A Multidisciplinary Investigation of the Influence of the Built Urban Environment on Driver Behaviour and Traffic Crash Risk

Karen Lesley Stephan
BSc(Hon), MPH, GradCertBiostats

A thesis submitted for the degree of Doctor of Philosophy at
Monash University in 2015
Monash University Accident Research Centre, Monash Injury Research Institute

COPYRIGHT NOTICE

ABSTRACT

Road trauma is a major source of death and disability which can be addressed by improving the safety of roads. Despite higher crash rates on urban roads than rural roads, less is known about the characteristics of the built urban environment (the road, surrounding environment and road user activity) that influence crash risk. It is essential to identify risk factors for crashes and to understand the underlying mechanisms driving risk before effective countermeasures can be developed. The overall aim of this thesis was to develop and apply a multidisciplinary approach to identify and understand the aspects of the built urban environment that influence crash occurrence.

Research Component 1 of this thesis sought to identify characteristics of the built environment that were associated with crashes on complex urban roads. Beyond the effect of traffic volume and intersections, there was a lack of strong evidence regarding the influence of the built urban environment on crash risk, in particular, the effect of the surrounding environment was largely neglected.

A comprehensive list was developed of characteristics of the built urban environment that were potential risk factors for crash occurrence. A cross-sectional study was conducted using a novel phased modelling approach. It identified that, in addition to traffic exposure and road design, the roadside environment and facilities and amenities were associated with the frequency of multi-vehicle, single-vehicle and pedestrian-vehicle crashes on strip shopping centre road segments in metropolitan Melbourne. Risk factors differed by crash type.

Research Component 2 of this thesis comprised a case study to demonstrate how behavioural research methods may be employed to investigate the behavioural mechanisms underlying crash risk. Driving simulation was used to investigate the effect of roadside parking (identified as a risk factor for multi-vehicle crashes in Component 1) and speed limit on driver behaviour.

Drivers chose a lane position further away from the kerb and weaved less within their lane as the number of cars parked on the roadside increased. Perceived risk, discomfort, task difficulty and effort also increased. Increasing the speed limit of the road caused increases in perceived risk, discomfort, task difficulty, mental effort and physical effort. An increase in speed beyond that preferred in a given parking

environment led to significant increases in the effort required to drive at that speed. No such threshold relationship with speed was discovered for ratings of risk, discomfort or task difficulty. As the number of parked cars varied, neither drivers' change in behaviour nor their choice of preferred speed served to maintain a stable level of risk or workload. Recommendations for countermeasures to address crash risk on urban roads with roadside parking were made.

This thesis demonstrates a rigorous scientific process for applying two complementary methodological approaches to identify risk factors for crashes and understand their mechanisms. The innovative contribution of this thesis was the synergistic combination of cross-sectional modelling and driving simulation to identify and further investigate risk factors for crashes. Implications for future road safety research and practice were discussed.

DECLARATION

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

[redacted] 1/12/2015

PUBLICATIONS DURING ENROLMENT

Stephan, K. & Newstead, S. (2011). Measuring the influence of the road and roadside in the safe system. *Proceedings Australasian Road Safety Research Policing Education Conference*, 6-8 November, Perth, Australia.

Stephan, K. & Newstead, S. (2012). Towards safer urban roads and roadsides: Factors affecting crash risk in complex urban environments. *Proceedings Australasian Road Safety Research Policing Education Conference*, 4-6 October, Wellington, New Zealand.

Stephan, K., Lenné, M.G, Newstead, S., Edquist, J. & Rudin-Brown, C. (2013). Using driving simulation to investigate behavioural mechanisms underpinning real world crash risk due to roadside parking. *Proceedings of the Road Safety and Simulation Conference*, 23-25 October, Rome, Italy.

Stephan, K.L & Newstead, S.N. (2014). Characteristics of the road and surrounding environment in metropolitan shopping strips: Association with the frequency and severity of single vehicle crashes. *Traffic Injury Prevention, 15*, S74-S80.

ACKNOWLEDGEMENTS

I begin by acknowledging my supervisors, Stuart Newstead and Michael Lenné for their guidance and advice. I am grateful for the trust and respect they showed in letting me work independently through this multidisciplinary PhD research. Thank-you also to Ian Johnston for being my co-supervisor at the start of my PhD prior to Stuart taking over when I changed research topics.

Thank-you to Stuart for becoming my PhD supervisor at a critical juncture, for always having time to discuss my project and for giving it his full attention when we met no matter what other projects required his attention. His open-door policy was much appreciated as were the late nights he spent reviewing thesis chapters.

Thank-you to Mike for his support throughout my candidature. I particularly valued his comments on my thesis that harked back to advice he had received from his PhD supervisor, the late Tom Triggs. Given that Tom's undergraduate psychology lectures in Human Factors and Statistics had such a strong influence on my career choice, I felt like I was receiving gems of wisdom from Tom himself.

I express my strong gratitude to the NRMA-ACT Road Safety Trust for awarding me a Postgraduate Research Scholarship and for their generosity and patience. Eddie Wheeler (the secretary) did a wonderful job organising the postgraduate scholarships showcase events that were held in Canberra for the awardees to present their research to a wide range of Australian road safety professionals. These were a highlight of my candidature. The Trust's generous support meant that I was able to attend a number of overseas conferences which were invaluable for allowing me to meet and learn from international experts in road safety. For over 20 years the NRMA-ACT Road Safety Trust has played an important role in providing support for research and training in road safety in Australia. Unfortunately, the Trust will soon cease to operate which is an enormous loss to the field in this country.

Thank-you to Peter Vulcan, the founding Director of MUARC, who was the scientific advisor for the NRMA-ACT Road Safety Trust and the external committee member for my confirmation and final reviews. His feedback was incisive and his endless enthusiasm for rigorous yet practical road safety research is inspiring. My thanks also go to Max Cameron for being a committee member for my final review. His advice has strengthened the thesis. For their academic advice and support, thank-you to Lesley Day, Rod McClure and Joan Ozanne-Smith and thank-you to Samantha Bailey, Glenda Cairns, Jude Charlton and Jennie Oxley for their assistance negotiating through the maze of PhD research administration.

My appreciation goes to Bruce Corben, Nimmi Candappa (MUARC) and the experts from the MUARC Baseline Project Advisory Committee (representatives of VicRoads, Transport Accident Commission, Department of Justice and Victoria police) who provided expertise in the planning stage of the thesis.

People from a number of organisations provided data or information regarding data sources. From VicRoads: Jas Bhatt, Graz Starc and Amy Stebbing. From ARRB Group Ltd: Blair Turner, Chris Jurewicz, Simon Barlow and the rating team. From MUARC: Jason Thompson, Hafez Alavi, Jim Scully, Laurie Budd and Linda Watson.

Acknowledgements to those people who provided assistance designing and conducting the driving simulation study. First, I express my appreciation to the participants who took part in the research. Thank-you to Ashley Verdoorn for programming the scenarios, Glenn Watson for writing the software so I could extract my data and Nebojsa Tomasevic for his advice. Thank-you to Eve Mitsopoulos-Rubens and Kristie Young for our discussions on simulator protocols and measures. Thank you to Jessica Edquist and Missy Rudin-Brown for assistance in initial design and piloting of scenarios, and to Jess for running several participants. Thank-you to Miranda Cornelissen for helping with the final piloting of the experiment. Thank-you to Amanda Stephens for (more recently) sharing your expertise regarding traffic psychology and models of driver behaviour.

I am grateful to my good friends who helped proof-read and format my thesis: Marilyn Johnson, Sean Smith, Di Magliano and Christine Mulvihill.

The PhD experience is enriched by the bond you share with your fellow students and colleagues. I thank them all for their support and intelligent conversation. For their friendship and the many entertaining distractions, special mention goes to: Carlyn Muir, Christina Ekegren, Christine Mulvihill, Lyndal Bugeja, Marilyn Johnson, Megan Bohensky, Shannon Gray, Jason Thompson, Angela Clapperton, Virginie Etienne, Irene Castelnovi, Jessica Killian, Trang Vu, Adam McKinnon, Gemma Read, Hafez Alavi, Maggie Trotter and Miranda Cornelissen.

Enormous thanks to all my family and friends who were my link to life outside the PhD and who helped provide much-needed balance between work and play. Thank-you to Samantha Rizak and my brother, Jim, for hosting me for interstate writing retreats. Last but definitely not least, much love and gratitude to my parents for the opportunities they gave me and for being a constant source of support.

TABLE OF CONTENTS

Copyright notice i
Abstract ii
Declaration iv
Publications during enrolment v
Acknowledgements vi
Chapter 1. Introduction 1
1.1 Focus and structure of the thesis 1
1.2 Research & countermeasure development cycle for injury prevention. 3
1.3 The public health burden from road trauma 4
1.3.1 Improving the safety of roads for road users 5
1.3.2 The case for a multidisciplinary approach 6
1.3.3 Urban road safety 8
1.4 Conceptual frameworks for road safety and crash risk 11
1.4.1 Defining crash phases and system components: Haddon's matrix. 13
1.4.2 Measuring risk: Road trauma chain 14
1.4.3 Driver behaviour and crash risk 18
1.5 Theories of driver performance and behaviour 21
1.5.1 Driving tasks and risk 22
1.5.2 Information processing approach: performance 23
1.5.3 Motivational (cognitive) approach: behaviour 23
1.6 Conceptual framework for the thesis 27
1.7 Aims 29
Chapter 2. Methods for identifying risk factors 31
2.1 Study designs 31
2.1.1 Descriptive study designs 31
2.1.2 Analytical observational study designs 32
2.1.3 Interventional study designs 35
2.2 Interpreting study results 37
2.3 Data analysis methods 40
2.3.1 Multiple regression 40
2.3.2 Non-parametric modelling approaches 50
2.4 Implications for this study 51
Chapter 3. The built urban environment and crash frequency 53
3.1 Potential risk factors 54
3.2 Literature search 55
3.3 Methodological review of previous research 56
3.3.1 Studies at the level of the road segment or corridor 57
3.3.2 Studies at the level of the area 67
3.3.3 Summary 71
3.4 Relationship between the urban environment and crashes 72
3.4.1 Exposure 74
3.4.2 Roadway 83
3.4.3 Roadside 99
3.4.4 Land use and amenities 101
3.4.5 Sociodemographic 107
3.5 Implications for this study 114
3.6 Aims of Research Component 1 116
Chapter 4. Methods: Site selection and data 117
4.1 Site selection 117
4.1.1 Definition of a strip shopping centre 118
4.1.2 Identification and selection of study road segments 119

4.2 Data requirements ... 121
4.2.1 Outcomes ... 121
4.2.2 Identification of potential risk factors ... 122
4.3 Data sources and methods of collection ... 126
4.3.1 Existing data ... 126
4.3.2 Data collected specifically for the project ... 126
4.4 Data items and collection methods ... 129
4.4.1 Outcome data ... 129
4.4.2 Exposure ... 130
4.4.3 Speed zone ... 133
4.4.4 Road cross-section ... 136
4.4.5 Accesses and intersections ... 139
4.4.6 Medians ... 142
4.4.7 Road type and traffic mix ... 143
4.4.8 Roadside parking ... 144
4.4.9 Public transport and bicycle facilities ... 146
4.4.10 Pedestrian facilities ... 148
4.4.11 Roadside development ... 149
4.4.12 Road Pavement condition ... 150
4.4.13 Height Clearance ... 151
4.4.14 Static law enforcement ... 151
4.4.15 Amenities/facilities ... 152
4.4.16 Sociodemographic data ... 154
4.5 Data availability and implications for this thesis ... 156
Chapter 5. Methods: Data analysis ... 167
5.1 Outcomes and units of analysis ... 167
5.2 Model development ... 168
5.2.1 Phased model building approach ... 168
5.2.2 Model diagnostics ... 173
5.3 Summary of data analysis ... 180
Chapter 6. Results: Identification of risk factors ... 181
6.1 Descriptive statistics ... 181
6.2 Multi-vehicle crashes ... 182
6.2.1 Selecting the optimal method for specifying exposure ... 182
6.2.2 Model development and diagnostic tests ... 184
6.2.3 Model results and interpretation ... 192
6.3 Single-vehicle crashes ... 204
6.3.1 Selecting the optimal method for specifying exposure ... 204
6.3.2 Model development and diagnostic tests ... 205
6.3.3 Model results and interpretation ... 210
6.4 Pedestrian vehicle crashes ... 216
6.4.1 Selecting the optimal method for specifying exposure ... 216
6.4.2 Model development and diagnostic tests ... 217
6.4.3 Model results and interpretation ... 221
6.5 Overall summary: characteristics of the built environment associated with multi-vehicle crashes, single-vehicle crashes and pedestrian-vehicle crashes ... 228
Chapter 7. Discussion: Identifying risk factors ... 234
7.1 Achievement of primary aims ... 234
7.2 Summary and interpretation of results ... 235
7.2.1 Road user exposure and traffic distribution ... 235
7.2.2 Roadway ... 236
7.2.3 Roadside ... 249
7.2.4 Amenities and facilities ... 251

7.3 Methodological issues ... 253
7.3.1 Strengths ... 253
7.3.2 Limitations ... 255
7.4 Implications for future research and practice ... 260
Chapter 8. The built environment and driver behaviour ... 263
8.1 Behavioural research methods ... 264
8.1.1 Self-reported behaviour ... 264
8.1.2 Objective measurement of behaviour ... 265
8.1.3 Summary and implications for this study ... 269
8.1.4 Measurement of behaviour in the driving simulator ... 273
8.2 Selection of a risk factor for further investigation ... 279
8.2.1 Essential criteria ... 279
8.2.2 Desirable criteria ... 284
8.2.3 Choice of risk factor ... 286
8.3 Roadside parking: crash risk and driver behaviour ... 287
8.3.1 Crash risk ... 287
8.3.2 Observational studies ... 287
8.3.3 Experimental studies ... 288
8.4 Aims ... 291
Chapter 9. Methods: Driving simulator study ... 292
9.1 Participants ... 292
9.1.1 Inclusion criteria ... 292
9.1.2 Exclusion criteria ... 292
9.1.3 Recruitment ... 292
9.1.4 Participant sample ... 293
9.2 Materials ... 293
9.2.1 Equipment ... 293
9.3 Design ... 293
9.3.1 Scenario design ... 295
9.4 Procedure ... 297
9.5 Measures ... 298
9.5.1 Self-report ... 298
9.5.2 Driving performance ... 298
9.5.3 Response to secondary task ... 299
9.5.4 Response to unexpected safety-critical event ... 299
9.6 Data analysis ... 299
9.6.1 Self-reported risk, discomfort, task difficulty, effort and objectively measured driving performance ... 299
9.6.2 Reaction time to secondary task ... 299
9.6.3 Reaction time to safety-critical unexpected event ... 302
Chapter 10. Results: Driving simulator study ... 303
10.1 Self-reported measures ... 303
10.1.1 Self-reported risk ... 304
10.1.2 Self-reported discomfort ... 305
10.1.3 Self-reported task difficulty ... 306
10.1.4 Self-reported effort ... 307
10.1.5 Self-reported preferred and maximum travel speeds ... 309
10.2 Driving performance ... 311
10.2.1 Lateral position and control ... 311
10.2.2 Longitudinal control ... 313
10.3 Relationship between driving performance and self-reported effort 317
10.4 Secondary task performance ... 317
10.4.1 Analysis approach 1: GEE using mean correct reaction times ... 318
10.4.2 Analysis approach 2: Multiple time to event analysis ... 319

10.4.3 Distance travelled between stimulus onset and reaction 320
10.4.4 Comparison of results from GEE and multiple time to event analysis 321
10.5 Reaction time to the safety-critical unexpected event............................322
Chapter 11. Discussion: Understanding risk factors 324
11.1 Statement of results 324
11.2 Methodological strengths and limitations 328
11.3 Future research..331
11.4 Implications...332
Chapter 12. Summary, conclusions and implications 335
REFERENCES 341
Appendices 355
Appendix A...356
Appendix B...357
Appendix C..... 367

LIST OF TABLES

Table 1.1 Crash rates per million vehicle km travelled on Victorian road midblocks, by crash severity and area (data used with permission from Turner & Tziotis, 2006) 8
Table 1.2. Haddon's matrix, including targets for injury countermeasures, as applied to road trauma 14
Table 3.1 Risk factors (by broad category) considered in studies included in review of risk factors for crashes on urban roads: studies with the road segment as the unit of analysis 61
Table 3.2 Risk factors (by broad category) considered in studies included in review of risk factors for crashes on urban roads: studies with the area as the unit of analysis 69
Table 3.3 Direction of association and weight of evidence for relationship between exposure-related risk factors and urban multi-vehicle crashes (MVC), single-vehicle crashes (SVC), pedestrian-vehicle crashes (PVC) and bicycle-vehicle crashes (BVC) 82
Table 3.4 Direction of association and weight of evidence for relationship between roadway-related risk factors and urban multi-vehicle crashes (MVC), single-vehicle crashes (SVC), pedestrian-vehicle crashes (PVC) and bicycle-vehicle crashes (BVC) 96
Table 3.5 Direction of association and weight of evidence for relationship between roadside-related risk factors and urban multi-vehicle crashes (MVC), single-vehicle crashes (SVC), pedestrian-vehicle crashes (PVC) and bicycle-vehicle crashes (BVC) 100
Table 3.6 Direction of association and weight of evidence for relationship between land use and amenities and urban multi-vehicle crashes (MVC), single-vehicle crashes (SVC), pedestrian-vehicle crashes (PVC) and bicycle-vehicle crashes (BVC) 105
Table 3.7 Direction of association and weight of evidence for relationship between sociodemographic-related risk factors and urban multi-vehicle crashes (MVC), single-vehicle crashes (SVC), pedestrian-vehicle crashes (PVC) and bicycle-vehicle crashes (BVC) 112
Table 4.1 Characteristics of the built environment with the potential to influence crash risk, by broad category 124
Table 4.2. Definition of crash types 130
Table 4.3. Traffic exposure measures: Formulae and median, interquartile range (IQR) and range 131
Table 4.4. Speed zone changes that occurred on study road segments during the study period 135
Table 4.5 Distribution of speed zones across study road segments 135
Table 4.6 Distribution of carriageway width (mode, minimum and maximum) across study road segments 136
Table 4.7 Distribution of lane width (mode, minimum and maximum) across study road segments 137
Table 4.8 Distribution of mid-line types across study road segments 138
Table 4.9 Median, interquarterile range (IQR) and range of the number of signalised intersections per km across study road segments 139
Table 4.10 Median, interquarterile range (IQR) and range of the number of unsignalised intersections per km across study road segments 140
Table 4.11 Distribution of service roads (by type) across study road segments 140
Table 4.12 Median, interquarterile range (IQR) and range of the number of dedicated turning lanes per km across study road segments 141

Table 4.13 Distribution of median types across study road segments....142
Table 4.14 Distribution of mode, minimum and maximum median widths across study road segments....142
Table 4.15 Distribution of roadside parking configurations across study road segments....145
Table 4.16 Distribution of presence and type of clearway across study road segments....145
Table 4.17 Median, interquarterile range (IQR) and range of the number of different types of bus stops per km across study road segments....146
Table 4.18 Presence of tram routes, types of tram lanes and median, interquarterile range (IQR) and range of the number of different types of tram stops per km across study road segments....147
Table 4.19 Median, interquarterile range (IQR) and range of the number of different types of pedestrian crossings per km across study road segments....148
Table 4.20 Median, interquarterile range (IQR) and range of the number of different types of poles and trees per km across study road segments....149
Table 4.21 Distribution of offset distance between road edge and buildings (mode, minimum and maximum) across study road segments....150
Table 4.22 Categories of liquor licences and median, interquarterile range (IQR) and range of the number of different types liquor licences per km across study road segments....153
Table 4.23 Distribution of different types of amenities and facilities across study road segments....154
Table 4.24 Median, interquarterile range (IQR) and range of sociodemographic characteristics across study road segments....155
Table 4.25 Median, interquarterile range (IQR) and range of vehicle ownership and distribution of IRSAD deciles across study road segments....156
Table 4.26 Characteristics of the built environment with the potential to influence crash risk on urban midblock road segments that were and were not included in this research: data items, availability, measurement level and data sources/collection methods....160
Table 5.1 Built environment characteristics categorised into higher order classes....169
Table 6.1 Frequency and rate of crashes on Melbourne metropolitan strip shopping segments, 2005 to 2009, by crash type....182
Table 6.2 Comparison of different ways to specify traffic volume and segment length for modelling multi-vehicle crashes....183
Table 6.3 Comparison of different ways to specify traffic density and segment length for modelling multi-vehicle crashes....184
Table 6.4 Fit statistics for models predicting multi-vehicle crash (MVC) frequency....186
Table 6.5 Risk factors for multi-vehicle crashes (MVC) for models developed with traffic volume and traffic density ($p<0.01$)....193
Table 6.6 Additional risk factors for multi-vehicle crash (MVC) frequency ($p<0.05$ or $p<0.10$)....196
Table 6.7 Comparison of different ways to specify traffic volume and segment length for modelling single-vehicle crashes....205
Table 6.8 Comparison of different ways to specify traffic density and segment length for modelling single-vehicle crashes....205
Table 6.9 Risk factors for single-vehicle crashes....211
Table 6.10 Comparison of different ways to specify traffic volume and segment length for modelling pedestrian-vehicle crashes....216
Table 6.11 Comparison of different ways to specify traffic density and segment length for modelling pedestrian-vehicle crashes....217
Table 6.12 Risk factors for pedestrian-vehicle crashes....223

Table 6.13 Risk factors for multi-vehicle crashes (MVC), single-vehicle crashes (SVC) and pedestrian-vehicle crashes (PVC) 229
Table 7.1 The three most commonly occurring types of crashes for multi-vehicle crashes (MVC), single-vehicle crashes (SVC) and pedestrian-vehicle crashes (PVC) 255
Table 8.1 Applicable driver behaviour research methods for measuring the effect of built environment on observable driver behaviour: by category of risk factor 271
Table 10.1 Correlation between self-report measures: perceived risk, discomfort, task difficulty and effort 303
Table 10.2 Hazard ratios comparing the reaction times for different speed zones 319
Table 10.3 Hazard ratios comparing the reaction times for different parking conditions 320
Table A.1 Glossary of abbreviations 356

LIST OF FIGURES

Figure 1.1 Research and countermeasure development cycle for injury prevention........3
Figure 1.2 Victorian rural road environment, photograph used with permission from Hillard, P. 9
Figure 1.3 Victorian urban road environment (strip shopping zone), photograph used with permission from Edquist, J. 10
Figure 1.4. Road Trauma Chain: Pre-Crash and Crash. Adapted with permission from *Accident Data Analysis to Develop Target Groups for Countermeasures. Volume 1: Methods and Conclusions*, by Cameron, 1992 16
Figure 1.5. Road Trauma Chain: Crash and Post-Crash. Adapted with permission from *Accident Data Analysis to Develop Target Groups for Countermeasures. Volume 1: Methods and Conclusions*, by Cameron, 1992 17
Figure 1.6 Framework for safety and mobility, by Schepers et al., 2014, in Accident Analysis & Prevention, used under CC BY-NC-ND 3.0, references to sections of source paper removed from figure. 20
Figure 1.7 Conceptual framework for thesis: road system, surrounding environment, road user behaviour and crash risk 28
Figure 2.1 Poisson distribution: Probability of x (predicted values) for different mean values 46
Figure 4.1. Melbourne metropolitan area with the road segments included in this thesis highlighted in blue (map used with permission from Lawrence, B., created using ArcGIS® software by ESRI) 121
Figure 6.1 Probability that the number of multi-vehicle crashes (MVC)=x: observed proportion and predicted proportion from the MVC model with traffic volume 186
Figure 6.2 Probability that the number of multi-vehicle crashes (MVC)=x: observed proportion and predicted proportion from the MVC model with traffic density 187
Figure 6.3 Quantiles of the Anscombe residual plotted against quantiles of the normal distribution for the multi-vehicle crash model with traffic volume 188
Figure 6.4 Quantiles of the Anscombe residual plotted against quantiles of the normal distribution for the multi-vehicle crash model with traffic density 189
Figure 6.5 Locally weighted smoothed scatterplot of the Anscombe residual against the predicted mean number of multi-vehicle crashes (MVC) for MVC Model 1 (with traffic volume) and MVC Model 2 (with traffic density) 190

Figure 6.6 Probability that the number of single-vehicle crashes (SVC)=x: observed proportion and predicted proportion 206
Figure 6.7 Quantiles of the adjusted deviance residual plotted against quantiles of the normal distribution 208
Figure 6.8 Locally weighted smoothed scatterplot of the adjusted deviance residual against the predicted mean number of single vehicle crashes (SVC) 209
Figure 6.9 Probability that the number of pedestrian-vehicle crashes (PVC)=x: observed proportion and predicted proportion 218
Figure 6.10 Quantiles of the Anscombe residual plotted against quantiles of the normal distribution 219
Figure 6.11 Locally weighted smoothed scatterplot of the Anscombe residual against the predicted mean number of pedestrian-vehicle crashes (PVC) 220
Figure 7.1 Four-lane urban arterial in Melbourne with shared tram lane and roadside parking, photograph used with permission from Edquist, J. 245
Figure 7.2 Four-lane urban arterial in Melbourne with bicycle lane that shares space with roadside parking, photograph used with permission from Smith, S. 246
Figure 8.1 Methods for investigating the influence of the environment on road user behaviour 265
Figure 8.2 Behavioural research methods applicable for investigating driver behaviour in different stages of the research and countermeasure development cycle for injury prevention 273
Figure 8.3 Relationship between task demand, mental workload and performance. Reprinted with permission from *The Measurement of Drivers' Mental Workload* (p. 24), by de Waard, 1996 275
Figure 9.1 Experimental conditions: parking scenarios 295
Figure 10.1 Mean (95% confidence interval) ratings of risk (out of 20) across speed and parking conditions 304
Figure 10.2 Mean (95% confidence interval) ratings of discomfort (out of 20) across speed and parking conditions 306
Figure 10.3 Mean (95% confidence interval) ratings of task difficulty (out of 20) across speed and parking conditions 307
Figure 10.4 Mean (95% confidence interval) ratings of effort (maximum 150) across speed and parking conditions 309
Figure 10.5 Mean (standard error) self-reported preferred speed and maximum speed for each parking condition 311
Figure 10.6 Mean (standard error) lane position across parking conditions (zero= middle of the lane of travel, negative values indicate a position closer to the kerb and positive values indicate a position closer to the centre-line) 312
Figure 10.7 Mean (standard error) standard deviation of lane position across parking conditions 313
Figure 10.8 Mean difference (standard error) between the mean travel speed and the speed limit for each parking condition 314
Figure 10.9 Mean difference (standard error) between the mean travel speed and the speed limit for each speed zone 315
Figure 10.10 Standard deviation of speed (standard error) across parking conditions and speed zones 316
Figure 10.11 Standard deviation (standard error) of accelerator pressure by speed zone 316
Figure 10.12 Median (interquartile range) reaction time to peripheral stimulus by speed zone and parking conditions (incorrect responses excluded) 318
Figure 10.13 Distance travelled during median reaction time to peripheral stimulus if travelling at the speed limit 321
Figure 10.14 Median reaction time (interquartile range) to unexpected safety-critical event 323

Figure 12.1 Modified research and countermeasure development cycle for injury prevention showing research methods appropriate for each stage.................... 339

CHAPTER 1. INTRODUCTION

1.1 Focus and structure of the thesis

The research conducted for this thesis involved an investigation of the influence of the built environment (including the road, roadside and human activity) on casualty crash risk and driver behaviour in complex urban environments. The research applied a multidisciplinary approach with two research components that employed different, but complementary, methodological approaches.

In the first research component, statistical models were developed to identify characteristics of the built environment that were associated with casualty crash frequency on strip shopping centre road segments (an example of a highly complex urban environment) in metropolitan Melbourne, Australia. This component addressed the research question: What characteristics of the built urban environment are associated with casualty crash risk?

The second research component was an experiment conducted in Monash University Accident Research Centre's driving simulator to investigate the effect of modifying the road environment on driver behaviour. This component addressed the question: What effect do the built environment risk factors associated with crashes have on driver behaviour?

The thesis is structured as follows. The remainder of Chapter 1 provides a background and rationale for the research and introduces the broad aims of the study.

Research conducted for Component 1 (the statistical modelling component) is presented in Chapters 2 to 7. Chapter 2 is a methodological review of the study designs and analysis methods that have been used to investigate the relationship between the built environment and crash risk that informs the choice of methods and analyses to be used in this thesis. In Chapter 3, previous multivariable studies of the built environment and crashes in urban areas, controlling for traffic exposure, are critically reviewed to identify gaps in knowledge. The specific aims of Component 1 of the thesis are presented at the end of Chapter 3. Chapters 4 and 5 present the methods and data analysis plan for the research conducted for Component 1 of this thesis. The aspects of the built environment that were found to be associated with crashes on strip shopping centre road segments in metropolitan Melbourne are presented in Chapter 6 and discussed in Chapter 7 along with the limitations and implications of the research.

The research conducted in Component 2 (the behavioural research component) is presented in Chapters 8 through 11. In Chapter 8, methods for studying the effect of the road environment on road user behaviour are reviewed and driving simulation is chosen as the method to be used in this thesis. The rationale for the choice of risk factor (from those identified in Component 1) to be investigated in terms of its effect on driver behaviour is presented. Previous studies of this risk factor and its effect on driver behaviour are critically reviewed and the aims of the study conducted for Component 2 are presented. Chapter 9 outlines the methods used, followed by the results (Chapter 10) and discussion (Chapter 11).

Finally, Chapter 12 draws together the main findings of the thesis. The contribution of the thesis to developing a multidisciplinary approach to studying the influence of the built environment on crash risk and driver behaviour is highlighted and recommendations for how to incorporate such an approach in future research are made. A glossary of the abbreviations used in this thesis is provided in Appendix A.

The remainder of Chapter 1 is organised as follows. First, the research and countermeasure development cycle for injury prevention is presented, including where this thesis sits within the cycle (Section 1.2). In Section 1.3, the size of the public health burden from road trauma is described and the design of the road and roadside is identified as a key target for improving road safety. Understanding the effect of the environment on road user behaviour is underscored as a critical requirement to understanding crash risk and for effective countermeasure development. The lack of knowledge about the characteristics of the environment that affect crashes in complex urban areas is also outlined. In Section 1.4, three conceptual frameworks for road safety are reviewed with respect to their utility for conceptualising the road, roadside and road user behaviour within the environment and their interactive contribution to crash risk. A major gap in the conceptual frameworks is identified in terms of the contribution to risk of road user behaviour during the trip. Theories of driving performance and behaviour with relevance to the influence of the environment on behaviour and risk are reviewed in Section 1.5. The three frameworks are then integrated and augmented with an explicit consideration of behaviour and crash risk to form the over-arching conceptual framework for the thesis (Section 1.6). Finally, the broad aims of the thesis are presented in Section 1.7.

1.2 Research and countermeasure development cycle for injury prevention

This thesis seeks to effectively identify and investigate risk factors for traffic crashes in order to inform development of future countermeasures. The research conducted for this thesis sits within the second phase of the research and countermeasure development cycle for injury prevention (Sleet, Hopkins, & Olson, 2003), which is a rigorous scientific framework for the investigation and remediation of public health problems, including traffic-related injury (refer to Figure 1.1). The phases in the cycle are described briefly below.

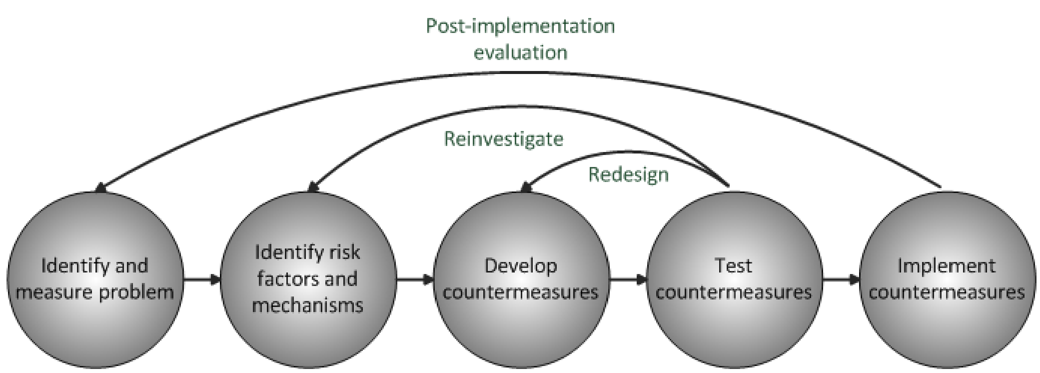


Figure 1.1 Research and countermeasure development cycle for injury prevention

The first step of the process concerns identifying the problem and measuring its size. As well as counting the number of crashes and injuries, this includes describing the problem, for example, who was injured (e.g. the age, sex and type of road users), the vehicles that were involved, and where the crash occurred (e.g. geographical locations, urban vs. rural, different road types, different road layouts).

The second step involves investigating the determinants of crashes. This includes both identifying risk factors and further investigation to understand the reasons, or mechanisms, behind the change in crash risk. Understanding why crash risk is increased is essential before countermeasures can be developed. A poor understanding of the reasons for a change in crash risk can lead to development of countermeasures that are ineffective, or worse, that have unintended consequences that increase crash risk. Risk factors and associated mechanisms can involve the road user, the vehicle, the environment, or interactions between these components, for example, how road users perform within a particular road environment. They can also come from various levels of the system, from road user behaviour, to road design and maintenance, to government policy.

Once risk factors and the underlying mechanisms are understood, countermeasures can be designed and tested. The process is not necessarily linear. If countermeasures are found to be ineffective, then they should be redesigned. It may be necessary to go back and re-investigate the problem and/or risk factors prior to redesign. Once a countermeasure has been tested and proven to be effective, it can be implemented more broadly and ongoing outcome evaluation performed in order to determine if the problem has been reduced (that is, crashes decreased). Therefore, this process represents a cycle, with ongoing surveillance of the size of the problem contributing to the evaluation of longer term effectiveness of interventions. The need for multidisciplinary effort in preventing crashes becomes apparent when the different levels for investigation and countermeasure development and implementation are considered.

By identifying aspects of the built urban environment that are associated with crash risk and seeking to determine how risk is increased, the research conducted as part of this thesis addresses the second step of the cycle for research into injury prevention. In the next section, the problem will be defined, in terms of the public health burden from road trauma, the need to improve the safety of roads and the rationale for focusing on urban crashes.

1.3 The public health burden from road trauma

Worldwide, traffic crashes are a major source of death and disability. In 2010, it was estimated that traffic crashes resulted in more than 1.33 million deaths and over

75 million healthy life years lost due to death and disability (Global Road Safety Facility & The World Bank, 2014). Traffic crashes are the eighth most common cause of death globally, and are predicted to rise to the fifth leading cause by 2030 (World Health Organization, 2013).

While much of the road trauma burden is in low to middle income countries (Global Road Safety Facility & The World Bank, 2014; World Health Organization, 2013), road safety is still an important problem in developed countries despite advances that have led to large reductions in fatalities over the last 40 years. In Australia, the road toll reduced from a peak of 3,798 in 1970 (30.4 per 100,000 population or 8.0 per 10,000 registered vehicles) to 1,367 in 2010 (6.1 per 100,000 persons or <1.0 per 10,000 registered vehicles), against a backdrop of a population that almost doubled in size and vehicle ownership rates that approximately quadrupled (Australian Bureau of Statistics, 2012a, 2012b; Lee, 2003). The number of serious injuries is much higher than the number of deaths: the most recent data show that 52,066 people were hospitalised as a result of a traffic crash in Australia in 2006 to 2007 (Henley & Harrison, 2009). While the total number of deaths and injuries has decreased over time, traffic crashes are still a significant source of death, injury and disability, with costs in Australia in 2006 of at least $27 billion per year estimated using the willingness to pay approach (Bureau of Infrastructure Transport and Regional Economics, 2009). Crashes are eminently preventable and even one death or serious injury is too many.

1.3.1 Improving the safety of roads for road users

Despite the reductions in road trauma it is recognised that there are still further great gains to be made by improving the safety of roads (Australian Transport Council, 2001), both for preventing crashes occurring and preventing and reducing injuries if a crash does occur (World Health Organization, 2010). The current Australian Road Safety Strategy (Australian Transport Council, 2011) has two aims that are particularly pertinent to safer roads: the assessment of risk for targeting high risk roads and setting speed limits appropriately for the road and roadside. For these reasons, the safety of roads is the broad area of research in this thesis.

Systematically identifying high risk road sections and locations presents an enormous challenge. In Victoria, Australia, the Transport Accident Commission Black Spot program treated high risk roads and locations that were identified according to their crash history. The program was successful in reducing crashes (Newstead &

Corben, 2001) however, the approach was reactive, not proactive, in that crashes must have already occurred for a location or route to be deemed a black spot. There are roads that do not (yet) have the required crash history to be classified as a black spot that can be deemed high risk due to the presence of known risk factors. In addition, blackspot treatments are generally chosen from a narrow range of traditional engineering based countermeasures to address the environment without fully understanding the breadth of factors affecting risk at the site. Thus, blackspot treatments may not be optimal. To achieve the aims of the Australian Road Safety Strategy, it is essential to be able to systematically identify high risk road sections and locations and the factors that drive this risk.

1.3.2 The case for a multidisciplinary approach

To successfully improve the safety of roads requires an appreciation of how road users act within the road environment. The interaction between the road user and the road environment and the impact on crash risk is an aspect of the road system which would benefit from targeted multidisciplinary research, rather than the current “silo” approach of considering road users, vehicles and environments separately. Rothengatter (1998) stated that “road design remains a key tool to accident prevention where understanding drivers’ perceptions, expectations, state and propensities can make a substantial impact” (p. 5). Roadways should be designed to support road users and reduce the probability and consequences of driver error (Theeuwes, 2001). This highlights an important role for the discipline of human factors, which is concerned with the design and evaluation of systems, taking into account the characteristics, needs, abilities and limitations of the human operator, who is considered a part of the system (Oppenheim & Shinar, 2011). This combined approach is a key component of the novelty of this thesis.

A large body of research has identified aspects of the road and roadside that were associated with crashes, particularly on highways and rural roads (Anastasopoulos & Mannering, 2009; Anastasopoulos, Mannering, Shankar, & Haddock, 2012; Anastasopoulos, Tarko, & Mannering, 2008; Chang, 2005; Gross & Donnell, 2011; Gross & Jovanis, 2007; Ivan, Wang, & Bernardo, 2000; Karlaftis & Golias, 2002; Lee & Mannering, 2002; Miaou, 1994; Milton & Mannering, 1998; Qin, Ivan, & Ravishanker, 2004). Although these studies have identified road-related risk factors for crash occurrence, they are powerless to describe the mechanisms behind the changes in crash risk (Elvik, 2006). On occasion, researchers propose changes in driver behaviour as post-hoc explanations of why certain aspects of the road environment increase risk but supporting evidence is usually limited to the authors' opinions or anecdotes and further scientific investigation is rare. This is because the studies to identify risk factors often approach the problem from a single-disciplinary focus rather than addressing the environment and behaviour together. It is concerning though, that researchers often recommend countermeasures after identifying risk factors, without fully understanding why risk is increased.

Behavioural research methods can be used to investigate how driver behaviour changes in different road environments, however human factors and traffic psychology researchers have not yet tapped into the vast literature from engineering and statistics that has identified aspects of the road environment that increase crash risk. There is, however, great potential to generate hypotheses about driver behaviour in the presence of identified risk factors that are based on accepted frameworks and theories of human behaviour rather than opinions or anecdotes. These hypotheses can then be tested in behavioural experiments or observational studies to shed light on the mechanisms by which risk is increased. Once the reasons behind increases in crash risk are determined, evidence-based countermeasures can be proposed and tested. There is therefore a pressing need for an approach that integrates both the initial identification of risk factors, followed by subsequent investigation of why risk is increased.

Evidence from previous research in terms of the aspects of the road environment that increase crash frequency and severity has been used to develop engineering standards and risk assessment tools for identifying high risk roads and prioritising treatments, such as the International Road Assessment Program (iRAP), The Australian Road Assessment Program, AusRAP (McInerney & Smith, 2009) and the Highway Safety Manual (AASHTO, 2010). Currently, engineering road design standards and risk assessment tools are based on a mixture of evidence and expert opinion in terms of the factors that are likely to make a road safe or unsafe. Often, it seems that road user behaviour is not explicitly considered, and this can mean that unintended effects might occur when these standards are applied in the real world. For example, the expected reduction in crashes might not eventuate if the road users change their behaviour to ameliorate the effects of the change in road design. Risk assessment tools and engineering standards have been criticised for not giving enough consideration to road user behaviour, particularly behavioural adaptation to road safety countermeasures (Hauer, 1999; Noland, 2013). Hence behavioural research methods should be part of the early process of investigating risk factors and developing countermeasures rather than simply being used to determine why a countermeasure may not be having its intended effect after it is implemented.

1.3.3 Urban road safety

To date, there has been relatively little research on the aspects of the environment that affect risk on urban road segments (rather than intersections). Factors that increase risk on rural roads do not necessarily have the same effect on urban roads (Abdel-Aty, Devarasetty, & Pande, 2009; Avelar, Dixon, Brown, Mecham, & Van Schalkwyk, 2013; Ewing & Dumbaugh, 2009; Lee, 2000). Hence, risk assessment tools for urban roads are rare—the United States Highway Safety Manual is one exception (AASHTO, 2010). It is vitally important to address health problems (including traffic injury) in urban areas because more than 50% of the population of the world resides in cities and the figure is increasing over time (Editorial, 2010). The figure is even higher in Australia where 60% of the population live in capital cities and a further 30% live in urban areas, which includes regional cities and towns (Australian Bureau of Statistics, 2013)

One of the reasons that past research has focused on identifying characteristics of the road and roadside that impact crash risk on rural roads is because rural roads have a higher fatality rate than urban roads. This is partly due to the higher speeds at which people travel in rural areas and hence the increased injury severity when a crash occurs (Department for Transport, 2014; Turner & Tziotis, 2006). Despite the higher fatality rate on rural roads, however, the overall crash rate is significantly higher on urban roads. In the United Kingdom in 2013, there were 781 crashes per billion miles travelled on urban roads compared to 342 crashes per billion miles travelled on rural roads (Department for Transport, 2014). On midblock road segments in Victoria, Australia, both the overall crash rate and the serious injury crash rate are higher on urban roads than rural roads (refer to Table 1.1; data from Turner & Tziotis, 2006).

Table 1.1 Crash rates per million vehicle km travelled on Victorian road midblocks, by crash severity and area (data used with permission from Turner & Tziotis, 2006)

Area	Fatal crashes	Injury crashes	All crashes
Urban	0.41	25.90	26.31
Outer urban	0.66	21.91	22.57
Rural	0.81	16.14	16.96

Another reason that highways and rural environments may have been studied more frequently to identify aspects of the environment that increase risk is that they are relatively simple environments to characterise. Figure 1.2 shows a typical Victorian rural road environment. Environmental factors affecting risk comprise those from the natural environment (e.g. roadside trees) and the built environment, which in this situation are restricted almost exclusively to the road system itself (e.g. road type and width, curvature and grade, traffic separation, shoulder type and width, pavement surface, roadside signs and vehicular traffic). Rural roads have fewer intersections than urban roads. It is thus relatively easy to identify and measure potential risk factors for rural roads in order to assess their relationship to crashes. Data on many of the potential risk factors in rural environments are available from existing databases held by road authorities.

Figure 1.2 Victorian rural road environment, photograph used with permission from Hillard, P.

Figure 1.3 shows an example of a Victorian urban environment, specifically, a strip shopping centre zone. Clearly there are many more aspects of the built environment that can impact crash risk in urban environments than in rural environments. These include the road system (road design) and also the broader built environment. There are more intersections. Roadside development and parking increase complexity. Land use, in terms of the presence of amenities and facilities, influences the number and type of road users. There are many more road users, including pedestrians and cyclists, and facilities for these road users, such as pedestrian crossings and cycle lanes. There are also a large array of objects, unrelated to driving, that can capture drivers' attention, such as shopfronts and roadside advertising. Much of the complexity in urban environments is related to the built environment outside of the road environment and is impossible to capture using traditional data sources held by road authorities. Consequently, obtaining high quality data to measure potential risk factors and characterise complex urban environments is a huge challenge.

Figure 1.3 Victorian urban road environment (strip shopping zone), photograph used with permission from Edquist, J.

