Modelling socioeconomic disadvantage and road trauma using a modified model to overcome data challenges

Pyta[a], V

[a]VicRoads/ARRB Group Ltd, Victoria[1]

Abstract

A key challenge identified in the *Australian National Road Safety Strategy 2011 to 2020* is to reduce the incidence of serious casualties within Indigenous communities and other disadvantaged communities.

This Austroads initiated study identifies variables that contribute to the relationship between socioeconomic status of an area and risk of being killed or seriously injured (KSI). It does this using a statistical technique that offers a more flexible approach to the analysis of crash data when expressed as counts. The intermediary variables identified may be useful in designing more targeted behavioural and infrastructure-based road safety interventions for disadvantaged communities.

The analysis examines whether road users living in disadvantaged areas were more likely to be killed or serious injured in a crash compared with road users from less disadvantaged areas. It was important to establish whether the relationship remained after controlling for factors related to crash risk and disadvantage such as remoteness, road environment, individual and behavioural factors.

A modified negative-binomial regression of KSI rates per population group was conducted using South Australian crash data (2001 to 2010), and socioeconomic and population data from the ABS. The usefulness of this method for modelling KSI rates is discussed, as well as steps that could improve the accuracy of the modelling and reliability of the results.

As expected, socioeconomic status was associated with the KSI rate per population even after controlling for remoteness. Explanatory variables added to the model (especially high alcohol hours and road environment variables) explained a large part of the relationship.

Introduction

In Australia, approximately 4 people are killed and 90 people are seriously injured in road crashes every day. Australia's response to this significant public health issue is laid out in the Australian *National Road Safety Strategy 2011 to 2020*. One of the key challenges identified in the Strategy is to reduce the incidence of serious casualties within Indigenous communities and among other disadvantaged people (ATC 2011).

Socioeconomic disadvantage has long been acknowledged as a correlate and potential contributing factor to poorer outcomes in the fields of health and education. In the 1980s, researchers in the UK turned their attention to the relationship between deprivation and road safety. They observed that deprivation was clearly associated with greater incidence of road transport injury, especially for child pedestrians (RoSPA 2012).

The World Health Organisation (WHO) has estimated that 91% of road deaths occur in low and middle income countries, and that within high income countries, people from lower socioeconomic backgrounds are at greater risk than people from higher socioeconomic backgrounds (WHO 2012).

---

[1] The work reported here was commenced as an Austroads project while the author was working at ARRB. It formed the basis of the author's Masters thesis and was completed after the author joined VicRoads. The views expressed are the author's and do not reflect the views of the organisation.

In Australia, it has been observed that:

- Persons from local government areas (LGAs) with higher rates of disadvantage[2] generally have higher risk of crash involvement per head of population, although the group with the highest level of risk is not necessarily the most disadvantaged group (Clapperton & Cassell 2010).

- The form of the relationship between disadvantage and crash risk varies by age group (Clapperton & Cassell 2010).

- Remoteness of crash location and low socioeconomic status of a person are associated with significantly increased risk of death (Chen et al. 2010).

To date there has been no single Australian study into the relationship between socioeconomic status (SES) and crash involvement that controls for age, gender and remoteness (in terms of crash location or the casualties' residential addresses). This is important because failing to properly account for potentially confounding or masking factors means that it is not possible to determine whether confounding factors alone would have been sufficient to explain the disparities between socioeconomic groups.

Crash risk is known to increase as distance from major population centres increases (ATC 2011). At the same time, SES tends to decrease as remoteness increases (ABS 2000). Young age tends to be associated with lower SES because younger people have generally not yet had the opportunities to accumulate wealth and progress in the workforce, and gender distributions can vary with the predominant local employment type. Age and gender are known correlates of crash risk, with male drivers in the 15 to 24 years age group at the greatest risk (ATC 2011).

To provide road safety stakeholders with potentially effective targets for interventions, more information is needed about the factors that mediate the relationship between disadvantage and road trauma. It has been found, for example, that fatalities among lower SES people are more likely to be associated with higher posted speed limits, fatigue and driving an older vehicle (Chen 2010).

Studies from overseas have also found that road environment variables such as curves and unsealed roads are important contributing factors that may mediate the relationship between SES and crash risk (Dissanayake, Aryaija & Wedagama 2009; Jones et al. 2008; Huang Abdel-Aty & Darwiche 2010; Morency et al. 2012).

