983***	.098	.000	--	(.0791; 1.175)
Score Errors	.861***	.121	.000	--	(.0624; 1.097)
Score Ordinary Violations	.274***	.055	.000	--	(.167; .381)
Score Aggressive Violations	1.109***	.133	.000	--	(.848; 1.370)

***, **, * Significance at 1%, 5%, 10% level, respectively.

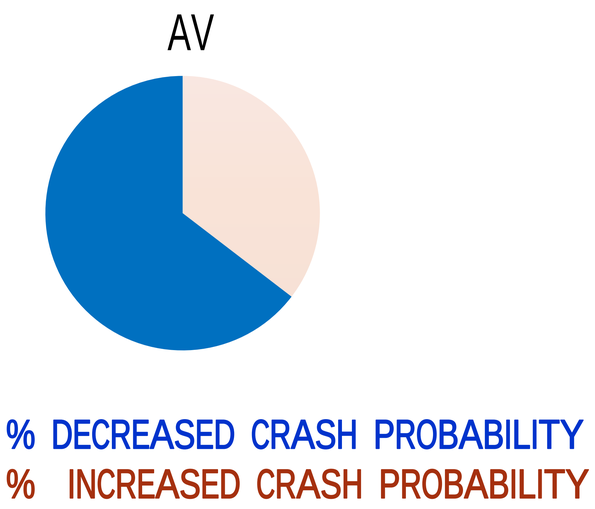
- The three DBQ scores were statistically significant random parameters



% DECREASED CRASH PROBABILITY
% INCREASED CRASH PROBABILITY

Results and Analysis

- Separation between Aggressive and Ordinary Violation
- DBQ version based on a widespread version of the DBQ, but different
- Expected that high AV would increase the likelihood of crashes
- Hostile behavior not imply higher involvement in risky situations, or the risks may be compensated by other factors

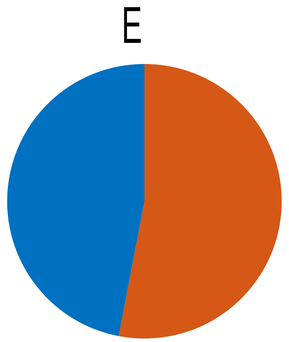


Results and Analysis

Variable	Estimated value	Standard Error	P[Z>z]	Marginal Effects	Confidence Intervals
Nonrandom parameters					
Age	-.023***	.006	.000	-.003***	(-.034; -.012)
Sex	-.362***	.125	.004	-.052***	(-.607; -.118)
Exposure 02 - two to three times a week	.269	.216	.214	.038	(-.155; .692)
Exposure 03 - four to five times a week	.357*	.213	.093	.051*	(-.060; .774)
Exposure 04 - six or more times a week	.394**	.197	.046	.056**	(.008; .779)
Means for random parameters					
Constant	-1.068***	.263	.000	--	(-1.585; -.552)
Score Errors	.065	.151	.669	.009	(-.232; .361)
Score Ordinary Violations	.263***	.087	.003	.037***	(.093; .433)
Score Aggressive Violations	-.415***	.130	.002	-.059***	(-.670; -.159)
Scale parameters for dists. of random parameters					
Constant	.983***	.098	.000	--	(.0791; 1.175)
Score Errors	.861***	.121	.000	--	(.0624; 1.097)
Score Ordinary Violations	.274***	.055	.000	--	(.167; .381)
Score Aggressive Violations	1.109***	.133	.000	--	(.848; 1.370)

***, **, * Significance at 1%, 5%, 10% level, respectively.

- Unobserved heterogeneity cancels out positive and negative effects
- Errors are unintentional, occur randomly

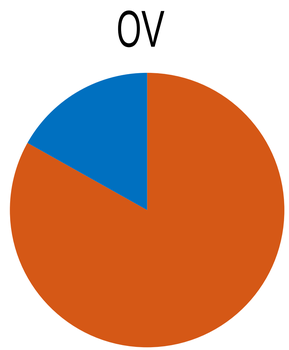


% DECREASED CRASH PROBABILITY
% INCREASED CRASH PROBABILITY

Results and Analysis

- Ordinary Violation scores that were statistically significant for the Spearman correlation and the standard logit model also
- OV even reduced the likelihood of crashes in some cases
- This result can be explained by external factors not included in the study, such as the vehicle technology available

	Age	Sex	Exposure	Score E	Score OV	Score AV
Spearman	-	-	+	+	+	
Stantard Logit Model	-	-	+		+	
RP Logit Model	-	-	+	+	+	-



% DECREASED CRASH PROBABILITY
% INCREASED CRASH PROBABILITY

Conclusion

- Despite several studies repeatedly tested DBQ x crashes, this study adds to the body of knowledge since it considers individual heterogeneity
- Crash causes are intrinsically heterogeneous and assuming a fixed and unique coefficient for all observations might result in misleading inferences for subgroups that do not hold the same relationship/magnitude
- The DBQ scores have a heterogeneous effect on crash probability, likely because not all relevant variables are included in the study

EVALUATING INDIVIDUAL HETEROGENEITY IN THE PROBABILITY OF CRASH INVOLVEMENT

M. Bobermin & S. Ferreira

Thank you

mariane.bobermin@fe.up.pt

U.PORTO

FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Citta
FEUP UNIVERSITY OF PORTO
FCTUC UNIVERSITY OF COIMBRA