Therefore, in order to improve the safety of roads, there is a need to integrate research methods to identify characteristics of the road and surrounding environment that impact crash risk and to further investigate why these factors influence risk (e.g. through changes in driver behaviour). There is a lack of research into crash risk in urban areas, despite the larger population and higher crash rate per kilometre travelled compared to rural areas. It is likely this is partly due to the difficulty in characterising and measuring the urban environment. This sets the scene for the focus of research in the current thesis.

The following sections introduce and discuss several frameworks that contribute to conceptualising road safety, in particular, the role of the road user and their behaviour within the built environment in influencing crash risk. These frameworks will be combined and augmented to form the conceptual framework for investigating risk in relation to the built urban environment used in this thesis.

1.4 Conceptual frameworks for road safety and crash risk

The mechanism of injury resulting from traffic crashes is that there is a transfer of energy between colliding entities. Injury results when the amount of energy transferred to the humans involved in the crash exceeds the biomechanical tolerance of the human body (Haddon, 1970). The injury outcome of a crash thus depends on the amount of energy transferred relative to the tolerances of the individual concerned. Crash prevention, at the most basic level, involves preventing the occurrence of a crash (energy transfer) whilst injury mitigation involves preventing the transfer of energy to the human above their tolerance when a collision does occur. Crash risk is a measure of the probability of the crash occurring while injury risk describes the probability of injury when a crash occurs. This basic physical description of energy exchange as the mechanism of injury belies the complexities involved in successfully preventing crashes and injuries in the road transport system.

The purpose of an effective road transport system is to provide safe and efficient mobility for all road users. Traffic crashes, therefore, are an indicator of system failure and it is the road users who suffer the consequences of this failure. The systems approach to safety is used across many safety-critical systems such as aviation and nuclear power and has recently been embraced in road safety (Larsson, Dekker, & Tingvall, 2010). The current road safety paradigm in many developed countries is that no-one should be killed or seriously injured while using the road transport system within the defined boundaries for use; for example, Sweden's Vision Zero (Tingvall & Haworth, 1999) and the Netherlands' Sustainable Safety (Wegman, Dijkstra, Schermers, & van Vliet, 2006). Proponents of the safe system approach recognise that the road system is comprised of several elements: the road users, the vehicles they operate and the roads on which they travel. These elements operate within a broader physical environment and under the practices, policies, laws and regulations of the transport system planners, designers, maintainers, managers, regulators, and policy makers, and indeed, within the norms, priorities and values of society as a whole. Deficiencies occurring at different levels of the system can contribute to crash occurrence, not just at the level of the road user. Importantly, it is also understood that preventing crashes and injuries involves addressing all of these components, and the interactions between them. This, however, has not always been the case.

Early in-depth investigations of crashes reported that humans contributed to the probable cause in the vast majority of crashes. For example, Treat et al. (1979) cited human factors as the probable cause in 93% of crashes, whereas environmental factors and vehicle factors were cited as the probable cause in a much smaller proportion of crashes (34% and 13%, respectively). The predominant response to a crash was to blame the individual when a crash occurred and consequently, much effort to prevent crashes focused on improving driver behaviour (Larsson et al., 2010). The high proportion of crashes that involve human error, however, is not surprising when you consider that at the time of the crash, the human was the only active component of the road system until relatively recently when active crash prevention technologies were developed for vehicles. Even now, police crash investigations often focus on the behaviour of road users because it is easier to identify failures on behalf of the road user in the moments prior to the crash than failures of other parts of the system (e.g. road designers, vehicle manufacturers). To focus entirely on the road user to remediate the problem, however, fails to recognise that the system should be designed to optimise human performance in order to achieve safe and efficient transport. This idea is not new. Even as early as the 1930s, it was recognised that humans make errors as a result of poorly designed vehicles and road environments (Toops & Haven, 1938). Over time, this safe system paradigm has become more accepted.

Despite the current widespread recognition of the importance of taking a systems approach, much of the research being conducted in road safety at present is focused on only one component of the system. Some examples are: driver behaviour research that ignores the contribution of the environment or the vehicle; vehicle safety feature development that does not consider human responses to, and use of, new technologies; and research into the aspects of road design that impact crash risk without explicitly considering how humans behave in the road environment. There is much scope to conduct programs of road safety research that cross disciplinary boundaries in order to address the interaction between different components of the road system.

There is no single over-arching theoretical framework for understanding road safety but there are several frameworks that assist in conceptualising road safety in terms of the phases of the crash, the components of the system and their interactive contribution to exposure and risk. The rest of this section presents a review of three conceptual frameworks for road safety and crash prevention that are useful for conceptualising the influence of how the road and surrounding environment influence road safety. These will be combined and augmented to develop the over-arching conceptual framework for this thesis.

1.4.1 Defining crash phases and system components: Haddon's matrix

William Haddon Junior was an injury epidemiologist who modified the successful preventive framework of infectious disease epidemiology and applied it to injury epidemiology. Haddon identified three distinct phases of the crash at which efforts to prevent injury can be targeted: before the crash (primary prevention/crash prevention), during the crash event (secondary prevention/injury prevention) and after the crash (tertiary prevention/injury treatment). Haddon also recognised the road system as being comprised of four main components that contribute to crash occurrence and injury and could thus be targets for countermeasures in the three crash phases. These were the human, the vehicle, the physical environment and the social environment. The development of Haddon's matrix (refer to Table 1.2) to represent the risk factors for crashes and injury across the time-course of a crash represents a defining moment in the field of injury prevention in terms of specifying a range of targets for countermeasures (Haddon, 1972). Haddon's work led to a paradigm shift in injury prevention where the focus expanded from simply concentrating on human behaviour in the moments prior to the crash to a recognition of the role of humans, vehicles and the environment on a broader time-scale. (Williams, 1999) Haddon's matrix has subsequently been adapted to many different areas of injury prevention. Table 1.2 shows a generic Haddon's matrix for traffic crash prevention identifying aspects of the human, vehicle and environment (physical and social) at the different phases that could be targeted with countermeasures to prevent crashes and injuries.

Table 1.2. Haddon's matrix, including targets for injury countermeasures, as applied to road trauma

	Human	Vehicle	Physical environment	Social environment
Pre-crash	Training, education, behaviour, attitudes of road users	Roadworthy vehicles, crash avoidance technologies (passive and active)	Safer roads and roadsides (e.g. road design, traffic signals, maintenance)	Cultural norms for behaviour (e.g. speeding, drink driving), legislation, enforcement
Crash	Seat-belt use, helmet use	Crash mitigation technologies (e.g. vehicle structure, airbags)	Safer roads and roadsides (e.g. roadside barriers, speed limits)	Policy and legislation for speed, crash mitigation technologies, promotion of safe vehicles
Post-crash	Training, education of emergency service personnel	Post-crash automatic emergency notification systems	Access for emergency services, congestion	Effective trauma systems

Haddon's matrix identifies both the components of the road system and the stages at which countermeasures can be targeted in order to prevent injuries. It is thus an appropriate conceptual framework to apply when attempting to identify and understand the factors that influence crash and injury risk. An essential consideration when measuring risk, however, is the concept of exposure measurement. This is not explicitly addressed by Haddon's matrix.

1.4.2 Measuring risk: Road trauma chain

Cameron's (1992) Road Trauma Chain is a useful adjunct to the Haddon matrix for defining different levels of exposure to crashes and consequent injury and thus denominators for measuring risk. Effective countermeasures act by breaking the chain of exposure to risk. Like the Haddon matrix, the Road Trauma Chain comprises the pre-crash, crash and post-crash phases (Figure 1.4 & Figure 1.5). Prevention of road trauma can be achieved by either preventing the occurrence of a crash through controlling the exposure to risk or preventing injury once a crash occurs by controlling exposure to energy transfer and thus crash consequence. The Road Trauma Chain outlines the different ways that risk can be measured in terms of relating the number of crashes or injuries to exposure as the denominator of a rate, starting from the general population level and moving to more specific event-based exposures.

At the broadest level, the Road Trauma Chain (like Haddon) recognises that entities, both human and equipment/infrastructure, exist in the road system (Figure 1.4). Risk (in terms of crash rate) can be measured at this broad level as the number of crashes per head of population (Risk D: public health risk).

The next step in the chain is that entities become eligible for road use. Formally, this refers to the processes of driver licensing, vehicle registration, and the opening of newly built roads; levels of exposure that allow calculation of crash rates per licensed driver, per registered vehicle or per km of the road network (Risk C). For pedestrians and cyclists, however, there is no formal or legal assessment of eligibility for road use. While there may be informal social processes for determining “eligibility” prior to pedestrians and cyclists using the road as individuals (e.g. parents teaching their children how to cross the road, or ride their bicycle), there is no formal census of eligible pedestrians and cyclists and so the risk per unit of exposure is difficult to measure at this level for pedestrians and cyclists.

The next level of exposure in the Road Trauma Chain is related to actual road use, at which level crash rates can be measured per distance travelled or per time spent on the road network (Risk B: transport risk). Such denominators for risk can be estimated from traffic volumes, other road user based exposure metrics (e.g. distance walked, time spent cycling) and length of roads. Transport risk (per km of roadway or per vehicle km travelled) is commonly measured in studies of the association between the road and the environment on crashes.

Next follows exposure related to the accumulation of energy with the potential to be transferred if a crash were to occur. This relates specifically to the speed and absolute or relative mass of moving entities within the system. Moving closer to the crash event, the next level of exposure reflects the exposure to crashes, hazards or hazardous situations known to increase crash risk. This level comprises hazard-specific exposures, for example, the number of pedestrians crossing roads as a denominator for pedestrian-vehicle crashes (Risk A).

Finally at the end of this phase of the road trauma chain comes the crash event itself; every level of exposure that precedes the crash identifies both a way of defining and measuring risk and a point at which effective countermeasures could be targeted to prevent crash occurrence.

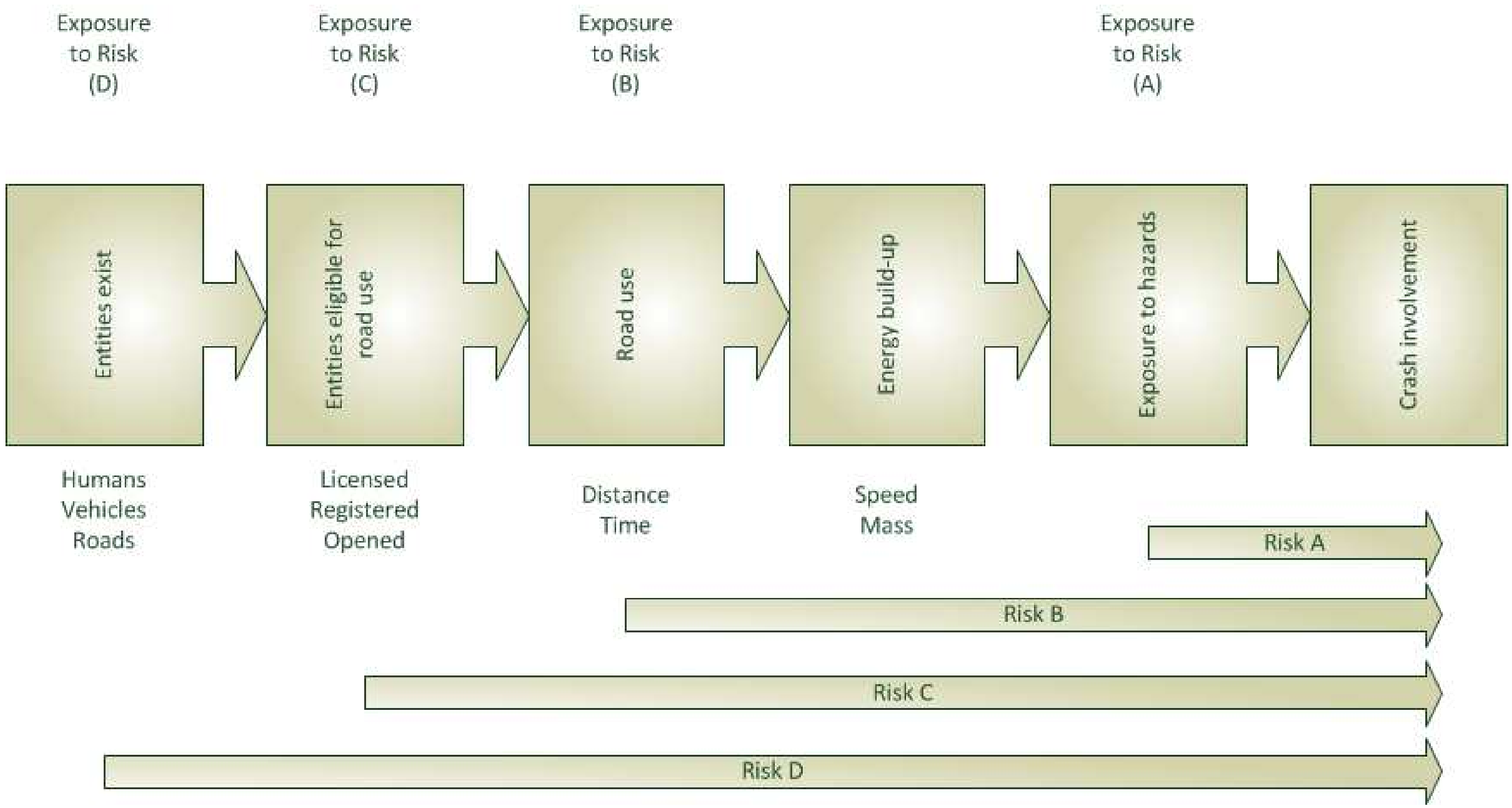


Figure 1.4. Road Trauma Chain: Pre-Crash and Crash. Adapted with permission from *Accident Data Analysis to Develop Target Groups for Countermeasures. Volume 1: Methods and Conclusions*, by Cameron, 1992

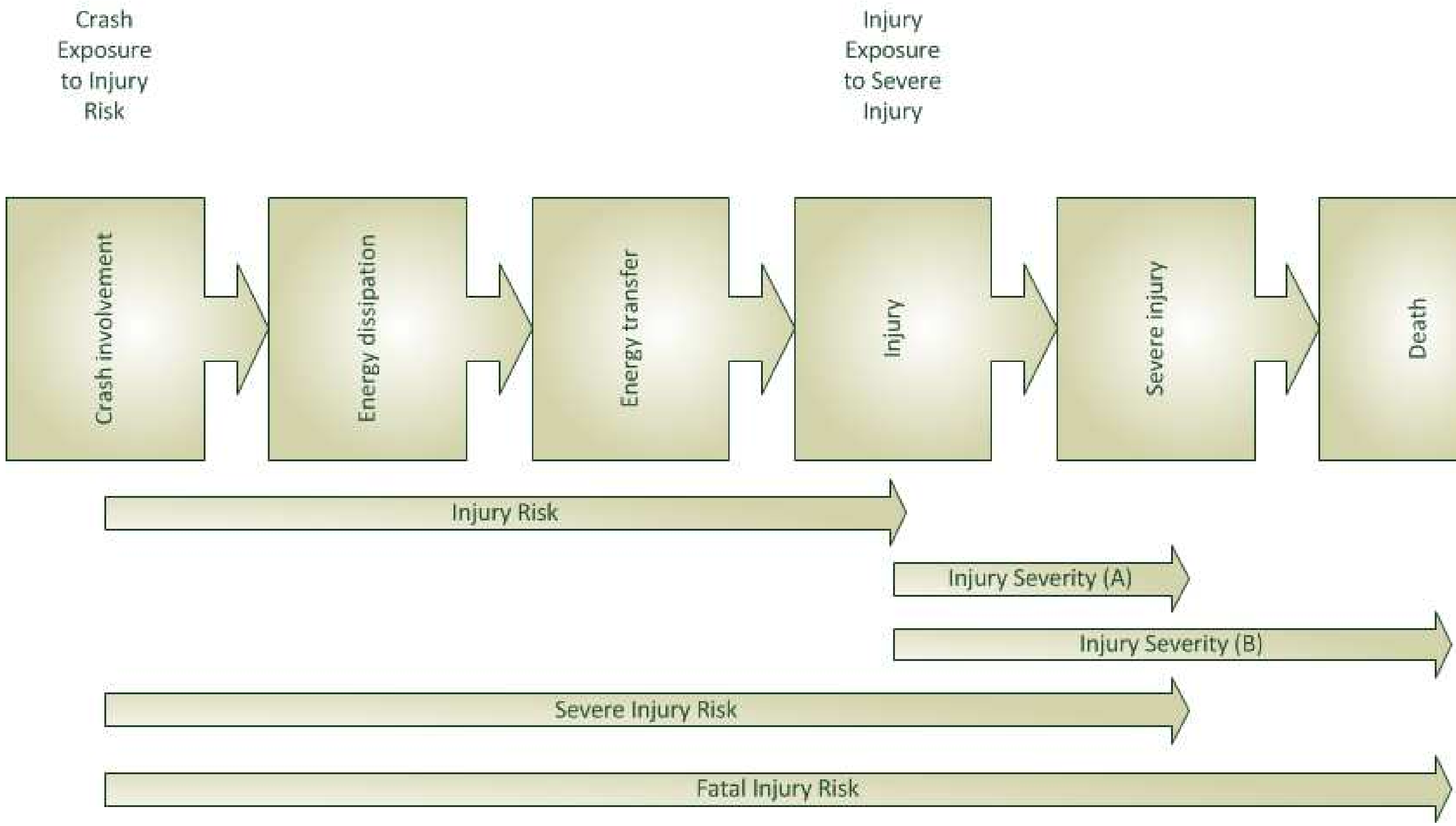


Figure 1.5. Road Trauma Chain: Crash and Post-Crash. Adapted with permission from *Accident Data Analysis to Develop Target Groups for Countermeasures. Volume 1: Methods and Conclusions*, by Cameron, 1992

The second phase of the Road Trauma Chain involves the path from crash involvement to injury outcomes of increasing severity (Figure 1.5). After a crash, energy is dissipated and may be transferred to humans, both vehicle occupants and those outside the vehicle, at levels high enough to cause injury of varying severities. The number of injuries at each severity level divided by the number of crashes that occurred is a measure of risk of that level of injury, given a crash occurred (injury risk, severe injury risk and fatal injury risk). In addition, we can derive the risk of severe injury (injury severity A) or death (injury severity B) given an injury has occurred.

Often, crash data collected by road authorities only include casualty crashes (that is, crashes that only result in property damage are not included). In studies that analyse casualty crash data, the risk being measured involves a combination of the risk of crash occurrence and the risk of injury. For example, the risk measured in a study of the number of casualty crashes that occur per km of road is a combination of Risk B (transport risk: the number of crashes per km of roadway) and injury risk (the risk of injury given a crash occurred), which combine to measure the transport injury risk.

The Road Trauma Chain is thus an extension of Haddon's matrix that is a linear representation of the chain between exposure, crashes and injuries that recognises that humans, vehicles and the environment are involved in exposure and risk. Injury prevention is achieved by targeting countermeasures at various points in the chain of risk and exposure to prevent crashes occurring, and prevent injuries in the event that a crash occurs. Despite the Road Trauma Chain being represented linearly, however, there is no implication that there is a linear chain of events that leads to crashes, rather, that exposure can be considered from the broad level to the more specific in relation to crash events and resulting injuries.

1.4.3 Driver behaviour and crash risk

The safe system approach recognises that the road system comprises several components and that the interaction between these components is an important contributor to crash risk. The first point in the Road Trauma Chain at which the components physically interact to contribute to exposure and risk is represented at the level of road use (Cameron, 1992). From then onwards, the interaction between road users, the vehicles they drive and the environment they operate in is broadly represented in terms of exposure to risk. An explicit representation of the role of human behaviour in influencing exposure to risk within the road system is lacking in Haddon's matrix (Haddon, 1972) and the Road Trauma Chain (Cameron, 1992). It is important to consider how the interaction between human behaviour and other system components can be conceptualised in terms of influencing exposure to risk at the different levels of the road trauma chain. Of particular interest for this thesis is the understanding of how human behaviour is influenced by the environment and the resulting influence on risk.

In the next sections, the contribution of road user behaviour to risk will be considered in terms of the decision to make a trip, the choice of travel mode and behaviour whilst making a trip.

1.4.3.1.1 Decisions to travel and mode choice

Schepers, Hagenzieker, Methorst, van Wee & Wegman (2014) developed a framework for road safety and mobility to explain crashes and injuries in terms of risk and exposure, with explicit consideration of the decision to travel and the choice of mode (refer to Figure 1.6). Similar to the frameworks already presented (Cameron, 1992; Haddon, 1972), the model recognises road users, infrastructure and vehicles as components of the road system and that interactions between these components contribute to crash and injury risk. Behavioural exposure to risk is conceptualised as resulting from the decision to make a trip and the choice of mode by which to travel. The decision to make a trip is related to the motivations, needs, opportunities and abilities of the road user and is a trade-off between the benefits obtained by travelling to a location to perform an activity and the expected costs of that travel, in terms of effort, time and cost (travel resistance). Travel resistance is further influenced by perceived risk.

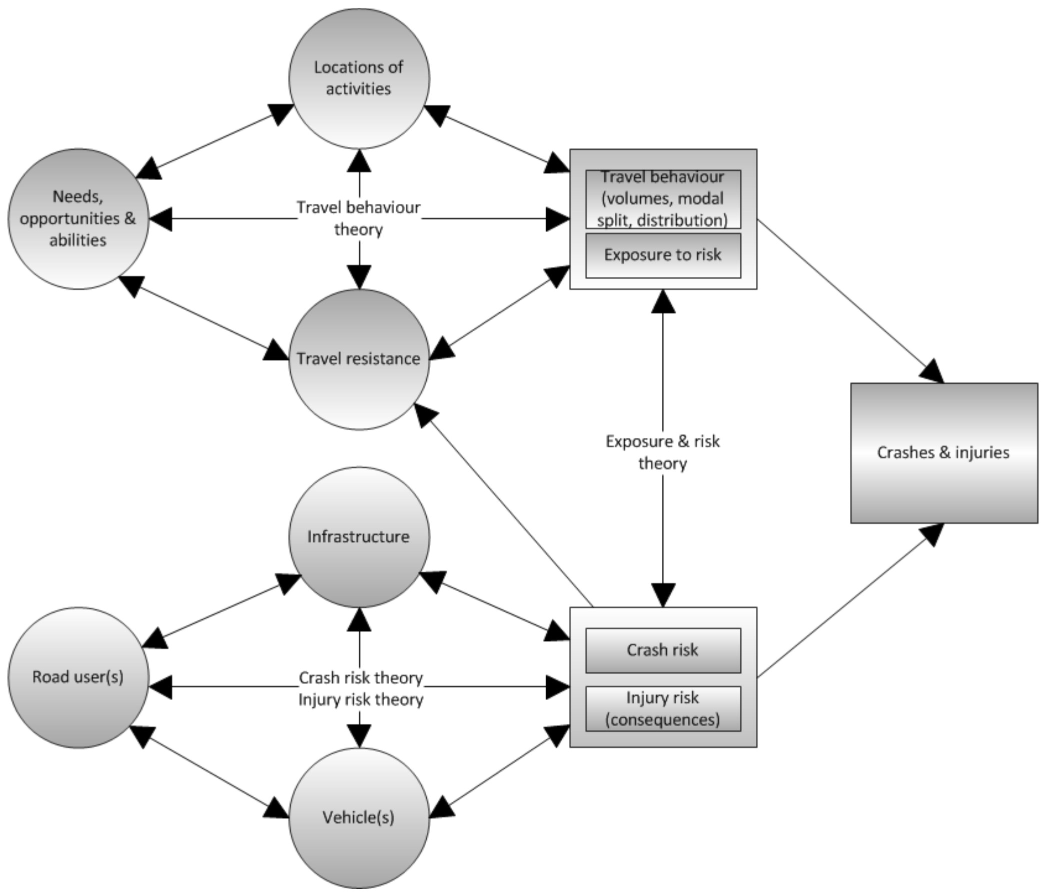


Figure 1.6 Framework for safety and mobility, by Schepers et al., 2014, in [Accident Analysis & Prevention], used under [CC BY-NC-ND 3.0], references to sections of source paper removed from figure.

While the treatment of exposure and risk by Schepers et al. (2014) is not as comprehensive as by the Road Trauma Chain (Cameron, 1992), the strength of this model is that it explicitly considers the influences behind a road user's decision to use the road system, their choice of travel mode and the resulting contribution to exposure to risk. It also shows how the perceived risk (which can be affected by the road environment) impacts on the decision to travel. Specifically, these factors influence the exposure to risk at Cameron's level of road use (Cameron, 1992). The motivations, needs, opportunities and abilities of the road user may also influence exposure at the level of becoming eligible to use the road (that is, in gaining a licence or registering a vehicle) and energy build-up (in terms of the mass and structure of the chosen vehicle).

1.4.3.1.2 Behaviour during a trip

Schepers et al. (2014) explicitly consider the contribution of behaviour to risk at the pre-trip stage in terms of the decision to make a trip and the choice of mode. A road user's behaviour during a trip has a large influence on exposure to risk yet this aspect of behaviour is not considered in the framework for safety and mobility (Schepers et al., 2014). Behaviour during the trip specifically affects exposure to risk at the levels of energy build-up and exposure to hazards. A driver's choice of speed, for example, influences energy-build-up because higher speeds result in greater levels of kinetic energy with the potential to be transferred in a crash. Speed choice is influenced by the vehicle, the road environment, the behaviour of other road users, road rules, enforcement and social norms.

Exposure to hazards will impact a road user's route choice, the time at which they travel and behaviour en route. Perceived risk and mobility affect route choice, trip timing and travel behaviour and the balance between risk and mobility will vary between drivers. Some road users prioritise getting to their destination quickly while others choose routes to both avoid risk and enhance mobility (e.g. drivers avoiding uncontrolled right turns at busy intersections, pedestrians avoiding uncontrolled road crossings). Schepers et al. (2014) identified perceived risk as a factor in resistance to travel, and it is also a factor in influencing behaviour on the road. For example, drivers decrease speed in more complex road environments (Edquist, Rudin-Brown, & Lenne, 2012). Road user behaviour can increase exposure to hazards (e.g. risk-taking) as well as reduce exposure to hazards and in addition, road user actions can prevent crashes occurring (by evasive actions, crash avoidance).

In order to explain more fully how the built environment and road user behaviour can interact to influence risk, it is necessary to consider how road users interact with their environment when making a trip. For this reason, theories of driver performance and behaviour that incorporate the influence of the environment on behaviour are reviewed in the next section.

1.5 Theories of driver performance and behaviour

Theories of driver performance and behaviour have been developed for various purposes, from theories of individual skills to theories of general driver performance and behaviour. The former are too specific to be applied to the current problem so the latter are the focus of this review. Specifically, this review will focus on those theories that can contribute to deriving hypotheses about how the environment influences driver performance and behaviour.

1.5.1 Driving tasks and risk

Michon (1985) defined a three-level hierarchy of driving tasks: strategic, tactical and operational. These levels can be considered in terms of behavioural contributions to crash risk. Strategic decisions are chiefly made prior to a trip regarding the route and mode of travel; how these travel and mode choice decisions contribute to risk is the focus of the Schepers et al. (2014) framework for road safety and mobility. Tactical and operational control occur during the trip. Tactical control relates to manoeuvring the vehicle according to constraints of the environment and other road users, for example, turning and overtaking. Operational control refers to tasks that are predominantly automatic vehicle control tasks, like steering during normal driving situations. These behaviours contribute to risk at the level of behaviour during the trip and are strongly influenced by the environment. Processing is fastest at the operational (automatic) level and slowest at the strategic level. Normally, goals from one level filter down to affect behaviour at the lower level; for example, the choice to accelerate and overtake a vehicle at the manoeuvring level is affected by the goals at the strategic level (for example, reaching the destination as quickly as possible).

Rasmussen's (1983) framework for human performance classifies behaviour as skill-based, rule-based and knowledge-based. Skill-based behaviour involves largely automatic performance of well learned actions, rule-based behaviour concerns the application of learned rules to recognisable problems, and knowledge-based behaviour is characterised by conscious decision-making in unfamiliar situations. Once drivers have learnt to control a vehicle, most operational control will occur at the skill-based level. Tactical control, or manoeuvring, is mostly automatic (e.g. turning at a familiar intersection) or rule-based (for example, negotiating a busy unsignalised intersection). In unfamiliar situations, manoeuvring may require knowledge-based reasoning, for example, trying to recover after vehicle control has been lost on an unsealed road shoulder. Strategic control can be rule-based (for example, choosing the most efficient route to take at different times of the day) or knowledge-based (devising a route to a location that has not been visited before).

Experienced drivers spend most of their driving time performing operational control tasks at the skill-based level (Weller, 2010), followed by rule-based manoeuvring tasks. When a potentially hazardous situation occurs, drivers will initially seek a rule-based solution. If, however, the situation is not recognisable and they cannot produce an appropriate rule, they will apply knowledge based reasoning, which generally takes longer before a decision is made and action occurs. In terms of safe road design, this means that roads designed to take advantage of automatic processes and rule-based behaviour are preferable to those that require knowledge-based reasoning. Consistent road and roadside design will therefore enhance familiarity and expectation, and should assist the driver in responding quickly and appropriately to potentially hazardous situations. It has been suggested that information aimed at tasks in the different levels of the hierarchy should not be presented simultaneously, for example by locating directional signs (strategic information) where it is likely that important tactical decisions will be need to be made (Sagberg, 2003).

1.5.2 Information processing approach: performance

It is essential to understand the perceptual, motor and cognitive capabilities of road users so that roads are designed with these in mind. Some argue that the driving task primarily involves information processing and decision making in response to the design of the road, the roadside, weather and traffic conditions (Heger, 1998). Driver skills, perceptual abilities and current state affect their task performance (Heger, 1998). There is no doubt that driving is a predominantly visual task and the road environment contains information relevant for operational and tactical control (Theeuwes, 2001). In the built urban environment there is also a vast amount of visual information competing for attention that is unrelated to driving, for example, roadside advertising (Edquist, 2009) and general human activity. It is essential that driving related information attracts drivers' attention, for example, with the use of conspicuous objects like red stop signs, or flashing lights to draw attention (e.g. to pedestrian crossings) (Theeuwes, 2001).

Gibson and Crooks (1938) claim that elements of the visual field important for driving will stand out to the driver compared to other elements. In extremely visually complex environments, like strip shopping centre road segments, identifying driving-relevant information can be challenging and it has been found that elements have to match driver expectations to be seen (Theeuwes, 2001). Expectancy, attention and mental workload greatly influence driving task performance (Heger, 1998). Roads should be designed so drivers' expectations are not violated, for example, it is not desirable to place a sharp bend after a series of gentle bends (Sagberg, 2003).

1.5.3 Motivational (cognitive) approach: behaviour

While designing for performance capabilities is important, it is insufficient to simply design for what humans are capable of doing. Performance capabilities inform us how people *can* perform, not how they *actually* behave. How people behave in any given environment (and their exposure to risk as a result of their behaviour) is contingent upon both their performance capabilities and their motivations and goals. It is therefore essential to take these into account.

Motivational models (which include a cognitive component) focus on driver behaviour rather than driving skill (or performance). The driving task is said to be self-paced; that is, drivers control task demand through their behaviour, for example, by varying travel speed. Behaviour is also assumed to be strongly dependent on motivations and goals, for example, the goal to reach a destination on time. The motivational models are characterised by the notion of a subjective state of risk (or similar) as the main driver of behaviour. The motivational models differ according to whether they propose that feelings of risk are continually monitored in order to maintain a target level (e.g. Risk Homeostasis Theory; Wilde, 1976) or range (e.g. Risk Allostasis Theory; Fuller, 2011) or whether there is a threshold level of risk (zero-risk theory; Näätänen & Summala, 1974) or safety margins (multiple comfort zone model; Summala, 2007) that must be reached before drivers feel risk. A series of experiments designed to test the hypotheses of driver behaviour models supported the notion that a threshold level of risk drives behaviour, rather than continuous monitoring of risk (Lewis-Evans, 2012; Lewis-Evans, De Waard, & Brookhuis, 2010; Lewis-Evans & Rothengatter, 2009).

Each of the motivational models states that the driver's perception of the environment will influence their feelings of risk and/or task demands. Yet, how this occurs is rarely discussed in sufficient detail to understand or predict the effects of particular road or roadside features on behaviour. An exception is the multiple comfort zone model (Summala, 2007) which is based on the idea of safety margins, or the field of safe travel (Gibson & Crooks, 1938). These concepts and the multiple comfort zone model will be discussed in more detail in the next sections.

1.5.3.1 The field of safe travel

Gibson and Crooks (1938) described driving in terms of locomotion through a space, akin to walking, along a path chosen to avoid obstacles and collisions, predominantly guided by the visual system. The vehicle is a tool for locomotion that allows the human to travel from one location to another more efficiently.

Within the boundaries of the road, drivers perceive the existence of a field of safe travel which comprises all possible paths which the vehicle may take. To avoid crashes, the perceived field of safe travel must be within the objective safe travel zone.

This field of safe travel influences driver behaviour. Drivers steer to direct the vehicle into the middle of the field of safe travel. Drivers accelerate in order to reach their goal (the destination) but when their field of safe travel is compromised they respond by decelerating. Many things can affect (and reduce) the field of safe travel: obstacles in the path of motion, moving obstacles with the potential to enter the path of motion, factors that affect visibility (obstacles, darkness, weather conditions, glare and the topography of the road) and road geometry that affects lateral acceleration (e.g. curves). The field of safe travel is also affected in the forward direction by the minimum stopping distance (based on speed, road condition and vehicle braking capability) and impacted by signs and signals and laws (stop signs, traffic signals, brake lights of other vehicles). Therefore, the built environment and activity within it influences the field of safe travel for a road user.

Although the field of safe travel is normally bounded by the road, in emergency situations, it may extend to the roadside (e.g. road shoulders, adjacent land) if it is necessary to avoid an obstacle that has appeared on the roadway. Thus it can be seen as the field of *safest* travel (Gibson & Crooks, 1938). Obstacles may be stationary (e.g. trees, poles, parked vehicles) or moving (e.g. other road users). They can also be visible obstacles (in the line of sight) or potential obstacles (e.g. potentially hidden behind a parked vehicle, or around a corner). Visible obstacles tend to generate a stronger response than potential obstacles. In the presence of moving obstacles, the driver gauges (unconsciously) the field of safe travel based on their own motion in reference to the predicted motion and pathway of the other road users. Some road users are less predictable than others and the field of safe travel is adjusted accordingly, for example, the movements of pedestrians and cyclists (who can walk or cycle anywhere) are less predictable than trams (which only travel on tram tracks).

The field of safe travel therefore describes the boundaries which affect travel and the driver's response in manoeuvring through the field of safe travel to their destination through concurrent lateral and longitudinal control of the vehicle. Though the theory relies mainly on visual perceptual processes, it also recognises that drivers' goals and motivations to reach a destination influence travel behaviour. The concept of the field of safe travel forms the basis for how the environment influences behaviour in the multiple comfort zone motivational model of driver behaviour, which is described in the following section (Summala, 2007).

1.5.3.2 The multiple comfort zone model

Summala proposed the multiple comfort zone model (Summala, 2007) for driver behaviour which is an extension of the previous risk-threshold models advanced by the same author (e.g. zero-risk theory; Näätänen & Summala, 1974). Under the multiple comfort zone model, driver behaviour is explained by the trade-off between motivations for mobility that can promote higher speeds and hazardous behaviour (e.g. to save time, conserve effort, maintain speed, reach destination) and the constraints of safety margins, which are related to the boundaries for safe travel as defined by Gibson and Crooks (1938). When drivers are driving within the safety margins, they are within their comfort zone and do not perceive they are at risk. A breach of safety margins engenders an emotive response which prompts a change in behaviour to negate the breach. If the driver does not modify their behaviour, a feeling of discomfort ensues. Other aspects apart from safety margins can also affect comfort zones, such as the smoothness of travel (affected by vehicle vibrations and road condition), avoiding fines, and achieving progress on the trip. The driver will attempt to keep all of these aspects within their comfort zone (hence the multiple comfort zone terminology).

Safety margins influence the amount of time available to the driver to perform tasks (both driving related and non-driving related): a reduction in safety margins reduces the amount of time available to the driver (Summala, 2007). For example, when a lead vehicle brakes heavily, the amount of time available to either brake or steer away is reduced. Reducing the time available to perform tasks increases the demands of the driving task. Task demands are also increased by increased complexity which can be impacted by a number of factors, for example the presence of multiple potential obstacles (or road users). The difficulty of the task is contingent upon task demands (including complexity) but is also related to an individual's capabilities, their state and the strategies they use to achieve their goals, which affect the mental effort they expend to perform the task (de Waard, 1996). Mental workload, the information processing capacity used to perform a task, is therefore also impacted by task demands (de Waard, 1996). Although the driver's goals do have an impact, task demands are predominantly due to external factors (de Waard, 1996). Task difficulty, effort and mental workload, however, involve the interaction between the task demands and how the driver responds to the demands to perform the tasks required to achieve their goals.

Like previous motivational models, the multiple comfort zone model defines driving tasks in terms of Michon's three-level hierarchy of control: strategic, tactical and operational (Michon, 1985). Safety margins and available time affect the demands of tasks at the operational and tactical level and can sometimes also cause drivers to reassess higher level goals (Summala, 2007). Thus, operational and tactical control are intrinsically related to task demands, task difficulty, mental effort and workload.

The multiple comfort zone model is the only motivational model that explicitly considers how the environment affects behaviour which makes it the most relevant model for framing how the environment and behaviour affect risk in this thesis. The characteristics of the built environment, in terms of the road, roadside and road user activity, are hypothesised to influence driver behaviour and mental workload through their effect on perceived safety margins. These, along with the driver's goals (such as arriving at a destination within a certain time, or driving as fast as possible) interact to affect behaviour during the trip in terms of operational and tactical vehicle control (e.g. lateral position, choice of travel path and speed choice) and can sometimes also impact higher level strategic decisions.

1.6 Conceptual framework for the thesis

The three conceptual frameworks that were reviewed in Section 1.4 each contribute a piece to the puzzle of understanding the interactive influence of the road user and the environment on crash risk. Haddon's matrix (Haddon, 1972) contributes to defining the components of the system and the stage of the crash as targets for countermeasures, the Road Trauma Chain (Cameron, 1992) contributes different levels of exposure and denominators for measuring risk while the framework for road safety and mobility (Schepers et al., 2014) contributes an explicit consideration of how the decision to travel and choice of travel mode influence risk at the pre-trip stage. What is missing, however, is an explicit representation of how a road user's behaviour during the trip (which is influenced by the built environment) contributes to risk. The three frameworks were therefore combined and augmented with an explicit consideration of risk related to behaviour during the trip to form the conceptual framework for this thesis. The conceptual framework behind the crash prevention research in this thesis is presented in Figure 1.7: an onion diagram comprising the fundamental aspects of Haddon's matrix (Haddon, 1972), Cameron's Road Trauma Chain (Cameron, 1992) and Schepers et al.'s framework for road safety and mobility (Schepers et al., 2014), supplemented with behavioural determinants of exposure to risk during the trip.

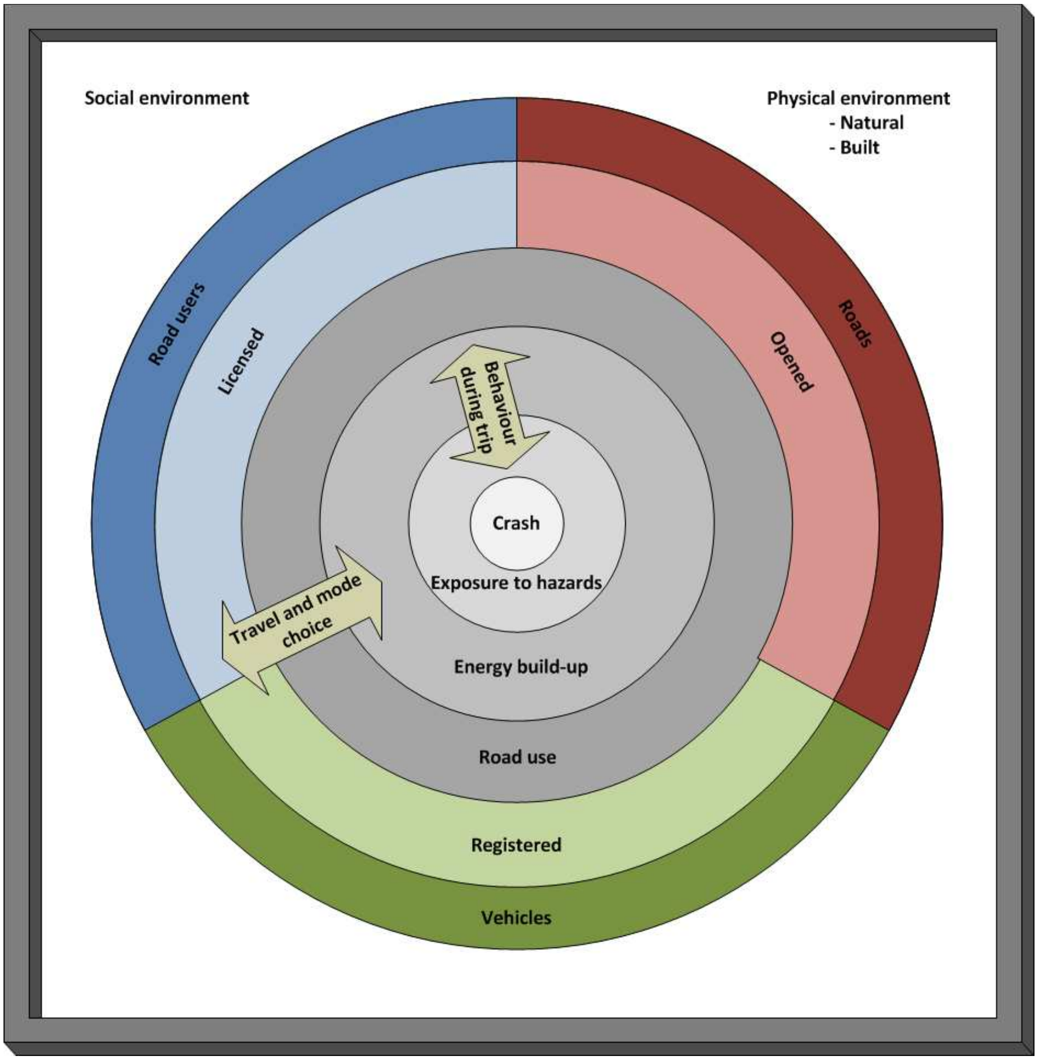


Figure 1.7 Conceptual framework for thesis: road system, surrounding environment, road user behaviour and crash risk

Briefly, the road system exists within broader society and the applicable laws, rules, regulations, policies and norms of that society (equivalent to Haddon's (1972) social environment). The road system also exists within the broader physical environment which includes the natural environment and the built environment. Health Canada defines the built environment as encompassing "…all buildings, spaces and products that are created or modified by people", recognising the link between the social and the built environments (Health Canada, 2002 cited in Srinivasan, O'Fallon, & Dearry, 2003, p 9.). The link between the built and social environments is elaborated by the Handy, Boarnet, Ewing and Killingsworth (2002) definition of the built environment that explicitly recognises patterns of human activity (e.g. traffic and other road user movements) as part of the built environment. That is, the social environment can influence individual behaviour (through social norms and values), and patterns of human activity (or behaviour) are part of the built environment.

The concentric circles in Figure 1.7 represent the road system and the different levels of exposure from the road trauma chain (Cameron, 1992). An onion diagram was chosen rather than the linear representation of the original road trauma chain in order to emphasise the systems approach and remove the implication that there is a linear chain of events leading to crashes (which Cameron did not intend). Road users, roads and vehicles are represented as distinct entities in the two broadest levels of exposure (level 1: entities exist and level 2: eligibility for road use). In contrast, from the level of road use to the crash itself, the components are not represented separately because it is the combined interaction between the entities that contributes to crashes from this point on. Behavioural influences are overlaid onto exposure. The decision to travel, route choice and vehicle choice, contribute most strongly to exposure at the level of road use and are located closer to the interface between the road user and the vehicle to represent mode choice. Behaviour during the trip is located close to the interface between the road user and the environment to represent the importance of the interactive contribution of these components to exposure in terms of energy build-up and exposure to hazards. The conceptual framework underscores the importance of taking a multidisciplinary approach to investigating crash risk.

1.7 Aims

The preceding discussion of the public health burden from traffic crashes, the lack of knowledge about factors affecting risk on urban roads, and the need for a multidisciplinary approach encompassing the interactions between the components of the road system and the broader social and physical environment, led to the aims of the research program to be presented in this thesis.