Other studies have found an association between SES and blood alcohol concentration of crashed drivers (Braver 2003), and drink driving convictions (Leal, King & Lewis 2006). Senserrick et al. (2010) found that the raw relationship between SES and pre-licensed driving was explained by other factors including 'risky behaviours', alcohol and drug use.

This project aims to (1) expand the evidence base in this area of road safety by modelling the relationship between socioeconomic disadvantage and road trauma; and (2) address some of the limitations of previous studies by simultaneously controlling for remoteness of residence, and age and gender of the casualty.

[2] measured using the ABS Socioeconomic Indexes for Areas

Method

Data sources

This project used data from two sources, the Australian Bureau of Statistics (ABS) and the South Australian Department of Transport, Planning and Infrastructure (DTPI):

- Census-based data downloaded from the ABS website:
  - 2006 SEIFA indexes by SLA (Statistical Local Area, ABS 2008a)
  - 2006 Remoteness Areas by SLA (ABS 2013a)
  - South Australian population by age, gender and SLA for each of the years 2001 to 2010 (ABS 2013b)
  - 2006 Postal Area Correspondences (ABS 2007)
- South Australian crash data made available by the South Australian DTPI for the Austroads project: ST1761 Safety for Disadvantaged Groups
  - Crash data
  - Traffic unit data
  - KSI data.

The crash data was restricted to killed and seriously injured persons from South Australian SLAs between 2001 and 2010. Crash data was aggregated by SLA, and the SLAs were ranked on the ABS Index of Relative Socioeconomic Disadvantage (IRSD) and Remoteness Index. The IRSD ranks areas based on weighted combinations of their scores on a range of variables. These variables cover barriers to participation in education, employment and social activities, access to resources and demands on those resources, and the burden of chronic illness and disability. The first quintile of IRSD scores represent the 20% most disadvantaged of areas on this scale, and the fifth quintile of scores represent the least disadvantaged 20% of areas.

Table 1: Population by Remoteness Area and Index of Relative Social Disadvantage (IRSD)

Statistic	Remoteness	IRSD Quintile					
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
Count	Major Cities	194,054	200,221	165,628	205,485	370,451	1,135,842
	Inner Regional	18,159	51,137	24,666	62,007	24,468	180,437
	Outer Regional	70,122	55,909	20,941	32,760	0	179,732
	Remote	0	18,020	13,721	9,240	4,186	45,167
	Very Remote	8,469	2,758	124	0	0	11,351
	All	**290,804**	**328,045**	**225,080**	**309,492**	**399,105**	**1,552,529**
% within IRSD	Major Cities	66.7%	61.0%	73.6%	66.4%	92.8%	73.2%
	Inner Regional	6.2%	15.6%	11.0%	20.0%	6.1%	11.6%
	Outer Regional	24.1%	17.0%	9.3%	10.6%	0.0%	11.6%
	Remote	0.0%	5.5%	6.1%	3.0%	1.0%	2.9%
	Very Remote	2.9%	0.8%	0.1%	0.0%	0.0%	0.7%
	All	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**
% within Remoteness Area	Major Cities	17.1%	17.6%	14.6%	18.1%	32.6%	100.0%
	Inner Regional	10.1%	28.3%	13.7%	34.4%	13.6%	100.0%
	Outer Regional	39.0%	31.1%	11.7%	18.2%	0.0%	100.0%
	Remote	0.0%	39.9%	30.4%	20.5%	9.3%	100.0%
	Very Remote	74.6%	24.3%	1.1%	0.0%	0.0%	100.0%
	All	**18.7%**	**21.1%**	**14.5%**	**19.9%**	**25.7%**	**100.0%**

Table 2: KSI and mean and variance of KSI within group by remoteness area and IRSD

Remoteness	IRSD	KSI	Mean	Variance
Major cities	Q5 - Least disadvantaged	2,362	1.98	1.825
	Q4	1,097	2.00	1.922
	Q3	1,391	2.56	3.202
	Q2	1,748	2.81	5.058
	Q1 - Most disadvantaged	1,740	2.68	5.529
	All	**8,338**	**2.35**	**3.417**
Inner regional	Q5 - Least disadvantaged	289	1.76	1.287
	Q4	801	2.11	2.598
	Q3	324	1.83	1.232
	Q2	645	2.27	3.280
	Q1 - Most disadvantaged	279	3.24	6.210
	All	**2,338**	**2.14**	**2.769**
Outer regional, remote and very remote	Q5 - Least disadvantaged	43	1.43	0.599
	Q4	515	1.68	1.752
	Q3	426	1.38	0.478
	Q2	1,035	1.71	1.209
	Q1 - Most disadvantaged	873	1.98	1.843
	All	**2,892**	**1.71**	**1.365**
All levels of remoteness	Q5 - Least disadvantaged	2,694	1.94	1.744
	Q4	2,413	1.96	2.111
	Q3	2,141	2.08	2.323
	Q2	3,428	2.27	3.424
	Q1 - Most disadvantaged	2,892	2.46	4.347
	All	**13,568**	**2.14**	**2.830**

Missing data on key variables

Residential postcodes are essential for linking data about casualties to data on remoteness and disadvantage. However, postcodes were not available for passengers in the South Australian crash data. Analysis of Victorian crash data over the same period showed that the IRSD quintile of the passenger was usually found to be similar to that of the driver unless the passenger was in a public transport vehicle (including taxis, buses, coaches, trams and trains). So, for the purpose of this analysis, the following rule was applied:

- If the passenger was in a taxi, bus, tram or train, they were excluded from the analysis because they cannot be assigned to an IRSD quintile.
- If the passenger was in any other car or truck, they were assigned to the driver's IRSD quintile.

The key variables on which the crash data were matched to the ABS data were age group, gender, postcode and crash year. After excluding KSI from interstate and overseas, and assigning passengers to their driver’s quintile, there remained:

- 832 KSI with unknown age (6.1%)
- 4 with unknown or uncategorised sex

- 84 with unknown postcode.

Because the program used to model the data can only handle complete cases, leaving information on key matching variables missing would have resulted in 871 South Australian KSI being excluded from the analysis (some KSI had missing data for more than one variable). This was not considered to be acceptable, because most of these cases had complete or near complete data on all of the variables of interest. Missing data for age and sex were replaced with predicted values using multiple imputation.

Choice of model

The effects of disadvantage, remoteness and crash and casualty variables on KSI likelihood were analysed using a zero-truncated negative binomial regression model through the VGAM package in R (R Development Core Team 2013; Yee 1996, Yee 2013). The dependent variable in this study was the number of KSI per statistical local area of residence, adjusted for population. KSI were divided into demographic segments based on:

- Remoteness Area (3 groups)
  - Major cities
  - Inner Regional
  - Outer regional, remote and very remote
- Index of Relative Social Disadvantage (IRSD) quintile
  - Q1 (the least disadvantaged group)
  - Q2
  - Q3
  - Q4
  - Q5 (the most disadvantaged group)
- Crash Year (10 years: 2001 to 2010)
- Sex (male and female)
- Age Group
  - 0 to 14 years
  - 15 to 24 years
  - 25 to 39 years
  - 40 to 54 years
  - 55 to 69 years
  - 70 years and older).

The distribution of counts of KSI within these demographic segments is shown Figure 1.