The broad aim of the research was to develop and apply a multidisciplinary approach to identify and understand the aspects of the built environment in urban areas that are associated with crash occurrence. The research was focused at the second step of the scientific approach to injury prevention and countermeasure development presented in Figure 1.1; identifying risk factors and mechanisms for traffic crashes in urban areas. It sits within the conceptual framework presented in Figure 1.7 as an investigation of the road, the built environment, and road users' behaviour within the environment and how these contribute to casualty crash exposure and risk.

The research was conducted in two components that employed different, but complementary, methodological approaches. The first component involved developing a statistical modelling framework to identify the aspects of the road and surrounding built environment associated with crash risk in complex urban environments.

The second component was conducted to demonstrate how behavioural research techniques can be used to investigate the reasons for increased risk in the presence of road-related risk factors and hence inform the development of effective countermeasures.

The broad aims of the two components of the thesis were:

- Aim of Component 1: to identify aspects of the built environment that are associated with the number of crashes occurring on road segments in complex urban environments (further elaboration of the aims of Component 1 are presented at the end of Chapter 3)
- Aim of Component 2: to investigate how driver behaviour changes in the presence of an aspect of the built environment that has been shown to be related to crash frequency in Component 1

The remainder of the thesis is structured around these two research components. First, Component 1 is presented (Chapters 2 to 7), followed by Component 2 (Chapters 8 to 11). Finally, Chapter 12 brings together the findings of the thesis and highlights the contributions, both methodological and practical.

CHAPTER 2. METHODS FOR IDENTIFYING RISK FACTORS

The first research component of this thesis is concerned with identifying built environment-related risk factors for crashes on urban roads. This chapter presents a critical review of methods for identifying the aspects of the built environment that are associated with crash occurrence or crash frequency. First, suitable study designs are introduced and then appropriate data analysis techniques are described. The objective of this chapter is to propose an appropriate study design and analysis methods for identifying aspects of the built environment that are associated with crash risk.

2.1 Study designs

Various study designs can be employed to identify the aspects of the built environment that are associated with crash risk. The choice of study design is primarily driven by the specific research question. The types of data that are available also inevitably constrain the study design, and thus the research questions than can be answered. Different study designs will be reviewed in terms of the research questions they can answer, along with their relative strengths and weaknesses.

2.1.1 Descriptive study designs

Descriptive studies, as the name suggests, involve describing the size and nature of a problem and potential contributing factors. Descriptive study designs include case studies and case series. Of these, case series studies have been used to identify aspects of the road and surrounding environment that may contribute to crashes.

2.1.1.1 Case series

The research question that can be answered using a case series study is:

What proportion of crashes involve various risk factors?

This question can be answered by studying a sample of crashes and identifying the number in which potential risk factors were present when and where the crash occurred. These studies are known as descriptive case series, because a series of cases (crashes) is described in terms of potential risk factors. Some examples of descriptive case series studies are: an in-depth study of cyclist fatalities in London over a 6 year period to determine the prevalence of environmental, road user and vehicle risk factors (Keigan, Cuerden, & Wheeler, 2009) and an investigation of fatal motorcycle crashes involving roadside barriers in Australia and New Zealand to describe the riders, the locations and the types of crashes (Jama, Grzebieta, Friswell, & McIntosh, 2011).

Though case series are valuable for identifying potential sources of risk, these purely descriptive studies cannot provide evidence of a relationship between risk factors and crashes without further information about the population from which the crashes were drawn. For example, if 80% of crashes in a case series occurred on four lane roads, then it is impossible to know if the number of lanes influences the risk of having a crash without knowing the proportion of all roads that have four lanes. Thus while case series studies are useful for generating hypotheses about potential risk factors than can be tested using more analytical study designs, a case series study will not achieve the objectives of this thesis (to identify aspects of the built environment that are associated with crashes).

2.1.2 Analytical observational study designs

The research question that can be answered using an analytical observational study is:

What factors are associated with crash occurrence and/or frequency on a road segment or location?

Studies in which the researcher observes the object of study (without intervening) with the aim of estimating (or analysing) an association between potential risk factors and crash occurrence or frequency are known as analytical observational study designs. These can be considered as "natural experiments" in which the researcher has no control over the risk factors being investigated, or other contributing factors. As such, they can be prone to issues of bias and confounding if potential risk factors that are associated with both the outcome and the predictors of interest are ignored. In addition, they do not provide definitive evidence for a causal relationship between risk factors and crashes. They are, however, valuable for identifying risk factors that can be further investigated in relation to causation and mechanism using other types of study designs. The two main types of analytical observational study designs used in road safety are case-control studies and cross sectional studies. These are reviewed in the following sections.

2.1.2.1 Case-control studies

The research question that can be answered with a case-control study is:

What factors are associated with the occurrence of crashes on a road segment or location?

Case-control studies were developed in the field of epidemiology as an efficient analytical method to retrospectively investigate the association between multiple risk factors and rare diseases (e.g. cancer) by selecting cases (people with a disease) and controls (people without a disease) and comparing the two groups in terms of previous exposure to risk factors (Hennekens & Buring, 1987). The efficiency of case-control studies for studying rare outcomes and multiple risk factors makes them appealing for investigating risk factors for crashes, which are rare events. Case-control studies in road safety that investigate infrastructure-related risk factors define locations or road segments (not people), as the units of analysis. Locations where crashes occur are compared to a sample of locations where crashes did not occur, in terms of the infrastructure present at the case and control sites. The odds of exposure amongst the cases is compared to the odds of exposure amongst controls; or, the odds of a factor being present at sites where crashes occurred is compared to the odds of the factor being present at sites where crashes did not occur. Other potentially confounding factors not primarily of interest as risk factors can be controlled during the study design stage (e.g. through matching cases and controls on the confounder) or analysis stage (e.g. by using multiple logistic regression). The odds ratio (which can be estimated from the parameters of the logistic regression) is the measure of association derived from case-control studies.

For example, Gross and Jovanis (2007) used a case-control approach to establish that increased shoulder width was associated with crash reductions, adjusting for confounding factors. Estimates of association derived from the case-control study were shown to be similar to those derived in the previous literature using other techniques, demonstrating the validity of the case-control study design in the measurement of the association between infrastructure-based risk factors and traffic crash occurrence.

The case-control study design is more often used when the crash experience of the individual road user is of interest (e.g. Stevenson, Jamrozik, & Spittle, 1995). Case-control studies are not commonly used in studies where the road segment is the unit of analysis (i.e. to investigate the effect of infrastructure on crash risk). The main disadvantage of using case-control studies for this purpose is that the outcome of interest is whether or not one or more crashes occurred at a location, not the number of crashes that occurred. Therefore, a location that has multiple crashes is not treated any differently to a location with only one crash. For this reason, a case-control study design will not be used in this thesis because potentially important information is lost by using the occurrence, not the frequency, of crashes as the outcome of interest.

2.1.2.2 Cross sectional studies

The research questions that can be answered with a cross-sectional study are:

What is the prevalence of various risk factors in the road transport system?

What risk factors are associated with the frequency of crashes at a location or on a road segment?

These questions can be answered using analytical cross-sectional studies in which a snapshot of data are collected at one point in time, or over a specified period of time, for a sample of the units of analysis (e.g. road segments, intersections or areas). Data are collected on the outcome (e.g. the number of crashes that occurred) and potential risk factors (e.g. traffic volume, characteristics of the road and surrounding environment). Statistical techniques are then used to establish if there is an association between the potential risk factors and crash frequency. The incidence rate of crashes can be compared between locations with different characteristics, which gives a measure of relative risk associated with the different characteristics.

Cross-sectional studies are the most commonly used method for investigating the association between characteristics of the road, surrounding environment and crash frequency (e.g. Brown & Tarko, 1999; Milton & Mannering, 1998). They are useful for investigating associations between multiple risk factors and crashes in a relatively efficient manner because they can make good use of existing administrative data. The advantage of cross-sectional studies over case-control studies is that the frequency, and not just the occurrence, of crashes is the outcome measure. Thus a cross-sectional study design is appropriate for addressing the aim of Component 1 of this thesis.

2.1.3 Interventional study designs

The research question that can be answered with an interventional study is:

What is the effect of a treatment (or treatments) on crashes?

In interventional studies, the researcher or controlling authority (such as the local road authority) controls whether or not an intervention (or treatment) is applied to the road segments. The effect of the intervention on the outcome of interest (e.g. crashes) is measured by comparing the number of crashes that occurred before the intervention to the number that occurred after the intervention. In many situations there is also a comparison or control group of sites that do not receive the treatment, and the change in crashes over time for the treatment group is compared to the control group. The main interventional study designs used in road safety are introduced briefly in this section.

2.1.3.1 Before-After studies

The effect of a treatment can be estimated using before-after studies, in which the number of crashes at a site or set of sites is counted both before and after a treatment is installed. Simple before-after studies are prone to bias and confounding because a change in the number of crashes over time could be due to either the treatment or some other factor that also changed over time. An observed reduction in crashes after installation of a treatment is not enough evidence that the treatment caused the improvement. If treatment sites are chosen because of their crash history (selection bias) then regression to the mean can become an issue in interpreting any change in crash frequency. Other factors that changed over that time period may also have been responsible for the difference (confounding), for example seasonal effects or concurrent road safety programs.

2.1.3.2 Quasi-experiments

If a comparison, or control, group of sites is included in a before-after study then the change in crashes after treatment installation at the treatment sites can be compared to the change in crashes over the same time period at the control sites. If the control sites are carefully chosen to be similar to the treatment sites in all ways except for the installation of the treatment, then a reduction in crashes at the treatment sites relative to the control sites provides stronger evidence than an uncontrolled study that the change is likely to be due to the treatment.

While the addition of a control group provides some control of confounding, quasi-experimental studies still suffer from the potential for bias and confounding. Regression to the mean can be a problem if treatment sites were chosen on the basis of crash history and control sites were not. If control sites differ to treatment sites in other respects relevant to crash occurrence, then confounding can occur. In addition, if the treatment leads to drivers choosing another route, then there might be fewer crashes at treatment sites simply because of a reduction in traffic (i.e. a reduction in exposure). These issues can be addressed through rigorous study design: e.g. not selecting treatment sites based on crash history, selecting the control sites on the same basis as the treatment groups, using multiple baseline measurements, measuring potential confounders (e.g. traffic volumes) and/or the use of empirical Bayes methods which use information on the crash history on the treatment sites and the crash frequency expected at a similar group of sites (Hauer, 1992; Persaud, Retting, & Lyon, 2004).

2.1.3.3 Experiments

The pure experiment, or randomised controlled trial (RCT) is the gold standard study design for measuring the effect of a treatment, or intervention. The main difference between RCTs and quasi-experiments is that units of analysis (e.g. road segments) are randomly assigned to be either part of the treatment group (that has the treatment installed) or the control group (which does not). Random assignment of units to treatment and control groups ensures that the units differ only according to the treatment being studied and are thus equivalent with respect to both known and unknown confounders. The change in the outcome in the treatment group is compared to the change in the outcome in the control group and the difference in the change in outcome between groups gives an estimate of the effect of the intervention.

In other fields (e.g. medicine), RCTs are often double-blinded; the researchers and the participants do not know which units are in the treatment group and which are in the control group. While random allocation of units to groups lessens the differences between groups at the beginning of a study, blinding is performed to ensure a lack of bias in how treatment and control groups are dealt with during the trial, including outcome ascertainment and analysis (Karanicolas, Farrokhyar, & Bhandari, 2010). Double-blind trials are very difficult to conduct in road safety, however, as it is often impossible to conceal if a location has had a particular treatment. This can lead to unintended effects like behavioural adaptation, for example, drivers may choose to avoid the treated road section. While this could lead to a decrease in crashes at the treatment site, crashes may increase on the other roads, in which situation there would be no net effect on road safety.

Overall, RCTs are rare in road safety, due to factors such as the expense of installing treatments or the desire of road authorities to apply treatments where they believe they are most needed. Consequently RCTs are essentially non-existent in the evaluation of the safety effects of modifying the design of the road or the roadside environment. Though the conduct of pure experiments is seldom practically possible in road safety, there are those who strongly support attempts to overcome institutional barriers to conducting RCTs (Bonneson & Ivan, 2013).

Most evaluations of treatments, or a set of treatments, in road safety are before-after studies; stronger evidence for an effect comes from those with a control group or some other way of addressing confounding and regression to the mean, e.g., the empirical Bayes approach.

A limitation of interventional studies is that only one or two factors can usually reasonably be manipulated at any one time. For this reason, the current thesis cannot use an interventional study design.

2.2 Interpreting study results

The ultimate objective of research into the safety of roads is to identify aspects of the built environment that *cause* a change in the risk of crashes. The discovery of a statistically significant relationship between a risk factor and crash frequency, however, does not necessarily mean that the relationship is causal, or even valid. There are several competing explanations for a statistically significant result. These are:

- a true causal relationship
- a chance finding
- bias
- confounding

The probability that a statistically significant result is due to chance is controlled by the application of statistical criteria for hypothesis testing; for example, setting the level for statistical significance at 0.05 means that there is a 1 in 20 chance that any statistically significant result is due to chance. Bias is defined as systematic error, for example, in the way that road segments were selected (selection bias), or in how risk factors were measured (measurement bias) that can be controlled through rigorous scientific protocols for selection and measurement. Confounding occurs when the observed relationship can be explained by other, unmeasured, risk factors that are associated with both the outcome and the risk factor being studied. Hence experimental designs (or RCTs) are the gold standard in controlling confounding, as random assignment of units to treatment and control groups ensures that the units differ only according to the treatment under investigation and thus are equivalent with respect to both known and unknown confounding factors.

As noted previously, however, the conduct of RCTs in road safety is seldom practically possible, even for new treatments. Various factors (e.g. feasibility, practicality and budget) contribute to the decision of where to install new infrastructure and it is unusual for researchers to have the opportunity to select a sample of locations and randomly allocate them into treatment and control groups. In addition, there are many types of road design and road infrastructure currently being used on the road that have not been evaluated experimentally to ascertain their effect on road safety outcomes. An investigation of the effect of these existing components on road safety must therefore take an observational approach to compare existing roads with and without these features. Taking an observational research approach, however, has implications for the conclusions that can be drawn, particularly with regard to causation.

The strongest level of evidence for causation comes from experimental designs; however the majority of studies in road safety are observational. Epidemiologists have grappled with the problem of inferring causation from observational studies and in 1965, the epidemiologist Austin Bradford Hill proposed nine criteria for assessing the evidence for a causal relationship between an exposure and a health outcome (Hill, 1965). These criteria are listed below (where relevant they have been re-interpreted in terms of road safety):

1. strength of association (stronger associations provide more evidence for cause and effect than weak associations)
2. consistency of association (are there consistent results from different studies conducted with different designs in different places at different times?)

3. specificity of association (is the association limited to specific types of sites and crash types?)
4. temporality (did the proposed cause precede the outcome?)
5. biological gradient (is there evidence for a dose-response relationship?)
6. plausibility (is the relationship biologically, or physically, plausible?)
7. coherence (the results should not contradict what is already known about the natural history and biology of disease, or physics of crashes)
8. experiment (is there any experimental or quasi-experimental evidence?)
9. analogy (is there a similar risk factor that is known to lead to a similar outcome?)

Hill stated that none of the nine criteria provide indisputable evidence for a cause and effect relationship and, indeed, gave examples for each criterion of where a cause and effect relationship may not meet the criterion. For example, Hill cautioned against dismissing cause-and effect simply because of a small effect (criterion 1) and that the results of one study may be so important as to warrant immediate action (criterion 2). Hill warned against believing that "we can usefully lay down some hard-and-fast rules of evidence that must be obeyed before we accept cause and effect" (Hill, 1965, p. 299). Therefore these form a set of guiding principles rather than a set of rigid criteria.

More recently, Lucas and McMichael (2005) questioned the application of Hill's principles to modern day problems in environmental health (including injury) due to the multi-factorial contributions of proximal and distal risk factors at different levels of the system. They also argue that a continued focus on simple criteria for causation for observational studies constrains the research that is proposed and funded, to the detriment of studies to investigate more distal and indirect risk factors. They state "contemporary environmental epidemiology confronts non-homogeneous health outcomes… likely to have multiple etiologies. Exposures can be difficult to quantify and even to define (e.g. socioeconomic status and urban design) as well as to link temporally and spatially to the …outcome… " (Lucas & McMichael, 2005, p. 794).

In terms of road safety-specific criteria for causation, Elvik (2011) developed criteria for causation that were predominantly based on Hill's principles and will therefore not be fully described here. Importantly though, Elvik put great emphasis on three aspects in particular: controlling for confounding; scientifically evaluating the mechanism for the change in risk, and having plausible theoretical explanations for a change in risk, in terms of the laws of physics or behaviour. This supports the approach of this thesis, in which an analytical observational study will be conducted to identify potential risk factors and subsequent behavioural studies will be performed to investigate potential mechanisms for the increase in risk, taking into account theories of road user behaviour in order to examine evidence for causation.

2.3 Data analysis methods

Accident prediction models (APMs, sometimes also known as Safety Performance Functions, or SPFs) are statistical models developed to model the relationship between crashes and a set of predictor variables (or risk factors), for example, traffic volume and road geometry, for a set of road segments or intersections. APMs are usually developed using cross-sectional studies. They are usually multivariable, that is, more than one risk factor is included in the set of predictors. Models can be developed for two main purposes: to predict the number of crashes on a given road segment or to establish the relationship between one or more risk factors and traffic crashes, adjusted for the other risk factors included in the model. The outcome of interest is the frequency of crashes on a particular road segment or intersection. Where exposure is different between road segments, the outcome can also be expressed as a rate with crash frequency as the numerator and some measure of exposure as the denominator. Commonly used measures of exposure are segment length (crashes per km), traffic volume (crashes per thousand vehicles), combined segment length and traffic volume (crashes per thousand vehicle km) or traffic density (crashes per thousand vehicles per lane, crashes per thousand vehicle km per lane).

2.3.1 Multiple regression

Previous researchers in the field have used a number of different types of regression techniques to develop accident prediction models. In these models, the random variable Y_i represents the number of crashes observed on the *i*th road segment over a period of time, with X_{i1} through X_{in} denoting the values of the n predictor variables (for example, traffic volume, characteristics of the road environment) for road segment *i*. β_o (the intercept) and β_1 through β_n are the regression coefficients estimated during the modelling process. Under the Generalised Linear Model (GLM) framework, the regression techniques differ according to the link function that relates the number of crashes to the predictor variables and the related distributional assumptions.

2.3.1.1 Linear regression

Early statistical modelling work in the field of road safety used multiple linear regression to develop accident prediction models. In these models, the number of crashes, Y, is continuous and normally distributed with mean μ and variance θ^2:

$Y_i \sim N(\mu,\theta^2)$

The expected number of crashes (E[Y], or μ) is modelled as an additive linear function of the predictors with identically distributed errors with a mean of zero and constant variance, as shown in equation 1.

$$E[Y_i] = \mu = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_n X_{in} = \boldsymbol{\beta}\, \boldsymbol{X_i} \quad , \varepsilon \sim i.i.d.\ N(0, \theta^2) \tag{1}$$

Linear regression, however, is not appropriate for modelling crashes because crash frequency (count) data do not meet the distributional assumptions necessary for linear regression models. Count data are discrete (not continuous), they are always zero or positive integers (that is, the crash count cannot be negative) and are often highly skewed. Under the central limit theorem, count data distributions will approach normality as the mean increases in size, (Hilbe, 2012), however, the mean number of crashes on a set of road segments is rarely high enough for this to occur. Because crash frequency data do not display the distributional properties necessary for normal linear regression models the resulting estimates are untrustworthy for describing the relationship between a risk factor and crash frequency.

2.3.1.2 Regression models for count data

Regression models for count data, such as Poisson regression and negative binomial (NB) regression have consistently been demonstrated to be more appropriate than linear regression for modelling crash frequency (count) data. For example, Miaou and Lum compared linear regression and Poisson regression techniques and found that the latter led to a superior model fit (Miaou & Lum, 1993).

2.3.1.2.1 Poisson regression

The Poisson distribution is used to model the number of independent events occurring in an interval of time. If the discrete random variable Y_i (the number of crashes on the ith road segment over a period of time) is Poisson distributed, $Y_i \sim Pn(\mu_i)$, then;

$$\Pr[Y = y] = \frac{e^{\mu}\, \mu^{y}}{y!}, \qquad y = 0,1,2 \ldots, where\ E[Y] = V[Y] = \mu$$

In Poisson regression, the log link function is used such that the natural log of expected number of crashes is a function of the linear predictor, as represented in equation 2 (or alternately, in equation 3). The form of the equation, with the natural log transformation, ensures that the predicted number of crashes will never be negative.

$$\ln(E[Y_i]) = \ln(\mu) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_n X_{in} = \boldsymbol{\beta X_i} \quad (2)$$

$$E[Y_i] = \mu = e^{(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_n X_{in})} = \boldsymbol{e}^{(\boldsymbol{\beta X_i})} \quad (3)$$

One limitation of using the Poisson distribution to model crash frequencies is that the variance is constrained to be equal to the mean (μ). When the variance is less than the mean, data are said to be underdispersed, likewise, when the variance is greater than the mean, the data are overdispersed. Crash frequency data are often overdispersed. If overdispersion is present in the data, then the regression coefficient parameter estimates obtained using Poisson regression will be unbiased, however, the standard errors of the parameter estimates will be biased. This has implications for statistical testing because the p-values and 95% confidence intervals (CIs) will be inaccurate (Cameron & Trivedi, 2013; Hilbe, 2012).

Various characteristics of the data can cause overdispersion. Overdispersion in crash frequency data will occur if the variation between counts is in excess. It can occur if the crash counts on different segments are not independent of one another, that is, if there is a correlation between outcomes. This can happen when data are clustered (Hilbe, 2012), for example if a set of intersections were chosen on two particular routes, then the crash counts at intersections on the same route may be correlated . Overdispersion may result if the outcome is an aggregate count of a mix of different Poisson distributions. For example, overdispersion can occur if the outcome to be predicted is a combination of different crash types (e.g. total crashes) rather than separate crash types if they have different distributions and predictors (Jonsson, 2005), thus some researchers advocate for modelling different crash types separately (Qin, Ivan, Ravishanker, & Liu, 2005).

An ill-fitting model can also lead to apparent overdispersion, that is, data can appear to be overdispersed due to a less than optimal model, rather than the actual data distribution (Hilbe, 2012). A Poisson regression model can appear to be overdispersed if an important predictor variable is omitted: this results in variance that is unaccounted for in the model. This may be especially important in road safety research where data for important risk factors can be difficult to measure or may be unavailable from traditional data sources. Historically, many accident prediction models in road safety were only concerned with establishing the relationship between traffic volumes and crashes, and omitted other sources of risk which could have potentially led to an apparently overdispersed model. It is not unreasonable to assume that there are other important risk factors for crashes apart from traffic volume, and so these models would be expected to show apparent overdispersion due to omitted variables. Apparent overdispersion can also occur if there are outliers in the data, if the model excludes important interactions between variables, if a predictor variable requires transforming or if the link between the outcome and the linear predictor (e.g. the log link in Poisson regression) is misspecified.

If overdispersion is found to be present after fitting a Poisson model, it is advisable to first check for apparent overdispersion (Hilbe, 2012), for example, by adding potentially important predictors, interactions between variables, checking for the need to transform predictor variables or the outcome, adjusting for outliers and testing whether the link function is appropriate. This process is important because if overdispersion is assumed without checking for these possibilities, and corrective techniques are used, the resulting standard error estimates will be incorrect. If overdispersion remains, other modelling approaches are warranted (Hilbe, 2012). For example, scaled standard error estimates or robust, bootstrap or jack-knife standard error estimates can be used. Alternatively, different model forms can be employed to deal with the overdispersion.

2.3.1.2.2 Negative binomial regression

The most common model form used to deal with overdispersion is the negative binomial regression model. The traditional negative binomial distribution (NB2) is similar to the Poisson distribution but includes a gamma-distributed error term to remove the constraint that the variance must equal the mean (equation 4).

$$E[Y_i] = \mu = \boldsymbol{e}^{(\boldsymbol{\beta X_i})} \qquad \mathrm{Y}_i \sim NB(\mu_i, \alpha) \tag{4}$$

The mean of the negative binomial distribution equals μ while the variance ($\mu + \alpha\mu^2$) includes an overdispersion parameter (α) that can be estimated using maximum likelihood techniques. Thus the Poisson distribution can be considered a specific case of the negative binomial distribution where overdispersion (α) is equal to zero. The likelihood ratio test is used to determine whether α (the overdispersion parameter) is significantly different to zero, and thus, whether Poisson or negative binomial regression is more appropriate for modelling the data. A large number of accident prediction models have been developed using negative binomial regression, however, it is unclear whether the overdispersion was real or apparent (possibly due to omitted predictors).

2.3.1.2.3 Zero-inflated models

Other situations that could lead to apparent overdispersion in models for count data are an excess number of observations with a zero count, a lack of observations with a zero count, censored or truncated data (Hilbe, 2012). Both Poisson regression models and negative binomial regression models can suffer from these problems.

Of these problems, it is common in road safety research for a large proportion of sites to have no crashes occur over the observation period of a study. Zero-inflated models (Zero-inflated Poisson (ZIP) and Zero-inflated negative binomial (ZINB)) can be used to model data where there are an excess number of zero observations. The underlying theoretical rationale for the use of zero-inflated models is that there are two processes responsible for generating counts of zero. As an example, imagine a study that aims to investigate the factors associated with the number of fish caught by visitors to a national park. There are two reasons that a visitor to a national park would not have caught any fish. The first is that they did not go fishing. The second is that they might have gone fishing but were not successful. Thus there are two processes generating the zero counts (the number of people who did not catch any fish). To put this in road safety terms, there are assumed to be two reasons for zero crashes to be observed on a road segment. One reason is that the road segment is inherently safe (Shankar, Milton, & Mannering, 1997) and that no crashes will occur (that is, it is in a zero crash state). The other reason is that the road segment is not inherently safe, but that by chance, due to the rarity of crash events, no crashes have been observed on this segment. It has also been suggested that the process behind the zero accident state is not that the segment is truly safe, but that the crashes that occur are of lower severity and the excess zeroes reflect under-reporting of lower severity crashes (Miaou, 1994; Shankar et al., 1997).

Zero-inflated models allow the modelling of both data-generating processes in one model and thus enable the identification of factors that are associated with a road segment being inherently safe (through modelling a binary outcome of safe/not safe using a logistic regression component) at the same time as identifying factors associated with the frequency of crashes (with a component of the model employing regression for count data). Many studies have employed zero-inflated models for predicting crashes. Interpreting zero-inflated models can, however, be difficult (Miaou, 1994). For example, a ZINB model for run-off road crashes found that as vertical curve length increased, the frequency of crashes decreased but the probability of the road segment being inherently safe (that is, in a zero count state) also decreased (Lee & Mannering, 2002). It is difficult to conceptualise how a risk factor could decrease the frequency of crashes while also making a road segment more likely to have crashes – this could be a sign of poor model specification.

It is debatable, however, whether or not applying zero-inflated models to road safety problems is theoretically sound (Lord, Washington, & Ivan, 2005). The assumption that some road segments are inherently safe (in a zero-crash state) and that crashes will never be observed on that road segment is hard to justify. Crashes are an outcome of a complex combination of factors relating to the road users, the vehicles and the environment and thus it is highly contentious to classify a road segment as inherently safe. Because crashes are rare events, a count of zero crashes on a road segment is not unexpected, even over a relatively long time period. Under the Poisson distribution, when the mean is low a large proportion of units are predicted to have a zero count. For example, with a mean of 1, 37% of the observations are predicted to be zero, while when the mean is 2, 14% of observations are predicted to be zero (refer to Figure 2.1). Thus, a large proportion of zero counts can be predicted under the Poisson (and negative binomial) distributions without resorting to zero-inflated models. For these reasons, the use of zero-inflated models was only considered in the current study if it was clear that the number of zero counts was in excess of that predicted by the Poisson or negative binomial models. The observed probabilities can be compared to the predicted probabilities as a guide to see whether the number of zero counts is in excess of that predicted and the Vuong test can be used to statistically test whether the zero-inflated model provides better fit to the data than the model without zero inflation. This approach is in accord with the advice of a previous study that compared Poisson, negative binomial and ZIP models for modelling crashes and found that the estimates derived from maximum likelihood estimation were similar for the three approaches, and that ZIP should only be used if overdispersion is too high for traditional count models (Miaou, 1994).

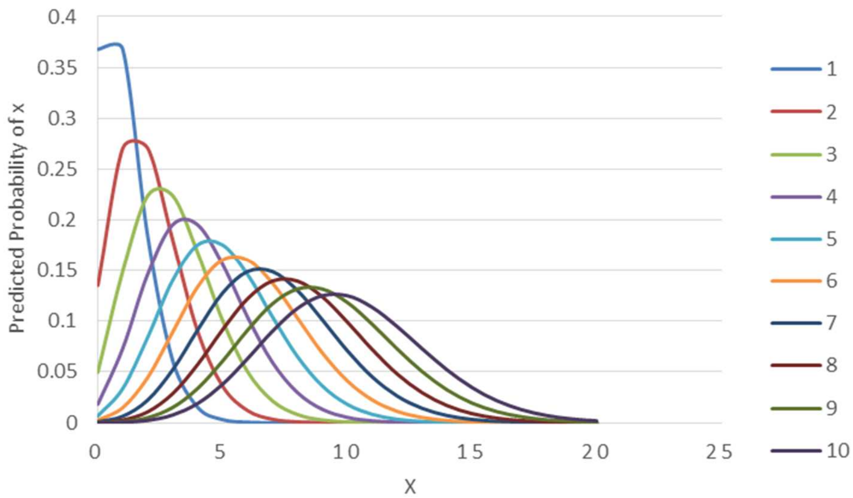


Figure 2.1 Poisson distribution: Probability of x (predicted values) for different mean values

2.3.1.2.4 Random parameters and mixed effects models

The coefficients estimated using traditional Poisson and negative binomial regression models describe the average effect of a risk factor on the outcome across road segments. That is, the coefficients are fixed and cannot vary across road segments (nor does the model intercept). It is possible however, that the effect of a risk factor may vary across different road segments. It is suggested that this could occur because of unmodelled environmental effects, or due to factors related to the drivers or vehicles (which are not included in the models). Random parameters count models allow parameter estimates to vary across sites, either just one parameter, for example the intercept (El-Basyouny & Sayed, 2009), some parameters (Anastasopoulos & Mannering, 2009; El-Basyouny & Sayed, 2009), or all parameters where the effect of the risk factor at particular sites can be determined. While this approach seems potentially useful, the process is computationally intensive, complex (Lord & Mannering, 2010) and requires a range of assumptions about the distribution of the parameters (Hilbe, 2012). The resulting model may not generalise to other datasets (Lord & Mannering, 2010) which suggests the possibility that the model may be overfitted.

The practical outcomes of random parameter models in terms of specifying the relationship between risk factors and crashes require consideration. Generally, after fitting a random parameters model, the researchers report the proportion of road segments for which a risk factor increases risk and the proportion for which it decreases risk (Anastasopoulos & Mannering, 2009; El-Basyouny & Sayed, 2009). For example, in one study, median barriers were found to decrease crash frequency on

98% of the road segments and increase crash frequency on 2% (Anastasopoulos & Mannering, 2009). Inconsistencies across road segments raise the question of why a road feature might affect risk on road segments in different ways. What was the difference between the road segments on which medians were protective and those on which medians increased crashes? It is possible that the presence of another road feature might be responsible for the difference in risk? For example, perhaps roadside parking might influence the effectiveness of medians and the proportion of the segment with roadside parking may be impacting on the estimates of median effectiveness. If so, then including an interaction term in the model between roadside parking and medians would explicitly model this relationship and potentially remove the need for complex random parameters modelling techniques. It is also possible that the different groups of drivers or vehicles on the different road segments affect the risk in some way. If, however, the road segments are likely to have similar driver and vehicle populations, then this effect would be similar across segments. Thus, the need for random parameters may simply be an indication that there are omitted variables or that the model is misspecified. If the goal of the study is to lead to practical outcomes in terms of identifying risk factors for crashes, then estimating the overall effect of a risk factor over a group of road segments is a sensible first approach. Further studies can then be carried out to determine how a factor increases risk, thus leading to targets for countermeasures.

2.3.1.2.5 Modelling incidence rates

The outcome (or response variable) in count data models is the frequency of some occurrence, in this case, traffic crashes. The rate of occurrence per unit of exposure (e.g. time, segment length, traffic volume) can also be modelled by including an offset term in the model. To demonstrate, consider the Poisson regression model. If the exposure is measured in t units, then the Poisson rate model is shown in equations 5 to 7 (which are equivalent).

$$ln(\mu/t) = \boldsymbol{\beta}\, \boldsymbol{X_i} \tag{5}$$

$$ln(\mu) = \boldsymbol{\beta}\, \boldsymbol{X_i} + \boldsymbol{ln(t)} \tag{6}$$

$$\mu = t\,(\boldsymbol{e}^{(\boldsymbol{\beta} \boldsymbol{X_i})}) = \boldsymbol{e}^{(\boldsymbol{\beta} \boldsymbol{X_i} + \boldsymbol{ln(t)})} \tag{7}$$

To model an incidence rate, the natural log transformation of the exposure variable is thus entered into the model with its coefficient set to equal one. This means the model is effectively treating crashes per unit of exposure as the outcome, therefore assuming that there is a directly proportional relationship between crashes and the exposure.

It is worth mentioning that a small number of studies have used Tobit regression to directly predict crash rates, e.g. crashes per vehicle miles travelled (Anastasopoulos et al., 2012; Anastasopoulos et al., 2008; Xu, Kouhpanejade, & Saric, 2013). In Tobit regression, the crash rates are treated as continuous data that are left censored at zero. Tobit regression predicts a latent dependent variable that is observed only if it is greater than zero and assumes an additive linear model structure with normally distributed errors not unlike linear regression for the observed range. As yet, there are no published studies comparing Tobit regression with regression models based on the Poisson or negative binomial distributions, however, the authors of one paper did briefly mention that analysis of the same data using both techniques led to similar variables being identified as significantly associated with crash risk (Anastasopoulos et al., 2008). Until formal comparative studies are conducted, it would seem prudent to choose a model form that has consistently been shown to be appropriate for crash data, that is, Poisson or negative binomial regression, and to use the offset term to model incidence rates if desired.

2.3.1.2.6 Interpretation of regression coefficients

The coefficients (β_1 through β_n) describe the relationship between the particular road environment characteristics, or risk factors (X_1 through X_n), and the frequency of crashes. Interpreting the relationship between the predictors and the outcome in terms of the natural log of the counts is not very intuitive. Instead, it is easier to conceptualise the effect of a predictor on the outcome in terms of how a change in the predictor affects the incidence rate.

An incidence rate ratio can be derived that compares the incidence rate for one level of the risk factor to the incidence rate for another (reference) level of the risk factor. Consider a binary risk factor (e.g. the absence/presence of a median) where the absence of the median is indicated by a zero and the presence of a median is indicated by a one. A Poisson model with this as the only risk factor is shown in equation 11.

$$ln(u)=\beta_0+\beta_1(X_{1:}\text{median indicator}) \tag{11}$$

As such, when there is no median present,

$$ln(u_{no\ median}) = \beta_0 \qquad (12)$$

Similarly when there is a median present, then

$$ln(u_{median}) = \beta_0 + \beta_1. \qquad (13)$$

The difference in the natural log of the expected counts between when there is a median present and when there is not is shown in equation 14;

$$\beta_1 = ln(u_{median}) - ln(u_{nomedian}) = ln(u_{median}/u_{nomedian}) \qquad (14)$$

Therefore, exponentiation of the coefficient gives the ratio of the incidence of crashes when there is a median present compared to the incidence of crashes when there is no median present, that is, the incidence rate ratio (IRR) (equation 15);

$$e^{\beta 1} = u_{median}/u_{nomedian} = IRR \qquad (15)$$

Similarly, for continuous variables, the IRR is interpreted as the difference in the incidence of crashes with a one unit increase in the predictor variable. In a multivariable model, the IRR is an estimate of the relative incidence rate when all other factors in the model are held constant, and is therefore invariant to the level of other factors which is useful for interpreting the association between a risk factor and crash incidence.

Given that it is a ratio, an IRR of one indicates that the incidence rate in each group (e.g. roads with and without medians) is the same. An IRR greater than one shows that the incidence rate in the group with the risk factor (e.g. the segments with medians) is greater than that in the group without the risk factor (the reference group, e.g. the road segments without medians) and the magnitude of the IRR indicates how much greater the incidence rate is. For example, an IRR of 2.0 shows that the incidence rate amongst the group with the risk factor is double the incidence rate amongst the group without the risk factor, while an IRR of 1.5 can be interpreted as a 50% increase in incidence. If the IRR is less than one, then the incidence rate is lower when the risk factor is present. For example, an IRR of 0.5 indicates the incidence rate in the group with the risk factor is half, or 50% less, than the group without the risk factor. In this situation, the risk factor is said to be protective. A 95% CI provides evidence of the precision of the IRR estimate. If the 95% CI includes one, then there is no statistical evidence (at the 5% level) that there is a significant difference between the incidence rates in the two groups. If, however, the 95% CI does not include one, then there is a significant difference (at the 5% level) in the incidence rates between the two groups.

There are other ways to express the effect of a predictor on the outcome in regression models for count data. Econometricians, including many researchers in road safety, express the effect in terms of an elasticity, rather than an IRR (e.g. Abdel-Aty & Radwan, 2000). For continuous variables, the elasticity indicates the percentage change in the outcome that occurs with a 1% change in the log of a continuous predictor. For categorical variables (e.g. a binary variable indicating the presence or absence of some feature) a pseudo-elasticity is calculated.

For the purposes of presentation and interpretation of the results in this thesis, the coefficients will be transformed into IRRs. Where comparison is to be made with previous research that reported elasticities, IRRs will be calculated from the regression coefficients reported in previous studies.

2.3.2 Non-parametric modelling approaches

Other non-parametric approaches have been used to investigate the effect of traffic volumes and road geometry on crashes. The advantage of non-parametric methods is that they do not require pre-definition of the structure of the relationship between predictors and outcomes.

One such approach is the use of regression tree based methods (Chang & Chen, 2005; Karlaftis & Golias, 2002). In this approach, the independent variable (or risk factor) that explains most of the variance in the outcome is selected. The next step is to determine how to split this variable into two categories in order to achieve the maximum reduction in variability. The process is repeated iteratively; that is, the questions are again posed; "What is the next variable to achieve the maximum reduction in variability, and how should it be categorised to do this?" The result is a hierarchical tree diagram which can be used to predict crashes on other road segments. A disadvantage of tree based regression methods is that they rely on the binary grouping of segments according to independent variables and therefore lose valuable information for determining the relationship between a continuous variable and crashes. The independent effects of a risk factor are also difficult to determine. In addition the method could be unwieldy when considering a large number of risk factors.

Neural networks have also been used to analyse crash frequencies (Chang, 2005). Neural networks consist of a number of input nodes representing risk factors (e.g. traffic volume or road geometry characteristics) and a number of output nodes (e.g. the number or crashes) which are connected by hidden nodes. The input, hidden and output nodes have weights associated with their connections. The network learns the relationship between the inputs (road characteristics) and the outputs (number of crashes) by feeding in the characteristics of a set of sites, comparing the output (the number of crashes predicted at those sites) to the observed number, and then training the network by feeding back through the network and adjusting the weights between units until the desired outcomes are obtained. Then the performance of the network in predicting crashes is tested on an entirely new set of sites, by feeding in the site characteristics and comparing the resulting output with the actual observed number of crashes. This main downfall of the neural networks is that they are computationally irreducible. The programmer has no knowledge of what is occurring at the hidden unit level, and cannot interrogate the network to investigate specific associations or test their statistical significance. The effects of a predictor on the outcome can be estimated by removing that input unit from the trained model and determining the effect on crash frequency, however, this method does not allow for the assessment of interactions between variables, something that may be of interest in this thesis.

Genetic programming is another non-parametric approach that has been applied to predict crash frequency (Das, Abdel-Aty, & Pande, 2010; Das & Abdel-Aty, 2011) which the authors claim strikes a good balance between accurate prediction and the interpretation of the effect of predictors. Many interactions are identified using this technique, which raises the possibility of overfitting. This technique has only been applied in road safety by a handful of researchers and software to perform the technique is not readily available so it will not be considered further for this thesis.

2.4 Implications for this study

Analytical cross-sectional studies are the most tractable and hence frequently used study design for identifying the aspects of the road and surrounding environment that are associated with crash frequency. The advantages of this design are that it affords the estimation of associations between multiple risk factors and crash outcomes, while using information on the number, and not just the occurrence of crashes. Regression models for count data are the most appropriate statistical technique for analysing crash frequency data because of their application to discrete non-negative count outcomes. The models will be developed using fixed parameters, however, a sensitivity analysis will be conducted on the final models to determine if random parameters are required (that is, to determine whether there is evidence that the baseline risk differs across sites, or whether the relationship between the risk factors and crash frequency varies across sites). Depending on the distribution of the data, Poisson regression, negative binomial regression or their zero-inflated counterparts (ZIP, ZINB) may be the most appropriate model form, and this can be ascertained using diagnostic statistical tests. Control of confounding is possible with the use of multivariable models and estimates of the association between individual risk factors and crash frequency can be derived, unlike some of the non-parametric approaches. Therefore, the statistical modelling component of this thesis will comprise a cross-sectional study of the association between characteristics of the built urban environment and crash frequency that will be analysed using regression models for count data. The choice of the most appropriate type of regression model for count data will be informed by diagnostic testing.

CHAPTER 3. THE BUILT URBAN ENVIRONMENT AND CRASH FREQUENCY

This chapter contains a targeted review of previous multivariable studies that investigated the relationship between the built environment and crash frequency on urban roads. The review was conducted to ascertain the current state of knowledge and identify the gaps that exist to inform the design of this study by:

- determining which methods have been most commonly used in studies that specifically investigated the relationship between the built urban environment and crashes
- identifying common methodological limitations of previous research with the potential to affect internal and external validity
- investigating what aspects of the built environment have been studied in the past (in relation to crash risk) and what factors have been neglected
- ascertaining the level of evidence available to support the presence (including direction and magnitude) of a relationship between various aspects of the built environment and crash risk.

First, the broad aspects of the environment with the potential to influence crashes are proposed in terms of the conceptual framework presented in Chapter 1. Next, the literature search is described and the relevant literature is reviewed, first in terms of the methods used and then in terms of the results obtained. The methods of the studies that have been conducted are described, with respect to the study design, modelling approaches and types of crashes that were included. Similarities between study methods are identified and study limitations are discussed. In particular, limitations that are common to many of the studies that are likely to affect the validity of the results will be discussed and will help to inform the design of the research conducted for this thesis. Then, each of the previously studied risk factors are presented, classified by type (exposure, roadway, roadside, land use and sociodemographic factors). Attention is given to how the risk factor was measured in previous research, along with a summary of the results of studies that have included that factor. Conflicting results are identified. Finally, the results are summarised according to the factors that have been investigated and the weight of evidence for, or against, an association with crash frequency. Other potential risk factors that have not been previously studied in multivariable studies of crash risk in urban environments are proposed.

3.1 Potential risk factors

In order to identify the range of risk factors of interest for the targeted literature review, general aspects of the environment with the potential to influence urban crashes were identified with reference to the conceptual framework for the thesis that was presented in Chapter 1. The road system sits within the broader environment, both social and physical. Aspects of both the social and physical environment have the potential to influence the number of crashes that occur on a road segment.

The social environment encompasses laws, regulations, practices and policies, however, the effect of these on crashes are not usually considered in studies at the level of the road segment. This is because road segments are usually sampled from areas under the same legal or policy jurisdictions and over time frames within which these remain stable. Likewise, they will not be included in this review or in this study. The social environment also includes the norms, priorities and values of society which influence the behaviour of individuals. Norms, priorities and values, and hence behaviour, have the potential to vary across different population groups (e.g. by age, sex, socioeconomic status). Behaviour, or human activity within the built environment is considered to be part of the built environment (Handy et al., 2002). In addition, population size and density may be considered as indirect measures of exposure to crash risk. Hence sociodemographic characteristics of the local area are considered as potential risk factors for urban crashes because of their relationship with behaviour and exposure.