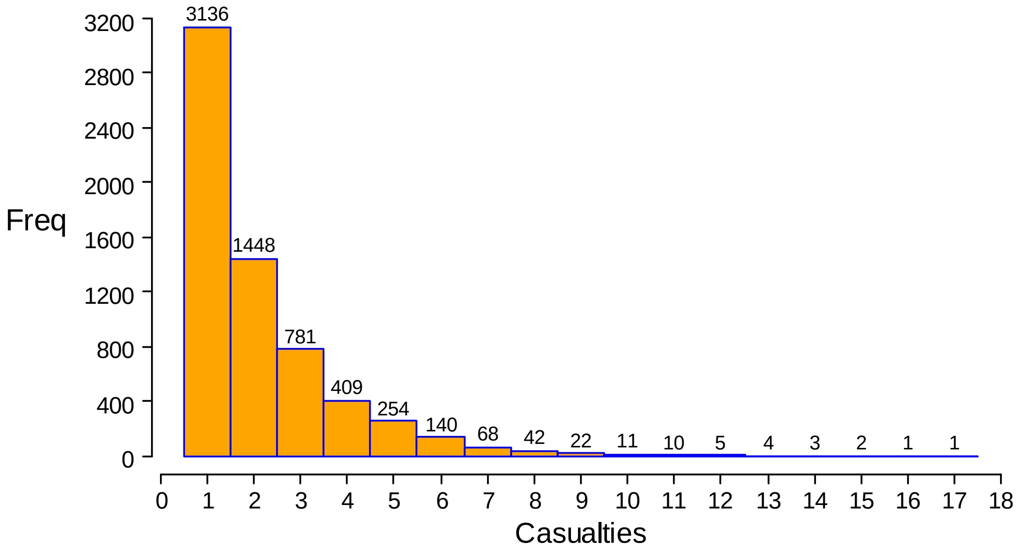

Figure 1: KSI per age, gender, year and SLA segment

Count and rate data are often analysed using Poisson regression. However, Poisson regression is only appropriate when the data meet the model's assumptions. One such assumption is that the variance is equal to the mean. Proceeding with a Poisson model in the presence of over-dispersion can result in poor estimation of parameter coefficients and their confidence intervals, potentially invalidating the conclusions drawn from the model. The negative binomial model, which includes an additional parameter that takes account of the dispersion in the data, can be used to overcome this problem (Hilbe 2011).

The current study-design violates another important assumption of Poisson regression. The count of KSI within an age, gender and SLA cohort per year commences at one (and not zero) because inclusion in the crash database requires a crash to have occurred. Attempting to fit a standard Poisson regression to this data would result in a model that attempts to predict a distribution that includes zeros. A zero-truncated model can be used to model count data in which the value zero cannot occur (Cameron and Trivedi 1998).

Grogger and Carson (1991) demonstrated that over-dispersion in truncated Poisson regression models is likely to result in biased and inconsistent parameter estimates, undermining the reliability and validity of the model and any inferences that are drawn from it. In their study, the truncated negative binomial model performed much better, resulting in consistent and unbiased estimates. For this reason, a zero-truncated negative binomial model was adopted in this study.

Figure 2 illustrates that the distribution of counts differs from what would be expected from a Poisson distribution with lambda=2.1 (the mean of the distribution, obtained from Table 2). At the left tail of the distribution, the absence of counts of zero in the observed data pulls the distribution of the counts of KSI away from a Poisson distribution, as does the large quantity of counts greater than approximately 7. Thus, applying a standard Poisson regression model to the data would have been expected to produce a poor fit between the model and the data.

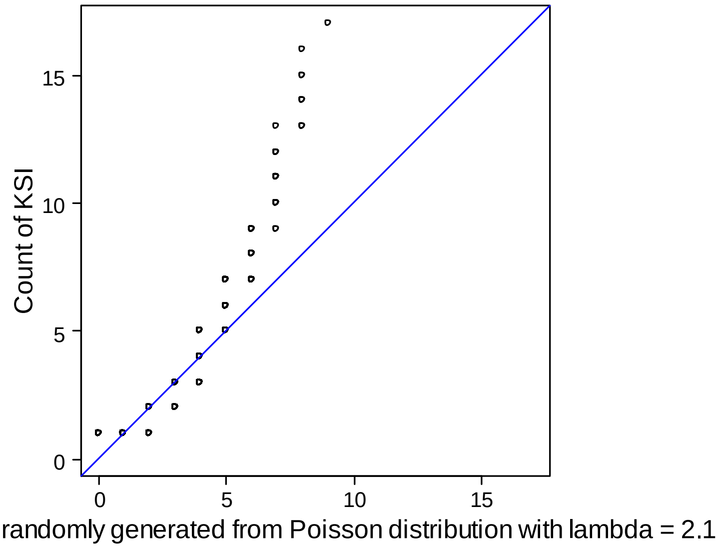

Figure 2: Quantile-Quantile plot of the count of KSI per demographic group against a theoretical Poisson distribution

Results

Model fit and overall results

Table 3 shows the overall results of the model fitting process. The empty model had an Akaike's Information Criterion (AIC) of 18,200 and -2 log-likelihood value of 18,196 on 12,672 degrees of freedom. Each group of variables made a significant improvement to the model fit when added ($p < .001$). The change in Akaike's Information Criterion (AIC) shows that after the confounding variables (model 1), the condition variables (model 5) had the greatest impact on the model fit, followed by the road condition variables (model 4) and the individual/driver variables (model 7). However, the diagnostic plots in Figure 3 and Figure 4 indicate that there is still room for improvement in the fit of the final model.