In urban areas, the physical environment consists mainly of the built environment and the human activity that occurs within it. Human activity within the road environment, as measured by traffic volumes and mix and activity of other road users, is a direct measure of exposure to risk. Other broad categories of the built environment with the potential to influence urban crash risk are the design of the road, the roadside environment and land use. Land use is indicative of environmental complexity and is also correlated with human activity in the broader built environment because it influences the number and type of road users and their movements. Relative to the built environment, the natural environment is not a strong contributor to risk in urban environments. Roadside vegetation could be considered a part of the natural environment but in urban areas it is more likely to be planned landscaping than naturally occurring vegetation. As such, roadside vegetation is treated as part of the roadside built environment for the purposes of this research. Weather is sometimes considered as a risk factor but if the road segments are sampled from within a limited geographic area, then weather will vary little across the units of analysis and from year to year. As the focus of this thesis is a defined geographical area (the Melbourne metropolitan area), the influence of weather was not investigated further.

3.2 Literature search

A literature search was conducted using the engineering database Compendex, the life sciences database Medline, PsycINFO (which covers psychology, social, behavioural and health sciences) and the Australian Transport Index (ATRI), which includes published and unpublished Australian studies of transport. Broad search criteria were used to identify multivariable studies that investigated the relationship between the social and physical environment and urban traffic crashes, with emphasis on traffic, infrastructure (road and/or roadside), land use and/or sociodemographic factors. Abstracts were examined and full-texts of articles were downloaded into EndNote (Version X6) and inspected to establish that they were multivariable studies of the relationship between the built environment and traffic crash occurrence or frequency in urban areas. Because the focus of this study was on crash occurrence or frequency on complex urban road segments, the following types of articles were excluded:

- Research into the factors associated with crash severity in the event of a crash
- Articles that focused exclusively on signalised intersections or relatively simple urban road segments (e.g. limited access roads, freeways, highways, motorways)

- Studies that did not include, or control for, traffic volume (because traffic volume is such a strong predictor of crash frequency)

Initially the literature search was restricted to studies that focused on risk at the level of the road segment but only one of these studies included an investigation of the effect of sociodemographic risk factors on urban crashes (Alavi, 2013). Rather, sociodemographic factors were more commonly investigated in studies with the area (not the road segment) as the unit of analysis. Therefore, in order to assess the evidence for a relationship between sociodemographic factors and crash frequency, studies of area-level (macro-level) influences were included in the review but only if they also included data for exposure (e.g. traffic demand) and some measure of the built environment (e.g. the road network).

Studies that included both urban and rural roads were only included in the review if separate statistical models were developed for urban roads. This resulted in the exclusion of several studies that included both urban and rural roads but did not report the results for urban and rural areas separately (Abdel-Aty & Radwan, 2000; Hamann & Peek-Asa, 2013; Milton & Mannering, 1998; Pande & Abdel-Aty, 2009; Shankar et al., 1997).

Studies focused at the level of the road segment (or location) that did not distinguish between midblock and intersection crashes were also excluded — three case-control studies that investigated infrastructure-related risk factors for crashes involving vulnerable road users on urban roads were excluded for this reason (Harris et al., 2013; Roberts, Norton, Jackson, Dunn, & Hassall, 1995; Stevenson et al., 1995).

Twenty-two studies were identified and met the criteria for inclusion in this targeted review of the association between the built environment and urban crashes: the unit of analysis was the road segment in 15 of the studies and the area was the unit of analysis in seven studies.

3.3 Methodological review of previous research

This section contains a review of the methods used in the 22 studies that met the inclusion criteria for the targeted literature review. The purpose of this section is to critically evaluate the methods in order to identify major limitations of previous research and to assess the potential implications for the validity of results of previous studies.

The methods are summarised briefly in this section and a detailed summary is provided in Table B1 in Appendix B. For each study, the detailed summary includes information about the authors, the year the study was published, where the study was conducted, the study design, time period, data sources, units of analysis (including information on road type), the types of crashes included, the modelling approach (including model form and details of diagnostic testing), the variables (or risk factors) that were considered, including whether or not the factor was significantly associated with crashes, and further information relevant for critically evaluating the study. Section 3.3.1 contains a review of the 15 studies with the road segment as the unit of analysis, while in Section 3.3.2 the seven studies with the area as the unit of analysis are reviewed.

3.3.1 Studies at the level of the road segment or corridor

All of the fifteen studies in which the unit of analysis was the road segment or corridor used a cross-sectional design to investigate environmental risk factors for crash frequency on urban midblock road segments (Abdel-Aty et al., 2009; Alavi, 2013; Avelar et al., 2013; Bonneson & McCoy, 1997; Brown & Tarko, 1999; Greibe, 2003; Jackett, 1993; Jonsson, 2005; Lee, 2000; Manuel, El-Basyouny, & Islam, 2014; Potts, Harwood, & Richard, 2007; Sawalha & Sayed, 2001; Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, & Teng, 2013) or on urban road segments clustered into corridors (El-Basyouny & Sayed, 2009). It is important to remember that cross-sectional studies can only be used to determine if there is an association between a risk factor and crashes, hence the conclusions that can be drawn from these studies are limited to correlation and not causation. Since cross-sectional studies are observational, results must be evaluated in terms of the potential for confounding (omitted variable bias) which is of particular concern in studies that only include a small number of potential risk factors.

3.3.1.1 Road type

It is important to note what type of roads were included in the studies because the effect of the built environment may vary according to road type. Most (13/15) of the studies focused on arterial roads or road links (Abdel-Aty et al., 2009; Avelar et al., 2013; Bonneson & McCoy, 1997; Brown & Tarko, 1999; El-Basyouny & Sayed, 2009; Greibe, 2003; Jackett, 1993; Jonsson, 2005; Lee, 2000; Potts et al., 2007; Sawalha & Sayed, 2001; Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013) while two included collector roads (Jackett, 1993; Manuel et al., 2014). One study included all roads in the Central Business District (CBD) in Melbourne, Australia (Alavi, 2013). The focus of this thesis is on strip shopping centre road segments on arterial roads in metropolitan Melbourne, hence the results of the previous research should be relevant for this particular road type.

3.3.1.2 Data sources

In all of the 15 studies data were sourced from existing administrative sources held by government departments — most commonly from traditional sources such as crash databases and road inventories held by road authorities (e.g. Departments of Transport). Though existing sources are extremely valuable for conducting cost-effective research, the research question is inevitably limited by the data available and omitted variable bias can be a problem if data are not available to control for potential confounders. For this reason, ten studies supplemented the existing data with data collected by the researchers: in six cases, the extra data were collected from Government-owned videos of roads or aerial images, e.g. Google Earth (Avelar et al., 2013; Bonneson & McCoy, 1997; Brown & Tarko, 1999; Manuel et al., 2014; Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013), while in four cases, the researchers collected data on-site (Alavi, 2013; Greibe, 2003; Jonsson, 2005; Sawalha & Sayed, 2001). The amount of data that can be collected specifically for the project is also limited by project funds and timelines, therefore, many studies only included a limited range of risk factors so the potential for confounding and omitted variable bias in the estimation of the models remains.

3.3.1.3 Crash type and severity

Studies differed according to the severity of crashes included, which was constrained by the types of crashes that were included in the available crash databases. Many crash databases only include crashes that result in injury. If a crash database does include property damage only crashes, there is usually a threshold cost that must be reached for the crash to be included. It is important to know what types of crashes were included so as to determine what type of risk is being measured as defined by the Road Trauma Chain (Cameron, 1992). In addition, property damage only crashes are more common than injury crashes and risk factors may differ between crashes of different severity. From a public health perspective, it is arguably more important to prevent crashes that result in injury than those that only result in property damage.

Three studies developed separate models for crashes of different severities. Abdel-Aty et al. (2009) developed two separate models to identify the predictors of the frequency of all types of crashes and severe (incapacitating and fatal) crashes, while Brown and Tarko (1999) developed separate models for all crashes, fatal crashes, injury crashes and property damage only (PDO) crashes. Potts et al. (2007) developed separate models for all crashes, fatal & injury crashes and PDO crashes.

Nine studies developed models for all crashes that occurred on the road segments of interest. Of these, two included both injury and property damage only crashes (Bonneson & McCoy, 1997; Greibe, 2003) while the other seven did not provide information as to whether the crashes included property damage crashes as well as injury crashes (Avelar et al., 2013; El-Basyouny & Sayed, 2009; Jackett, 1993; Manuel et al., 2014; Sawalha & Sayed, 2001; Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013). This omission is unfortunate, for the reasons outlined above.

It is likely that there are different risk factors for different crash types. Most studies investigated crashes aggregated by type, while other studies developed separate models to identify risk factors for different crash types. Abdel-Aty et al. (2009) reported the results of analyses for all crashes and rear-end crashes separately, while Potts et al. (2007) reported different estimates for all crashes, MVC and SVC. Separate models for injurious MVC, SVC, BVC and PVC were developed by Jonsson (2005). Lee (2000) concentrated solely on run-off road SVC, although it was unclear whether property damage crashes were included or not. One study concentrated exclusively on pedestrian crashes (Alavi, 2013), with separate models developed for crashes involving pedestrians resulting in injury that occurred during weekdays and weeknights.

3.3.1.4 Risk factors investigated

Table 3.1 provides details on the risk factors included in each study with the road segment or corridor as the unit of analysis, categorised by the characteristic of the environment that they describe (the relationship between each risk factor and different types of crashes is presented in Section 3.4). Cells shaded grey indicate that particular study did not include any risk factors from that broad category. It is clear that the majority of studies focused mainly on traffic volumes and specific characteristics of the road (e.g. number of lanes, speed limit, intersection density, median type or roadside parking) with little detailed consideration given to other aspects of urban environments that might affect risk.

Despite the different types of vehicles, road users and facilities for vulnerable road users that are common in urban environments, these factors were not commonly investigated in relation to urban crash risk. Of the fifteen studies, only four included data on public transport facilities (Alavi, 2013; Greibe, 2003; Manuel et al., 2014; Sawalha & Sayed, 2001), three considered bicycle facilities (Alavi, 2013; Greibe, 2003;

Jonsson, 2005), two included data on pedestrian and cyclist volumes (Alavi, 2013; Jonsson, 2005) and the distribution of traffic was only included in one study (Alavi, 2013). Four included data on the presence or rate of pedestrian crossings (Alavi, 2013; El-Basyouny & Sayed, 2009; Jonsson, 2005; Sawalha & Sayed, 2001) and two considered the presence or width of footpaths (Alavi, 2013; Manuel et al., 2014). Only two studies considered roadside hazards. Nine studies did include broad classifications of land use (Bonneson & McCoy, 1997; Brown & Tarko, 1999; El-Basyouny & Sayed, 2009; Greibe, 2003; Jackett, 1993; Jonsson, 2005; Sawalha & Sayed, 2001; Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013), while one study had more detailed data on a large range of different types of land use and amenities in terms of capacity and floor space (Alavi, 2013).

A lack of existing data and the difficulty and cost of collecting data specifically for the project are one possible reason for the omission of a large range of factors that may influence crash risk in urban areas. Another possible reason is that the researchers did not consider that characteristics of the built urban environment beyond the design of the road itself had the potential to influence crashes. This is a major limitation of the previous research and constitutes a large gap in knowledge.

Table 3.1 Risk factors (by broad category) considered in studies included in review of risk factors for crashes on urban roads: studies with the road segment as the unit of analysis

Reference	Length	Traffic volumes (tv)/ density	Cross-section, speed limit	Intersections	Public transport, bicycle & pedestrian facilities	Roadside	Land use & amenities	Sociodemographic	Other
(Abdel-Aty et al., 2009)	ln(length)	ln(tv)	# lanes, speed limit						
(Alavi, 2013)	length	tv, # crossing pedestrians	# lanes, median type, grade, direction of travel, orientation, speed limit	# minor roads and alleyways, # driveways	Bus lane, # bus routes, # bus stops, # tram routes, # tram stops, tram stop type, taxi rank, bicycle lane, pedestrian crossings, pedestrian lights, footpath width	Parking, clear zone	shop density, distance from railway station, car parks, specific types of land use (area or capacity) within 14 buffer zones	population density, employment density	% truck traffic
(Avelar et al., 2013)	ln(length)	ln(tv)	# lanes, median type	# driveways (diff types)					
(Bonneson & McCoy, 1997)	ln(length)	ln(tv)	median type	density of unsignalised approaches, driveways		parking	land use		

Component 1: Identifying risk factors for crashes

Reference	Length	Traffic volumes (tv)/ density	Cross-section, speed limit	Intersections	Public transport, bicycle & pedestrian facilities	Roadside	Land use & amenities	Sociodemographic	Other
(Brown & Tarko, 1999)	offset	offset	# lanes, median type, outside shoulder, speed limit	density of unsignalised approaches, driveways, accesses. Proportion of accesses signalised, channelised, with right turn lanes			land use		
(El-Basyouny & Sayed, 2009)	ln(length)	ln(tv)	# lanes	density of unsignalised intersections	pedestrian crossing density		land use		
(Greibe, 2003)	offset	ln(tv)	# lanes, central island, direction of traffic, road width, speed limit, speed reducing measures	# minor crossings, exits, side-roads	bus stops, cyclist facilities, footpath	parking	land use		

Reference	Length	Traffic volumes (tv)/ density	Cross-section, speed limit	Intersections	Public transport, bicycle & pedestrian facilities	Roadside	Land use & amenities	Sociodemographic	Other
(Jackett, 1993)	offset	offset	# lanes, median type, speed limit	# intersections (diff types)			land use		road class
(Jonsson, 2005)	offset	ln(tv) ln(pedestrian volumes) ln(cyclist volumes)	# lanes, visibility, median type, speed limit	# intersections (diff types), distance between intersections, exits	bicycle separation, vulnerable road user crossings	parking	land use		road function, average speed, crossing behaviour

Component 1: Identifying risk factors for crashes

Reference	Length	Traffic volumes (tv)/ density	Cross-section, speed limit	Intersections	Public transport, bicycle & pedestrian facilities	Roadside	Land use & amenities	Sociodemographic	Other
(Lee, 2000)	N/A	tv/lane	lane width, median width, shoulder width, shoulder length, # & length of vertical curves, vertical grade, guardrail length & height, distance from shoulder to guardrail, bridge length, speed limit	# at grade intersections		# catch basins, # culverts,* ditch depth, *fence length,* # miscellaneous fixed objects,* # utility poles,* # sign supports, *# light poles, *# tree groups, * # isolated trees, distance from shoulder to all marked* & side slopes			
(Manuel et al., 2014)	ln(length)	ln(tv)	presence of curve,	access point density	bus route, footpath				

Component 1: Identifying risk factors for crashes

Reference	Length	Traffic volumes (tv)/ density	Cross-section, speed limit	Intersections	Public transport, bicycle & pedestrian facilities	Roadside	Land use & amenities	Sociodemographic	Other
			presence of midblock road width change, road size						
(Potts et al., 2007)	offset	ln(tv)	lane width, shoulder width			parking, roadside hazard rating			
(Sawalha & Sayed, 2001)	ln(length)	ln(tv)	# lanes, median type	density of unsignalised intersections, driveways	bus stops/km, pedestrian crosswalks/km	parking	land use		
(Xu, Kouhpanejade, et al., 2013)	denominator for rate and length as predictor	denominator for rate, tv/lane	median type, media opening density, speed limit	driveway density			land use		average speed
(Xu, Kwigizile, et al., 2013)	denominator for rate and length as predictor	denominator for rate, tv/lane	median type, media opening density, speed limit	driveway density			land use		average speed

3.3.1.5 Data analysis

The methodological review in Chapter 2 established that regression models for count data were the most appropriate analysis method for analysing data from cross-sectional studies of the relationship between the environment and crash frequency. Regression models for count data were used to analyse the data in the majority (13/15, 87%) of the studies in this targeted review. All of these studies provided a rationale for the application of regression models for count data to crash frequencies but few reported conducting investigations to choose the most appropriate model form for their data, e.g. whether to use Poisson or negative binomial regression and whether zero-inflated models were warranted (e.g. Alavi, 2013; Lee, 2000). Negative binomial regression was the most common regression technique applied to the crash frequency data (Abdel-Aty et al., 2009; Avelar et al., 2013; Bonneson & McCoy, 1997; Brown & Tarko, 1999; Lee, 2000; Manuel et al., 2014; Potts et al., 2007; Sawalha & Sayed, 2001). The authors of three studies conducted analyses using Poisson regression (Alavi, 2013; Greibe, 2003; Jackett, 1993), while Jonsson (2005) used a quasi-Poisson model with scaling to address overdispersion. El-Basyouny and Sayed (2009) compared three different models: Poisson lognormal, Poisson lognormal with a random intercept and Poisson lognormal with random parameters to determine if there was evidence for parameter estimates to vary significantly by segment. A ZIP model was also used in one study (Alavi, 2013). Finally, two studies that did not use models for count data included travel speed as a predictor and used statistical methods to account for the endogeneity between travel speed and crashes; one used a tobit model with endogenous variable (Xu, Kouhpanejade, et al., 2013) and the other used random coefficient simultaneous equations to predict both travel speed and crashes (Xu, Kwigizile, et al., 2013).

When developing statistical models, it is important to test how well the model fits the data and whether distributional assumptions are met. The results of such tests should be reported in the journal article or report. Despite their importance, diagnostic testing and reporting of model fit were not commonly reported; only four studies reported conducting analyses of residuals, four reported goodness of fit and one study reported both. Without this information it is difficult to assess the adequacy of the models that were developed and the likely influence on the validity of the results.

3.3.1.6 Other issues

When summarising the results, it is important to take into account that some of these studies are based on similar, or the same, data and would thus be expected to lead to similar conclusions. The studies by El-Basyouny and Sayed (2009) and Sawalha and Sayed (2001) both used data collected from arterial roads in Vancouver, Canada from 1994-1996, although the first study also included data from Richmond, Canada that did not appear to be used in the subsequent study. Thus the estimates from these two studies are not completely independent of each other. Two other studies that were published concurrently in different journals used different statistical techniques to analyse almost the same data-set of crashes on divided arterial roadways in Las Vegas, USA collected between 2003 and 2005 (Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013). While the use of different techniques on almost the same data did identify some of the same risk factors, some risk factors were identified in one model but not the other; therefore the evidence for these risk factors being associated with crashes is less convincing. The related studies will thus need to be considered when assessing the weight of evidence for or against a risk factor being associated with crash frequency.

3.3.2 Studies at the level of the area

A cross-sectional study design was used in all of the seven studies with the area (rather than the road segment) as the unit of analysis. (Dumbaugh & Li, 2011; Dumbaugh, Li, & Joh, 2013; Gruenewald et al., 1996; Hadayeghi, Shalaby, & Persaud, 2003; Haynes et al., 2008; LaScala, Gerber, & Gruenewald, 2000; LaScala, Johnson, & Gruenewald, 2001) but the studies differed markedly in terms of the overall aims and the types of crashes that were included. Haynes et al. (2008) studied the influence of road curvature on fatal crashes at the territorial local authority level in New Zealand separately for urban and rural areas. Hadayeghi et al. (2003) investigated risk factors for injury and property damage only traffic crashes in traffic zones in Toronto. Dumbaugh and colleagues focused on the relationship between the built environment and different types of motorist crashes, pedestrian crashes and bicycle crashes in one study (Dumbaugh & Li, 2011) and pedestrian and bicycle crashes in another (Dumbaugh et al., 2013). It was unclear whether these two studies included only injury crashes or whether property damage crashes were also included. The role of neighbourhood characteristics including alcohol outlet density was of interest in three USA studies of alcohol-related crashes that focused on single vehicle night-time crashes (Gruenewald et al., 1996), pedestrian involved crashes with and without alcohol involvement (LaScala et al., 2000) and pedestrian injuries suffered in collisions, with and without alcohol involvement (LaScala et al., 2001). The last study focused entirely on injuries but it is unknown whether property damage crashes were also included in the other two studies focusing on alcohol-related crashes.

The area-level studies were included in the review in order to assess the evidence for a relationship between sociodemographic factors and crashes, controlling for traffic exposure and road network. Given there were only seven studies and the studies included different types of crashes, the review is only likely to find weak evidence (if any) between sociodemographic factors and specific crash types.

3.3.2.1 Data sources

Each of the seven studies that used the area as the unit of analysis obtained data from existing government data sources, including crash databases and traffic data from road authorities. A broader range of sources were used compared to those with the road segment as the unit of analysis, for example, sociodemographic details were obtained from the population census and information on alcohol outlets was sourced from regulatory authorities. Two studies also obtained data from a population-based telephone survey (Gruenewald et al., 1996; LaScala et al., 2001). This highlights the opportunity for this thesis to seek data from a range of different sources beyond those traditionally used in road safety research.

3.3.2.2 Risk factors investigated

Table 3.2 provides further detail on the risk factors included in each study with the area as the unit of analysis, categorised by the characteristic of the environment that they describe (the relationship between each risk factor and different types of crashes is presented in Section 3.4). Cells shaded grey indicate that particular study did not include any risk factors from that broad category. Exposure was accounted for in all of the models through traffic related measures such as vehicle miles travelled, traffic flow or arterial mileage and measures of the size of the area or length of the road network. All seven studies included data on the sociodemographic characteristics of the areas and five also investigated the association between broad land use and crashes. Six of the seven studies included data relating to the complexity of the road network as measured by the number of intersections, intersection density or network density. In all but one, these were the only variables relating to the road system that were included. The exception was the study conducted in New Zealand by Haynes et al. (2008) which also comprised data describing curves, topography, temperature and precipitation. This introduces the potential for confounding if the road and roadside infrastructure vary across different sociodemographic groups (e.g. if areas of high socioeconomic rank have different infrastructure to areas of low socioeconomic rank).

Table 3.2 Risk factors (by broad category) considered in studies included in review of risk factors for crashes on urban roads: studies with the area as the unit of analysis

Reference	Exposure: length, area and/or traffic	Road network	Sociodemographic	Land use & amenities	Other
(Dumbaugh & Li, 2011)	Vehicle miles travelled per km, arterial mileage, freeway mileage, total area of block	# 3 leg intersections, # 4 leg intersections	Population density	# big box stores, # strip commercial uses, #	
(Dumbaugh et al., 2013)	Vehicle miles travelled per km, arterial mileage, total area of block	# 3 leg intersections, # 4 leg intersections	Median household income, population density, % population aged 5–17, % aged 65+	# big box stores, # strip commercial uses, # pedestrian scaled retail uses	
(Gruenewald et al., 1996)	Local traffic flow, highway traffic flow. All measures expressed per km of roadway	Network density	Population density, average age, % single, % male, % income <$20,000, % income>$60,000, % unemployed, % white	alcohol outlet (bars, restaurants, off-premise) density	Self-reported drinking behaviour of sample
(Hadayeghi et al., 2003)	Major road km, minor road km, total road km, area, traffic demand in morning peak (volume:capacity, in-flow, out-flow, vehicle km travelled, total flow)	Posted speed, # intersections, intersection density	Total population, population density, # households, household density, full-time employment, part-time employment, total employment, employment density, # vehicles, # vehicles/household,		
(Haynes et al., 2008)	ln(vehicle km)	Measures of road curvature: bend density, detour ratio, cumulative angle	Population, % aged 15–24, % aged 75+, % overseas tourists, % households with 1+ or 2+ cars, % drive to work, % driven		Topography (mean and standard deviation of altitude) for whole

Reference	Exposure: length, area and/or traffic	Road network	Sociodemographic	Land use & amenities	Other
		turned, mean angle turned	to work, socioeconomic status (deprivation level).		area and area with roads. Weather: mean daily temperature, precipitation, wind speed, minimum daily temperature
(LaScala et al., 2001)	Local traffic flow, highway traffic flow. All measures expressed per km of roadway	Cross-street density	Population density, average age, % single, % divorced/ widowed, % male, % income <$20,000, % income>$60,000, % unemployed, % white, % black, % Hispanic, % with college education, average number of persons/ household <18	alcohol outlet (bars, restaurants, off-premise) density	Self-reported drinking behaviours of sample
(LaScala et al., 2000)	Average daily traffic flow/ km of roadway in area (usually measured for high volume roads or high crash locations extrapolated to all other intersections, weekday only), all measures expressed per km of roadway	# cross-streets	Population, % aged 0–15, % aged 16–29, % aged 55+, % unemployed, % never married, median income, % males, % high school graduate or higher. Ethnic group (excluded from final analysis)	alcohol outlet (bars, restaurants, off-premise) density	

3.3.2.3 Data analysis

Negative binomial regression models were used to analyse the data in four of the studies (Dumbaugh & Li, 2011; Dumbaugh et al., 2013; Hadayeghi et al., 2003; Haynes et al., 2008), while the other three studies used multiple linear regression techniques and accounted for spatial correlation between units (Gruenewald et al., 1996; LaScala et al., 2000; LaScala et al., 2001). Although spatial correlation was accounted for, the use of multiple linear regression for count data is problematic as described in Chapter 2. In one of the studies, the outcome was log transformed in an attempt to overcome the problem (LaScala et al., 2000), however, this can lead to biased estimates (O'Hara & Kotze, 2010) so regression models specifically formulated for count data are preferred. Only three studies reported the results of diagnostic testing for model fit therefore it is difficult to assess model adequacy for the other four studies.

3.3.2.4 Other issues

Dumbaugh published two studies with colleagues on crashes that occurred between 2003 and 2007 in the San Antonio-Bexar county metropolitan region, USA (Dumbaugh & Li, 2011; Dumbaugh et al., 2013). Both studies included analyses of pedestrian crashes and bicycle crashes, although they did differ slightly in that the second study assessed a wider range of sociodemographic factors than the first. The latter study will be used when evaluating the evidence for or against a risk factor being associated with PVC or BVC.

3.3.3 Summary

A major methodological limitation of previous studies (both at the level of the road segment and the area) is the limited range of risk factors included in the models. In general, the studies that investigated risk factors for crash frequency on road segments tended to include traffic volumes and a small number of road design factors but the roadside environment and human activity have been relatively neglected. Likewise, the area-level models only included broad estimates of road network complexity (e.g. intersection density) that are unlikely to capture the variation in road design and features across different areas. These omissions raise the possibility of confounding and biased estimates of association due to omitted variables.

A more general comment is warranted with regard to the inconsistent reporting of methods and results in these studies. Certain details about the methods and results must be included in a scientific journal article or report in order for a reader to be able to critically evaluate the research. Several of the studies lacked detail of the crashes that were studied, in terms of their severity level, which has implications for being able to define the risk being measured. There was a serious lack of information provided about whether the models fitted the data well, whether the model form was appropriate and the results of other diagnostic tests. In most cases it was unclear whether any diagnostic testing had been performed at all. This makes it difficult for the reader to assess the validity of the results.

3.4 Relationship between the urban environment and crashes

In this section, the results of previous studies of the association between aspects of the built urban environment and crashes are reviewed. Only the studies that considered the road segment (or corridor) as the unit of analysis are reviewed when considering the built-environment risk factors that are specific to the road segment, for example, exposure (length and traffic volume), traffic mix, road cross-section, speed limit, intersections, roadside parking, facilities for bicycles and pedestrians and public transport. The one exception is that the area-level study by Haynes et al. (2008) is included in the discussion of the influence of curves on crashes because that was the main focus of their study. Except for road network length and broad estimates of travel exposure and network or intersection density, studies that considered the area as the unit of analysis rarely included road and roadside-related risk factors. The importance of these area-focused studies, however, will become apparent when discussing the relationship between sociodemographic and land use characteristics of the urban area and crashes.

The risk factors that have been previously studied and the evidence for a relationship with urban crashes are presented by broad category. Meta-analysis is a statistical technique whereby the results of previous studies are combined to derive an overall estimate of the relationship between a particular risk factor and an outcome (e.g. frequency of crashes). Meta-analysis is appropriate when it is sensible to investigate the average effect across studies (Borenstein, Hedges, Higgins, & Rothstein, 2009). For various reasons, however, it was judged not to be appropriate to obtain combined estimates of effect from the previous studies reviewed in this thesis. In many situations, there was only one study of the effect of a risk factor on that particular crash type so a meta-analysis was not possible. Where there was more than one study of the effect of a risk factor, the estimates of association were not strictly comparable. The studies differed in terms of the types of crashes that were included and the specification of the risk factors (e.g. treating the number of the lanes as a continuous variable vs. a categorical variable). All studies differed in terms of the other risk factors included in the models. Given that the estimates of association derived from a regression model are adjusted for the other factors included in the model, this means that the results were not directly comparable. Therefore, meta-analyses were not conducted.

The results of previous research were therefore summarised qualitatively. At the end of each section there is a table (Table 3.3, Table 3.4, Table 3.5, Table 3.6, Table 3.7) that contains a summary of the associations between the different risk factors in that category and urban crashes, by crash type. The tables also include a judgement of the weight of evidence for each relationship. The weight of evidence was judged according to the number of studies (or estimates) that supported the association. For studies that developed multiple estimates, only independent estimates were counted. For example, Alavi (2013) developed models for PVC that occurred during the day and the night. These were counted as two separate estimates, because the data used to develop the two models were mutually exclusive. If an association was found in only one study, the evidence was rated as weak; if there were two studies that supported the association, the evidence was rated as weak to medium; three studies in agreement gained a rating of medium; and if there were four or more studies with a consistent association, evidence was rated as strong. The direction of association had to be consistent across all studies that investigated that factor, otherwise, the relationship was judged as equivocal. The one exception to this was if there were 6 or more studies of a risk factor and at least four found a consistent and significant association, the evidence was rated as medium-strong as long as the other studies did not find an association in the opposite direction.

The cells in the tables are shaded to aid interpretation; positive relationships between risk factors and crashes are shaded in green, with the shade becoming stronger as the weight of evidence for an association strengthens. Negative relationships are shaded in blue: again, the shading becomes stronger as the weight of evidence strengthens. Non-significant associations are shaded in grey, with the shade deepening as the evidence against the existence of a relationship strengthens. Black cells indicate that there were no studies that investigated that particular association. White cells indicate equivocal results.

An issue to keep in mind when presenting the results of past studies is that some researchers developed multiple models and therefore have multiple estimates of the effect of risk factors for different road types, crash types or crash severities. In most cases, it is impossible to determine if these estimates differ significantly from each other, that is, to determine if there is truly a difference in the effect of a risk factor across different crash types. For example, Potts et al. (2007) attempted to develop 90 separate models; one for each of five arterial road types (2 lane undivided, 3 lane with two-way left turn lane (TWLTL) median, 4 lane divided, 4 lane undivided and 5 lane TWLTL median) by 3 crash types (all crashes, multi-vehicle crashes (MVC) and single-vehicle crashes (SVC)) by 3 severity levels (all crashes, KSI crashes and PDO crashes) for 2 locations (Minnesota and Michigan). Discussion of this study is limited to attempts to discover consistent patterns in the results.

Before presenting the findings from previous research it is essential to explain that the coefficients describing the relationship between a predictor variable and crash frequency estimated using regression models must be interpreted as the association between that variable and the outcome, holding all other variables in that model constant (providing there are no interactions with that variable and others). Therefore estimates from models that include different sets of predictors may not be directly comparable, particularly if confounders have been included in one model and not the other.

Each of the different categories of risk factors are now presented, and the evidence for an association between risk factors and crash frequency (overall and by crash type) is assessed.

3.4.1 Exposure

Exposure is one of the key variables in the integrated conceptual framework for this thesis and it is essential to account for factors affecting exposure to risk in a study of crash occurrence. One method of accounting for exposure involves selecting all of the units of analysis to be equal in terms of the exposure. Another is to adjust for exposure during the analysis process. Factors affecting exposure to risk include time, traffic volumes, other road user volumes (e.g. vulnerable road users like pedestrians and bicyclists) and road segment length.

3.4.1.1 Time period

In most studies, the time period over which crashes are tallied is the same for each road segment. If, however, the time period for counting crashes differs across road segments, time is almost always included in regression models for count data as an offset term (that is, with the natural log of the exposure entered into the model with the coefficient set to equal one, as explained in Chapter 2). The rationale is that if crashes are independent random events the number of crashes should be directly proportional to the number of days over which the crashes are counted. Some models included time (e.g. number of years) as an offset regardless of whether the time period differed across units, so that the predicted number of crashes can be expressed as an incidence rate per year which makes it easy to compare across studies that measure the same risk.

3.4.1.2 Segment length

Road segments are sometimes selected to be of equal length (Lee, 2000) or selection can be based on other factors, for example, the length between signalised intersections (Bonneson & McCoy, 1997; Sawalha & Sayed, 2001; Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013), the length over which the road design characteristics were homogeneous (Avelar et al., 2013; Brown & Tarko, 1999; Jackett, 1993) or homogenous road sections between signalised intersections (Jonsson, 2005).

If segments are the same length, then there is no need to adjust for segment length in the analysis. If not, then length must be included in the modelling process. Five studies that used regression models for count data controlled for segment length by including it as part of the offset term (Brown & Tarko, 1999; Greibe, 2003; Jackett, 1993; Jonsson, 2005; Potts et al., 2007) effectively modelling the frequency of crashes per km (or per vehicle km where length is combined with traffic volume as a measure of exposure). The assumption that there was a direct linear relationship between crash frequency and segment length was only tested explicitly in one of these five studies (Brown & Tarko, 1999).

All but one (Alavi, 2013) of the remaining studies with road segment or corridor as the unit of analysis included the natural log of the segment length as a predictor of crash frequency in a regression model for count data (Abdel-Aty et al., 2009; Avelar et al., 2013; Bonneson & McCoy, 1997; El-Basyouny & Sayed, 2009; Manuel et al., 2014; Sawalha & Sayed, 2001). The resulting formulation of the model assumes that crash frequency is proportional to segment length raised to some power (which is the coefficient to be estimated — see equation 1 for an example).

$$\mu = e^{\boldsymbol{\beta 1 ln}(\boldsymbol{length}))} = \text{Length}^{\beta 1} \qquad (1)$$

If the power (β_1) is not significantly different to one, then there is no evidence that crash frequency is not directly proportional to segment length.

Crash frequency increased as segment length increased in all of these studies but it is unclear whether the relationship was directly proportional since this was only tested in one study (Brown & Tarko, 1999). Estimated coefficients from models ranged from 0.36 (Avelar et al., 2013) to 1.24 (Abdel-Aty et al., 2009). The majority of estimates lay between 0.80 and 1.0.

Alavi (2013) entered untransformed segment length into regression models for count data. The number of pedestrian injury crashes increased exponentially with every extra metre in segment length on weekdays and weeknights. This form of the relationship between length and crashes is different to what is usually modelled; however, it is possible that the relationship is different in the CBD (the location for his study) where pedestrian volumes are much higher than in other metropolitan areas. Alavi hypothesised that the exponential increase in risk as segment length increased might be due to vehicles having the opportunity to travel at higher speeds when there is a longer distance between intersections in the CBD. It could also be that pedestrians were less likely to use designated pedestrian crossings on longer blocks (Ewing & Dumbaugh, 2009).

If crashes are random independent events, then we would expect crash frequency to increase in a directly proportional relationship with segment length. Longer segments between intersections, however, may afford drivers the opportunity to drive faster while providing fewer designated places for pedestrians to cross. This could lead to an exponential rise in crashes with segment length. The majority of studies where the effect of segment length on crash frequency has been explicitly estimated have found estimates that are close to one, but there is not enough evidence to determine if the number of crashes is directly proportional to segment length in urban areas. Although the mechanism behind this non-linear relationship (if it exists) is not yet established, it indicates that segment length should not be included in the offset term of regression models for count data without first establishing whether there is a directly proportional relationship to crash frequency.

3.4.1.3 Traffic exposure

3.4.1.3.1 Traffic volume

Traffic volume is an important contributor to exposure to risk and is usually measured as the annual average daily traffic volume (AADT) or expressed as a combined measure of traffic volume and segment length (e.g. vehicle km, or vehicle miles). While some studies have included real-time traffic data (e.g. Abdel-Aty, Pande, Uddin, Dilmore, & Pemmanaboina, 2005) or hourly traffic volumes (e.g. Qin, 2006), these data are rarely available. None of the studies that met the criteria for this review used real-time or hourly traffic volumes, so they will not be discussed further here.

While there are various ways in which the relationship between traffic volume and crash frequency can be entered in a regression model, results have consistently shown that as traffic volumes increase, so does the frequency of crashes.

Traffic volume is sometimes included as an offset, effectively modelling the number of crashes per unit of traffic volume (Brown & Tarko, 1999; Jackett, 1993). This approach assumes the relationship between the number of crashes and traffic volume is directly proportional. Brown and Tarko (1999) tested this assumption prior to including traffic volume in the offset term but Jackett (1993) did not. Treating traffic volume as an offset has intuitive appeal because as the number of vehicles rises, so does both the number of vehicles on the road that could have a single-vehicle crash and the opportunity for conflict between vehicles (MVC).

Previous research, however, has found that while crashes do increase with an increase in traffic volume, the number of crashes and traffic volume are not directly proportional. It is possible that as traffic volumes increase and congestion occurs, the opportunity for injury-causing conflicts decreases. Thus many researchers use a model of the form shown in equation 2, where β_1 is a coefficient to be estimated (Abdel-Aty et al., 2009; Avelar et al., 2013; Bonneson & McCoy, 1997; El-Basyouny & Sayed, 2009; Greibe, 2003; Haynes et al., 2008; Jonsson, 2005; Manuel et al., 2014; Potts et al., 2007; Sawalha & Sayed, 2001). One advantage of including the natural log of traffic volume in the model is that the crash frequency will equal zero if traffic volume equals zero, which has face validity.

$$\mu = e^{\beta 1 ln(AADT))} = \mathrm{AADT}^{\beta 1} \qquad (2)$$

If β_1 is not significantly different to one, AADT can be entered in the model as an offset in order to reduce the number of parameters to be estimated (e.g. Brown & Tarko, 1999). If, however, β_1 is significantly different from one traffic volume should not be included in the offset (e.g. El-Basyouny & Sayed, 2009; Greibe, 2003). Often, though, a test is not performed and there is not always enough information provided to be able to conduct the test post-hoc (Abdel-Aty et al., 2009; Avelar et al., 2013; Bonneson & McCoy, 1997; Manuel et al., 2014; Sawalha & Sayed, 2001).

In all studies that used regression models for count data and included the natural log of traffic volume as a predictor, crash frequency increased as traffic volume increased but the estimate of $\beta_{1,}$ varied widely, particularly between studies that aggregated all crashes together (range 0.48 to 1.93). Notwithstanding the possibility that confounders were unaccounted for, the reason for the large range was unclear. Results become more consistent, however, when crashes were disaggregated by type. In general, the coefficients for MVC were above one (Abdel-Aty et al., 2009; Dumbaugh & Li, 2011; Jonsson, 2005; Potts et al., 2007), whereas the coefficients for SVC were below one (Greibe, 2003; Jonsson, 2005; Potts et al., 2007). Although the evidence was limited to one study, the coefficients for crashes involving vulnerable road users were also below one (Jonsson, 2005). It could be that crashes are more likely to occur between multiple vehicles than for single vehicles or vulnerable road users as traffic volumes become more congested.

Similar to segment length, a directionally proportional relationship between traffic volume and crash frequency should not be assumed (that is, the traffic volume should not be entered into the model as an offset term) without first testing whether the assumption is correct.

3.4.1.3.2 Traffic density

It is possible that in urban areas where traffic volumes are high and congestion is common, traffic density (vehicles per lane or vehicle km per lane if the segments differ in length) may be a more important predictor of crashes than traffic volume, especially for MVC. For example, a two lane road with an AADT of 20,000 vehicles per day has, on average, double the traffic density of a four lane road with the same traffic volume. Denser traffic provides more opportunity for conflict between vehicles.

Despite the potential importance of traffic density for influencing crash frequency in urban areas, few studies of urban midblock road segments have assessed the role of traffic density. Only three of the 15 studies of road segments included traffic density as a potential risk factor. Lee (2000) did not find a significant relationship between run-off road SVC frequency on a principal arterial and AADT per lane in urban areas. Xu and colleagues found that an increase in AADT per lane was associated with a reduction in crashes on divided arterials in Las Vegas using tobit regression (Xu, Kouhpanejade, et al., 2013) but failed to find a significant relationship between AADT per lane and crashes using random coefficient simultaneous equations on almost the same data (Xu, Kwigizile, et al., 2013).

The lack of previous research means there is no strong evidence of the relationship between traffic density and crashes on urban midblock road segments. Traffic density, however, has been shown to be associated with increases in truck crashes on rural interstate highways in the USA (Miaou, 1994) and crashes on principal arterials that traverse both urban and rural areas in Washington State (Milton & Mannering, 1998). More evidence is needed regarding the relationship between traffic density and crashes in urban areas.

3.4.1.3.3 Traffic distribution

The distribution of traffic (e.g. percentage of traffic that is made up of trucks) as a potential risk factor has been largely overlooked in studies of urban roads. Alavi (Alavi, 2013) did not find a significant relationship between the percentage of truck traffic and the frequency of pedestrian-vehicle crashes (PVC) on weekdays or weeknights in the CBD of Melbourne, Australia. None of the other researchers included measures of traffic distribution in their models. It has, however, been shown to be negatively associated with reductions in truck crashes on rural interstate highways in Utah (Miaou, 1994) and all fatal, injury and non-injury crashes on interstate highways in Indiana (Anastasopoulos & Mannering, 2011). It is therefore worthwhile investigating if crash risk on urban roads is related to the percentage of truck traffic.

3.4.1.4 Pedestrian volume

Pedestrian volumes were measured by the two researchers who conducted studies of crash frequency on urban road segments that specifically focused on pedestrian crashes. Jonsson (2005) counted the number of pedestrians walking along each road segment and the number of pedestrians that crossed the road during one fifteen minute period for each road link. The product of these (number of pedestrians walking straight x number of pedestrians crossing) was found to be the best exposure-based predictor of pedestrian crashes as shown in equation 3.

$$\text{No. pedestrian crashes } \alpha \text{ (Pedestrians walking x Pedestrians crossing)}^{0.38} \quad (3)$$

Alavi (2013) counted the number of pedestrians crossing the road on midblock segments during 20 minute periods at eight different times during the day that corresponded to peaks and troughs in pedestrian activity in the Melbourne CBD. The best predictor (in terms of pedestrian volumes) of pedestrian injury crashes on weekdays was the number of pedestrians crossing the road at midblock locations multiplied by the AADT which was positively related to weekday pedestrian crash frequency. In contrast, pedestrian crossing volumes were not significantly associated with pedestrian injury crashes on weeknights.

While pedestrian volumes are likely to be an be an important predictor of pedestrian crashes in urban areas, consistent data for pedestrian volumes are rarely available from existing sources (Alavi, 2013). If pedestrian crashes are of interest and the research budget and scope does not allow for on-site collection of pedestrian volumes by the researchers, other methods to measure pedestrian activity should be sought.

3.4.1.5 Bicycle volume

Only one study at the level of the road segment measured bicycle volumes (Jonsson, 2005). Similar to the measurements for pedestrian volume collected for the same study, the number of cyclists riding straight and the number crossing the road were counted during one fifteen minute period on each road link. The estimated relationship with bicycle crash frequency is shown in equation 4. An increase in bicycle volumes led to an increase in bicycle crashes, however, the rate of increase diminished as bicycle volumes grew.

$$\text{No. cycle crashes } \alpha \text{ (cycles riding x cycles crossing)}^{0.35} \qquad (4)$$

The relationship between bicycle volumes and other crash types (apart from bicycle crashes) was not estimated.

Similar to pedestrian volumes, cyclist volumes are not consistently collected for all roads and are rarely available from existing administrative data. Again, estimates of bicycle exposure derived from other sources may be useful if data collection is not feasible, particularly if BVC are of interest.

3.4.1.6 Evidence for the influence of exposure on crash frequency

Table 3.3 contains a summary of the evidence for the influence of different types of exposure on crash frequency for each type of crash. There was evidence for a non-linear relationship between AADT and aggregated crashes (strong evidence), MVC (strong evidence), SVC (medium strength) and PVC and BVC (weak). The difference in the level of evidence reflects the fewer studies of urban midblocks that focused on crashes involving vulnerable road users. This non-linear relationship strongly suggests that the outcome measure in models of urban crashes should be crash frequency not a rate with the number of vehicles as the denominator. Similarly, AADT should not be entered as an offset term in regression models for count data unless the assumption of proportionality has been explicitly tested. Instead it should be entered as a covariate in the regression model with the consideration of a non-linear relationship with the outcome.

Other exposure-related variables were investigated so rarely that there is only weak or equivocal evidence for a relationship for certain urban crash types, and no evidence for other urban crash types. There is some evidence that pedestrian volumes are related to urban PVC, at least those that occurred in the daytime, and that bicycle volumes are related to urban BVC, but the effect of pedestrian and cyclist volumes on other urban crash types was not investigated. This is likely due to the lack of existing data (Alavi, 2013) and the time and cost associated with collecting pedestrian and cyclist data specifically for the research project. It is likely that this will be a barrier to obtaining pedestrian and cyclists volumes in this thesis, however, other cost-effective avenues for obtaining surrogate exposure data will be explored. Likewise, there is little evidence regarding the relationships between traffic density and traffic distribution (percentage trucks) and urban crashes. These data are more readily available than vulnerable road user volumes, and thus can be readily explored for their effect on urban crashes.

Table 3.3 Direction of association and weight of evidence for relationship between exposure-related risk factors and urban multi-vehicle crashes (MVC), single-vehicle crashes (SVC), pedestrian-vehicle crashes (PVC) and bicycle-vehicle crashes (BVC)

Risk Factor	Crash type	# studies	Direction of association (Weight of evidence)				
			All crashes	MVC	SVC	PVC	BVC
Traffic volume	All MVC SVC PVC BVC	8 4 3 1 1	Non-linear ↑ (strong)	Non-linear ↑ (strong)	Non-linear ↑ (medium)	Non-linear ↑ (weak)	Non-linear ↑ (weak)
Traffic density	All SVC	2 1	↓ or NS (equivocal)		NS (weak-medium)		
Pedestrian volumes	PVC	2				All ↑ Day ↑ Night NS (equivocal)	
Bicycle volumes	BVC	1					Non-linear ↑ (weak)
% trucks	PVC	1				NS (weak-medium)	

↑= positive relationship (shaded in green); ↓=negative relationship (shaded in blue); NS=association not statistically significant (shaded in grey); strength of shading indicates strength of evidence. Black cells indicate no evidence, white cells indicate equivocal evidence

3.4.2 Roadway

The design of the road clearly has the potential to influence crashes through the effect on the field of safe travel and safety margins (Gibson & Crooks, 1938; Summala, 2007). This section details the relationship between a broad range of roadway design characteristics and crash frequency on urban road segments. All of the studies with the road segment or corridor as the unit of analysis assessed the influence of roadway design on crash frequency, however, there was variation across studies in the specific characteristics that were investigated.

3.4.2.1 Number of lanes

Nine of the 15 studies of urban road segments evaluated the association between the number of lanes and crash frequency. One further study was restricted to roads of only two lanes (Manuel et al., 2014). Two studies failed to provide information on the range observed in their studies which prevents sensible interpretation of their results (Avelar et al., 2013; Brown & Tarko, 1999). For example, Avelar et al. (2013) discovered that crash frequency was reduced on roads with four lane roads compared to roads that did not have four lanes, however, this is meaningless without knowing how many lanes the other road segments had. Brown and Tarko (1999) found no significant association between the number of lanes and all crashes, KSI crashes or PDO crashes. As the authors did not provide information on the distribution of the number of lanes across their sample of segments, it is possible that the number of lanes did not vary enough across road segments to observe any effects.

Several studies found a significant increase in the aggregate number of crashes as the number of lanes increased (Abdel-Aty et al., 2009; El-Basyouny & Sayed, 2009; Sawalha & Sayed, 2001) while other studies failed to find a significant relationship (Greibe, 2003; Jackett, 1993). All studies controlled for traffic volumes. There was little evidence for an association between the frequency of different types of crashes and the number of lanes. The number of lanes was not significantly associated with rear end crashes (Abdel-Aty et al., 2009), MVC, SVC, BVC (Jonsson, 2005) or PVC (Alavi, 2013; Jonsson, 2005). Given the ambiguity of previous results, it will be important to consider the number of lanes as a potential risk factor for crashes on urban roads.

3.4.2.2 Direction of travel

Whether the road allowed for one-way or two-way travel was not significantly associated with crashes on Danish road links (Greibe, 2003) or pedestrian crashes in the Melbourne CBD (Alavi, 2013). None of the other studies mentioned whether there were any one-way roads in their samples. Given the lack of evidence, if there are both one-way and two-way roads in a sample of roads then this should be considered as a potential risk factor.

3.4.2.3 Vertical curves and grade

Grade (slope) and/or vertical curves (where two grades meet) and their relationship to crash frequency were only investigated in two of the studies included in this review. Neither the number of vertical curves, vertical curve length nor vertical grade were significantly related to run-off road SVC (Lee, 2000), nor was the grade of the road associated with pedestrian crashes in the Melbourne CBD (Alavi, 2013). There is little evidence to assess whether or not vertical curves or grade are associated with crashes, therefore, it would be interesting to include this as a potential risk factor in this thesis study if data are available.

3.4.2.4 Road shoulder

Only three studies investigated whether road shoulders were related to crash frequency. In Melbourne, Australia, shoulders on urban roads are rare because roads are predominantly bounded by kerbs. As such, it is possible that road shoulders may not have been present in the other urban environments studied.

Brown and Tarko (1999) estimated that the presence of an outside shoulder reduced all crashes on Indiana urban arterial streets by roughly half. Lee (2000) estimated that urban run-off road SVC decreased by 65% for every extra metre in the length of a road shoulder but found no association between shoulder width and run-off road SVC.

Potts et al (2007) assessed the effect of shoulder width on crashes of different severities and types on five different types of arterial roads in Minnesota and Michigan and found an increase in shoulder width was associated with a reduction in all crashes on all types of roads. Further analysis by crash type revealed that this was likely due to reductions in MVC as there was no association between shoulder width and SVC, however, the evidence was limited.

3.4.2.5 Road width

Road width was significantly associated with crashes on urban road links in Denmark (Greibe, 2003). The estimated relationship was non-linear as the risk was lower for roads of width 8 to 8.5m, compared to wider and narrower roads.

Manuel et al. (2014) evaluated the effect of a change in roadway width on a midblock road segment on two lane residential collector roads. When roadway width changed midway through the segment, the crash frequency more than doubled (IRR=2.28). None of the other studies considered roadway width or the change in roadway width as a risk factor. There is, therefore, little evidence upon which to base a sound judgement of whether road width is related to crashes in urban areas. Further evidence is required.

3.4.2.6 Lane width

The main aim of the research conducted by Potts et al. (2007) was to assess the effect of lane width on safety on urban and suburban roads that differed according to the number of lanes and type of median treatment. The preferred lane width according to highway design policies in the USA (where the study was conducted) was 12 feet.

As noted previously, the authors attempted to develop 90 models, although only 38 converged and fit the data well. Lane width was significantly associated with crash frequency in 37 of the fitted models (the exception being the model for all crashes on 4 lane divided roads in Michigan). Coefficients comparing each lane width category to the reference category were reported but there was no information about the precision of the estimates, so it was not clear whether the pairwise differences were statistically significant. With so many models, the results were difficult to interpret, but it was apparent that there was no evidence of a dose-response relationship—that is, risk did not change uniformly as lane widths changed. Instead there was some evidence of non-linear relationship that varied across road type. There was no evidence that the preferred lane width of 12-feet was associated with a reduction in crashes; in fact, evidence suggested that this was the one of the least safe lane widths for two lane undivided roads and three lane roads with a TWLTL median. Four lane undivided roads were the only road type for which 12-feet lanes were associated with a lower crash frequency compared to narrower or wider lane widths.

The only other study to include lane width was Lee (2000) who established that lane widths between 3.7 and 5.39m were associated with a run-off road SVC incidence more than five times that of roads with lane widths between 1.93 and 3.69m.

With only two studies investigating the effect of lane width, and the difficulty in interpreting the results from many different models in the Potts et al. (2007) study, (with no evidence that the estimates differed significantly between models), there is currently no strong evidence for the effect of lane width on crashes on urban roads. Further evidence is therefore needed.

3.4.2.7 Horizontal Curves

Only one of the 15 studies at the level of the road segment investigated whether horizontal curvature was related to crashes on urban roads. Manuel et al. (2014) found no significant association between the presence of curves and crashes on two lane residential collector roads in Edmonton, Canada.

The main focus of one of the area-level studies was to determine the influence of road curvature on fatal crashes in NZ and separate models were developed for urban and rural areas (Haynes et al., 2008). As the cumulative angle per km turned by the roads in an urban area increased, fatal crash frequency decreased. Crashes also decreased as the detour ratio (the sum of road lengths divided by sum of straight distances) increased. There was no significant association with the bends per km or the mean angle turned per bend.

The small number of studies that investigated the effect of road curvature on urban roads is surprising, as this risk factor is commonly included in studies of risk on rural roadways and limited access highways (e.g. Abdel-Aty & Radwan, 2000; Othman, Thomson, & Lannér, 2009; Schneider, Savolainen, & Moore, 2010; Shively, Kockelman, & Damien, 2010). Further evidence is required to determine the relationship between curves and crashes on urban roads.

3.4.2.8 Speed limit

The relationship between the speed limit of the road segments and crash frequency was considered for nine of the studies of urban road segments. The results were inconsistent.

Of the studies that included all crashes aggregated by type or MVC, despite having a similar range of speed limits, some found that crash frequency increased as speed limit increased (Jonsson, 2005; Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013), some found that crash frequency decreased as speed limit increased (Abdel-Aty et al., 2009; Jackett, 1993), while others found no significant association between speed limit and crash frequency (Brown & Tarko, 1999). No significant association was found between speed limit and SVC, PVC or BVC (Alavi, 2013; Jonsson, 2005; Lee, 2000).

The speed limit of roads in Victoria (and elsewhere) is set according to the characteristics of the road, roadside development, types of road users, likely road user movements and crash history (VicRoads, 2010). Hence, speed limit is correlated with roadway design, roadside development and road users. In a study that includes these variables as potential predictors of crash frequency, speed limit would not be expected to be an independent risk factor. It is arguably more important to determine the specific aspects of the road and environment that affect risk, rather than simply speed limit, to identify specific targets for intervention. Therefore, while it is important to consider speed limit as a potential risk factor, it is absolutely essential to include the aspects of the road and roadside that are related to speed limit and crash risk.

3.4.2.9 Accesses, intersections and driveways

Thirteen of the 15 studies investigated the influence of accesses, intersections and/or driveways on crash frequency. In most cases, the density (the number per km), was entered into the models as a potential predictor. One of the two studies that did not include any measure of intersections or accesses as potential risk factors was conducted on partially limited access roads, which may explain the omission (Abdel-Aty et al., 2009).

With few exceptions, there is overwhelming evidence that as the number of accesses and intersections increases, so does the frequency of crashes aggregated by type and MVC (Bonneson & McCoy, 1997; Brown & Tarko, 1999; El-Basyouny & Sayed, 2009; Greibe, 2003; Jackett, 1993; Jonsson, 2005; Manuel et al., 2014; Sawalha & Sayed, 2001). This makes intuitive sense as accesses and intersections are locations where conflict between vehicles can occur.

The relationship between SVC and accesses and intersections was less clear. One study found that the frequency of run-off road SVC increased along with the rate of accesses per km (Lee, 2000) while another found no evidence of a relationship to SVC (Jonsson, 2005). Prior studies also found no evidence for an association between accesses and intersections and PVC or BVC (Alavi, 2013; Jonsson, 2005).

Crash frequency (aggregated by type) was also found to be positively related to driveway density (Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013). This relationship may be modified by land use—Sawalha and Sayed (2001) estimated that crash frequency increased along with driveway density but only in areas classified as business land use. The results of Avelar et al. (2013) support this: for every extra commercial or industrial driveway per km, crash frequency increased, however, for every extra driveway of other types, crash frequency decreased. Driveways were positively associated with pedestrian crashes that occurred during the day in the Melbourne CBD, however, there was no relationship with PVC that occurred during the night (Alavi, 2013).

The weight of evidence strongly supports a positive relationship between accesses (including intersections and driveways) and urban crashes, particularly MVC, hence, it is important to include such measures in models of crash frequency on urban roads.

3.4.2.10 Medians

Nine studies investigated the influence of medians on crash frequency, while one study developed separate models for undivided and divided roads (Potts et al., 2007). Xu and colleagues (Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013) focused entirely on divided roads but measured the effect of the number of median openings on crashes. It is unusual that the remaining three studies did not consider whether medians influenced risk. It is possible that their samples of roads were restricted to those either with or without medians, however, not enough information was provided to be certain (Abdel-Aty et al., 2009; El-Basyouny & Sayed, 2009; Manuel et al., 2014).

Most (8) of the studies classified the road segments according to the presence and type of median. In four of these studies, the crash frequency significantly decreased when there was a median present (Brown & Tarko, 1999; Jackett, 1993; Sawalha & Sayed, 2001); Bonneson and McCoy (1997) found that the effect of median type differed according to land use and whether or not parking was present. For roads in business and office areas, medians were associated with a reduction in crash frequency but only where parking was allowed on the side of the road. In contrast, four studies found no association between median presence or type and crashes, either aggregated by type or for MVC, SVC, PVC or BVC (Alavi, 2013; Avelar et al., 2013; Greibe, 2003; Jonsson, 2005).

One study investigated the relationship between median width and crashes. Wider medians were associated with reductions in run-off road SVC; for every extra metre in median width, the frequency of run-off road SVC on a principal arterial in Washington State was 2% lower (Lee, 2000).

While Xu and colleagues (Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013) only included divided arterials in Las Vegas in their study, they found that as the density of median openings increased, so did the crash frequency.

Despite the contradictory results, there is enough evidence that there may be a relationship between the presence and type of medians and the frequency of crashes on urban roads that this relationship should be assessed in any study of urban crashes. In addition, consideration should be given to more detailed descriptors of medians than simply presence and type, e.g. median width or the proportion of a road segment with a median.

3.4.2.11 Roadside Parking

Only six studies considered the relationship between roadside parking and crashes. The results were variable. In studies that included all crashes aggregated by type, two studies found that crashes were significantly more frequent when roadside parking was permitted (Bonneson & McCoy, 1997; Greibe, 2003) but another study failed to find a significant association between parking and crashes (Sawalha & Sayed, 2001).

It is possible that roadside parking has a different effect on different crash types. Results, however, were also variable in studies that considered different crash types separately and may have also differed by road type. In a study that included all types of road links in Sweden, there was no association between roadside parking and MVC or SVC (Jonsson, 2005). A study conducted in the USA also failed to find an association between roadside parking and MVC frequency on most road types (3 or 5 lane roads with a TWLTL median, 2 lane undivided roads and 4 lane divided roads), however, MVC frequency was significantly higher on 4 lane undivided roads with parking compared to the same road type without parking. The same study found that roadside parking was significantly associated with SVC on two-lane and four-lane roads (Potts et al., 2007). No significant association was found between roadside parking and crashes involving vulnerable road users (Alavi, 2013; Jonsson, 2005).

The results of studies investigating the potential influence of parking on crashes are inconsistent and further evidence is needed to determine the relationship to urban crashes, particularly for different crash types. The possibility that the effect is modified by road type should also be borne in mind.

3.4.2.12 Pedestrian facilities

Pedestrian activity is more prevalent in urban areas than rural areas hence models of urban road crash risk might be expected to consider pedestrian facilities as potential risk factors; however, only six of the fifteen studies included such measures. Two of these studies also accounted for pedestrian volumes in their analyses (Alavi, 2013; Jonsson, 2005).

3.4.2.12.1 Footpaths

Whether or not there was a footpath (also known as a sidewalk or footway) present was not significantly related to crash frequency (all types considered together) on Danish road links (Greibe, 2003) or two lane collector roads in residential areas of Edmonton, Canada (Manuel et al., 2014). This result must be considered in the context of the high prevalence of footpaths in urban areas: 88% of the road length considered in the Danish study had footpaths, as did 93% of the road segments in the Canadian study. There was also no association between footpath width and pedestrian crash frequency in the Melbourne CBD, adjusted for pedestrian volumes (Alavi, 2013).

3.4.2.12.2 Pedestrian crossings

Two Canadian studies (that may have shared some data) found a positive association between the number of pedestrian crossings per km and crash frequency (El-Basyouny & Sayed, 2009; 2001). Pedestrian volumes were not adjusted for in these studies. Presumably, pedestrian volumes are higher on road segments with pedestrian crossings, so it is possible that the increased risk is due to confounding (omitted variable bias).

Alavi (2013) did control for pedestrian crossing volumes and found no association between the presence of pedestrian crossing facilities and pedestrian injury crashes in the Melbourne CBD. Jonsson (2005) also controlled for pedestrian crossing volumes and estimated that the presence of a crossing for vulnerable road users increased the frequency of pedestrian crashes on Danish road links by 54%. The method for measuring pedestrian exposure, however, was more thorough in the first study; the number of pedestrians crossing the road during a 20 minute period was counted eight times throughout the day (chosen to reflect periods of differing pedestrian activity). Thus, Alavi collected data on pedestrians crossing for 160 minutes on a typical weekday for each segment. In comparison, Jonsson only collected pedestrian volume data for 15 minutes on each road link. The discrepancy in the results of the two studies could therefore be due to better control of confounding in the study conducted in the Melbourne CBD, or it could be that the presence of pedestrian crossings is an independent risk factor for crashes on Danish road links but not Melbourne CBD midblock road segments.

In terms of other specific crash types, Jonsson (2005) failed to find any association between the presence of crossings for vulnerable road users and MVC, SVC or BVC.

It is important to consider the impact of facilities for pedestrians where pedestrian activity is likely to be high, however, one must be mindful of the potential confounding effect of pedestrian volumes, and the potential for an endogenous relationship between pedestrian crossings and crashes (i.e. pedestrian crossings might be installed in areas with a pedestrian crash problem).

3.4.2.13 Bicycle facilities

Only three studies measured the influence of bicycle facilities on crashes that occurred on urban midblock road segments. Greibe (2003) found no significant association between cyclist facilities (none vs. bicycle lane vs. bicycle track) and crashes of all types and severities on Danish road links. Alavi (2013) found no significant relationship between the presence of a bicycle lane and pedestrian injury crashes in the Melbourne CBD. Finally, Jonsson (2005) did not find any association between bicycle separation and MVC, SVC, PVC or BVC.

There is a lack of information of the effect of bicycle facilities on crashes and although previous studies suggest there is no relationship to crash risk, it is important to obtain further evidence. It is also possible that the width of a bicycle lane is related to crash frequency; something that was not investigated in previous studies.

3.4.2.14 Public transport

Only four studies included any measure related to the public transport network in their models of crash frequency on urban roads. The presence of a bus route was unrelated to crash frequency on two lane residential collector roads in Edmonton, Canada (Manuel et al., 2014). The number of bus stops did not influence crash the number of crashes on Danish road links (Greibe, 2003), likewise, the number of bus stops per km was not associated with the number of crashes on arterial roads in Vancouver and Richmond, Canada (Sawalha & Sayed, 2001).

In contrast, Alavi (2013) discovered several aspects related to public transport that were associated with the risk of pedestrian crashes in the Melbourne CBD. A combined measure of the number of bus stops multiplied by the number of bus routes within 100m of the road segment was associated with an increase in pedestrian crashes on weekdays, as was a combined measure of the number of tram stops times the number of tram routes within 700m. Increases in the distance to the nearest train station or the density of bus routes within 600m were associated with a reduction in pedestrian crash frequency on weekdays. The presence of bus lanes, the type of tram stops and the presence of taxi ranks were not related to weekday pedestrian crash frequency. Weeknight pedestrian crashes increased with increases in the density of tram stops within 400m but were unrelated to tram stop type, number of tram routes, bus stops or routes, bus lanes or taxi ranks.

In Melbourne, Australia, public transport routes and stops are a distinguishing characteristic of urban roads that can have a strong impact on the traffic mix, traffic movements and pedestrian activity. The results from the study of pedestrian crashes in the Melbourne CBD underscore the importance of considering these sources of risk. It is likely that public transport related risk factors also influence risk in other concentrated urban areas. It is surprising that so few other studies included these risk factors in their models.

3.4.2.15 Road type

In most of the studies, the samples of road segments studied were restricted to one type of road (e.g. arterials or collectors) and so road type was not assessed as a predictor of crash frequency. There were two exceptions to this. Jackett (1993) found no significant difference in crash frequency between collector roads, arterial roads or strategic routes on urban roads in NZ. For Danish road links, crash frequency was highest on road links in the city, followed by tangential road links, with the lowest crash frequency on thoroughfares and entrances (Greibe, 2003).

3.4.2.16 Evidence for the influence of the roadway on crash frequency

The characteristics of the roadway that were investigated with regard to their association with crashes in urban areas are presented in Table 3.4 with an assessment of the strength of the evidence for or against a relationship. Because there were so many roadway-related risk factors investigated, the results will be summarised by type of crash.

3.4.2.16.1 Crashes aggregated by type

The density of accesses and intersections was the only roadway-related factor for which there is strong evidence of a positive relationship with aggregate crashes. For almost all other aspects of the roadway, the evidence for an association with aggregate urban crashes can only be rated as weak, weak to medium or equivocal.

There is medium strength evidence that as the number of driveways (at least for some types of properties) and median openings increase, so do aggregate crashes. The strength of evidence is weak to medium that pedestrian crossings increase aggregate crash risk. There is weak evidence that a change in roadway width increases risk and that roadside shoulders decrease aggregate crashes. There is also weak evidence for a non-linear relationship between lane width, road width and aggregate crashes.

There is medium strength evidence that there is no relationship between public transport facilities and aggregate crashes. Weak to medium strength evidence suggests there is no relationship between aggregate crashes and footpaths and there is weak evidence that bicycle facilities, direction of travel and road type are not associated with aggregate crashes.

There was equivocal evidence that the number of lanes and the presence of parking may increase aggregate crash risk, however this was not consistent across studies, as some found no association. There was also some evidence for a decrease in aggregate crashes with certain measures of road curvature, but again, this was not consistent. For speed limit and medians, the evidence is truly equivocal: previous research has found positive relationships, negative relationships and no effect. No studies of aggregate urban crashes investigated vertical curvature or grade.

3.4.2.16.2 Multi-vehicle crashes

Few studies developed separate models for MVC so the strength of evidence for the influence of various factors on MVC is only rated as weak or weak to medium. There is weak evidence that MVCs increase as intersection density increases and as shoulder width decreases, and that there is a non-linear relationship between lane width and MVCs. There is also weak evidence that CBD roads have more MVCs than thoroughfares and entrances.

The evidence is weak-medium that the number of lanes and parking (on all types of roads) are not related to MVC, and weak evidence that medians, pedestrian crossings and bicycle facilities are unrelated to MVC.

There was no consistency in the relationship found between speed limit and MVC. There was some evidence for an increase in MVC with parking on undivided roads, however this was not consistent across studies.

None of the studies of MVC investigated the influence of driveways, public transport, footpaths, curvature (horizontal or vertical), direction of travel, median openings or road width.

3.4.2.16.3 Single-vehicle crashes

Similar to MVC, there were few studies that specifically focused on SVC so the evidence for associations between roadway related risk factors and SVC is not strong. There is weak evidence that parking on undivided roads increases SVC risk, and that shoulder length is negatively associated with run-off road SVC. There is also weak evidence for a non-linear association between lane width and SVC. Another study found a positive relationship between lane width and SVC but there were only two categories of lane width so a non-linear relationship could not be established.

There is weak-to medium evidence that speed limit and shoulder width are not associated with SVC. There is also weak evidence that the number of lanes, pedestrian crossings, bicycle facilities and vertical curvature and road type are unrelated to SVC frequency.

Intersection density and parking (on all roads) were found to increase SVCs in some studies but not others. Likewise, medians were related to a decrease in SVC, but not consistently in all studies. None of the studies of SVCs measured the association with driveways, public transport, footpaths, horizontal curves, direction of travel, median openings or road width.

3.4.2.16.4 Pedestrian-vehicle crashes

Few roadway-related factors were found to be associated with PVC. There was weak evidence that road type was related to PVC, with PVC frequency higher on CBD roads than thoroughfares and entrances. There was also weak evidence that public transport facilities (buses and trams) were related to PVC, with combined measures of the number of stops multiplied by the number of routes positively related to daytime PVC and bus route density and the distance to the nearest train station negatively related to daytime PVC. Tram stop density was positively related to night-time PVC.

There was medium strength evidence that PVC were not related to intersection density, number of lanes, speed limit, medians, roadside parking, or bicycle facilities.

There was weak to medium evidence that footpaths, the grade of the road and direction of travel were not related to PVC. The evidence was equivocal for the role of driveways which were related to increases in PVCs during the day but not during the night. Pedestrian crossings were associated with increases in PVC in one study but not another (which may have had better measures of pedestrian exposure). None of the studies investigated the influence of roadside shoulders, horizontal curves, median openings, lane width, or road width on PVCs.

3.4.2.16.5 Bicycle vehicle crashes

Only one study specifically looked at roadway-related risk factors for BVC, so the evidence can only be rated as weak. The only statistically significant relationship with roadway-related factors was that BVC frequency was significantly higher on CBD roads than thoroughfares and entrances. There was weak evidence that intersection density, number of lanes, speed limit, medians, roadside parking, pedestrian crossings or bicycle facilities are not related to BVC frequency. The influence of driveways, public transport, roadside shoulders, footpaths, curvature, grade, direction of travel, median openings, road width and lane width were not investigated.

Table 3.4 Direction of association and weight of evidence for relationship between roadway-related risk factors and urban multi-vehicle crashes (MVC), single-vehicle crashes (SVC), pedestrian-vehicle crashes (PVC) and bicycle-vehicle crashes (BVC)

Risk Factor	Crash type	# studies	Direction of association (Weight of evidence)				
			All crashes	MVC	SVC	PVC	BVC
Accesses/ intersections	All MVC SVC PVC BVC	7 1 2 2 1	↑ (strong)	↑ (weak)	↑ or NS (equivocal)	NS (medium)	NS (weak)
Driveways: All types	All PVC	3 1	↑ (medium)			Night NS Day ↑ (equivocal)	
Driveways: Commercial/business	All	2	↑ (weak-medium)				
Driveways: Other	All	1	↓ (weak)				
# lanes	All MVC SVC PVC BVC	7 2 1 2 1	↑, direction unknown or NS (equivocal)	NS (weak-medium)	NS (weak)	NS (medium)	NS (weak)
Speed limit	All MVC SVC PVC BVC	6 2 2 2 1	↑, ↓ or NS (equivocal)	↑, ↓ (equivocal)	NS (weak-medium)	NS (medium)	NS (weak)
Medians	All MVC SVC PVC BVC	6 1 2 2 1	↑, ↓ or NS (equivocal)	NS (weak)	↓ or NS (equivocal)	NS (medium)	NS (weak)

Risk Factor	Crash type	# studies	Direction of association (Weight of evidence)				
			All crashes	MVC	SVC	PVC	BVC
Parking: All roads	All MVC SVC PVC BVC	2 2 2 2 1	↑ or NS (equivocal)	NS (weak-med)	↑ or NS (equivocal)	NS (medium)	NS (weak)
Parking: Undivided roads	All MVC SVC	2 1 1	↑ or NS (equivocal)	↑ or NS (equivocal)	↑ (weak)		
Public transport	All PVC	3 1	NS (medium)			yes (direction depends on measure) (weak-medium)	
Shoulder presence	All	1	↓ (weak)				
Shoulder width	All MVC SVC	1 1 2	↓ (weak)	↓ (weak)	NS (weak-medium)		
Shoulder length	SVC	1			↓ (weak)		
Footpath	All PVC	2 1	NS (weak-medium)			Day/night NS (weak-medium)	
Pedestrian crossings	All MVC SVC PVC BVC	2 1 1 2 1	↑ (weak-medium)	NS (weak)	NS (weak)	All ↑ Day/night NS (equivocal)	NS (weak)
Bicycle facilities	All MVC SVC PVC BVC	1 1 1 2 1	NS (weak)	NS (weak)	NS (weak)	NS (medium)	NS (weak)

Component 1: Identifying risk factors for crashes

Risk Factor	Crash type	# studies	Direction of association (Weight of evidence)				
			All crashes	MVC	SVC	PVC	BVC
Horizontal curves	All	2	↓ or NS (equivocal)				
Vertical curves/grade	SVC PVC	1 1			NS (weak)	Day/night NS (weak-medium)	
Direction of travel	All PVC	1 1	NS (weak)			Day/night NS (weak-medium)	
Road type	All MVC SVC PVC BVC	1 1 1 1 1	NS (weak)	yes (weak)	NS (weak)	yes (weak)	yes (weak)
Median openings	All	3	↑ (medium)				
Travel speed	All	2	↑ (weak-medium)				
Lane width	All MVC SVC	1 1 2	Non-linear (weak)	Non-linear (weak)	↑, Non-linear (weak)		
Road width	All	1	Non-linear (weak)				
Road width change	All	1	↑ (weak)				

↑= positive relationship (shaded in green); ↓=negative relationship (shaded in blue); NS=association not statistically significant (shaded in grey); strength of shading indicates strength of evidence. Black cells indicate no evidence, white cells indicate equivocal evidence

3.4.3 Roadside

Despite roadside environments varying greatly in urban areas and the potential for the roadside to influence crash risk, only two of the fifteen studies investigated the influence of specific hazardous aspects of the roadside environment on crashes. The factors that were investigated and the evidence for an association with crash frequency of different types are shown in Table 3.5.

Lee (2000) investigated run-off road SVC on a principal arterial in Washington State that passed through rural and urban areas. Separate models were developed for each area. The frequency of a large number of roadside characteristics and the offset distances from the outside shoulder to the hazards were measured. Fence length was positively associated with run-off road SVC frequency in urban areas while crash reductions were related to isolated trees, fixed objects and sign supports: for each extra one of these hazards, run-off road SVC frequency decreased by between 8% and 9%. In addition, as the distance from the outside shoulder edge to the guardrail increased, so did the SVC frequency. Factors that were not significantly associated with run-off road SVC frequency in urban areas were: ditch depth, side slopes, the number of catch basins, culverts, utility poles, light poles, tree groups, the distance from the outside shoulder to these hazards or fences, guardrail length and height. It is possible that some of these risk factors were not present in urban areas but not enough information was provided on the distribution of hazards in rural and urban areas to be certain.

Road segments in the Potts et al. (2007) study were rated according to the level of roadside hazards where a rating of one represented the least hazardous roadside and a rating of seven represented the most hazardous roadside. Roadside hazard rating was unrelated to crash frequency (aggregated by type) and MVC frequency on three or five lane roads with a TWLTL median. On roads with two or four lanes, an increase in roadside hazard rating was associated with an increase in SVC frequency, but not MVC frequency.

The roadside environment in urban areas is much more complex and variable than in rural areas yet not many aspects of the roadside have been investigated to determine their influence on crash frequency. It is possible that the difficulty in characterising the urban roadside environment has contributed to the lack of evidence of its influence on urban crashes. More evidence is needed of the relationship between the roadside environment and urban crash frequency.

Table 3.5 Direction of association and weight of evidence for relationship between roadside-related risk factors and urban multi-vehicle crashes (MVC), single-vehicle crashes (SVC), pedestrian-vehicle crashes (PVC) and bicycle-vehicle crashes (BVC)

Risk Factor	Crash type	# studies	Direction of association (Weight of evidence)				
			All crashes	MVC	SVC	PVC	BVC
Hazard rating	All MVC SVC	1 1 1	↑ or NS (Undivided roads), NS (divided roads) (equivocal)	NS (all roads) (weak)	↑ (All roads) (weak)		
Fence length	SVC	1			↑ (weak)		
Isolated trees, fixed objects, sign supports	SVC	1			↓ (weak)		
Other hazards and distance to them	SVC	1			NS (weak)		

↑= positive relationship (shaded in green); ↓=negative relationship (shaded in blue); NS=association not statistically significant (shaded in grey); strength of shading indicates strength of evidence. Black cells indicate no evidence, white cells indicate equivocal evidence

3.4.4 Land use and amenities

Land use and land use mix can be used to describe an area, as well as being relevant to risk on the specific road segments located in that area. For this reason, both studies that used the road segment and the area as the unit of analysis are reviewed in this section.

3.4.4.1 Broad land use

Nine road segment-based studies investigated the relationship between broad measures of land use and crash frequency. There is reasonably strong evidence that crash frequency was higher in business, office or commercial areas compared to residential areas, and shopping zones, in particular, were found to be areas of high risk (Bonneson & McCoy, 1997; El-Basyouny & Sayed, 2009; Greibe, 2003; Jackett, 1993; Jonsson, 2005; Sawalha & Sayed, 2001). The results were generally consistent for all types of crashes. It is important to note that these relationships were independent of traffic volume, which was controlled for in the models. These types of land use attract more pedestrians and cyclists (which were not controlled for) which may be one explanation for the increased risk.

A study in the Melbourne CBD that controlled for pedestrian volume found that weekday pedestrian crashes on midblocks in the Melbourne CBD increased with increases in the density of shops, office floor space and entertainment and recreation areas but were negatively associated with the density of non-commercial accommodation within 600m (Alavi, 2013). There were different risk factors for weeknight pedestrian crashes which were positively associated with the density of non-commercial accommodation, entertainment and recreation areas and the capacity of cinemas, theatres, concert halls and stadiums (Alavi, 2013). Increased pedestrian volumes, therefore, cannot be the sole explanation for the increased risk in business, office and commercial areas. It is therefore important to try and understand the factors affecting crash risk in office, commercial and shopping zones.

3.4.4.2 Types of retail development

Two studies that analysed the same data (with a slightly different focus on crash types and sociodemographic risk factors) investigated the influence of specific types of retail development in census block areas in San Antonio Bexar county metropolitan region of the USA (Dumbaugh & Li, 2011; Dumbaugh et al., 2013). Dumbaugh and colleagues focused on the number of big box stores (retail buildings of greater than 50,000 square feet with more area devoted to surface parking than the building), the number of strip commercial uses (commercial and retail areas typically separated from the road by car parks with driveway access to the road) and the number of pedestrian scaled retail uses (retail areas where the buildings are located at the street-front). The pedestrian scaled retail uses are most similar to the standard strip shopping road segments in Melbourne, Australia.

An increase in the number of strip commercial uses was associated with increases in all crash types, while an increase in the number of big box stores was related to significant increases in all crash types except those involving fixed objects. In contrast, pedestrian scaled retail uses were associated with reductions in the frequency of most types of crashes. For each pedestrian scaled retail use in an area, there were fewer motorist crashes and MVC, and some evidence of a reduction in crashes involving fixed objects and parked cars. There was, however, no significant association between pedestrian scaled retail uses and PVC or BVC.

There is therefore some evidence that retail development is associated with crash frequency, and that big box stores and strip commercial uses are riskier than pedestrian scaled retail uses in this particular region of the USA. These studies did not, however, control for factors relating to the design of the road, the roadside, or vulnerable road user volumes, hence the increases in risk might be due to factors other than the type of retail development. While it appears that pedestrian scaled retail uses in the USA are not associated with crash frequency, Alavi (2013) found that shop density was related to increased pedestrian crash risk during weekdays in the Melbourne CBD, adjusted for the characteristics of the road, roadside and pedestrian activity. Shopping zones in the Melbourne CBD would be classified as pedestrian scaled retail areas according to the definition used by Dumbaugh and colleagues (Dumbaugh & Li, 2011; Dumbaugh et al., 2013). It is possible, therefore, that pedestrian scaled retail uses in the greater Melbourne area are also an environment where road safety is a concern.

3.4.4.3 Alcohol availability

Establishments that serve alcohol (e.g. restaurants, bars and cafes) are often located on strip shopping centre road segments in metropolitan Melbourne. Three studies by the same group of researchers in the USA concentrated specifically upon alcohol availability and crash rates. Restaurant density was positively associated with the frequency of SVC involving passenger vehicles that occurred at night-time (Gruenewald et al., 1996). There was, however, no association between the density of bars and off-premise outlets and night-time SVC. Two studies found that as the density of bars increased, so did the number of pedestrian crashes in which the pedestrian was judged to have been drinking by the attending police officer (whether or not they were judged as being impaired), but there was no influence on all pedestrian crashes (LaScala et al., 2000) or those where the pedestrian was judged not to have been drinking (LaScala et al., 2001). There was no significant association between the number of pedestrian crashes and either restaurant density or off-premise outlet density, whether or not the pedestrian had been drinking.

One limitation of these studies is that, apart from network density, characteristics of the road and roadside were not included in their models. The increased risk in areas with alcohol establishments could thus be due to confounding. If areas with alcohol establishments had poorer roads, more complex roadsides, more vulnerable road users, or an over-representation of other crash risk factors then the increase in risk could be due to those factors, not alcohol availability per se. Yet confounding by characteristics of the road, roadside or pedestrian activity cannot be the sole explanation. Pedestrian crashes that occurred in the Melbourne CBD were also related to alcohol availability even when characteristics of the road, roadside and pedestrian exposure were controlled for. The density of alcohol establishments (bars, taverns, pubs, nightclubs) within 100m of the road segment was associated with an increase in pedestrian injury crashes on weeknights, but not weekdays (Alavi, 2013).

Thus there is evidence that the availability of alcohol has an influence, independent of exposure, road design and the roadside environment, on SVC and PVC frequency. This may be due to an increased number of impaired road users, or, it may be that areas with alcohol establishments are different to those without in terms of some other factor that influences risk. It is therefore important to determine if alcohol availability is contributing to crash risk in complex urban areas such as shopping strips. It is also important to determine if alcohol availability is related to other crash types.

3.4.4.4 Evidence for the influence of land use and amenities on crash frequency

Table 3.6 presents the factors that were investigated and the evidence for an association with crash frequency of different types. Business and commercial land use was related to an increased number of some types of crashes: evidence ranged from medium-strong (aggregate crashes), to weak (SVC and BVC). There was an increase in all PVC and daytime PVC in business/commercial land use areas but not night-time PVC, thus the evidence was equivocal. There was weak evidence that land use was not associated with MVC. More entertainment venues in an area were associated with increased PVC during the day and night (weak-medium evidence). There was weak-medium evidence that the following were not related to PVC: car parks, gaming and amusement centres, commercial accommodation and food venues.

Investigations of specific types of retail development indicated weak evidence that MVC, some types of SVC, PVC and BVC increased as the number of big box stores increased, and all crash types increased as the number of strip commercial uses in an area increased. There was weak evidence that pedestrian scaled retail uses were associated with reductions in MVC and PVC, but had no significant effect on SVC and BVC.

Alcohol availability was investigated in the context of SVC and PVC. SVCs were influenced by restaurant density, while night-time PVC and PVC involving alcohol-affected pedestrians were impacted by the density of bars. None of the studies of aggregate crashes, MVC or BVC investigated the impact of alcohol availability.

Table 3.6 Direction of association and weight of evidence for relationship between land use and amenities and urban multi-vehicle crashes (MVC), single-vehicle crashes (SVC), pedestrian-vehicle crashes (PVC) and bicycle-vehicle crashes (BVC)

Risk Factor	Crash type	# studies	Direction of association (Weight of evidence)				
			All crashes	MVC	SVC	PVC	BVC
Business/commercial/ retail	All MVC SVC PVC BVC	8 1 1 2 1	↑ (medium-strong)	NS (weak)	↑ (weak)	↑ (all, day) NS (night) (equivocal)	↑ (weak)
Car parks, gaming/amusement centres, commercial accommodation, food venues	PVC	1				NS (weak-medium)	
Entertainment venue density	PVC	1				↑ (weak-medium)	
Restaurant density	SVC PVC	1 2			↑ (weak)	NS (whether or not alcohol involved) (weak-medium)	
Bar density	SVC PVC	1 3			NS (weak)	All, day NS Night/alcohol related ↑ (equivocal)	
Off-premise outlet density	SVC PVC	1 2			NS (weak)	NS (weak-medium)	
Retail development: Big box stores	MVC SVC PVC BVC	1 1 1 1		↑ (weak)	↑ or NS (equivocal)	↑ (weak)	↑ (weak)

Risk Factor	Crash type	# studies	Direction of association (Weight of evidence)				
			All crashes	MVC	SVC	PVC	BVC
Retail development: Strip commercial	MVC SVC PVC BVC	1 1 1 1		↑ (weak)	↑ (weak)	↑ (weak)	NS (weak)
Retail development: Pedestrian scaled	MVC SVC PVC BVC	1 1 1 1		↓ (weak)	NS (weak)	↓ (weak)	NS (weak)

↑= positive relationship (shaded in green); ↓=negative relationship (shaded in blue); NS=association not statistically significant (shaded in grey); strength of shading indicates strength of evidence. Black cells indicate no evidence, white cells indicate equivocal evidence

3.4.5 Sociodemographic

The social environment can influence crash occurrence (Haddon, 1972) through laws, rules and regulations and social norms. Norms, in particular, affect behaviour of individuals and may vary between different population subgroups. In addition, population size and density may be considered indirect measures of exposure to risk. As such, sociodemographic characteristics of an area may be related to crash frequency.

All of the studies that used the area as the unit of analysis and one of the studies with the road segment as the unit of analysis considered sociodemographic characteristics of the local area and their association with crashes in urban areas. Studies that used the area as the unit of analysis did not usually account for the design of roads, roadsides or vulnerable road user activity in their models, although motorist exposure and network density were controlled for. Hence, if sociodemographic factors are related to any of the aspects of the built environment that change crash risk, it may appear that the sociodemographic factors themselves are associated with risk (due to omitted variable bias). Another important point is that the road users who travel through (and crash in) an area, are not necessarily residents of that area, particularly motorists. Thus the relationships between sociodemographic factors and crash risk in an area are not necessarily expected to be strong, particularly for crashes that involve vehicles only.

3.4.5.1 Population size and density

Although road user volumes are more specific measures of exposure, population size and/or density may be reasonable measures of exposure when road user volumes are unavailable. Population size and/or density were, however, not associated with crashes aggregated by type (Hadayeghi et al., 2003; Haynes et al., 2008), SVC (Dumbaugh & Li, 2011; Gruenewald et al., 1996), MVC or BVC (Dumbaugh & Li, 2011; Dumbaugh et al., 2013). The relationship between population density and PVC was inconsistent across studies: two studies found that PVC increased along with population density (LaScala et al., 2000; LaScala et al., 2001) while another two found no association between population density and PVC (Alavi, 2013; Dumbaugh et al., 2013). There was weak evidence that the number of households was positively associated with aggregate crashes (Hadayeghi et al., 2003).

3.4.5.2 Age and sex distribution

Age and sex are risk factors for crash involvement with young drivers (Jonah, 1986), particularly males, being at higher risk of being involved in a crash (Turner & McClure, 2003). Children and the elderly are vulnerable to being involved in PVC, partly due to cognitive limitations (Oxley, Congiu, Whelan, D'Elia, & Charlton, 2007; Oxley, Ihsen, Fildes, Charlton, & Day, 2005), while the elderly are also physically frail and more prone to injury in a crash (Davis, 2001). Hence the age and sex distribution of the population living in an area may be associated with the frequency of different types of crashes.

The age distribution of the population was significantly related to some types of crashes but not others. The proportion of the population aged 15 to 24 or older than 75 was not significantly associated with fatal crashes in NZ (Haynes et al., 2008). PVC and BVC were more frequent when there were a greater proportion of children in a community (Dumbaugh et al., 2013; LaScala et al., 2000; LaScala et al., 2001). There is also some, albeit conflicting, evidence that a greater proportion of older adults in a community was associated with increased PVC frequency: two studies found a significant relationship (Dumbaugh et al., 2013; LaScala et al., 2001) although one study did not (LaScala et al., 2000). Conversely, night-time SVCs (a surrogate for alcohol-related crashes) are more frequent when there was a lower proportion of children or older adults in the community (Gruenewald et al., 1996). These results must be considered in terms of the potential for confounding by characteristics of the built environment that affect crashes.

The proportion of the population that were male was positively associated with the rate of pedestrian crashes in San Francisco but not the subset of crashes where the pedestrian had been drinking (LaScala et al., 2000). In another study of four communities in California, the sex distribution of the population was not related to night-time SVC (Gruenewald et al., 1996) or PVC, regardless of whether or not the pedestrian had been drinking (LaScala et al., 2001). There is therefore little evidence that the sex distribution of a population is related to crashes.

3.4.5.3 Socioeconomic status

Social norms and values, and therefore behaviour, may vary by socioeconomic status. Socioeconomic status may also be related to exposure to crashes, for example, areas with low rates of vehicle ownership may have higher pedestrian exposure. It is also possible that socioeconomic status may be correlated with other risk factors for crashes, such as the state of road infrastructure. Hence, socioeconomic status of an area may be associated with crash frequency.

3.4.5.3.1 Employment

Employment was not associated with crashes in Toronto traffic zones (Hadayeghi et al., 2003), PVC on weekdays or weeknights in the Melbourne CBD (Alavi, 2013), Californian night-time SVC or PVC (whether or not alcohol was involved) (Gruenewald et al., 1996; LaScala et al., 2001). LaScala et al. (2000), however, did find a positive association between the unemployment rate and the rate of pedestrian crashes in San Francisco; the relationship was stronger for crashes that involved pedestrians that had been drinking. There is, therefore, little evidence suggesting population-level employment is associated with crashes (apart from perhaps PVC) in urban areas.