Table 3: Model evaluation

Model	Variables	AIC (Δ)	Likelihood Ratio Chi-square Test
1	+ Year + Age + Sex + ≥10% pedestrians + ≥10% cyclists + ≥10% motorcyclists	-1561.81	1581.8, p < .001
2	… + IRSD (Q5 = Most disadvantaged, Q1 = Least disadvantaged)	-134.88	142.9, p < .001
3	… + Remoteness (Major Cities; Inner Regional; Outer Regional, Remote and Very Remote)	-154.29	158.3, p < .001
4	… + ≥10% curves + ≥10% crests + ≥10% unsealed roads + speed limit group (≤40, 41 to 50, 51 to 70, 71 to 90 and 91 to 110 km/h)	-863.93	879.9, p < .001
5	… + ≥10% during high alcohol hours + ≥10% in the dark + ≥10% in the rain	-1006.31	1012.3, p < .001
6	… + Median vehicle age (<2, 2 to <3, 3 to <4, 4 to <5, 5 or more years) + Median estimated impact speed (≤40, 41 to 50, 51 to 70, 71 to 90, 91 to 110, and 110+ km/h)	-74.06	82.1, p < .001
7	… + ≥10% probationary driver + ≥10% drunk driver + ≥10% unlicensed driver + ≥10% not wearing seatbelt or helmet + ≥10% inattention	-660.44	680.4, p < .001

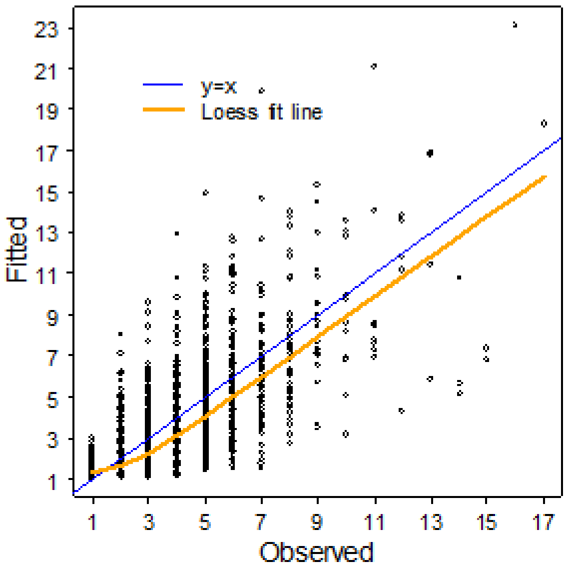

Figure 3: Quantile-Quantile plot comparing the distributions of the predicted KSI with the observed KSI

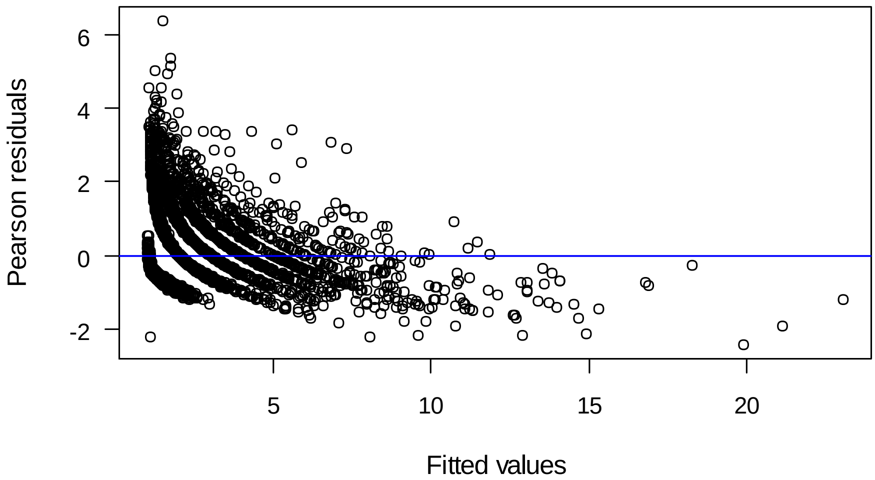

Figure 4: Raw Pearson residuals by fitted values

Risk ratios

Table 4 shows that all of the confounding variables made significant contributions to the initial model, with age group having the greatest effect on KSI risk.