3.4.5.3.2 Income and deprivation

Income was negatively related to PVC and BVC frequency (LaScala et al., 2000; LaScala et al., 2001) but there was no significant association with SVC (Gruenewald et al., 1996). Haynes et al. (2008) established there were significantly more fatal crashes in urban areas with increased levels of deprivation compared to areas with less deprivation. The authors postulated that this could be due to differences in the vehicle fleet or differences in the behaviour of road users. It is also possible that affluent areas have better designed and maintained roads. Hence, the relationship between income/deprivation and crashes could be confounded by crash risk factors related to the built environment.

3.4.5.3.3 Education

Two studies assessed the relationship between the education level of the community and pedestrian crash rate. The proportion of the population with at least a high school education was negatively associated with pedestrian crashes in San Francisco, however, this association disappeared for the subset of crashes in which the pedestrian was judged to have been drinking (LaScala et al., 2000). The proportion of the population with at least some college education was not associated with PVC in four communities in California regardless of whether the pedestrian had been drinking (LaScala et al., 2001). There was little evidence for an influence of education on crashes in the Californian communities in these studies.

3.4.5.4 Other

The relationship between a range of other sociodemographic factors and crashes was assessed but none were statistically significant. Marital status and ethnicity were unrelated to night-time SVC (Gruenewald et al., 1996) and PVC in four communities in California. The average number of people under 18 per household was also unrelated to PVC in these communities (LaScala et al., 2001). Vehicle ownership was not associated with crashes in areas of Toronto (Hadayeghi et al., 2003) or fatal crashes in NZ (Haynes et al., 2008).

3.4.5.5 Evidence for the influence of sociodemographic factors on crash frequency

Table 3.7 presents sociodemographic risk factors and the level of evidence for an association with crash frequency of different types. There was some evidence that the age distribution of the population in an area was related to some types of crashes. When there was higher proportion of children in a population, there was weak to medium strength evidence that PVCs were more frequent and weak evidence that BVCs were more frequent. The evidence was weak that the proportion of older adults in a population was not related to aggregate crashes or BVC. One study found an increase in PVCs when there were a greater proportion of older adults, although this was not consistent across studies. Evidence suggests there is no relationship between the proportion of the population that are young adults and the aggregate crash frequency. There was also weak evidence that the relationship between mean age and night-time SVCs is quadratic, with the highest risk when the mean age of the population was approximately 41. There is weak evidence that the proportion of males in the population is unrelated to night-time SVC, although there is equivocal evidence it may be related to increases in PVC. There is no information on the effect of the sex distribution of the population and aggregate crashes, MVC or BVC. There was weak evidence that no relationship exists between marital status and SVC or PVC.

There is no evidence for an association between employment density and aggregate crashes (medium-weak strength evidence), or PVC (weak-medium evidence). Equivocal evidence suggests that PVC might increase as the proportion of the population who are unemployed increases, but that it is unrelated to SVC. There was weak evidence that more deprived communities have more aggregate crashes and that income is negatively associated with BVC. One study found a negative relationship between income and PVC, but another did not. Evidence suggests there is no relationship between income and SVC. No information was available regarding the relationship between employment or income and MVC.

Finally, there was weak to medium evidence that vehicle ownership is unrelated to aggregate crashes, but no information about the relationship with the different crash types. Higher levels of education in a community were associated with decreases in PVC in one study but not another. None of the studies of aggregate crashes, MVC, SVC or BVC investigated the effect of education level.

Table 3.7 Direction of association and weight of evidence for relationship between sociodemographic-related risk factors and urban multi-vehicle crashes (MVC), single-vehicle crashes (SVC), pedestrian-vehicle crashes (PVC) and bicycle-vehicle crashes (BVC)

Risk Factor	Crash type	# studies	Direction of association (Weight of evidence)				
			All crashes	MVC	SVC	PVC	BVC
Older adults (>55)	All PVC BVC	1 2 1	NS (weak)			↑ or NS (equivocal)	NS (weak)
Young adults (15-29)	All PVC	1 1	NS (weak)			NS (weak)	
Children (<17)	PVC BVC	2 1				↑ (weak-medium)	↑ (weak)
Mean age	SVC PVC	1 1			Non-linear: highest risk for 41 year olds (weak)	Non-linear: highest risk for children and elderly (weak)	
Males	SVC PVC	1 2			NS (weak)	↑ or NS (equivocal)	
% single	SVC PVC	1 1			NS (weak)	NS (weak)	
Vehicle ownership	All	2	NS (weak-medium)				
Population size/ density, household density	All MVC SVC PVC BVC	4 1 2 4 1	NS (medium)	NS (weak)	NS (weak-medium)	↑ or NS (equivocal)	NS (weak)
# households	All		↑ (weak)				

Risk Factor	Crash type	# studies	Direction of association (Weight of evidence)				
			All crashes	MVC	SVC	PVC	BVC
Employment density or total number employed	All PVC	2 1	NS (weak-medium)			NS (weak-medium)	
% unemployed	SVC PVC	1 2			NS (weak)	↑ or NS (equivocal)	
Income	SVC PVC BVC	1 2 1			NS (weak)	↓ or NS (equivocal)	↓ (weak)
Deprivation	All	1	↑ (weak)				
Education level	PVC	2				↓ or NS (equivocal)	

↑= positive relationship (shaded in green); ↓=negative relationship (shaded in blue); NS=association not statistically significant (shaded in grey); strength of shading indicates strength of evidence. Black cells indicate no evidence, white cells indicate equivocal evidence

3.5 Implications for this study

There is a lack of knowledge about risk factors for crashes on urban roads, relative to the number of studies of the influence of the environment on crashes on rural roads or limited access roads. The studies that have been conducted in urban areas have often only investigated a limited subset of the potential risk factors in urban environments. The majority of studies have restricted their focus to traffic volume and aspects of road geometry as risk factors. This is perhaps due to the difficulty with collecting data to comprehensively characterise the built urban environment, including the roadway, the roadside and surrounding area, especially if researchers rely on existing administrative data. Another possible reason is that the researchers did not consider that characteristics of the built urban environment, beyond the design of the road itself, had the potential to influence crashes. As such, there is a lack of strong and unbiased evidence about the relationship between many potential risk factors and urban crashes. The majority of studies used crash data aggregated across all crash types so the evidence is particularly lacking for different crash types. In addition, many researchers failed to report (and perhaps even failed to test) whether their model form was appropriate or whether the model fitted the data well, which makes it difficult to assess the validity of the results.

The advantage of using existing data is that it is a very cost effective and time-saving way of conducting research into the crash risk of a large number of road segments or sites. The disadvantage is that the researchers are restricted to the data available within the databases unless other data are collected specifically for the study. Basing a study on what data are available constrains the research questions that can be addressed. There are also potential limitations in that the factors may not be measured in the way the researcher wishes. Different road authorities measure factors in different ways, which can make it difficult to compare the results of studies. However, the cost and time savings associated with conducting research using existing data far outweigh the disadvantages, as long as the researchers fully understand the limitations of the data they are using. The optimal study would make use of existing data where possible and also collect data to supplement what is already available in order to investigate a broad range of risk factors.

The multivariable models were required to include traffic volume as a criterion for inclusion in this review, thus it is not surprising that the strongest evidence is available for this exposure-related risk factor. There is reasonable evidence for a non-linear relationship between traffic volume and crashes. It was surprising, however, that so few studies considered traffic density as a contributor to risk in urban environments. Likewise, the distribution of traffic (in terms of the proportion that were heavy vehicles) was overlooked. It is possible that traffic density is a more important contributor to risk in urban environments than traffic volumes.

Pedestrian and bicycle volumes were related to crashes involving those particular road users, but the influence on other crash types was not investigated. Unlike traffic volumes, vulnerable road user volumes are not consistently collected. Other measures of road user exposure from more readily available data sources (e.g. travel surveys) should be explored, particularly in studies that focus on crashes involving vulnerable road users.

Although many roadway-related risk factors have been studied in the past, each study generally only uses a small subset of the possible risk factors, thus raising the question of whether important predictors are omitted. Few studies have investigated a large range of potential risk factors relating to the roadway and some potential risk factors have been overlooked altogether in multivariable studies of urban crashes, for example, road pavement condition. Likewise, the role of the roadside and surrounding environment has been relatively ignored, even though the roadside urban environment is highly changeable and complex.

Business and commercial areas have been consistently shown to have more crashes than other land use types in urban areas. They are also highly complex environments. Only one study focused on a concentrated urban business/commercial environment (Alavi, 2013), however, this study was restricted to pedestrian crashes in the CBD of Melbourne, Australia. The Melbourne CBD is a relatively uniform environment that is very densely developed and has high concentrations of pedestrians and public transport. The greater Melbourne metropolitan area has complex urban environments with very different characteristics to the CBD. It is important to understand what components of the built environment in business and commercial areas outside of the CBD influence the risk of different types of crashes. No previous studies have addressed this issue.

Alcohol availability was related to SVC and PVC, however, this risk factor was rarely considered in studies at the level of the road segment. There is also no information about alcohol availability and other crash types. It is important to determine if alcohol availability is related to different crash types controlling for exposure and roadway-related risk factors. Establishments that serve alcohol for consumption on premises are often concentrated in strip shopping/entertainment areas, so this is the ideal environment in which to measure their influence on crashes. It is also possible that the presence of other facilities and amenities (e.g. educational facilities, health care, community centres) are related to urban crash risk; this has not previously been investigated in multivariable studies of the influence of the built environment on urban crashes.

Finally, while there is some evidence for a relationship between sociodemographic factors and crashes, this evidence predominantly came from macro-level studies that did not control well for roadway-related risk factors. It is possible that there is some confounding or even mediation between the effect of the social and built environment on crashes, e.g., road infrastructure may differ along socioeconomic gradients. It is important to determine if sociodemographic factors are related to crash risk when exposure and roadway related risk factors are controlled for.

3.6 Aims of Research Component 1

This review highlighted the gaps in knowledge surrounding the characteristics of the built environment that are associated with crashes in urban areas and it is these gaps that have led to the conception of the aims for Research Component 1 of this thesis. The primary aims are:

- To create a list of characteristics of the built urban environment (including the road, roadside and human activity) that are potential risk factors for crashes
- To find sources for these data, or where existing sources are not available, develop efficient methods for data collection
- To identify the aspects of the built urban environment, including activity within that environment, that are associated with crash frequency in business and commercial areas; specifically, strip shopping road segments in the greater Melbourne metropolitan area, which are complex urban environments that are challenging to characterise
- To identify risk factors for specific types of crashes on urban strip shopping road segments: MVC, SVC, PVC and BVC

A secondary aim was to investigate whether traffic volume or traffic density is a better measure of exposure to crash risk in complex urban environments, and how the choice of exposure measure affects the other risk factors that are identified.

CHAPTER 4. METHODS: SITE SELECTION AND DATA

This chapter describes the methods employed for the statistical modelling component of the research. The processes of site selection, identification of data requirements, data sources, extraction/coding and linking of data to create the final database are described. The statistical modelling methods that were used are then described in Chapter 5.

4.1 Site selection

This study involved an investigation of the built environment-factors associated with crash frequency in complex urban environments, in particular, strip shopping centre road segments. The method for selecting road segments depends upon the specific research question and characteristics of the available data. For example, road segments were chosen to be of equal length in a study of roadway geometry and environmental factors on crash frequency on a rural interstate in Washington State. The rationale for choosing equal length segments was that road and weather conditions changed frequently, so if segments were chosen to be homogenous, the segments would have been very short (Shankar, Mannering, & Barfield, 1995). This was of particular concern because of potential error in reporting the location of crashes (crash location was reported to the nearest mile-post). When segmenting a long corridor into shorter parts for analysis (or several corridors segmented into parts), however, there is the potential for spatial correlation between crashes on adjacent sections (that is, crash frequency on one segment is likely to be related to crash frequency on adjoining segments). This contravenes the assumption of independence. For studies that are concerned with crash frequency in a particular type of environment, for example, outside schools, or on strip shopping road segments, it makes intuitive sense to define each road segment of interest as a unit of analysis. Consequently, the road segments are likely to differ in length, but spatial correlation is less likely than if the researchers choose to artificially segment the road into smaller sections.

For this reason, strip shopping centre road segments were selected as the entire length of road on which shops or business were present. First, however, strip shopping centres were defined, and then strip shopping centre road segments in metropolitan Melbourne were identified.

4.1.1 Definition of a strip shopping centre

There is no strict accepted definition of what constitutes a strip shopping centre. Simply, strip shopping centres are predominantly retail areas where establishments have direct access to the major road. Midson (2007; p3) stated that strip shopping centres;

"can loosely be defined as an attached row of stores or service outlets managed as a coherent retail entity, with most parking usually located on-street in front of the stores. Open canopies may connect the storefronts, but a strip shopping centre may not have enclosed walkways linking the stores. A strip shopping centre may be configured in a straight line, or have an "L" or "U" shape along a road corridor."

The Victorian Planning Provisions (Department of Environment Land Water & Planning, 2015, Clause 52.28) define a strip shopping centre as an area that meets all of the following requirements:

- zoned for commercial use
- at least two separate buildings on at least two separate and adjoining lots
- a significant proportion of the buildings are shops
- a significant proportion of the lots abut a road accessible to the public generally
- not included in the Capital City Zone in the Melbourne Planning Scheme

For the purposes of this research project, the strip shopping centre road segments were required to meet the following criteria to be included in the study:

- located on an arterial road in the Melbourne metropolitan area, excluding the Central Business District (because traffic volume data were only consistently available for arterial roads)
- predominantly retail/commercial buildings

- buildings must have direct frontage onto the main road (including service roads) on one or both sides of the road
- at least 200 m in length

4.1.2 Identification and selection of study road segments

A number of different methods were used to identify strip shopping centre road segments on arterial roads in the Melbourne metropolitan area. Initially, local council planning schemes were consulted. Not all planning schemes specifically identify strip shopping centres but they are sometimes named in the section that defines areas where gaming machines are prohibited. The websites of local councils also provided some information on the location of strip shopping centres within their municipality. Finally, the Melway street directory (2005, 2006, 2007, 2008, 2009) was consulted because road segments with shops are shaded in orange.

Once a list of potential strip shopping centre road segments was identified, Google Maps with Street View (Google, n.d.) was used to determine if each segment was, in fact, a strip shopping centre and to identify the location of the properties at the beginning and end of each segment. The street addresses and the geographical co-ordinates of the properties at the ends of each segment were determined by referring to the Land Victoria, Department of Sustainability and Environment on-line interactive maps (Department of Sustainability and Environment, n.d.). The approximate length of the road segment was also calculated. Those segments that did not have retail/commercial properties for at least 200m on at least one side of the road were excluded from the list.

Whether or not the road segment was on an arterial road was ascertained next. VicRoads, the Victorian State Road Authority, is responsible for the construction, maintenance and management of major arterial roads in Victoria, which are listed in the Register of Public Roads (VicRoads, 2008b). This document was consulted to determine if the strip shopping centre study road segments were on arterial roads. Those that were not were excluded.

Road segments or sections of road segments were excluded from the study if there were major roadworks performed between January 1st, 2005 and December 31st, 2009 (the time period of interest for this study). During this time, VicRoads conducted a major program of roadworks to improve road safety called the Safer Road Infrastructure Program (SRIP). Details of the location and type of works conducted under SRIP were obtained to determine if major safety related road works were undertaken on the study road segments during the study period (2005 to 2009). Thirty-two of the potential study segments had some form of safety-related roadworks performed under the SRIP program during 2005 to 2009. Five road segments were excluded from the study entirely because major works were carried out that affected most of the road segment. Eight road segments had variable speed limits implemented sometime in 2005; for these road segments the exposure period was reduced to include only the period after the speed limit was changed. Two road segments had variable speed limits implemented in 2008; for these segments, the exposure period for counting crashes was reduced to include only the period prior to the change in speed limit. For 17 road segments, roadworks were performed at one signalised intersection, usually located at or near the end of the study segment. If the intersection was located near the end of the road segment, the road segment was shortened to exclude the signalised intersection and any midblock segment between that intersection and the end of the strip shopping zone. If the intersection was located in the middle of the road segment, the intersection and a short section of adjacent midblock was excluded. Although it is difficult to be certain, it was judged that the roadworks at signalised intersections would have been unlikely to disrupt traffic flow or affect crashes on adjacent sections of road.

Road segments were also excluded if data essential for the conduct of the research were unavailable. Four road segments were excluded because either traffic volume or road geometry data were unavailable.

There were six strip shopping centre road segments where the traffic volume changed markedly either side of one signalised intersection on those road segments. For example, the AADT on one road segment differed by more than 20,000 vehicles per day on either side of the signalised intersection of another major road in metropolitan Melbourne. Because traffic volume is strongly related to crash risk, these segments were bisected at the point at which traffic volume changed appreciably, and considered as separate units for analysis. Of the road segments that resulted from bisecting the segments with heterogeneous AADT, three of the new resulting segments were too short to meet the criteria for inclusion and were thus excluded from the study.

At the end of this process, 142 road segments remained for inclusion in the study. The signalised (major) intersections on each road segment were identified from the Melway (Melway, 2005, 2006, 2007, 2008, 2009) street directories so they could be excluded from the analysis of midblock crashes. The road segments that were included in the study are displayed as blue lines on the map of metropolitan Melbourne in Figure 4.1.

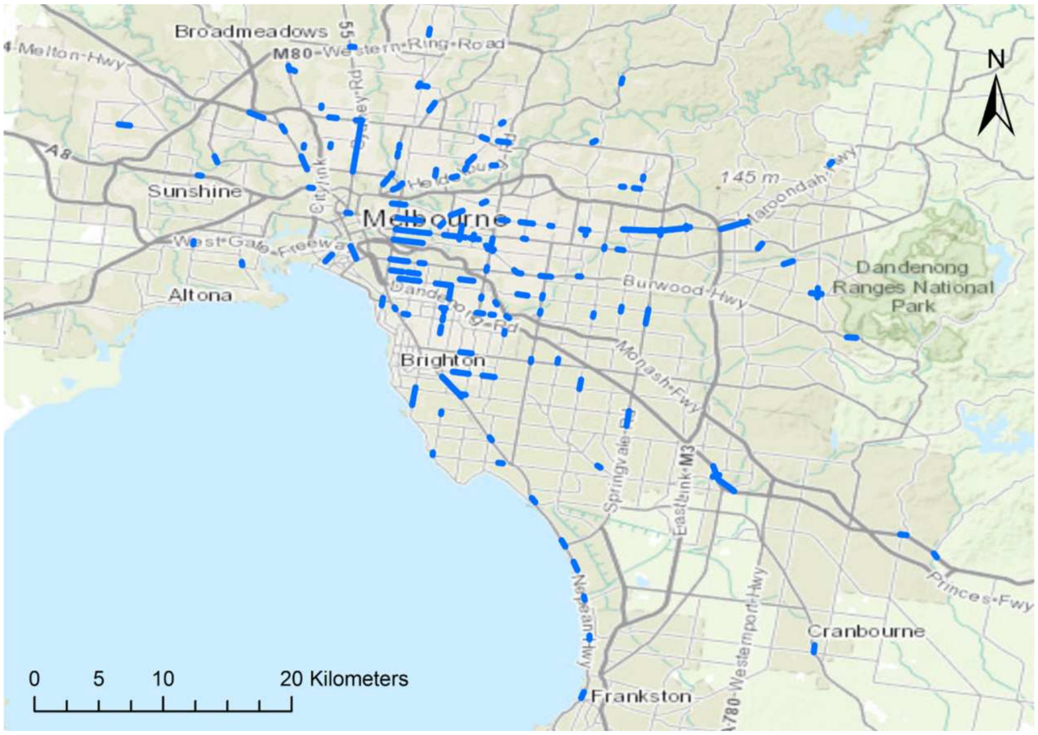


Figure 4.1. Melbourne metropolitan area with the road segments included in this thesis highlighted in blue (map used with permission from Lawrence, B., created using ArcGIS® software by ESRI)

4.2 Data requirements

The data required to undertake an investigation of the association between the characteristics of the built environment and the frequency of crashes can be simply categorised as that required for measurement of the outcomes, or dependent variables (frequency of different types of crashes) and that required for measurement of the risk factors, or independent variables (the built environment including the road, roadside and human activity).

4.2.1 Outcomes

Data were required on the cumulative frequency of crashes that occurred on each road segment during the period January 1st, 2005 to December 31st, 2009. Information was also required on the number and type of road users involved in each crash so that the risk factors for different types of crashes (SVC, MVC, BVC and PVC) could be investigated.

4.2.2 Identification of potential risk factors

Data requirements for measuring the influence of the built environment on crash risk were developed in two-stages. At this initial stage of the research, a broad view was taken and data requirements were developed for all road environments including rural roads, highways, motorways as well as urban roads and both midblocks and intersections. The reason for this was that there have been many more studies of the influence of the road and roadside in rural areas and highways and at intersections and it was felt that narrowing the focus to urban road midblocks at this point may have led to omission of potentially important risk factors that have not yet been studied in urban areas.

In stage one, a literature review was conducted to identify the characteristics of the environment that had previously been found to be associated with crash risk. The review included the research relating to urban midblock crashes, reviewed in Chapter 3, as well as studies conducted in rural areas, highways, midblocks and intersections. The literature search was performed using the MedLine, Compendex and Australian Transport Research Index databases in order to cover a range of relevant disciplines. Broad search terms were used in order to discover as many articles as possible. Abstracts were scanned individually to identify articles that focused on the identification of risk factors relating to the built environment and sociodemographic factors. The literature search focused mainly on multivariable modelling studies, although before-after studies were included where appropriate. The initial search focused on papers published in the peer-reviewed literature however because this thesis sought to derive an exhaustive list of environmental factors that may impact crash risk, a second phase also included searching the grey (unpublished or limited publication) literature and PhD dissertations to find high quality reports in the field. In addition, recent documents defining data requirements for evidence-based road safety engineering practice were also consulted (e.g. AASHTO, 2010; Stefan, Dietze, Marchesini, Louise, & Candappa, 2010). A list of factors related to the environment that may affect crash risk was compiled from these studies and reports (AASHTO, 2010; Abdel-Aty et al., 2009; Abdel-Aty & Radwan, 2000; Alavi, 2013; Aljanahi, Rhodes, & Metcalfe, 1999; Anastasopoulos & Mannering, 2009; Avelar et al., 2013; Bonneson & McCoy, 1997; Brown & Tarko, 1999; Chang, 2005; Chayanan, Nebergall, Shankar, Juvva, & Ouyang, 2003; Dissanayake, Aryaija, & Wedagama, 2009; Dumbaugh & Li, 2011; Dumbaugh et al., 2013; El-Basyouny & Sayed, 2009; Fildes & Lee, 1993; Greibe, 2003; Gross & Jovanis, 2007; Gruenewald et al., 1996; Hadayeghi et al., 2003; Haynes et al., 2008; Ivan et al., 2000; Jackett, 1993; Jonsson, 2005; Karlaftis & Golias, 2002; LaScala et al., 2000; LaScala et al., 2001; Lee, 2000; Manuel et al., 2014; Miaou, 1994; Milton & Mannering, 1998; Noland & Oh, 2004; Pande & Abdel-Aty, 2009; Potts et al., 2007; Roberts et al., 1995; Sawalha & Sayed, 2001; Shankar et al., 1997; Stefan et al., 2010; Stevenson et al., 1995; Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013).

Much of the focus of past studies was on simple environments and was restricted only to risk factors for which data were readily available from road authorities (e.g. traffic volume and road design). Sociodemographic determinants of crash risk were also considered in previous research. Consequently, a list of other factors with the potential to affect crash risk in the urban environment was developed. These included aspects of the built environment (road, roadside and human activity) with the potential to affect the field of safe travel, safety margins and the complexity of the urban environment. Most characteristics of the road with the potential to influence crashes had been covered in previous research, however, there was a lack of consideration of the roadside and surrounding environment and patterns of human activity (apart from traffic volumes). As such, the new factors that were added to the list mainly related to the surrounding built environment rather than the road itself, for example, development height, the presence of nature strips, and roadside poles and trees. In addition, the existence of amenities or facilities that could affect human activity in the environment in terms of the number and movement of road users were included, in particular, those that could increase the number of potentially vulnerable road users such as children (e.g. childcare centres, schools), the elderly (e.g. aged care facilities), or potentially impaired road users (e.g. healthcare facilities, licensed premises).

In stage two, the compiled list of potential risk factors was sent to Victorian road safety experts for comment. The group of experts comprised road safety professionals from VicRoads, the Victorian Transport Accident Commission (TAC), Victorian Department of Justice, Victoria Police and the Monash University Accident Research Centre. Between them, the professionals had expertise in a broad range of relevant areas, including engineering, human behaviour, and law enforcement. Experts were asked to comment on the items on the list, and to add anything they considered important for predicting crash risk, particularly in complex environments.

Table 4.1 displays the list of the characteristics of the built environment that were identified as potential risk factors for crashes during the two-stage process. The characteristics may be relevant to all types of roads (e.g. carriageway width) or particular types of roads (e.g. presence of trams is likely to be relevant only to urban areas). The characteristics are categorised according to the broader aspect of the road, surrounding environment or human activity that they describe. In addition, examples are given of specific data items, or variables, that could be collected to measure the characteristics.

Table 4.1 Characteristics of the built environment with the potential to influence crash risk, by broad category

Category	Characteristic	Example data items
AREA	Metropolitan/rural	
	Land use	Zoning
EXPOSURE	Traffic volume	AADT, peak hour volumes
	Cyclist volume	
	Pedestrian volume	Number of pedestrians crossing
	Segment length	
	Direction of travel	One way/two way
SPEED LIMIT	Speed zone	
	Variable speed zones	Presence, type (e.g. school zone, shopping centre zone), times of operation
ROAD CROSS-SECTION	Curves	Presence, number, curve angle, distance between curves
	Vertical curvature/grade	Presence, number, distance, % grade
	Carriageway width	Mode, minimum, maximum
	Lanes	Number of lanes
	Lane width	Mode, minimum, maximum
	Shoulder	Type, width, length
	Line marking	Presence, type
	Bridges	Presence, number, width
	Passing lanes	Presence, number, length
	Special lanes	Presence, type (e.g. transit), width, length
	Guardrails/barriers	Presence, length, type
	Speed management devices	Number, type (e.g. speed humps, chicanes)
ACCESSES AND INTERSECTIONS	Signalised (major) intersections	Number, type, number of arms, traffic signal phasing
	Unsignalised (minor) intersections	Number, type, number of arms, type of traffic control
	Roundabouts	Presence, type, number
	Service road	Presence, type, number of access points
	Driveways/laneways	Number, type
	Dedicated turning lanes	Presence, type, number

Category	Characteristic	Example data items
	Keep clear zones	Presence
	Skew angle of intersection	
MEDIANS	Divided road	Presence of median, proportion of the road with a median
	Median type	Paint/concrete/barrier
	Median width	
	Median accesses	Number
ROAD TYPE AND TRAFFIC MIX	Road function	Freeway, arterial (primary or secondary), municipal
	Heavy vehicle access	Over-dimensional vehicle route, approved for heavy vehicles
	Vehicle mix	% of traffic that are heavy vehicles
ROADSIDE PARKING	Parking zones	Presence, type, length
	Parking clearways	Presence, distance, hours of operation
	Loading zones	Presence, number
PUBLIC TRANSPORT AND BICYCLE FACILITIES	Railway level crossings	Presence, type
	Buses	Presence, number and type of stops, presence of bus lanes
	Trams	Presence, number and type of stops, type of tram-lane
	Bicycle	Presence of facilities, e.g. bicycle lanes (type and width), presence of facilities at intersections (e.g. advanced stop lines, storage boxes)
PEDESTRIAN FACILITIES	Pedestrian crossings	Presence, number, type
	Pedestrian barriers	Presence, length
	Footpath (sidewalk)	Presence, width
ROADSIDE DEVELOPMENT	Roadside poles/signs/trees	Presence, number, type (frangible/non-frangible)
	Nature strip	Presence, one or both sides
	Development/buildings	Number, height, offset
PAVEMENT CONDITION	Road surface	Sealed/unsealed, roughness, distress
HEIGHT CLEARANCE	Low clearance	Presence, type (reason), height
STATIC LAW ENFORCEMENT	Static enforcement cameras	Presence, type (red light/speed), number
AMENITIES AND FACILITIES	Retail	Trading hours, footpath trading (yes/no), indoor shopping centres
	Licensed venues	Number, type, trading hours
	Educational facilities	Presence, type, number

Category	Characteristic	Example data items
	Healthcare facilities	Presence, type, number
	Aged care facilities	Presence, type, number
	Places of worship	Presence, number
	Public transport	Presence of railway stations, bus stations
	Community centres	Presence, type, number
	Emergency services	Presence, type, number
	Parks/ sports fields/ leisure centres	Presence, type, number
SOCIODEMOGRAPHIC	Population data	Population density, age/gender distribution
	Socioeconomic status	e.g. Index of relative socioeconomic advantage and disadvantage (IRSAD)
	Access to vehicles	Vehicle ownership rates

4.3 Data sources and methods of collection

Once the list of possible risk factors was finalised, sources of good quality data for each of the risk factors were sought. Detailed information about each of the variables collected and their source is provided in the next section, but, first an overall summary of the data collection methods is provided.

4.3.1 Existing data

Given the large number of potential risk factors, preference was given to obtaining good quality data from existing data sources. Unlike much of the previous research however, this was not restricted to existing data held by road authorities. Existing data were obtained from VicRoads, the Victorian Department of Justice, the Victorian Department of Transport and the Australian Bureau of Statistics (ABS).

Data from existing sources were obtained for the time period 2005 to 2009, or the midpoint of that period (2007) where possible. While some aspects of the road environment or other risk factors may have changed over time, the chances of this were reduced by excluding road segments that had major road works over the study period.

4.3.2 Data collected specifically for the project

Where it was not possible to find an existing data source with good quality data for potential risk factors, consideration was given to identifying accurate, efficient and cost-effective means to collect data required for the study. Some data were collected from the Melway street directory (2005, 2006, 2007, 2008, 2009), on-line maps and on-line image sources.

A range of characteristics could not be collected from existing data sources or maps, in particular, those data items that required accurate distance measurements (e.g. carriageway width, lane width and offset distance). Site visits to collect data specifically for the project were not practical as there was no funding available and it would have been extremely time-consuming and potentially unsafe for the data collectors. In addition, data collection took place after the study period had ended and there was the possibility that characteristics of the road may have changed in the intervening time period.

Investigations identified that VicRoads owns digital video images of arterial roads that are collected every two years by the ARRB Group Ltd[1]. The ARRB Group has a fleet of Network Survey Vehicles equipped with a range of devices for measuring road geometry and pavement condition and a suite of four cameras that record images of the road environment (Roper, 2003). A specialised proprietary software program (Hawkeye) is used to view and analyse the images and take calibrated distance measurements. The ARRB Group have a team of video raters (or data coders) who are highly experienced in efficiently and accurately extracting information from the digital video images. The services of the ARRB Group were engaged to code data for the study road segments.

A data specification was provided to the ARRB group that identified the road segments of interest (including geographic co-ordinates) and a comprehensive description of the variables to be coded, including definitions and coding rules. Some variables with an underlying continuous distribution (e.g. distance measurements) were coded as categories according to those used for risk-based rating schemes such as the international road assessment program (iRAP), the Australian road assessment program (AusRAP) and NetRisk (Affum & Goudens, 2008; Australian Automobile Association, 2006; McInerney & Smith, 2009). While it would have been preferable to have continuous data for these variables, the coding team were experienced in coding them as categorical data, so the compromise was made to use categorical data instead of continuous data in the interests of maximising data quality. The ARRB Group's experienced rating team coded digital video images of the road segments of interest that were collected between January 2009 and February 2010 (corresponding to the

[1] https://www.arrb.com.au/home.aspx

end of the study period). The rating team coded the data by viewing images of the road segments using the Hawkeye software program and coding the presence/absence of particular road and roadside features and measuring required distances (e.g. lane width, median width) every 20m from the beginning to the end of each road segment. The data were provided by ARRB in excel format.

Once the data file was provided by ARRB, it was subjected to an extensive and iterative process of validation and quality checks. First, the geographic co-ordinates in the data file were used to ensure that the study road segments had all been coded as defined. Some errors were identified and these were subsequently rectified by the ARRB group.

Available means were used to check the quality of the data in the ARRB-coded data set. For example, on-line images (Google Street View and Google Earth) were used to check the number of lanes, presence of turning lanes, presence of medians, median types, presence of service roads, presence of bicycle lanes and parking controls. The Melway street directory (Melway, 2005, 2006, 2007, 2008, 2009) was used to check that all intersections were included, the intersection type was coded correctly and that public transport variables were correct (e.g. presence of trams and buses, location of tram stops, bus stops). Distance measurements could not be accurately checked without the Hawkeye software. Variables were also tabulated to identify inconsistencies in coding; e.g. where there was a median width coded but the median type was coded as none, or if tram stops were recorded when trams were coded as absent.

Once the data had been thoroughly checked and cleaned, summary data were derived and coded into an IBM SPSS Statistics for Windows (Version 20.0) database, with one row per road segment. Each road segment comprised multiple rows in the raw data file provided by ARRB (one row per 20m of segment length). Thus, each road segment comprised between ten and 200 rows in the raw data file. For each variable, careful consideration was given as to how to summarise the available information. Initially, data were summarised at the most detailed level possible, following that, less detailed variables could be derived. For example, some variables involved a count of the road feature (e.g. roundabouts), with a second variable simply indicating the presence or absence of roundabouts. For variables relating to aspects of the roadway or environment that could change over the road segment, the mode, minimum and maximum values were derived for each road segment (e.g. lane width). For some variables, the proportion of the road with the feature was also calculated (e.g. proportion of the road with a median).

4.4 Data items and collection methods

In the following sections, a detailed description of the data items and the methods used to collect each data item for the urban midblock road segments in this study are presented. All of the road segment characteristics in Table 4.1 are considered in this section, including those for which data could not be obtained. Those characteristics that were not relevant for urban midblock road segments are identified. In addition, at the end of each variable description, descriptive data are provided for the road segments used in the study, in the form of median (interquartile range (IQR)) or the number (%) in each category.

4.4.1 Outcome data

In Victoria, Australia, a database of police-reported casualty crashes is held by VicRoads that is publicly accessible via the CrashStats (VicRoads, 2008a) on-line searchable database[2]. The database contains details of police-reported crashes from 1987 onwards where the crash resulted in the death of any person within 30 days or where the police identified one of the road users as injured. To be included, the crash must have occurred on a public road (including footpaths) and to have involved one or more road vehicles in motion at the time of the crash.

The CrashStats database can be searched by location and the search results can be limited to a particular time span. Information is available on the crash circumstances, location, environmental conditions, road users involved and vehicles involved. Each crash has a unique crash number.

A map-based search of CrashStats was performed by selecting the road segment of interest. Data for crashes that occurred on that segment were extracted by requesting all data for the casualty crashes that occurred during the period January 1st, 2005 to December 31st, 2009. Each crash was identified using the unique accident number in the crash database. Information in the crash data files was used to determine the total number of crashes on each segment, as well as the number of different types of crashes. Specific definitions for different crash types were rarely provided in the previous literature. The definitions of different crash types (MVC, SVC,

[2] https://www.vicroads.vic.gov.au/safety-and-road-rules/safety-statistics/crash-statistics

PVC, BVC and total crashes) used for the purposes of this thesis are defined in Table 4.2. The different crash types used in this study were identified using the variables describing road user type and the number of vehicles involved in the crash. Crashes that occurred at signalised (major) intersections were excluded because the focus of this study was on crashes occurring on midblock road segments.

Table 4.2. Definition of crash types

Crash type	Abbreviation	Definition
Multiple Vehicle Crashes	MVC	Number of crashes involving more than one vehicle
Single-vehicle crashes	SVC	Number of crashes involving only one vehicle (excluding those where a pedestrian was also involved)
Pedestrian-vehicle crashes	PVC	Number of crashes involving a pedestrian
Bicycle-vehicle crashes	BVC	Number of crashes involving a bicycle
Total casualty crashes		Total number of casualty crashes

4.4.2 Exposure

4.4.2.1 Segment length

Road segment length was measured using the measurement tool available on the Land Victoria, Department of Sustainability and Environment on-line interactive maps (http://services.land.vic.gov.au/) and validated using the coded video data from ARRB. The median segment length was 402m (range 200m to 4.1km; IQR=403m).

4.4.2.2 Traffic volume

Data on the Average Annual Daily Traffic Volume (AADT) was obtained from VicRoads for the entire state of Victoria for 2005 to 2011. The file comprised the following information; year of measurement, AADT and the percentage of heavy vehicles, and the measurement location. AADT for all road segments in the study were extracted for the midpoint of the study period (2007). The bi-directional AADT for each study road segment was calculated by adding the AADT for each direction. In many cases, AADT was measured at more than one point on the segment of interest. For example, AADT was measured at five different locations between Church St and Yarra Boulevard on the Bridge Rd, Richmond strip shopping centre road segment. Where AADT were available for more than one point in the segment of interest, the average bi-directional traffic volume was recorded and entered into the summary database.

Other measures of traffic exposure were calculated using the average AADT. For ease of interpretation, all measures were divided by 1,000 and expressed as thousand vehicles per day. The traffic exposure measures calculated from average AADT are shown in Table 4.3.

Table 4.3. Traffic exposure measures: Formulae and median, interquartile range (IQR) and range

Traffic exposure measure	Formula	Median (IQR)	Range
Thousand vehicles per day	AADT/1000	24.02 (16.02)	5.7 to 77.5
Thousand vehicle km per day	(AADT/1000) x segment length	11.68 (11.35)	1.0 to 127.4
Thousand vehicles per lane per day	(AADT/1000) / number of lanes	6.40 (2.98)	2.6 to 17.3
Thousand vehicle km per lane per day	((AADT/1000) x length)/number of lanes	2.89 (2.87)	0.7 to 23.9

Measures of traffic volume at a finer resolution than AADT (e.g. hourly traffic volumes) were not available for this study.

4.4.2.3 Direction of travel

Direction of travel was obtained from the Melway street directory. All road segments in the study sample carried two-way traffic.

4.4.2.4 Pedestrian volume/activity

Unlike traffic volumes, pedestrian volumes are not routinely measured by VicRoads or by every local council (Alavi, 2013). Thus no direct measurements of pedestrian volumes were available for the study road segments.

The challenge was therefore to find a data source that allowed estimation of the pedestrian volumes on the study road segments. The Victorian Integrated Survey of Travel and Activity (VISTA) (Department of Transport, 2009) was the only data-source identified that contained data on pedestrian activity in the regions of interest in this study. VISTA 2007 was a study of the travel behaviour of people in the state of Victoria, conducted from May 2007 to June 2008 involving 11,400 randomly selected households in metropolitan Melbourne. All household members were requested to record all travel outside their home on one specified date. VISTA 2007 may be useful as an indicator of relative differences in pedestrian activity across road segments in different areas. Estimates of pedestrian activity in the local area in which each segment was located were calculated using data from VISTA 2007.

The Statistical Local Area (SLA) was the smallest geographical area for which trip data were available. SLAs correspond to local government areas, or part thereof (Australian Bureau of Statistics, 2007). For each SLA, the number of trips made on an average day were obtained for each method of travel, by distance travelled. The methods of travel were categorised by vehicle, walking, bicycle, motorcycle/scooter, taxi, school bus, public bus and tram, while the distances were recorded in one km bins (0–0.9 km, 1–1.9 km, 2–2.9 km, etc., up to >150km). Data were downloaded for three types of trips: those with an origin and a destination in the SLA (within-SLA trips), those with an origin in the SLA (and a destination anywhere) and those with a destination in the SLA (and an origin anywhere). The total number of trips within each SLA was calculated as the number of trips with an origin in the SLA plus the number of trips with a destination in the SLA minus the number of trips with both an origin and a destination in the SLA.

To estimate the distance travelled by walking in a SLA, the number of walking trips in each distance bin was obtained. On average, 83% of the walking trips in VISTA 2007 had an origin and a destination in the same SLA, for the SLAs in which the study road segments were located. For trips with an origin and a destination in the same SLA, it was assumed that 100% of the trip occurred within the SLA. The number of trips within the SLA was multiplied by the median value of the bin (e.g. 0.4km for the 0–0.9km bin) to estimate the total distance walked for trips within the SLA. For trips with only an origin or a destination in the SLA, it was assumed that half of the trip occurred in the SLA of interest — again, the median value was used (e.g. 1/2 x 0.4km for the 0–0.9km bin). This assumption was justified because a relatively small proportion (17%) of walking trips began or ended in a different SLA, and it is unlikely that many of these crossed more than two SLAs.

This calculation led to an estimate of the total distance walked in each SLA. This was divided by the area (in square km) of the SLA to calculate the distance (km) walked per square km for the SLA. Then, this was multiplied by the length of the segment (km) to adjust for segment length. The median value was 0.62 (range 0.02 to 34.6; IQR=1.04).

4.4.2.5 Cyclist volume/activity

Similar to pedestrian volumes, data on cyclist volumes are not collected routinely by state or local government and so consistently measured estimates of cyclist activity on the road segments of interest were not available. While data were available from VISTA 2007 regarding cycling activity in the SLAs in which road segments were located, they were not as useful for estimating cyclist exposure as they were for estimating pedestrian exposure. The reason for this was that cyclists travel longer distances than pedestrians per trip, therefore they are more likely to travel through more than one or two SLAs on a trip. There was not enough information available to make assumptions about the average distance that should be attributed to the SLA of interest for cyclist trips. As such, estimates of distance travelled within the SLA by bicycle were not calculated. Because no estimates of cyclist activity were available for this study, the decision was made not to investigate BVC as a separate crash type in this thesis.

4.4.3 Speed zone

The speed zone of the road on which the crash occurred is reported by police and is recorded in the VicRoads CrashStats database. Initial plans were to use the speed zone variable in the crash file to determine the speed zone on the road segments of interest. Unfortunately, this proved problematic because the data were inconsistent. For example, one study road segment with a known speed limit of 70 km/h was coded as 50 km/h, 60 km/h, 70 km/h and 80 km/h for different crashes on that road segment in the crash database. For some crashes, it is possible that the police officer recorded the speed zone of an intersecting street, however, this was not always the case: in the example given above, there are no streets with a speed limit of 80 km/h intersecting the road segment of interest. As a compromise, use of the mode value was considered, however, sometimes the data for a road segment were bimodal: one site had six crashes coded as occurring on a 50 km/h road and six crashes coded as occurring on a 60 km/h road. As a consequence, other sources for ascertaining the speed zone of the road segments were sought.

VicRoads have a map-based database (RoadNet) that contains the speed limits of roads in Victoria, amongst other data. A senior engineer at VicRoads organised for the author of this thesis to have access to the database and training in its use. The speed limit on each of the roads was obtained, including details of whether the speed limit was variable, and if so, the reason for the variable speed zone (school zone or strip shopping centre zone) and times of operation. In Victoria, variable speed limits were introduced on all roads outside schools prior to the commencement of this study. On some strip shopping centre road segments however, variable speed limits have been introduced in a staged approach over the last decade. For the study road segments that were identified as having a strip shopping centre variable speed limit using the RoadNet database, further investigation was undertaken by the Senior Engineer at

VicRoads to identify the date at which the variable speed limit was implemented. If the date was prior to 1st January 2005, then the site was classified as having a variable speed limit during the entire study period (2005 to 2009). If the date of implementation of the variable speed limit was after 31st December 2009, the site was classified as having the original speed limit it had prior to the variable speed limit being introduced.

During the period 1st January 2005 to 31st December 2009, two of the strip shopping centre road segments chosen for the study had the speed zone reduced permanently from 60 km/h to 40 km/h and eight road segments had the speed zone changed from permanent 60 km/h zones to variable speed zones that were 40 km/h for part of the day and 60 km/h for the remainder (refer to Table 4.4.). Eight of the speed zone changes occurred during the first half of 2005. The other two occurred toward the end of 2008. Considering these changes occurred at the beginning and toward the end of the five year (1826 day) study period, these sites were not excluded from the study. Instead, the period of observation for the cumulative crash count was reduced to include the longest period during which speed limit was unchanged. So, for road segments on which the speed zone changed during the first half of 2005, the period for counting crashes started on the date of the speed zone changed and ended on 31st December, 2009. For road segments on which speed zone changed at the end of 2008, the period for counting crashes started on the 1st January 2005 and ended on the date of the speed zone change. The reduced number of days for outcome ascertainment (counting the number of crashes) is also shown in Table 4.4, and this was accounted for in the analysis.

Table 4.4. Speed zone changes that occurred on study road segments during the study period

Segment	New speed limit	Date of change	Speed limit used for study	Days of observation for crash data extraction
Mt Alexander Rd, Moonee Ponds	Variable 60/40 km/h	16/3/2005	Variable 60/40 km/h	1752
Toorak Rd, Toorak (Punt Rd to Surrey Rd Nth)	Permanent 40 km/h	21/3/2005	40 km/h	1747
Commercial Rd, Prahran	Variable 60/40 km/h	21/3/2005	Variable 60/40 km/h	1747
Toorak Rd, Toorak (Tintern Ave to Grange Rd)	Permanent 40 km/h	21/3/2005	40 km/h	1747
Glenferrie Rd, Malvern (Wattletree Rd to High St)	Variable 60/40 km/h	31/5/2005	Variable 60/40 km/h	1676
Glenferrie Rd, Malvern (Dandenong Rd to Wattletree Rd)	Variable 60/40 km/h	31/5/2005	Variable 60/40 km/h	1676
Springvale Rd, Springvale	Variable 60/40 km/h	8/6/2005	Variable 60/40 km/h	1668
Pascoe Vale Rd, Glenroy	Variable 60/40 km/h	29/6/2005	Variable 60/40 km/h	1647
Glenferrie Rd, Hawthorn	Variable 60/40 km/h	27/10/2008	60 km/h	1395
Station St, Fairfield	Variable 60/40 km/h	17/11/2008	60 km/h	1416

Table 4.5 shows the distribution of speed zones for the study road segments.

Table 4.5 Distribution of speed zones across study road segments

Speed limit	Number of segments	% of segments
40 km/h	2	1.4
50 km/h	2	1.4
Strip shopping centre zone variable speed limit: 40km/h or 60 km/h depending on time of day	14	9.9
60 km/h	100	70.4
70 km/h	17	12.0
80 km/h	7	4.9

4.4.4 Road cross-section

4.4.4.1 Road curvature (horizontal and vertical)

Melway (2005, 2006, 2007, 2008, 2009) street directory maps revealed that the majority of the road segments in this study were completely straight. Considering the small proportion of sites with any type of curve, only the presence or absence of curve/s was measured for this study (125 (88.0%) had no curve, 17 (12.0%) had a curve).

No data sources were found with data on vertical curvature or grade for roads in metropolitan Melbourne, consequently, this characteristic was not investigated in this study.

4.4.4.2 Carriageway width and lane width

Carriageway width is defined as the width of the road, measured from kerb to kerb. Hence, for roads with a concrete median, the carriageway width reflects the width for one direction of travel only, while for roads without a concrete median, both directions of travel are included. Digital video images were used to obtain the mode, minimum and maximum carriageway width for each road segment. For two divided roads, the mode carriageway width on one side of the road was different to the mode on the other side. So as to avoid missing data, two variables were constructed to reflect the mode: one used the smaller of the two values for these two road segments (mode – min), the other used the larger of the two values (mode-max). Table 4.6 presents the distribution of carriageway widths across study road segments.

Table 4.6 Distribution of carriageway width (mode, minimum and maximum) across study road segments

Carriageway width	Mode –min N (%)	Mode –max N (%)	Minimum N (%)	Maximum N (%)
5.5–7m	1 (0.7)	1 (0.7)	6 (4.2)	0 (0)
7–8.5m	4 (2.8)	3 (2.1)	10 (7.0)	1 (0.7)
8.5–10m	7 (4.9)	6 (4.2)	6 (4.2)	3 (2.1)
>10m	130 (91.6)	132 (93.0)	120 (84.5)	138 (97.2)

The mode, minimum and maximum lane width on each road segment were obtained from the digital video images. For ten road segments, the lane width data were bi-modal. To avoid missing data, two variables were created where one classified the road segment according to the smaller of the two modes (mode-min) while the other classified the road segment according to the larger of the two modes (mode-max). Table 4.7 shows the distribution of lane widths across study road segments.

Table 4.7 Distribution of lane width (mode, minimum and maximum) across study road segments

Lane width	Mode –min N (%)	Mode –max N (%)	Minimum N (%)	Maximum N (%)
2.3–2.7m	8 (5.6)	8 (5.6)	12 (8.5)	2 (1.4)
2.7–3m	18 (12.7)	15 (10.6)	28 (19.7)	9 (6.3)
3–3.3m	67 (47.2)	61 (43.0)	70 (49.3)	60 (42.3)
3.3–3.5m	23 (16.2)	24 (16.9)	20 (14.1)	27 (19.0)
>3.5m	26 (18.3)	34 (23.9)	12 (8.5)	44 (31.0)

4.4.4.3 Number of lanes

Digital video images were used to determine the mode number of through lanes for each road segment. These data were checked using on-line image sources (Google Street View and Google Earth). The mode number of lanes ranged from one to eight, with a median of four (IQR=0).

4.4.4.4 Shoulder type and width

Shoulder type and width for each road segment were measured from digital video images. Only one of the road segments of interest had a road shoulder because the lanes for all other roads were directly adjacent to either roadside parking spaces or directly abutted the kerb leading to footpaths. Due to the lack of variation across road segments, this variable was not included in the analysis.

4.4.4.5 Line marking

The digital images were used to code the most commonly occurring midline marking on each road segment (refer to Table 4.8 for the distribution of mid-line types across study road segments). Single unbroken lines and medians (no lines) accounted for 94% of the study road segments. This variable provided no further information than the variable describing the presence or absence of a median and was therefore not used in the analysis.

Table 4.8 Distribution of mid-line types across study road segments

Predominant type of mid-line	Number of road segments	% of road segments
No midline	1	0.7
Single unbroken line	71	50.0
Single broken line	5	3.5
Median (no midline)	62	43.7
Double unbroken line	3	2.1

4.4.4.6 Bridges

The Melway street directory (2005, 2006, 2007, 2008, 2009) was used to ascertain that none of the road segments of interest crossed a bridge. While some of the road segments had bridges crossing over them, this was coded under the variable relating to low clearance (where applicable). Hence, the presence of bridges was not included in the summary database.

4.4.4.7 Passing lanes and transit lanes

The digital video images were used to establish that none of the road segments of interest had passing lanes or transit lanes. Consequently, these were not included in the summary database.

4.4.4.8 Guardrails/barriers

Guardrails and barriers installed to protect drivers from risky roadside features such as slopes, trees and drains are more common in rural areas than urban areas. The barriers used in urban areas are more likely to be for restricting pedestrian access, and this was included in the variables relating to pedestrian facilities (refer to section 4.4.10.2).

4.4.4.9 Speed management devices

Digital video images and the Melway street directory (2005, 2006, 2007, 2008, 2009) revealed that none of the road segments had speed management devices such as speed humps or chicanes. These speed management devices are commonly used on other types of urban roads in metropolitan Melbourne, but not arterial roads. As such, these were not included in the summary data.

4.4.5 Accesses and intersections

4.4.5.1 Signalised intersections

Signalised intersections were intersections that had traffic lights. The digital video image data were used to code the number of signalised intersections. The type of intersection was also recorded (3-arm, 4-arm or other). The Melway street directories (2005, 2006, 2007, 2008, 2009) were used to validate these data. The rate per km was calculated by dividing the number of signalised intersections by the segment length. The rate per km was used in the analysis. Table 4.9 shows the median, range and IQR of the rate of signalised intersections across study road segments.

Table 4.9 Median, interquarterile range (IQR) and range of the number of signalised intersections per km across study road segments

Signalised intersection type	Median number/km	IQR	Range
3 way	0	1.6	0–7.7
4 way	1.6	3.2	0–8.3
Other	0	0	0–2.8
Total	2.7	2.8	0–11.0

Digital video images were used to determine whether the traffic signals had a controlled turn phase: that is, right or left turn arrows. These data, however, were not used in the current research which focused on midblock crashes. Data were not readily available on other aspects of the traffic signal phasing systems used at the intersections in this study.

4.4.5.2 Unsignalised intersections

Unsignalised intersections were intersections that did not have traffic lights. To be defined as an intersecting minor road, the road was required to be part of the declared public road network (identified by having a name) as opposed to laneways or private access roads which do not have names. The number and type of unsignalised intersections was recorded from the digital video images. Data were validated using the Melway street directory (2005, 2006, 2007, 2008, 2009). The rate per km was also recorded and used in the analysis. It was not possible to determine the type of traffic control (e.g. stop sign, give-way sign) for all of the intersecting roads at the unsignalised intersections from the digital video images and so this was not included in the study. As can be seen in Table 4.10, which presents the median, IQR and range of the rate of unsignalised intersections across study road segments, the majority of unsignalised intersections on these road segments had three arms.

Table 4.10 Median, interquarterile range (IQR) and range of the number of unsignalised intersections per km across study road segments

Unsignalised intersection type	Median number/km	IQR	Range
3 way	7.9	6.7	0–24.8
4 way	0	0	0–7.1
Staggered T	0	0	0–5.5
Total	8.5	6.8	0–28.0

4.4.5.3 Roundabouts

The number of roundabouts on each road segment was counted using the digital video images and validated using the Melway street directory (Melway, 2005, 2006, 2007, 2008, 2009). This was then recorded in the summary data file as the number of roundabouts per road segment, the number of roundabouts per km and as a binary variable indicating the presence or absence of roundabout/s. The majority of segments (134: 94.4%) had no roundabouts. Seven segments had one roundabout and one segment had two roundabouts. The median number of roundabouts per km was zero (range 0 to 5.0, IQR=0)

4.4.5.4 Service roads

The presence and type of service roads was determined using the digital video images and validated using the Melway street directory (2005, 2006, 2007, 2008, 2009) and Google Street View on-line images. The distribution of service roads, by type, across study road segments is shown in Table 4.11.

Table 4.11 Distribution of service roads (by type) across study road segments

Type of service road	Number of road segments	% of road segments
None	115	81.0
One side without parking	2	1.4
One side with parking	11	7.8
Both sides with parking	14	9.9

In addition, the number of service road access points was counted using the digital video images for each road segment. The rate of service road accesses per km was also calculated (median=0, range 0 to 27.1, IQR=0).

4.4.5.5 Driveways/laneways

Driveways were defined as vehicle accesses for properties, while narrow intersecting lanes without road names were defined as laneways (common in the inner suburbs of Melbourne). Both provide a means of accessing the main road. Frequently, they are only wide enough for one vehicle to enter or exit at a time, and often there is limited visibility for (and of) traffic exiting the driveway/laneway. The number of driveways and laneways was counted using the digital video images for each road segment. The number of laneways was validated using the Melway street directory (2005, 2006, 2007, 2008, 2009) while the number of driveways was validated for a sample of the road segments using Google Street View. The rate of driveways and laneways per km was also calculated (median=17.0, range 0 to 90.5, IQR=20.8).

4.4.5.6 Dedicated turning lanes

Dedicated turning lanes were identified using the digital video images for each road segment and validated for a sample of road segments using Google Street View. The presence and type of turning lanes was determined for each road segment. Summary data included the number of each type of turning lane in each road segment and the rate per km, which was used in the analysis. Table 4.12 presents the median, IQR and range of the rate of dedicated turning lanes across study road segments.

Table 4.12 Median, interquarterile range (IQR) and range of the number of dedicated turning lanes per km across study road segments

Type of dedicated turning lane	Median number/km	IQR	Range
Right	2.6	7.1	0–15.3
Left	0	1.8	0–12.4

4.4.5.7 Keep clear zones

Zones with “Keep Clear” markings are frequently found adjacent to intersecting roads, laneways and some driveways to enable access for cross traffic when traffic is queued on a busy road. It is forbidden for drivers to stop in keep clear zones. The number of keep clear zones was counted using the digital video images for each road segment. The rate per km was also calculated and used in the analysis (median=1.4, range 0 to 15.6, IQR=4.1).

4.4.5.8 Skew angle

The skew angle of the intersections (the variation in intersection angle from 90 degrees) on the study road segments was not measured in this study, as the focus was on crashes on midblock road segments, not intersections.

4.4.6 Medians

The digital video images were used to determine whether there was a median present and if so, the type, width and how many median accesses there were (not including those accesses at signalised intersections). These data were validated using the Melway street directory (2005, 2006, 2007, 2008, 2009) and Google Street View.

The most common median type and the mode, minimum and maximum median width were coded in the summary data file (refer to Table 4.13 and Table 4.14). Because a road which was predominantly undivided could still have a maximum median width if there was a median or a traffic island anywhere on the road segment, the maximum median width variable was also used to generate a new variable that designated whether there was a median or traffic island anywhere on the road segment (that is, if the maximum median width was greater than zero, then at least one median or traffic island was present).

Table 4.13 Distribution of median types across study road segments

Median type	Number of road segments	% of road segments
None	82	57.8
Paint	10	7.0
Separated tram lane	10	7.0
Raised concrete	40	28.2

Table 4.14 Distribution of mode, minimum and maximum median widths across study road segments

Median width	Mode Number (%)	Minimum Number (%)	Maximum Number (%)
None	82 (57.8)	128 (90.1)	40 (28.2)
<1.2m	21 (14.8)	3 (2.1)	14 (9.9)
1.2–3m	25 (17.6)	2 (1.4)	62 (43.7)
>3m	14 (9.9)	9 (6.3)	26 (18.3)

Several other variables describing the medians of each road segment were derived for entry into the summary data file:

- If there was a median of any type present for more than half of the road segment (n=82), the road was categorised as divided.
- A stricter definition of divided was also used: if a median that provided a physical barrier to being crossed (e.g. concrete or separated tram-lane) was

present for more than half of the road segment, then the road met the strict definition for being divided (n=50).

- Variables were also generated to indicate the proportion of the road that was divided (any type of median), the proportion that was divided according to the strict definition, and the proportion with different median widths.

The number of midblock median accesses was also counted for each road segment and entered into the summary data file along with the rate per km which was used in the analysis (median=0, range 0–8.3, IQR=0).

4.4.7 Road type and traffic mix

4.4.7.1 Road type

To be included in the study, the strip shopping centres had to be located on an arterial road. Primary arterial roads are principal routes for moving goods and people in urban areas. Secondary arterial roads are also principally for traffic movement but may have a higher level of access than primary arterial roads (VicRoads, 2010). Information in the Melway street directory (2005, 2006, 2007, 2008, 2009) was used to distinguish whether the road segment was on a primary or secondary arterial road. The final dataset held two binary variables indicating road type: one variable that identified all roads that were entirely, or part, primary state arterial (114, 80.3%) and one variable that identified all roads that were entirely or part, secondary state arterial (29, 20.4%).

4.4.7.2 Heavy vehicle access

Whether or not heavy vehicle access was approved on the road segments of interest was determined using the Melway street directory (2005, 2006, 2007, 2008, 2009) and validated against historical VicRoads heavy vehicle access maps (VicRoads, 2004, 2007). The following categories of heavy vehicles and their mass limits were only permitted to travel on approved routes (VicRoads, 2004):

- double articulated (b-double) trucks ≤19m in length with road friendly suspension and mass between 50 and 57 tonnes (b-doubles of ≤19m in length and mass <50 tonnes were allowed to travel on any roads)
- b-double trucks between 19m and 25m in length
- prime mover and semi-trailers with tri-axle group and road friendly suspension at higher mass limits (>42.5 tonnes but ≤45.5 tonnes)

Heavy vehicle access was approved for the entire road segment for 131 (92.3%) of study segments.

Road segments were also classified according to whether they were on a route for over-dimensional (over-size and over-mass vehicles). These were defined as vehicles of up to 49.5 tonnes (depending on axle configuration) that were also more than 3.5m wide, 4.6m high or 12.5m long for rigid vehicles or 25m long for prime-movers and semi-trailers, extendable semi-trailers or low loader/dolly combinations (VicRoads, 2007). Road segments were also classified as to whether or not they intersected with an over-dimensional route. Seven road segments (4.9%) were on an over-dimensional route and a different seven road segments (4.9%) intersected with an over-dimensional route.

4.4.7.3 Percentage heavy vehicle traffic

Data describing the percentage of traffic that were heavy vehicles were obtained from VicRoads. Some study road segments had multiple measurements taken along their length. The average percentage of heavy vehicle traffic (median=4.6%, range 1.5–10.6%, IQR=3.0%) for the whole road segment of interest was calculated, along with the maximum percentage of heavy vehicle traffic (median=4.7%, range 1.5–11.1%, IQR=3.0%).

4.4.8 Roadside parking

4.4.8.1 Parking type

Digital video images were used to determine whether parking was permitted on each side of the road, and if so, what type of parking. Google Street View was used to validate these data for a sample of the road segments. Possible roadside parking types were: parallel parking in a through travel lane, parallel parking in a sheltered lane (that is, specifically allocated parking areas that were sheltered by extensions of the kerb, not a through lane) and angle parking on the roadside in a designated reserved space. Whether or not parking was permitted in the centre of the road was also recorded for divided roads. The distribution of roadside parking across road segments by type is shown in Table 4.15.

Table 4.15 Distribution of roadside parking configurations across study road segments

	Number of segments	% of segments
	Parking (any type)	
Parking not permitted	18	12.7
1 side of the road	16	11.3
2 sides of the road	105	73.9
2 sides of the road and in the centre median	3	2.1
	Parking (in-lane)	
None	29	20.4
1 side of the road	29	20.4
2 sides of the road	84	59.2
	Parking (sheltered)	
None	96	67.6
1 side of the road	20	14.1
2 sides of the road	26	18.3
	Parking (angle)	
None	117	82.4
1 side of the road	17	12.0
2 sides of the road	8	5.6

4.4.8.2 Parking clearways

Clearways are zones where parking is not permitted for certain time periods, for example, during peak traffic periods. Most commonly, a clearway would be present on road segments where parking is allowed during hours not covered by the clearway period. It is possible, however, to have a part-time clearway during peak traffic periods on a road where no parking is permitted at other times. While this seems counterintuitive, the penalties for parking in a clearway are higher than for parking in a no standing zone so there is a stronger disincentive to park during peak traffic periods. The digital video images were used to identify the presence and type (one-side, two-side, part-time or full-time) of clearways for each road segment. A sample of observations was validated using Google street view images. The number and proportion of road segments with different types of clearways are shown in Table 4.16

Table 4.16 Distribution of presence and type of clearway across study road segments

Type of parking clearway	Number of segments	% of segments
None	86	60.6
Part-time on 1 side of road	12	8.5
Part-time on 2 sides of road	44	31.0

4.4.8.3 Loading zones

Particular vehicle types may stop in a loading zone if they are dropping off or picking up goods or people. The number of loading zones per km was calculated for each road segment (median=0, range, 0–15.6, IQR=3.6).

4.4.9 Public transport and bicycle facilities

4.4.9.1 Railway level crossings

The presence or absence of a railway level crossing in each road segment was ascertained by referring to the Melway street directory (2005, 2006, 2007, 2008, 2009). There were twelve (8.5) study road segments with railway level crossings within the road segment, nine (6.3%) road segments with railway level crossings adjacent to the road segment and 121 (85.2%) road segments with no railway level crossings. No road segment in this study had more than one railway level crossing.

4.4.9.2 Buses, bus lanes and bus stops

The Melway street directory (2005, 2006, 2007, 2008, 2009) was used to determine that there were 101 (71.1%) road segments with one or more bus routes. It was determined from the digital video images that only two road segments had specialised bus lanes so this variable was not included in the analysis. The number per km and of different types (on-road, off-road) of bus stops on each road segment was calculated (refer to Table 4.17).

Table 4.17 Median, interquarterile range (IQR) and range of the number of different types of bus stops per km across study road segments

Bus stops per km	Median (IQR)	range
On-road (within a travel lane)	1.7 (5.2)	0–24.9
Off-road (sheltered from travel lane)	0 (0)	0–14.7
Total (any type of bus stop)	3.5 (6.9)	0–24.9

4.4.9.3 Trams, tram lanes and tram stops

Trams are an iconic feature of the Melbourne public transport system that are more common in the CBD and inner suburbs. Trams and vehicles are integrated on the road system in different ways on different roads. On some roads, vehicles and trams are required to share the traffic lane closest to the midline. This means that on roads with two lanes in either direction that also have roadside parking, vehicles are often forced to share the remaining free lane with trams when cars are parked on the roadside. On other roads, trams travel on a separate section in the middle of the road that vehicles are not permitted to enter. There are also some part-time tram-lanes where vehicles can enter during certain time periods, although these are rare.

Tram stops differ according to how trams are integrated into the road system. Where vehicles are required to share the middle lane with trams, the tram stops where travellers wait for the tram are generally on the roadside (kerbside tram stops). When the tram halts for boarding and alighting passengers, cars are legally required to stop behind the tram until passengers have boarded and/or alighted the tram and the tram doors close. In cases where trams have their own lane, the tram stops are generally located next to the tram lane, in the centre (median) of the road.

The presence of a tram route, the type of tram-lane and the number and type of tram stops on the study road segments were ascertained from the digital video images. The presence of a tram route was validated using the Melway street directory (2005, 2006, 2007, 2008, 2009) and the other data were validated for a sample of road segments using Google Street View. Data were summarised by road segment into the following variables: presence of trams, tram lane type, the number per km of tram stops (any type, kerbside and centre median) (refer to Table 4.18).

Table 4.18 Presence of tram routes, types of tram lanes and median, interquarterile range (IQR) and range of the number of different types of tram stops per km across study road segments

Categorical variables	Number of segments	% of segments
	Tram route on segment	
No	89	62.7
Yes	53	37.3
	Tram lane type	
None (no trams)	89	62.7
Shared with vehicles	43	30.3
Separate from vehicles	10	7.0
Continuous variables	Median (IQR)	Range
Number of kerbside tram stops per km	0 (3.3)	0–9.1
Number of median tram stops per km	0 (0)	0–11.0
Total number of tram stops per km	0 (3.8)	0–11.0

4.4.9.4 Bicycle facilities

The digital video images were used to code the presence of bicycle lanes on one or both sides of the road and the bicycle lane width (<1.2m or >1.2m) for each road segment. These data were validated using Google Street View. The road segments were classified according to the presence or absence of bicycle lanes (24, or 16.9% of segments had a bicycle lane) and the modal bicycle lane width for each segment (14 of the bicycle lanes were <1.2m wide while 10 were >1.2m wide). Another variable was also created which indicated whether or not the bicycle lane was separated from roadside parking spaces or whether it shared space with the marked parking spaces — 16 of the 24 bicycle lanes shared space with marked roadside parking bays.

4.4.10 Pedestrian facilities

4.4.10.1 Pedestrian crossings

The number and type of pedestrian crossings was ascertained from the digital video image data and validated using the Melway street directory (Melway, 2005, 2006, 2007, 2008, 2009) and Google Street View. The pedestrian crossings were coded as being either at or between signalised intersections and whether or not there was a stop and go traffic signal or if it was simply a signed pedestrian crossing with flashing lights. Almost every signalised intersection had a pedestrian crossing. The midblock pedestrian crossings were of more interest in this study of crashes on urban midblocks. The number of the different types of midblock pedestrian crossings per km per road segment are shown in Table 4.19.

Table 4.19 Median, interquarterile range (IQR) and range of the number of different types of pedestrian crossings per km across study road segments

Midblock pedestrian crossings per km	Median (IQR)	range
Stop and go traffic signal	1.1 (2.6)	0–10.0
Signed and flashing lights	0 (0)	0–4.3
Any type of midblock pedestrian crossing	1.2 (2.8)	0–10.0

4.4.10.2 Fencing at pedestrian crossings

The presence and location of pedestrian fencing at pedestrian crossings was determined from the digital video image data and validated for a sample of observations using Google Street View. Road segment were categorised according to whether there were no crossings (6 road segments, 4.2%) or pedestrian crossings where none (78 road segments, 54.9%), some (39 road segments, 27.5) or all (19 road segments, 13.4%) of the pedestrian crossings were fenced.

4.4.10.3 Footpath (or sidewalk)

Digital video images revealed that all of the road segments in this study had footpaths, so this variable was not included in the summary database or the analysis.

4.4.11 Roadside development

4.4.11.1 Nature strips

A nature strip is an area (usually grassed) that lies between the kerb and the footpath; common in suburban Melbourne. Some densely populated inner suburban streets, however, do not have nature strips. The presence of a nature strip on one or both sides of the road was determined from the digital video images and validated for a sample of observations using Google Street View. The majority of strip shopping centre road segments did not have nature strips (105, 73.9%). Sixteen study road segments (11.3%) had a nature strip on one side of the road and 21 (14.8%) study road segments had nature strips on both sides of the road.

4.4.11.2 Poles and trees

The number and type of poles and trees on the roadside and the median was counted using the digital video images, and validated using Google Street View for a sample of road segments. Utility poles, large trees and verandah poles were classified as non-frangible while small trees, parking signs, advisory signs and public transport stop signs were classified as frangible. This definition reflected the likely amount of damage to a vehicle if it was to collide with the pole or tree (non-frangible poles and trees were likely to result in significant damage). The number per km of frangible and non-frangible poles and trees was calculated for the roadside and the median for each road segment (refer to Table 4.20).

Table 4.20 Median, interquarterile range (IQR) and range of the number of different types of poles and trees per km across study road segments

Number of poles and trees per km	Median (IQR)	range
Roadside non-frangible	69.7 (36.4)	10.9–275
Roadside frangible	77.3 (33.5)	31.6–248.6
Total roadside	153.3 (55.0)	88.8–348.8
Median non-frangible	0 (12.4)	0–101.7
Median frangible	0.7 (17.8)	0–112.3
Total median	1.5 (39.0)	0–161.9

4.4.11.3 Development height

Development height was determined from Google Street View. For each road segment, the development height was classified as predominantly single storey (52, 36.6%), predominantly double storey (54, 38.0%), or a mix (which indicated there was no predominant development height; 36, 25.4%). In addition, the height of the highest building was classified as single storey (1, 0.7%), double storey (76, 53.5%) or three or more storeys (65, 45.8%)

4.4.11.4 Development type

This study focused on strip shopping centre road segments, so development type had to be retail/commercial for the road segment to be included in the study. Road segments were classified according to whether or not retail/commercial development was present on one (30, 21.1%) or both (112, 78.9%) sides of the road.

4.4.11.5 Offset distance

The offset distance between the edge of the road and buildings was categorised from the digital video images. The mode, minimum and maximum offset for each road segment are shown in Table 4.21.

Table 4.21 Distribution of offset distance between road edge and buildings (mode, minimum and maximum) across study road segments

Offset distance (m)	Mode* N (%)	Minimum N (%)	Maximum N (%)
0–3	51 (35.9)	92 (64.8)	18 (12.7)
3–5	64 (45.1)	39 (27.5)	33 (23.2)
5–10	3 (2.1)	4 (2.8)	26 (18.3)
>10	23 (16.2)	7 (4.9)	65 (45.8)

*Cells do not add to 142 because one road segment did not have a modal offset distance

The distance between the road edge and roadside trees and poles was also measured, however, the distance did not vary between road segments (all were categorised as between 0 to 3m) so this variable was not included in the analysis.

4.4.12 Road Pavement condition

Information on the condition of the road pavement on the study road segments was obtained from the VicRoads RoadNet database. Pavement roughness was measured using the International Roughness Index (IRI) which is related to the displacement of a vehicle's suspension per distance travelled (Cairney, 2008). Based on the results of community surveys, VicRoads define a road pavement as moderately rough when the average roughness over a 100m section is greater than 4.2 IRI and as very rough when the average roughness over a 100m section is greater than 5.3 IRI (VicRoads, personal communication, March 19th, 2012). For the purposes of this study, if part of a study road segment was designated as being moderately rough then the whole segment was classified as having moderate roughness present (41 road segments, 28.9%), likewise, if any part of a study road segment was classified as being very rough then the whole segment was classified as having very rough pavement present (52 road segments, 36.6%). Forty-nine (34.5%) road segments did not have any rough pavement present according to this definition.

The presence or absence of distressed road pavement on a road segment was also ascertained using the RoadNet database for each of the road segments. VicRoads classify a section of pavement as distressed if at least 30% of the segment has more than 10mm lane rutting in addition to at least 10% cracking (VicRoads, personal communication, March 19th, 2012). Lane rutting is measured as the maximum depth from a line drawn from high points on either side of the lane (Cairney, 2008). According to these criteria, just over half of the road segments had distressed pavement present (76 road segments, 53.5%).

4.4.13 Height Clearance

Information on low clearance on the road segments of interest was available from the Melway street directory (2005, 2006, 2007, 2008, 2009) and the VicRoads heavy vehicle access maps (VicRoads, 2004, 2007). There were two variables relating to low clearance.

The first related to situations where low clearance was a result of the presence of a bridge (or bridges) at one point on the road segment, and was classified as being no low clearance (134, 94.4%), clearance of between 4.3m and 4.6m (1, 0.7%), or clearance of 4.3m or less (7, 4.9%)

The second variable related to the presence of low overhead electrical wires used to power trams via a pantograph system (<4.6m) on the road segment. Low tram wires were present on five (3.5%) of the study road segments (the other road segments with trams had wires higher than 4.6m).

4.4.14 Static law enforcement

The presence of static speed and red light enforcement cameras at one or more signalised intersections on the road segment was ascertained from the Melway street directory (2008, 2009) and validated against information provided to the Monash University Accident Research Centre by the Victorian Department of Justice (J. Scully, personal communication, November 1st, 2010). There were static speed and red light enforcement cameras present at one location in 22 (15.5%) study road segments and at two locations in 3 study road segments (2.1%). There were no static speed and red light enforcement cameras on the remainder of the study road segments (117, 82.4%).

4.4.15 Amenities/facilities

4.4.15.1 Liquor licences

Information on the number of establishments with liquor licences and the type of liquor licence for each strip shopping road segment was obtained from the Victorian Department of Justice on-line database of liquor licences (Department of Justice, n.d.). Retrospective data were not available. Data were obtained in February, 2011, just over one year after the study period ended. It was assumed that the number of liquor licences in a strip shopping centre road segment would not have changed markedly in the intervening period. In addition, the relative difference between road segments was unlikely to have altered over the relatively small time period.

Postcodes were used to search for all licences held in that postcode. For each individual licence within the postcode of interest, a hyperlink was followed to obtain the address of the establishment and ascertain whether it was located within any of the road segments of interest in that postcode and to identify the type of licence. The different types of liquor licence in Victoria were classified into four higher-level categories for the purposes of this study and the number of each type of liquor licence per km for each segment was calculated (refer to Table 4.22).

Table 4.22 Categories of liquor licences and median, interquarterile range (IQR) and range of the number of different types liquor licences per km across study road segments

Category used in this study	Department of Justice liquor licence type	Description	Median number per km (IQR) [range]
BYO*	Bring Your Own (BYO)	BYO licences (usually cafes or restaurants) allow patrons to bring their own alcohol (usually restricted to wine) for consumption with the meal purchased at the restaurant/cafe.	2.0 (4.5) [0–22.1]
Late night	Late night: general Late night: on-premises	Allowed to sell alcohol for consumption on premises until after 1am	0 (1.8) [0–15.9]
Other on – premises*	General Restaurant and cafe On-premises Full club Restricted club	Allowed to sell alcohol for consumption on premises until 1am (many premises close earlier).	8.8 (13.8) [0–43.6]
Take-away	Packaged liquor	Allow purchase of alcohol but no consumption on premises.	2.9 (3.5) [0–14.9]

* These categories were also pooled to generate a higher level category of non-late night liquor licences for consumption on premises (median=11.7, range 0–59.2, IQR=14.0)

4.4.15.2 Other amenities/facilities

For the purposes of this study, an amenity or facility was defined as being on the road segment of interest if it was located either on the road segment or on the corner of the intersection at the end of the road segment. The one exception to this rule was for railway stations which were also classified as being on the road segment if they were accessible via a short street or lane that was only accessible from the study road segment.

The Melway street directory (2005, 2006, 2007, 2008, 2009) was used to determine the presence and number of various amenities and facilities on the road segments of interest. It was rare that there was more than one of each amenity or facility on any road segment so only the presence of the amenity/facility was included in the analyses. The one exception was that it was common to have multiple off-street parking facilities per road segment so the rate per km was calculated and used in the analyses (median=2.1, range 0–15.6, IQR=5.0). The number and percentage of strip shopping road segments with different types of amenities and facilities is shown in Table 4.23.

Table 4.23 Distribution of different types of amenities and facilities across study road segments

Presence of amenity/facility	Number of road segments	% of road segments
	Primary/Secondary School	
No	119	83.8
Yes	23	16.2
	Tertiary education institute	
No	136	95.8
Yes	6	4.2
	Facilities for young children	
No	117	82.4
Yes	25	17.6
	Hospitals/nursing homes	
No	137	96.5
Yes	5	3.5
	Place of worship	
No	109	76.8
Yes	33	23.3
	Community facilities	
No	112	78.9
Yes	30	21.1
	Parks/sports facilities	
No	110	77.5
Yes	32	22.5
	Railway stations	
No	92	64.8
Yes	50	35.2
	Emergency services	
No	125	88.0
Yes	17	12.0
	Indoor shopping centres	
No	109	76.8
Yes	33	23.2
	Petrol stations	
No	81	57.0
Yes	61	43.0

Tertiary education institutions included universities and colleges of technical and further education (TAFE); facilities for young children included kindergartens, childcare centres and maternal health centres; community facilities included community centres, libraries, neighbourhood houses, senior citizens centres and town halls; and emergency services included ambulance, fire and/or police stations.

4.4.16 Sociodemographic data

The SLA is defined under the Australian Standard Geographical Classification and corresponds roughly to Local Government Areas although some larger Local

Government Areas have more than one SLA (Australian Bureau of Statistics, 2007). The SLA is the smallest area for which sociodemographic and socioeconomic data are available from the ABS. Data for the smallest possible area were chosen to uncover any variability between study road segments that were located near each other.

Sociodemographic data for the SLA in which each road segment was located were obtained from the ABS for the year 2007. Data obtained were the population density (total population/square km), proportion of the population aged over 75, proportion of the population aged 15 to 24 (Australian Bureau of Statistics, 2008b), proportion of the population that were male, proportion of the population that were male and aged 15 to 24 (Australian Bureau of Statistics, 2008a). Table 4.24 presents the distribution of sociodemographic characteristics across the study road segments. Several of the study road segments formed the border between two different SLAs—for these segments, the average of the two SLAs was taken for all variables in Table 4.24.

Table 4.24 Median, interquarterile range (IQR) and range of sociodemographic characteristics across study road segments

Sociodemographic characteristic	Median (IQR)	range
Population density	2500.3 (1420.0)	141.7–6972.1
% male	49.0 (0.7)	47.4–52.0
% aged >75	7.3 (2)	1.5–10.7
% aged 15 to 24	13.9 (2.4)	10.4–42
% male and aged 15 to 24	6.9 (1.4)	5.0–20.1

Socioeconomic data for the SLA in which each road segment was located were also obtained from the ABS. Total vehicle ownership rates per 1,000 population for all vehicles, passenger vehicles and motorcycles were obtained for the year 2007 (Australian Bureau of Statistics, n.d.). Where the road segment formed the border between two SLAs, the average vehicle ownership was recorded. Socioeconomic Indexes for Areas rank geographic areas according to the characteristics of people, families and dwellings using data from the national census. The 2006 census Index of Relative Socioeconomic Advantage and Disadvantage (IRSAD) was calculated by the ABS using a range of variables reflecting advantage and disadvantage, e.g. the distribution of education level, employment, job categories, income, rent/mortgage payments, vehicle ownership and dwelling size. An increase in ranking reflects an increase in advantage and a decrease in disadvantage. The 2006 census IRSAD decile ranking within Australia for the SLA in which each road segment was located (Australian Bureau of Statistics, 2008c) was recorded. In cases where the road segment formed the border between two SLAs, the minimum IRSAD decile was recorded. It can be seen from Table 4.25 that the majority of strip shopping centre road segments in metropolitan Melbourne are located in areas ranked relatively highly in terms of advantage (and low in terms of disadvantage).

Table 4.25 Median, interquarterile range (IQR) and range of vehicle ownership and distribution of IRSAD deciles across study road segments

Vehicle ownership per 1,000 population	Median (IQR)	Range
Passenger vehicles	568.0 (71.0)	476.0–1142.0
Motorcycles	15.0 (4.5)	11.0–38.0
All vehicles	644.0 (102.0)	551.0–1397.0
SEIFA : IRSAD decile ranking within Australia (minimum)	Number of segments	% of segments
Decile rank 1	0	0
Decile rank 2	7	4.9
Decile rank 3	0	0
Decile rank 4	3	2.1
Decile rank 5	6	4.2
Decile rank 6	6	4.2
Decile rank 7	20	14.1
Decile rank 8	25	17.6
Decile rank 9	35	24.7
Decile rank 10	40	28.2

4.5 Data availability and implications for this thesis

In section 4.2.2, a list of aspects of the road and roadside environment with the potential to influence crash risk in any road environment (urban, rural, highway; midblock, intersection) were presented (Table 4.1). Sections 4.4.1 to 4.4.16 presented the data items collected for this study and the sources of data, or methods for collection. Table 4.26 provides a summary of the subset of those characteristics presented in Table 4.1 that were deemed relevant to urban midblocks and therefore considered for inclusion in this study. Also summarised in Table 4.26, for each data item, are the availability of data, the level at which the data item was measured (e.g. segment or area), data sources or data collection methods, and whether or not the data item was used in the final analysis. Data items for which no source was discovered and no data were collected are also distinguished.

Overall, most of the data required to characterise the road and roadside environment were available at the level of the road segment for the Melbourne metropolitan strip shopping centre midblocks included in this study. The majority of these data were specifically collected for this study using digital video images (held by

ARRB Group and owned by VicRoads) and the Melway street directory (2005, 2006, 2007, 2008, 2009), including data describing the road cross section, accesses and intersections, medians, road type, roadside parking, public transport facilities, bicycle facilities, pedestrian facilities, roadside development, height clearance, static enforcement cameras and amenities and facilities. Existing data sources held by VicRoads were found for AADT, percentage heavy vehicles, speed zone, and pavement condition. Liquor licensing data were also available from the State Government (Department of Justice).

Some data items were unavailable for the specific road segments of interest, but were available (or could be estimated) from existing sources for the area in which the road segment was located. Pedestrian activity in the local area was estimated from a State Government survey of travel activity (VISTA 2007). Unfortunately, recent research suggests that travel survey data are not useful for predicting the number of pedestrians crossing at the road segment level and that pedestrian crossing counts at the locations of interest are necessary (Alavi, 2013). This result had not been reported when data were being collected for this study and even if it had been, the current study did not have the funds available to undertake pedestrian counts at all of the study road segments. Hence, the VISTA travel survey data were the only data available for describing pedestrian activity in this study.

Sociodemographic and socioeconomic data for the local areas in which the road segments were located were available from the ABS. It is uncertain whether sociodemographic and socioeconomic data measured at the area level were representative of the users of the strip shopping centre arterial road segments in this study. Arterial roads function to transport people and goods and the people who live in the surrounding area may not represent the road users travelling through it. Thus it remains to be seen whether these macro-level characteristics will be useful in a study of crash risk at the road segment level.

There were some data items for which existing data sources could not be found and the only way to collect the data was to visit the sites, but funding for site visits was not available for this study. As mentioned in the previous paragraph, there were no available data for pedestrian volumes. The lack of pedestrian volume data has implications for this research, in particular, for the development of a model of pedestrian crash frequency. Unless the pedestrian exposure estimated using the VISTA travel survey data proves to be a reasonable surrogate for pedestrian volumes, this will mean that the pedestrian crash frequency model will not include pedestrian volume. The omission of such a (potentially) important predictor is likely to mean that other characteristics of the road and roadside that are correlated with pedestrian volume may be included in the models. For example, facilities and amenities that attract large numbers of pedestrians may appear to be related to pedestrian crash risk. It will not be possible to determine whether these factors are related to crash risk independently of pedestrian volumes, without having pedestrian volume data.

Neither bicycle volume data nor reasonable estimates of the distance travelled by bicycle in a local area were also unavailable therefore risk factors for BVC will not be investigated in this study. While it would have been desirable to have bicycle exposure or activity data and to investigate risk factors for BVC, its absence should not have a major effect on the other crash types being studied here (MVC, SVC, and PVC). In future, however, with health authorities encouraging people to choose more active transport options (e.g. walking and cycling), it would seem prudent for governments to launch an organised program to consistently collect data on vulnerable road user volumes in order to be able to evaluate the effect of travel mode choice on health outcomes, including road trauma.

Traffic exposure data were available as AADT. There were no traffic exposure data available at a finer level of resolution, e.g. hourly traffic volumes or peak-hour volumes. This is not a major issue for the current thesis because the analysis will focus on the number of crashes overall (adjusting for the number of days follow-up). This study may identify some risk factors, however, whose effect might be expected to be more prominent at certain times of the day. For example, the presence of schools would be expected to affect crash risk around school hours, whereas late night liquor licences would be expected to affect night-time crash risk. Future research to investigate such hypotheses would require traffic volume data at a more detailed level than simply AADT. It is likely that more detailed traffic volume data could be extracted from the current traffic measurement system in Victoria, however, this issue was not pursued with the road authority because it was beyond the scope of the current thesis.

Finally, there were no data available on vertical curvature or grade of the road segments. Previous multivariable studies of crashes in urban areas found that vertical curvature and grade were not significantly associated with run-off road SVC (Lee, 2000) or PVC (Alavi, 2013). Thus it would seem, from the limited evidence available, that the lack of data on vertical curvature and grade for the current study is not of great concern.