When IRSD was entered at model 2, lower IRSD score was associated with higher KSI risk. However, when remoteness was added at model 3, the gaps between the lower (Q2 to Q5) and the highest (Q1) socioeconomic groups reduced.

The risk ratios for Q2 to Q5 reduced further when road environment variables (including speed limit) were added at model 4, and reduced again for all quintiles (relative to Q1) at each subsequent model. Of the road environment variables entered at model 4, speed limit made the greatest contribution to crash risk. Groups with a median speed limit more than 50 km/h had approximately three times the KSI risk of groups with a median speed limit of 40 km/h or less. This effect reduced slightly when other explanatory variables were added in subsequent models, but remained approximately twice the risk in the final model.

Curves and unsealed roads were associated with a two-fold increase in risk, reducing to a 30% increase in risk after other variables were taken into account in subsequent models.

After the confounding variables (year, age, sex and road user groups), the condition variables (darkness, rain and high alcohol hours) made the biggest contribution to the model when added. This is mostly due to the influence of groups for whom at least 10% of KSI occurred during high alcohol hours. The risk for these groups was 2.2 as high as the risk for other groups when first entered, and reduced to 1.8 times in subsequent models.

After controlling for all other variables in the model, groups with probationary drivers accounting for at least 10% of KSI had a KSI risk of 1.3, groups with at least 10% of KSI involving unlicensed drivers had a KSI risk of 1.2 and groups with at least 10% of KSI found not to have been wearing a seatbelt had a KSI of 1.5.

Table 4: Exponents of the beta paramaters (risk ratios)

	Null model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Mu(NB)	1.001*	3.32E+47*	8.03E+47*	1.62E+46*	7.20E+32*	1.19E+23*	5.64E+23*	3.7E+2*
Size(NB) (disp. par.)	1.001*	0.002*	0.002*	0.003*	0.005*	0.013*	0.013*	0.018*
Year		0.943*	0.943*	0.944*	0.958*	0.969*	0.968*	0.968*
Agegp: 0-14 (REF)		1	1	1	1	1	1	1
Agegp: 15-24		5.886*	6.158*	6.245*	4.846*	3.119*	3.063*	2.733*
Agegp: 25-39		3.768*	3.890*	3.901*	3.139*	2.095*	2.080*	2.021*
Agegp: 40-54		2.880*	3.009*	2.988*	2.545*	1.975*	1.958*	1.915*
Agegp: 55-69		0.920	0.955	0.984	1.048	1.063	1.062	1.097
Agegp: 70+		1.816*	1.942*	2.065*	2.125*	2.227*	2.221*	2.240*
Males		1.573*	1.574*	1.558*	1.374*	1.181*	1.193*	1.166*
Pedestrians		2.695*	2.481*	2.561*	2.696*	1.858*	1.899*	1.520*
Cyclists		1.984*	2.128*	2.288*	2.199*	2.158*	2.147*	2.092*
Motorcyclists		2.447*	2.321*	2.305*	1.583*	1.379*	1.407*	1.272*
Q5 (REF)			1	1	1	1	1	1
Q4			1.121*	0.986	0.979	0.916*	0.902*	0.910*
Q3			1.330*	1.276*	1.317*	1.148*	1.127*	1.095*
Q2			1.519*	1.361*	1.277*	1.080*	1.052	0.992
Q1 - Most disadv.			1.586*	1.450*	1.356*	1.131*	1.103*	1.011
Major cities (REF)				1	1	1	1	1
Inner regional				1.721*	1.633*	1.687*	1.672*	1.665*
Outer, remote & v. remote				1.373*	1.647*	1.898*	1.914*	1.976*
>10% at curves					1.867*	1.425*	1.419*	1.248*
>10% at crests					1.418*	1.246*	1.194	1.017
>10% at slopes					1.368*	1.199*	1.195*	1.122*
>10% on unsealed rds					1.644*	1.436*	1.428*	1.276*
SpdLimGp: 0-40 km/h (REF)					1	1	1	1
SpdLimGp: 41-50 kmh					1.868*	1.723*	1.675*	1.589*
SpdLimGp: 51-70 kmh					3.134*	2.582*	2.432*	2.227*
SpdLimGp: 71-90 kmh					3.075*	2.740*	2.617*	2.465*
SpdLimGp: 91-110 kmh					2.680*	2.546*	2.451*	2.316*
>10% during hi alc hrs						2.164*	2.114*	1.836*
>10% in the dark						1.198*	1.183*	1.095*
>10% in the rain						1.231*	1.230*	1.107*
VehAgeGp: 0-<2 yrs (REF)							1	1
VehAgeGp: 2-<3 yrs							1.502*	1.472*
VehAgeGp: 3-<4 yrs							1.735*	1.687*
VehAgeGp: 4-<5 yrs							1.614*	1.559*
VehAgeGp: 5+ yrs							1.232*	1.210
VehSpdGp: 0-40 km/h (REF)								1
VehSpdGp: 41-50 kmh								1.116*
VehSpdGp: 51-70 kmh								1.060
VehSpdGp: 71-90 kmh								1.017
VehSpdGp: 91-110 kmh								0.969
VehSpdGp: >110 kmh								0.955
>10% probationary drivers								1.303*
>10% drunk drivers								1.205*
>10% unlicensed drivers								1.216*
>10% not wearing seatbelt								1.234*
>10% inattention								1.478*

* $P < .05$
(REF) denotes the reference group for a categorical variable

Discussion

There was a positive association between socioeconomic disadvantage of an area and KSI risk after controlling for age, gender, year and vulnerable road users. This association remained statistically significant for the three most disadvantaged quintiles after controlling for remoteness.

However, the addition of further explanatory risk factors mediated the apparent relationship. In particular, road environment variables such as higher speed limits, curves and unsealed roads, and condition variables including high alcohol hours resulted in noticeable changes in the relationship between disadvantage and KSI risk.

This supports the hypothesis that groups from lower socioeconomic areas do have higher KSI risk. However, the reduction in risk ratios that occurred when explanatory variables were added suggests that the association can mostly be explained by road environment, conditions and alcohol.

The preliminary finding of an association between socioeconomic status and risk is consistent with previous studies both in Australia (Clapperton & Cassell 2010; Chen et al. 2010) and overseas (RoSPA 2012; WHO 2012).

The finding that road environment variables such as speed limit, curves and unsealed roads are important contributing factors that may mediate the relationship between socioeconomic status and risk is also consistent with findings from the UK, USA and Canada (Dissanayake, Aryaija & Wedagama 2009; Jones et al. 2008; Huang Abdel-Aty & Darwiche 2010; Morency et al. 2012).

The mediating influences of unlicensed drivers and riders, alcohol, and inexperience (captured by age and probationary licence status) are also consistent with previous studies (e.g. Braver 2003; Leal, King & Lewis 2006; Senserrick et al. 2010).

The zero-truncated model may prove to be a useful tool for road safety researchers, as information that is pertinent to crashes is often only collected when a crash occurs, and in many jurisdictions, information about involved persons is only available for people who have been killed or injured.

A limitation of the modelling approach was that detailed person-level information had to be aggregated to allow it to be paired with essential information that was only available at the population level (remoteness and socioeconomic status). It would be valuable to explore the possibility of a multilevel model to overcome this limitation.

Conclusions

The results provided support for the hypothesis that groups from lower socioeconomic areas do have higher KSI risk. However, the association was largely explained by the road and condition variables (including high alcohol hours) entered in subsequent models. These mediating variables indicate potential areas for intervention.

It is important to remember that the aggregate characteristics of groups, and the relationships observed between more and less disadvantaged groups in this study cannot be assumed to hold true for individuals. Individuals can and do differ from the group average.

Acknowledgements

The initial stage of this project was completed as part of an Austroads project at ARRB. The latter stages were completed as part of my Masters thesis under the supervision of Dr Lyndon Walker at Swinburne University.

References

ABS 2007, 'SLA 2006 from POA 2006', data cube from *ABS Postal Area Concordances, August 2006*, released 2 July 2007, accessed 3 January 2013, <http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/2905.0.55.001Aug%202006?OpenDocument>.

ABS 2008a, 'SEIFA, Statistical Local Areas, Data Cube only, 2006', data cube from *Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia*, Cat.No. 2033.0.55.001, released 26 March 2008, accessed 3 January 2013, <http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/2033.0.55.0012006?OpenDocument>.

ABS 2008b, 'Explanatory Notes' from *Population by Age and Sex, Australia, 2006*, Cat. No. 3235.0, updated 18 August 2008, accessed 10 September 2013, <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/3235.0Explanatory%20Notes12006?OpenDocument>.

ABS 2013a, 'RA from 2006 SLA Correspondence', data cube from Australian Standard Geographical Classification (ASGC) Remoteness Area Correspondences, 2006, Cat. No. 1216.0.15.003, released 20 December 2011, accessed 3 January 2013, <http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1216.0.15.0032006?OpenDocument>.

ABS 2013b, 'Population Estimates by Age and Sex, Australia by Statistical Geography [ASGC 2011], 2001 to 2011', SRD data cube from *Population by Age and Sex, Regions of Australia*, Cat. No. 3235.0, released 30 August 2013, accessed 10 September 2013, <http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3235.02012?OpenDocument>.

ATC 2011, *National Road Safety Strategy*, Department of Infrastructure and Regional Development, Canberra.

Braver, ER, 2003 Race, Hispanic origin, and socioeconomic status in relation to motor vehicle occupant death rates and risk factors among adults, *Accident Analysis and Prevention, 35*, 395-309.

Cameron, AC & Trivedi, PK 1998, *Regression Analysis of Count Data*, Cambridge University Press, Cambridge, UK.

Chen, HY, Senserrick, T, Martiniuk, ALC, Ivers, RQ, Boufous, S, Chang, HY & Norton R 2010, 'Fatal crash trends for Australian young drivers 1997-2007: Geographic and socioeconomic differentials', *Journal of Safety Research*, vol. 41, pp. 123-128.

Clapperton, A & Cassell, E 2010, 'The impact of socioeconomic inequity on serious injury in Victoria', *Hazard*, no. 70, Victorian Injury Surveillance Unit (VISU), Monash University, Clayton, Victoria.

Dissanayake, Aryaija & Wedagama 2009, 'Modelling the effects of land use and temporal factors on child pedestrian casualties', *Accident Analysis & Prevention*, vol. 41, pp. 1016-1024.

Hilbe, JM 2011, *Negative Binomial Regression*, 2nd Edition, Cambridge University Press, Cambridge, UK.

Huang, H, Abdel-Aty, MA & Darwiche, AL 2010, 'County level crash risk analysis in Florida: Bayesian spatial modeling', *Transportation Research Record: Journal of the Transportation Research Board*, No. 2148, pp. 27-37.

Jones, AP, Haynes, R, Kennedy, V, Harvey, IM, Jewell, T & Lea, D 2008, 'Geographical variations in mortality and morbidity from road traffic accidents in England and Wales', *Health and Place,* vol. 14, pp. 519-535.

Leal, N, King, M, & Lewis, I 2006, 'Profiling drink driving offenders in Queensland', in *Proceedings 2006 Australasian Road Safety Research, Policing and Education Conference,* Gold Coast, Australia.

Morency, P, Gauvin, L, Plante, C, Fournier, M & Morency, C 2012, 'Neighbourhood social inequalities in road traffic injuries: The influence of traffic volume and road design', American Journal of Public Health, vol. 102, no. 6, pp. 1112-1119.

Pink, B 2008, *Information Paper: An Introduction to Socio-Economic Indexes for Areas (SEIFA),* ABS Catalogue No. 2039.0, Australian Bureau of Statistics (ABS), Canberra, Australia.

R Development Core Team 2013, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Senserrick, T, Chen, HY, Boufous, S, Ivers, R, Stevenson, M & Norton, R 2010, 'Demographic factors associated with pre-licensed driving in a NSW young driver cohort: the DRIVE study', *2010 Australasian Road Safety Research, Policing and Education Conference,* 31 August – 3 September 2010, Canberra, Australian Capital Territory.

World Health Organisation (WHO) 2012, *Road traffic injuries,* fact sheet no. 358, September 2012, viewed 23 October 2012, <http://www.who.int/mediacentre/factsheets/fs358/en/>.

Yee, TW and Wild, CJ 1996, 'Vector Generalized Additive Models', *Journal of Royal Statistical Society*, Series B.

Yee, TW 2013, VGAM: Vector Generalized Linear and Additive Models, R package version 0.9-2, <http://CRAN.R-project.org/package=VGAM>.