Table 4.26 Characteristics of the built environment with the potential to influence crash risk on urban midblock road segments that were and were not included in this research: data items, availability, measurement level and data sources/collection methods

Characteristic	Data item	Available	Measurement level	Existing data source	Data collection method	Included in analysis?
EXPOSURE						
Distance	Segment length (m)	✓	Segment	N/A	Interactive on-line maps; ARRB digital images	Yes
Traffic	AADT	✓	Segment or smaller	VicRoads	N/A	Yes
Traffic	Hourly traffic	✗		None found	Not collected	N/A
Traffic	Peak-hour traffic	✗		None found	Not collected	N/A
Pedestrian volumes	Pedestrian crossing volumes	✗		None found	Not collected	N/A
Pedestrian activity	Estimate of distance (km) walked per square km of SLA multiplied by segment length	✓	Area in which segment located	VISTA	N/A	Yes
Cyclist volumes	Daily cyclist volumes	✗		None found	Not collected	N/A
Cyclist activity	Distance cycled per square km of SLA	✗		None found	Not collected	N/A
Direction of travel	One-way/two-way	✓	Segment	None found	Melway street directory	No (all segments were two-way)
SPEED LIMIT						
Speed zone	Speed zone	✓	Segment	VicRoads RoadNet database	N/A	Yes
Variable speed zones	Presence of strip shopping centre variable speed zone	✓	Segment	VicRoads RoadNet database	N/A	Yes

Characteristic	Data item	Available	Measurement level	Existing data source	Data collection method	Included in analysis?
Variable speed zones	Presence of school variable speed zone	✓	Segment	VicRoads RoadNet database	N/A	No (collinear with variable describing presence of school)
Variable speed zones	Hours of operation	✓	Segment	VicRoads RoadNet database	N/A	No (all were active during business hours for that segment)
ROAD CROSS SECTION						
Horizontal curves	Presence of curves	✓	Segment	None found	Melway street directory	Yes
Horizontal curves	Type of curve	✓	Segment	None found	Melway street directory	No (sample size not large enough for each type)
Horizontal curves	Measures of curve angle, distance, etc.	×		None found	N/A	N/A
Vertical curves/grade	Any (presence, number, distance, % grade)	×		None found	N/A	N/A
Carriageway width	Mode, minimum, maximum	✓	Every 20m of segment	None found	ARRB digital images	Yes
Lanes	Number of through traffic lanes	✓	Every 20m of segment	None found	ARRB digital images	Yes
Lane width	Mode, minimum, maximum	✓	Every 20m of segment	None found	ARRB digital images	Yes
Shoulder	Type, width	✓	Every 20m of segment	None found	ARRB digital images	No (only one segment had a road shoulder)
Line marking	Presence, type	✓	Every 20m of segment	None found	ARRB digital images	No (collinear with median type)
Bridges	Presence of bridges	✓	Segment	None found	Melway street directory	No (none present)

Characteristic	Data item	Available	Measurement level	Existing data source	Data collection method	Included in analysis?
Special lanes	Transit lanes	✓	Every 20m of segment	None found	ARRB digital images	No (none present)
Speed management devices	Presence, type	✓	Segment	None found	Melway street directory	No (none present)
ACCESSES AND INTERSECTIONS						
Signalised intersections	Number, type, rate	✓	Every 20m of segment	None found	ARRB digital images; Melway street directory	Yes
Unsignalised intersections	Number, type, rate	✓	Every 20m of segment	None found	ARRB digital images; Melway street directory	Yes
Roundabouts	Number, type, rate	✓	Every 20m of segment	None found	ARRB digital images; Melway street directory	Yes
Driveways/Laneways	Number, rate	✓	Every 20m of segment	None found	ARRB digital images	Yes
Service roads	Presence, type, number of accesses	✓	Every 20m of segment	None found	ARRB digital images; Melway street directory	Yes
Dedicated turning lanes	Number, type, rate	✓	Every 20m of segment	None found	ARRB digital images	Yes
Keep clear zones	Number, rate	✓	Every 20m of segment	None found	ARRB digital images	Yes
MEDIANS						
Median	Type, width, proportion of segment with median	✓	Every 20m of segment	None found	ARRB digital images	Yes
Median accesses	Number, rate	✓	Every 20m of segment	None found	ARRB digital images	Yes

Characteristic	Data item	Available	Measurement level	Existing data source	Data collection method	Included in analysis?
ROAD TYPE AND TRAFFIC MIX						
Road type	Type of arterial	✓	Segment	None found	Melway street directory	Yes
Heavy vehicles	Proportion of traffic	✓	Segment or smaller	VicRoads	N/A	Yes
Heavy vehicle access	Yes/no	✓	Segment	VicRoads		Yes
Over-dimensional vehicle access	Yes/no	✓	Segment	VicRoads	Melway street directory	Yes
Intersect with over-dimensional vehicle route	Yes/no	✓	Segment	VicRoads	Melway street directory	Yes
ROADSIDE PARKING						
Parking	Presence, type, one or both sides	✓	Every 20m of segment	None found	ARRB digital images	Yes
Clearways	Presence, type	✓	Every 20m of segment	None found	ARRB digital images	Yes
Loading zones	Number, rate	✓	Every 20m of segment	None found	ARRB digital images	Yes
PUBLIC TRANSPORT AND BICYCLE FACILITIES (ON THE ROADWAY)						
Railway level crossings	Presence, type	✓	Every 20m of segment	None found	Melway street directory	Yes
Buses	Presence of route/s	✓	Every 20m of segment	None found	ARRB digital images; Melway street directory	Yes
Bus stops	Number, type, rate	✓	Every 20m of segment	None found	ARRB digital images	Yes
Bus lanes	Presence	✓	Every 20m of segment	None found	ARRB digital images	No (only two road segments had designated bus lanes)

Component 1: Identifying risk factors for crashes

Characteristic	Data item	Available	Measurement level	Existing data source	Data collection method	Included in analysis?
Trams	Presence of route/s	✓	Every 20m of segment	None found	ARRB digital images; Melway street directory	Yes
Tram stops	Number, type, rate	✓	Every 20m of segment	None found	ARRB digital images	Yes
Bicycle	Bicycle lane presence, width	✓	Every 20m of segment	None found	ARRB digital images	Yes
PEDESTRIAN FACILITIES						
Midblock pedestrian crossings	Number, type, rate	✓	Every 20m of segment	None found	ARRB digital images; Melway street directory	Yes
Pedestrian fencing/barriers at crossings	Presence	✓	Every 20m of segment	None found	ARRB digital images	Yes
ROADSIDE DEVELOPMENT						
Roadside poles/signs/trees	Number, type, rate	✓	Every 20m of segment	None found	ARRB digital images	Yes
Nature strip	Presence, one or both sides	✓	Every 20m of segment	None found	ARRB digital images	Yes
Development type	Retail on one or both sides of the road	✓	Every 20m of segment	None found	ARRB digital images	Yes
Development height	Predominant height, maximum height	✓	Every 20m of segment	None found	Google Street View	Yes
Offset	Between road and buildings, other objects	✓	Every 20m of segment	None found	ARRB digital images	Yes
PAVEMENT CONDITION						
Pavement roughness	Presence	✓	Segment	VicRoads RoadNet database	N/A	Yes

Component 1: Identifying risk factors for crashes

Characteristic	Data item	Available	Measurement level	Existing data source	Data collection method	Included in analysis?
Pavement distress	Presence	✓	Segment	VicRoads RoadNet database	N/A	Yes
HEIGHT CLEARANCE						
Low clearance	Presence, reason	✓	Segment	None found	Melway street directory	Yes
STATIC LAW ENFORCEMENT CAMERAS						
Speed/red light cameras	Presence, number	✓	Segment	None found	Melway street directory	Yes
AMENITIES AND FACILITIES						
Liquor licences	Number, type, rate	✓	Segment	Department of Justice	N/A	Yes
Off-street parking facilities	Number, rate	✓	Every 20m of segment	None found	ARRB digital images	Yes
Educational facilities	Number, type	✓	Segment	None found	Melway street directory	Yes
Facilities for young children	Number, type	✓	Segment	None found	Melway street directory	Yes
Healthcare/aged care facilities	Number, type	✓	Segment	None found	Melway street directory	Yes
Places of worship	Number	✓	Segment	None found	Melway street directory	Yes
Railway station	Presence	✓	Segment	None found	Melway street directory	Yes
Community facilities	Number, type	✓	Segment	None found	Melway street directory	Yes
Emergency services	Presence, type,	✓	Segment	None found	Melway street directory	Yes
Parks/ sports fields/ leisure centres	presence, type, number	✓	Segment	None found	Melway street directory	Yes

Characteristic	Data item	Available	Measurement level	Existing data source	Data collection method	Included in analysis?
Indoor shopping centres	Presence	✓	Segment	None found	Melway street directory	Yes
Petrol stations	Number	✓	Segment	None found	Melway street directory	Yes
SOCIODEMOGRAPHIC						
Population data	Population per square km	✓	Area	ABS	N/A	Initial stages (refer to Chapter 6)
Age distribution	% older adults, % young adults	✓	Area	ABS	N/A	Initial stages (refer to Chapter 6)
Sex distribution	% male	✓	Area	ABS	N/A	Initial stages (refer to Chapter 6)
Vehicle ownership	Vehicles owned per 1,000 population (all, passenger, motorcycle	✓	Area	ABS	N/A	Initial stages (refer to Chapter 6)
Socioeconomic ranking	IRSAD decile ranking	✓	Area	ABS	N/A	Initial stages (refer to Chapter 6)

CHAPTER 5. METHODS: DATA ANALYSIS

5.1 Outcomes and units of analysis

The outcomes of interest for each analysis was the number (count) of crashes that occurred during the five year period from 1st January to 31st December 2009 on strip shopping centre midblock road segments. Hence, road segments were the units of analysis. The crash types that were the focus for this thesis were the frequency of MVC, SVC and PVC on midblock road segments. Thus SVC were mutually exclusive from MVC and PVC, however, MVC and PVC were not mutually exclusive, as it was possible for a crash to involve multiple vehicles and at least one pedestrian.

For the 10 road segments on which the speed limit changed to a variable speed limit at some time during the study period, the time period for counting the frequency of crashes was reduced to account for changes to the speed limits. These changes occurred near the beginning (June 2005 or before; 8 sites) or the end (October 2008 or after; 2 sites) of the study period. The number of days over which the crash frequency was counted was included in the statistical models as part of the offset term. Hence, the outcome was modelled as a rate (crashes per day). The assumption that the relationship between the number of days and the number of crashes was directly proportional was tested by fitting a negative binomial regression model with the number of crashes as the outcome and ln(days) as a predictor and determining whether or not the coefficient for ln(days) was significantly different to one (that is, to test for evidence of non-linearity). For all three crash types, there was no evidence that the relationship between the number of days over which the crashes were counted and the number of crashes was non-linear, therefore the number of days was included in the offset term of the regression models.

5.2 Model development

The models for MVC, SVC and PVC were developed separately. Models for count data were used to establish the relationship between road segment characteristics and crash frequency. The suitability of these models for modelling the relationship between the built environment and traffic crashes was established in Chapter 3. In addition, the impact of changing the way that traffic volume was included in the models was also assessed.

Models were developed using negative binomial regression because of the tendency for crash frequency data to display overdispersion. Once the set of predictor variables was identified, if the likelihood ratio test revealed no evidence that the overdispersion parameter (α) was significantly different from zero, then the model was re-fitted using Poisson regression. The observed and predicted probabilities that the number of crashes was equal to x (where x= 0 to the maximum number of crashes observed) were plotted to inspect whether the model fitted the data, particularly whether the number of road segments with zero crashes was in excess of that predicted. If an excess number of zeros were observed, then the Vuong test was used to determine if a zero-inflated model was necessary (Hilbe, 2012). Tests of model specification, fit and other diagnostic tests were conducted on the resulting regression model.

5.2.1 Phased model building approach

Due to the limited number of units of analysis (142 midblock road segments) and the relatively large number of variables including some that were potentially correlated, the model was built using a phased approach. The phased process was developed independently for this research but shares many of the features of purposeful selection as described for logistic regression (Hosmer, Lemeshow, & Sturdivant, 2013). The phased process was used instead of standard stepwise regression because it allowed more control over the investigation of potential correlation and confounding. In addition, in practice, simple stepwise regression was unable to deal with the large number of variables at once and failed to converge to a solution. The main phases of the model building approach are described below.

5.2.1.1 Phase 1—categorisation of predictors

First, the measured characteristics of the built environment were characterised according to one of 15 higher-order classes. The classes were measurements or indicators of exposure, road cross-section, accesses and intersections, medians, road type and traffic mix, roadside parking, public transport and bicycle facilities, pedestrian facilities, roadside development, pavement condition, height clearance, presence of enforcement, amenities/facilities, speed zone and sociodemographic characteristics of the local area. Variables that described similar characteristics of the built environment were therefore classified as members of the same higher order class (refer to Table 5.1).

Table 5.1 Built environment characteristics categorised into higher order classes

Class	Variables
Exposure	Traffic volume (and derived variables), pedestrian activity, segment length
Road cross-section	Curvature, number of through lanes, carriageway width, lane width
Accesses and intersections	Signalised intersections, unsignalised intersections, roundabouts, driveways/laneways, dedicated turning lanes, service roads, keep clear zones
Medians	Presence, type, width, proportion of segment with a median, etc.
Road type and traffic mix	Road type, heavy vehicle access, percentage of heavy vehicle traffic
Roadside parking	Presence and type of roadside parking, presence of clearways, loading zones
Public transport and bicycle facilities	Presence of railway level crossings, presence and type of public transport, number and type of bus stops, tram stops, presence, type and width of bicycle lanes
Pedestrian facilities	Pedestrian crossings and fencing
Roadside development	Development height and type, nature strips, offset, roadside poles/trees, median poles/trees
Pavement condition	Distress, roughness
Height clearance	Low clearance due to bridges, tram wires
Enforcement	Presence of static red light/speed cameras
Amenities/facilities	Presence, type of amenities and facilities including establishments with liquor licences
Speed zone	Speed zone, presence and reason for variable speed limits
Sociodemographics of local area	Population density, age/sex distribution, vehicle ownership rates, IRSAD decile ranking

5.2.1.2 Phase 2—identification of best exposure measures

It is essential to account for exposure (time, segment length and traffic) when modelling crash frequencies. As described previously, the number of days over which the crashes were counted was included in the regression models as part of the offset term. The other main measures of exposure in this study were segment length and traffic exposure. The second phase of the modelling process involved investigating the relationship between traffic exposure variables, segment length and crash frequency to determine which form of each of the traffic exposure variables was most strongly related to the particular crash type. Traffic exposure was measured as traffic volume (AADT/1000), traffic density (AADT/1000 per lane) or a combined measure of traffic exposure and segment length (AADT/1000 x segment length or AADT/1000 per lane x segment length).

The objective of Phase 3 (the next phase in the model) was to fit two multivariable models for each crash type (MVC, SVC and PVC): one with the best combination of exposure variables including traffic volume and one with the best combination of exposure variables including traffic density. To determine the best combination of exposure variables including traffic volume or traffic density for each crash type, in Phase 2 a series of six negative binomial regression models were fitted with the possible combinations of AADT and segment length and the model with the lowest Akaike information criterion (AIC) value informed how traffic volume and length were to be included in Phase 3. The six possible combinations were:

- AADT/1000 and segment length
- AADT/1000 and natural log of segment length
- Natural log of AADT/1000 and segment length
- Natural log of AADT/1000 and natural log of segment length
- Combined AADT/1000 and segment length (AADT/1000 x segment length)
- Natural log of combined AADT/1000 and segment length (ln(AADT/1000 x segment length))

The same process was conducted to choose the optimal combination of variables for the model including traffic density. Six models were fitted with the possible combinations of AADT/1000 per lane and segment length and the model with the lowest AIC was chosen to inform the inclusion of traffic density and length in Phase 3. The six models were analogous to those listed above for AADT, except that AADT/1000 per lane was used as the traffic exposure variable.

If the best fitting model had the natural log of an exposure variable included, the coefficients and 95% CI were inspected to determine if the coefficient was significantly different to one. If the coefficient was not significantly different to one, then that exposure variable was included as part of the offset term in subsequent modelling phases. If the coefficients were significantly different to one, then the variables were retained as part of the linear predictor.

Finally, if any of the exposure variables (length, traffic volume or traffic density) or their transformations were not significantly associated with a crash type, then that variable was not included in the subsequent modelling phases until it was re-assessed during post-hoc testing (as described in the next section).

5.2.1.3 Phase 3—Relationship between the built environment and crash frequency

In Phase 3, up to two models were developed for each crash type to assess the relationship between the built environment and crash frequency: each using one of the aforementioned methods of controlling for traffic exposure (traffic volume or traffic density). This allowed investigation of whether the way traffic exposure is modelled affects the aspects of the built environment that are found to be associated with crashes. That is, do the characteristics of the road segment that are associated with crash frequency differ according to whether traffic volume or traffic density are used as exposure measures? In addition, the number of signalised intersections per km was adjusted for in all models, to account for the exclusion of signalised intersection crashes from the crash count. This was necessary because as the number of signalised intersections per km increased we would expect the number of midblock crashes to decrease, simply because a smaller proportion of the strip shopping centre road segment was classified as midblock.

First, the exposure variable was chosen (e.g. traffic volume, traffic density or the combined traffic exposure-segment length measure; see section 5.2.1.2). Next, a separate multi-variable model was developed for each higher order class of variables (see Table 5.1), adjusting for the rate of signalised intersections, segment length and traffic exposure (where appropriate) to identify those variables in each category that were associated with crash frequency. Variables that described similar characteristics were never included in the same model, for example the mode of carriageway width was never included with maximum or minimum carriageway width. If separate models showed that variables describing a similar characteristic were both associated with crash frequency, then the variable that led to the model with the best fit (judged using the AIC value) was identified as the best predictor. The p-value for identifying potential risk factors was set at $p<0.10$. This process resulted in a list of variables in each category that were significantly ($p<0.10$) associated with crash frequency and served to narrow down the number of variables for inclusion in a full multivariable model.

Once the variables in each category that were independently associated with crash frequency were identified ($p<0.10$), all of the category-specific significant predictors were entered into an all-category backwards stepwise regression model to determine which were independently associated with crash frequency when adjusting for other variables from all categories. Variables were retained if they were associated with crash frequency at a p-value of less than 0.1. The one exception was the number of signalised intersections per km, which was retained in all the models even if it was not a significant predictor, to adjust for the crash-selection process.

Once the significant predictors from all categories were identified, a post-hoc test process was conducted to determine that all important predictors were included in the multivariable model. This involved adding each of the other variables (even those that were not associated with crash frequency in the category-specific models) to the multivariable model, one by one, to test if they were significantly related to crash occurrence and if they improved model fit (indicated by a lower value of AIC). If so, then they were retained in the model, if not, they were removed. This process continued iteratively until all variables had been assessed to see if they should be included in the multivariable model. The resulting multivariable model identified the significant main effects, that is, those variables that were independently associated with crash occurrence. At this final stage, the p-value for inclusion in the model was reduced to the standard level for statistical significance of $p<0.05$.

During this process it became apparent that the models were less stable when sociodemographic predictors were included, particularly the variables describing socioeconomic advantage and disadvantage for the local area (IRSAD decile ranking). When these variables were included, substantially more iterations were required to develop the model and in some cases, a stable model was not found. Other variables dropped out and re-entered multiple times during the iterative process (hence the instability). If sociodemographic variables were excluded, however, the models converged towards a solution in fewer iterations. Further investigation revealed that although the correlations were not strong in a statistical sense, several of the other variables were associated with the socioeconomic ranking of the local area. For example, railway level crossings were more prevalent in road segments located in areas with the lowest observed decile of socioeconomic disadvantage (29% had level crossings) compared to road segments that were located in other areas (7% had level crossings). This is not surprising as the removal of railway level crossings across metropolitan Melbourne through grade separation was sometimes politically motivated. It is therefore possible that road infrastructure is related to socioeconomic characteristics of the area. Moreover, the characteristics of road segments in the inner suburbs are different to those in the outer suburbs and socioeconomic indices also vary along these grounds. The sociodemographic variables were also potentially problematic because they described the greater area in which the road segment was located, and were not specific to the road segment itself. Furthermore, the road users of a particular road segment do not necessarily live in the local area, particularly drivers on arterial roads. As such, it was not certain that the sociodemographic variables were appropriate to describe the road users of a particular road segment. It was therefore decided that because the long-term purpose of this research was to identify modifiable risk factors for crash occurrence, that sociodemographic variables would be excluded from model development.

5.2.2 Model diagnostics

While diagnostic tests for multiple linear regression are described in many text-books, it is surprisingly difficult to find information about diagnostic tests that are applicable to models for count data which may explain their omission from previous studies. The text by Cameron and Trivedi (2013)and the Statalist[3] forum for Stata users were valuable sources of information about diagnostic testing for models for count data.

Diagnostic testing was conducted to assess model specification, fit and assumptions once the final all-category main effects model was developed. First, the decision of whether negative binomial regression or Poisson regression should be used to fit the final model was determined using a likelihood ratio test of whether the overdispersion parameter (α) was significantly different from zero. If the test was significant ($p<0.05$) then the negative binomial regression model was favoured, whereas if the test was not significant, then the model was re-fitted using Poisson regression. The deviance statistic divided by the degrees of freedom, and the Pearson

[3] http://www.statalist.org/

statistic divided by the degrees of freedom also provided information about the degree of dispersion (ideally, these should be equal to one).

5.2.2.1 Model specification and goodness of fit

The link test, goodness of fit tests and plots of the observed against predicted probabilities that the number of crashes were equal to x, were used to assess model specification and fit.

The link test is a simple procedure to ascertain if there is evidence that the link function of the GLM is specified appropriately (Pregibon, 1979 cited in StataCorp, 2013). The rationale behind the link test is that if the link function and the linear predictor are specified correctly then no other variable should be significantly related to the outcome, unless by chance. The link test conducts a regression of the number of crashes on the predicted values from the regression model (hat) and on the square of the predicted values (hat^2). If the appropriate link function and linear predictor are specified, then the predicted values should be significantly associated with the outcome, but the squared predicted values should not be (that is, hat^2 should not have any explanatory power). A significant link test (where hat^2 is significantly associated with the outcome) indicates that the link function is inappropriate, or that the linear predictor is misspecified (assuming that the link function is appropriate).

To determine if the model form fitted the data well, two goodness of fit $\chi 2$ tests were conducted; one using the Pearson residuals and the other using the deviance residuals. These tests compare the observed and predicted values, and if the test is statistically significant, indicate that the model form does not fit the data well. These tests could only be conducted after Poisson regression (not negative binomial regression). To further investigate the fit of the model, the observed and predicted probabilities were also plotted against the observed values of the number of crashes.

5.2.2.2 Linearity

The optimal form of the relationship between the continuous variables in the model and the crash frequency was assessed. To investigate the relationship between a continuous predictor and the outcome in the multivariable model, a number of methods were considered. One method involved categorising the continuous variables into quartiles, fitting the regression model with the categorical variable instead of the continuous variable, plotting the coefficients and inspecting for linearity. However, this method is only sensitive to gross departures from linearity and will not detect, for example, departures within a quartile.

A more sensitive method is the Box and Tidwell (Royston & Ambler, 1999) power transformation (or exponential transformation) regression model to estimate the optimum value for the power to which a continuous variable should be raised to achieve linearity of the relationship with the outcome. To test whether the data are significantly non-linear, a test is conducted to determine if the deviance difference between a straight line and the Box Tidwell model is significant. If the test is not statistically significant, then there is no evidence to support the hypothesis that the relationship between the predictor and the outcome is non-linear. If the test is significant, then transforming the variable by raising it to the optimum power should lead to a linear relationship between the predictor and the outcome in that multivariable model.

The Box and Tidwell power and exponential transformation regression models were used to investigate if there was evidence that the relationship between the continuous predictor variables and the outcome was non-linear. If there was evidence for non-linearity, then the continuous variable was transformed and the regression was refitted with the transformed variable. To determine whether or not the increase in complexity with the transformed variable was warranted, the fit of the two models was compared using the AIC. In addition, the relationship between the predictor variable and the outcome was plotted for a representative range of the predictor variable, holding other variables in the model constant at their mean (continuous variables) or reference (categorical variables) values, for both the transformed and untransformed variable. The plots were inspected visually to assess the differences between the plots, and consideration was given to other methods of modelling the continuous variable that would lead to a similar shaped relationship as that of the optimal transformation. The final decision as to which version of the continuous variable should be included in the final model was made according to considerations of model fit and practical interpretation of the coefficients.

5.2.2.3 Multicollinearity

Multicollinearity is defined as when the predictor variables in the regression model are correlated. If the objective of developing a regression model is purely for predictive purposes, then multicollinearity between predictor variables is not an issue for concern. However, if the objective of the model is to estimate and interpret the association between an outcome and a set of risk factors, then multicollinearity between risk factors will impact on the ability to distinguish and interpret the independent effect of each risk factor on the outcome.

The bivariate correlation between variables can be assessed using Pearson's correlation coefficient for two continuous variables; chi-square tests or Fisher exact tests for two categorical variables; and t-tests or analyses of variance for continuous and categorical variables. In multivariable regression, however, the bivariate correlations between variables do not tell the whole story. Instead, how each variable is related to all others in the model is of interest.

The variance inflation factor (VIF) is a measure of multicollinearity that can be derived for each coefficient in the model (Belsley, Kuh, & Welsch, 1980). Separate multiple linear regressions are conducted using the variables in the model (where the number of variables in the model=K) where each variable is, in turn, considered the outcome, and the others are considered the predictors. Thus K separate regressions will be fitted. The equation for the VIF of the jth variable is:

$$VIF_j = \frac{1}{(1 - R_j^2)}$$

where R^2_j is the coefficient of determination for the model with *j* as the outcome and the other independent variables as the predictors. The VIF indicates the effect that multicollinearity has on the estimated variance of the coefficient for a particular variable. A VIF greater than 10 is generally considered evidence of problems with multicollinearity. The argument for using the VIF (derived from multiple linear regression) to diagnose collinearity for continuous predictors is sound. Linear regression, however, is usually inappropriate for other types of outcome data, such as categorical variables. In this context, however, it is appropriate to use VIF to diagnose issues involving the linear association between a categorical outcome and a set of predictors, because that is the actual linear association in the model for which multicollinearity needs to be identified.

The VIF therefore enables determination of whether there are multicollinearity problems for a particular variable, or variables, but does not inform which variables contribute to the multicollinearity or whether the overall model suffers from multicollinearity. For this purpose, the condition number can be calculated. First, a principal components analysis is performed which entails transformation of the K predictor variables so as to create a set of K variables that are a linear combination of the K predictors which are uncorrelated and have maximum variance. Eigenvalues (λ) are calculated as the variances of the principal components. The condition index for a principal component (*j*) is equal to the square root of the largest eigenvalue divided by the eigenvalue for that principal component.

$$CI_j = \sqrt{\frac{\lambda_{max}}{\lambda_j}}$$

The condition number (CN) for a model equals the square root of the largest eigenvalue divided by the smallest eigenvalue.

$$CN = \sqrt{\frac{\lambda_{1max}}{\lambda_{min}}}$$

A large CN indicates potential issues with multicollinearity yet the threshold for defining a large CN is debated. Belsley et al. (1980) suggest that CNs between five and ten indicate weak correlation while CNs between 30 and 100 suggest moderate to strong correlation. For the purposes of this thesis, a CN of greater than 30 will be considered as indicative of potential problems with multicollinearity. Along with the CN, the variance decomposition proportions provide a tool to diagnose multicollinearity and identify the variables responsible. The variance decomposition proportions indicate the proportion of the variance of each predictor's regression coefficient that is associated with each principal component. If two or more predictors load highly (that is, have a variance proportion greater than 0.5) on a principal component with a high condition index, this indicates multicollinearity between variables.

In this study, bivariate correlations between variables were assessed prior to model development mainly to identify which variables were essentially measuring the same characteristic of the built environment so as not to include them in the same models during the variable selection phase of model development (Phase 3). After the models were developed, multicollinearity was assessed using the VIF and the CN along with the variance decomposition proportions.

5.2.2.4 Residual plots

The values predicted from the regression models were compared to the observed values using residuals. The simplest residual to conceptualise (and calculate) is the raw residual which equals the difference between the fitted value and the observed value. The raw residual is the most commonly used residual for diagnostics in linear regression, in which situation the raw residual has the desirable properties of a mean of zero, constant variance, and a symmetrical distribution. Unfortunately, the raw residual does not have these properties for count data: the variance is not constant and the distribution is not symmetrical (Cameron & Trivedi, 2013).

There are several residuals that have more desirable properties than the raw residual for count data, although none have a mean of zero, constant variance and a symmetrical distribution. The raw, Pearson, standardised Pearson, deviance, standardised deviance, adjusted deviance and Anscombe residuals were calculated (Cameron & Trivedi, 2013) and their distributions inspected to find the residual with the least skew and kurtosis. This residual was subsequently used for diagnostic testing of the model.

A smoothed scatterplot of the residual against the predicted means was inspected to determine if the residuals clustered around zero and had constant variance. This plot also indicates whether the fit is poor for any of the predicted values. If the model is adequate then the residuals should also be normally distributed. This was assessed by plotting the quantiles of the chosen residual against the quantiles of the normal distribution.

Residuals were also used to identify observations with the potential to disproportionately influence the regression estimates. Observations with high leverage were identified as those with a diagonal entry in the hat matrix of more than three times the average (where the average value is the number of parameters in the model divided by the number of observations). The influence of these observations on the regression estimates was determined by fitting the regression without the observations with high leverage and comparing the coefficients to those obtained with the full sample. Coefficients were inspected to determine if the associations remained statistically significant and whether the strength of the association as indicated by the parameter estimate changed markedly (e.g. by more than 10%). If so, this was considered evidence that the association between those variables and the outcome were largely a reflection of the observations that were excluded and hence might not be generalisable to other samples.

5.2.2.5 Overfitting

An overfitted model is too complex and dependent on the data with which it is developed and hence might not be able to be generalised to other samples. That is, an overfitted model might not be useful for predicting crashes on other samples of urban roads. A limitation of this study is the relatively few units of analysis in comparison to the number of risk factors (or predictors) being assessed which leads to the potential to overfit the model. One way to avoid overfitting is to split the dataset into two, fit the model using half of the data, and test the predictive validity on the other half of the data. There were too few units of analysis available in this study to take this approach.

Bootstrap resampling is another way to determine if there is evidence that the fitted model may not be generalisable to other samples (Babyak, 2004). It is used to estimate the standard error, 95% CI and p-values of the coefficients by taking repeated samples of size n (in this case, n=142) and fitting the regression model a large number of times. To take repeated samples of size n from n road segments, sampling with replacement is employed. That is, a unit of analysis is randomly selected to be included in the model, and then replaced in the pool of units for selection before the next unit of analysis is randomly selected. This process is repeated until 142 road segments have been selected. The regression model is then fitted a large number of times; each time a new subset is sampled with replacement. The final estimate of the standard errors is derived through the variability in the estimates obtained during the bootstrapping process.

The number of bootstrap replications was chosen such that stable estimates of the standard errors were obtained. This was investigated by running the bootstrap resampling again with the same number of repetitions and determining if the standard error estimates were similar to those obtained the first time. If they were the same to approximately three decimal points, then the result was deemed stable. If they differed, then the number of replications was increased.

The bootstrapped p-values were inspected to see if any of the variables that were significantly associated with the outcome in the model were no longer significantly associated with the outcome once the bootstrapped standard error estimates were used. If so, this was taken as evidence that the association was not robust and may not generalise to other samples.

5.2.2.6 Interactions

With so many potential risk factors and a relatively small sample size it was not considered prudent to exhaustively investigate whether variables had an interactive effect on crash frequency. Interactions were only investigated if the link test was significant, which can indicate that the linear predictor is misspecified and that relevant variables (or interactions) may be excluded from the model.

5.3 Summary of data analysis

A phased data analysis approach was developed to estimate the association between characteristics of the built environment and crash frequency, which took into account the large number of potential risk factors and the potential for correlation between them. This enabled the investigation of a broader range and larger number of risk factors than previous research in the area. Another major limitation of previous research was the lack of diagnostic testing of model specification and fit which made it difficult to assess the validity of the results. A range of diagnostic tests were conducted on the models developed for each crash type. The statistical analysis software package Stata (Version 11) (StataCorp, 2009) was chosen for conducting the analyses because of its capability to fit regression models for count data and a broad range of diagnostic tests.

CHAPTER 6. RESULTS: IDENTIFICATION OF RISK FACTORS

6.1 Descriptive statistics

The frequency of crashes that occurred in the five-year period from 2005 to 2009 on each of the 142 Melbourne metropolitan strip shopping road segments was obtained from the police-reported crash database for Victoria, Australia (refer to Table 6.1).

In total, 1,954 police-reported casualty crashes occurred during the five-year period of interest. The number (and rate) of crashes on each segment ranged from zero to 180 (0 to 65.1 crashes per km per 5 years). Almost two-thirds were MVC (crashes involving two or more vehicles). More than a quarter of the crashes involved at least one pedestrian (PVC). Three (2.1%) of the 142 road segments had no crashes at all, six (4.2%) had no MVC, 62 (43.7%) had no SVC and 26 (18.3%) had no PVC.

Table 6.1 Frequency and rate of crashes on Melbourne metropolitan strip shopping segments, 2005 to 2009, by crash type

Crash type	Crash frequency				Crashes/km*	
	Total crashes across sites (% of total)	Range	Median	Interquartile range (IQR)	Range	Median (IQR)
Multi-vehicle	1,291 (66.1%)	0 to 125	5	8	0 to 54.5	12.9 (12.1)
Single-vehicle	170 (8.7%)	0 to 16	1	2	0 to 11.0	1.4 (3.8)
Pedestrian-vehicle	519 (26.6%)	0 to 43	3	4	0 to 37.4	5.4 (5.7)
Total crashes	1,954 (100%)	0 to 180	9.5	12	0 to 65.1	20.8 (15.2)

* adjusted for different follow-up time across segments.

Models were developed using the phased approach described in Chapter 5. The outcomes of the modelling process are described separately for MVC, SVC and PVC in the next sections. Briefly, up to two models were developed for each crash type. The two models differed according to how the two main exposure variables (traffic exposure and segment length) were included in the models. First, a model was developed using the best combination of traffic volume and segment length, as judged by the lowest AIC value (but only if traffic volume was significantly associated with crash frequency). Second, a model was developed using the best combinations of traffic density and segment length as judged by the lowest AIC value (again, only if traffic density was significantly associated with crash frequency). If neither traffic volume nor density were significantly associated with crash frequency, the model was developed with segment length as the only exposure variable. Where more than one model was developed for a particular crash type, resulting models were compared to determine how changing the way traffic exposure was included in the models affected the selection of variables, and which model fitted the data better.

6.2 Multi-vehicle crashes

6.2.1 Selecting the optimal method for specifying exposure

A series of negative binomial regression models were fitted to determine the optimal way to specify segment length and traffic exposure for MVC. An offset term was included to account for the different number of days for the collection of crash data across road segments. The number of signalised intersections per km was also included as a predictor in the model, to adjust for the exclusion of crashes at signalised intersections.

Table 6.2 displays the different specifications of traffic volume and length, and the AIC value for each MVC model. The best model fit (lowest AIC value) resulted from the model that included the natural log of traffic volume and the natural log of segment length. This model showed that the natural log of the traffic volume was significantly related to MVC (β=0.41, 95%CI 0.19–0.63). The 95% CI for the coefficient did not include one, which indicated that the relationship between traffic volume and MVC was not directly proportional so the traffic volume variable was retained in the linear predictor. The natural log of segment length was also significant related to MVC (β=1.06, 95%CI 0.93–1.22) and the coefficient was not significantly different to one. Hence, segment length was included as part of the offset term during subsequent development of MVC model 1, using the natural log of traffic volume as the traffic exposure variable.

Table 6.2 Comparison of different ways to specify traffic volume and segment length for modelling multi-vehicle crashes

Traffic volume	Segment length	AIC
AADT/1000	length	5.667
AADT/1000	ln(length)	5.541
ln(AADT/1000)	length	5.636
ln(AADT/1000)	ln(length)	5.508 (Lowest)
thousand vehicle km per day		N/A*
ln(thousand vehicle km per day)		N/A*

* Coefficients for traffic volume and length were significantly different when entered as separate variables, so there was no justification for pooling as one combined measure of exposure.

Table 6.3 shows the different specifications of traffic density and segment length. The best model fit (lowest AIC) was achieved with the natural log of the number of vehicles per lane and the natural log of segment length. The natural log of the number of vehicles per lane was significantly related to MVC (β=0.48, 95%CI 0.19–0.78) and the relationship between traffic density and MVC was not directly proportional, as the 95% CI did not include one. Therefore the traffic density variable was retained in the linear predictor. The natural log of segment length was also significant related to MVC (β=1.10, 95%CI 0.95–1.25). Because the 95% CI for this coefficient included one, the segment length was included as part of the offset term when developing MVC model 2, using the natural log of vehicles per lane as the traffic exposure variable.

Table 6.3 Comparison of different ways to specify traffic density and segment length for modelling multi-vehicle crashes

Traffic density	Segment length	AIC
AADT per lane/1000	length	5.685
AADT per lane/1000	ln(length)	5.541
ln(AADT per lane/1000)	length	5.671
ln(AADT per lane/1000)	ln(length)	5.527 (Lowest)
thousand vehicle km per lane per day		N/A*
ln(thousand vehicle km per lane per day)		N/A*

* Coefficients for traffic density and length were significantly different when entered as separate variables, so there was no justification for pooling as one combined measure of exposure.

6.2.2 *Model development and diagnostic tests*

6.2.2.1 Model development

The models developed using the natural log of traffic volume (MVC model 1) and the natural log of traffic density (MVC model 2) were very similar. During model development when the p-value for inclusion in the model was set at $p<0.10$, the traffic density model and the traffic volume model differed by only one variable (apart from the traffic exposure variables). In the model with the traffic density as the traffic exposure variable, the number of lanes was significantly associated with MVC frequency, however, the number of lanes was not significantly associated with MVC frequency when traffic volume was included as the traffic exposure variable. This is not surprising, given that the number of vehicles per lane is equal to the number of vehicles divided by the number of lanes. This indicates that including the traffic volume in the model for MVC frequency is roughly equivalent to including the number of vehicles per lane and the number of lanes.

For both models, the variables that were most strongly associated with MVC frequency (as assessed by the size of the p-value) entered the model either during the initial stage of model development (when the variables that were significantly associated with MVC frequency in the category-specific models were considered together in an all-category model the first time), or during the first iteration of post-hoc testing to see if any other variables were significantly associated with MVC frequency once all categories were considered together. At the end of the first iteration of post-hoc testing, all variables except for one (the last one to enter the model) were associated with MVC frequency at the $p<0.01$ level. At this point in model development, MVC Model 1 (traffic volume) included 14 predictor variables with $p<0.01$, and MVC Model 2 (traffic density) included 15 predictor variables with $p<0.01$. The overdispersion parameter was significantly different to zero, which indicates that negative binomial regression was preferred over Poisson regression. Further iterations of post-hoc testing discovered a further eight variables that were associated with MVC frequency at the $p<0.10$ level, (almost all at the $p<0.05$ level). Once these extra variables entered the model, the overdispersion parameter was no longer significantly different from zero, so Poisson regression was more appropriate.

The fact that the strongest associations were discovered in the earlier stages of model fitting is worth noting because of the potential for overfitting the model through excessive post-hoc testing. With 142 strip shopping road segments (or units of analysis) in the study, it was judged there were sufficient data to support the models with 14 or 15 variables, but the inclusion of an extra eight variables was likely asking too much of the data available. Consequently, the models that were subjected to tests of fit and model diagnostics were those where the p-value for inclusion was set at $p<0.01$. The implications of including the extra variables are assessed when investigating the potential for overfitting (Section 6.2.2.6).

6.2.2.2 Model fit

Table 6.4 displays several measures of model fit for MVC Model 1 and MVC Model 2 (p-value for inclusion set at $p<0.01$). For both models, the overdispersion parameter (α) was significantly different from zero, which indicates that negative binomial regression was more appropriate than Poisson regression for the data. The log likelihood was closer to zero for MVC Model 2 than MVC Model 1, which indicates better fit. MVC Model 2, however, had two more parameters than MVC Model 1 (which were the coefficients for the categorical variable representing the number of lanes). The AIC and Bayesian information criterion (BIC) values carry a penalty for an increased number of parameters and therefore are biased towards more parsimonious models. The AIC and BIC were lower for MVC Model 1 than MVC Model 2, so MVC Model 1 was preferred as the more parsimonious model. The deviance statistic divided by the number of degrees of freedom was slightly closer to one for MVC Model 1, as was the Pearson statistic divided by the degrees of freedom (that is, MVC Model 1 was less overdispersed than MVC Model 2).

Table 6.4 Fit statistics for models predicting multi-vehicle crash (MVC) frequency

Fit statistic	MVC Model 1: Traffic volume	MVC Model 2: Traffic density
α (over dispersion)	0.0317021	0.0289242
Likelihood ratio test of α=0	$\chi^2(1)$=7.17, p=0.004	$\chi^2(1)$=6.25, p=0.006
Number of parameters	21 (including constant & α)	23 (including constant & α)
Log Likelihood	-339.089	-338.636
AIC	5.072	5.093
BIC	78.523	87.528
Deviance/df	1.19	1.22
Pearson/df	1.09	1.12

Model fit was inspected visually by plotting the observed probabilities that the number of MVC was equal to x (where x=0 to 99) from the raw data, and the probabilities predicted from the negative binomial regression models (refer to Figure 6.1 for MVC Model 1 and Figure 6.2 for MVC Model 2). Figure 6.1and Figure 6.2 reveal that both models were a reasonable fit to the data, despite some noise. Of particular interest was that both models predicted the number of road segments with zero crashes well, therefore, the use of zero-inflated models was not necessary for these data.

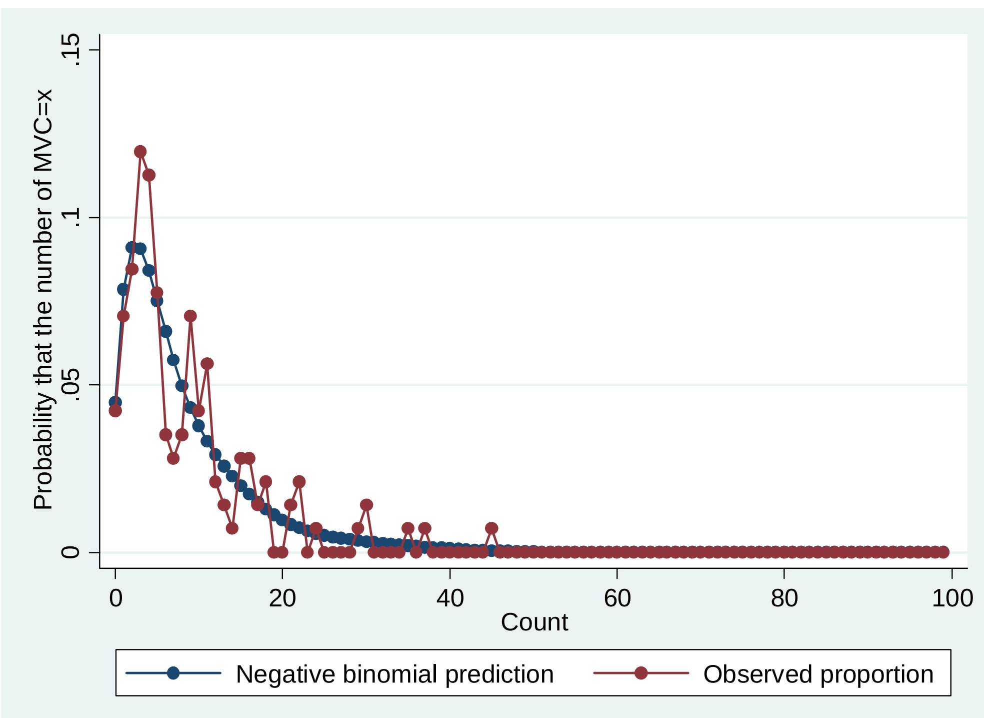


Figure 6.1 Probability that the number of multi-vehicle crashes (MVC)=x: observed proportion and predicted proportion from the MVC model with traffic volume

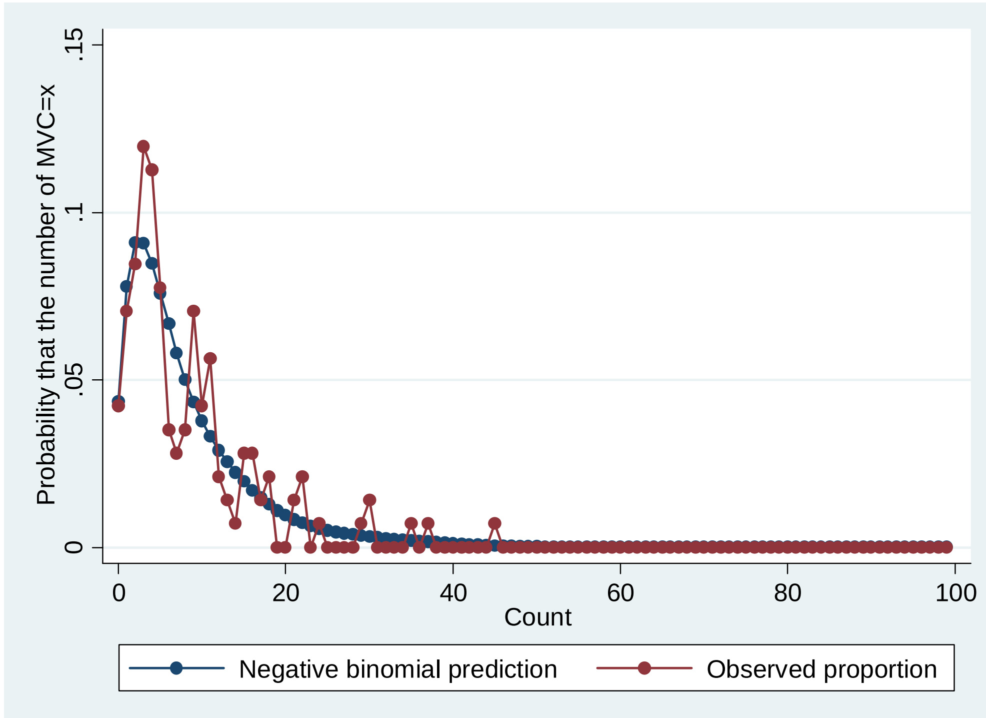


Figure 6.2 Probability that the number of multi-vehicle crashes (MVC)=x: observed proportion and predicted proportion from the MVC model with traffic density

There was no evidence for misspecification of the link function or that any relevant variables were excluded in either model (p-values for link test: MVC Model 1, p=0.116; MVC Model 2, p=0.078). As such, the investigation of interactions between variables was not warranted.

6.2.2.3 Tests of linearity

The Box Tidwell power and exponential transformation regression models showed no evidence that the relationship between the continuous variables and MVC frequency was non-linear in either model (all p-values>0.12).

6.2.2.4 Tests for multicollinearity

There was no evidence that the predictors were multicollinear. The VIFs ranged from 1.10 to 2.14 for MVC Model 1 and 1.12 to 2.18 for MVC Model 2; values well below the level for concern (VIF=10). The CN was 39.42 for MVC Model 1 and 30.21 for MVC Model 2. The CNs were above 30, which is the threshold for identifying potential problems with multicollinearity. For both models, there was one principal component where more than one variable loaded highly. In both MVC Model 1 and MVC Model 2, the constant and the traffic exposure variable loaded highly on the last principal component. Given that the constant is essentially a nuisance variable and that two predictor variables did not load highly on any principal component, multicollinearity was judged not to be of concern.

6.2.2.5 Analysis of residuals

For both models, the Anscombe residual was chosen for diagnostic testing of the model because it displayed less skew and kurtosis than the other residuals. Plots of quantiles of the Anscombe residuals for MVC Model 1 (Figure 6.3) and MVC Model 2 (Figure 6.4) against quantiles of the normal distribution showed no obvious departures from normality.

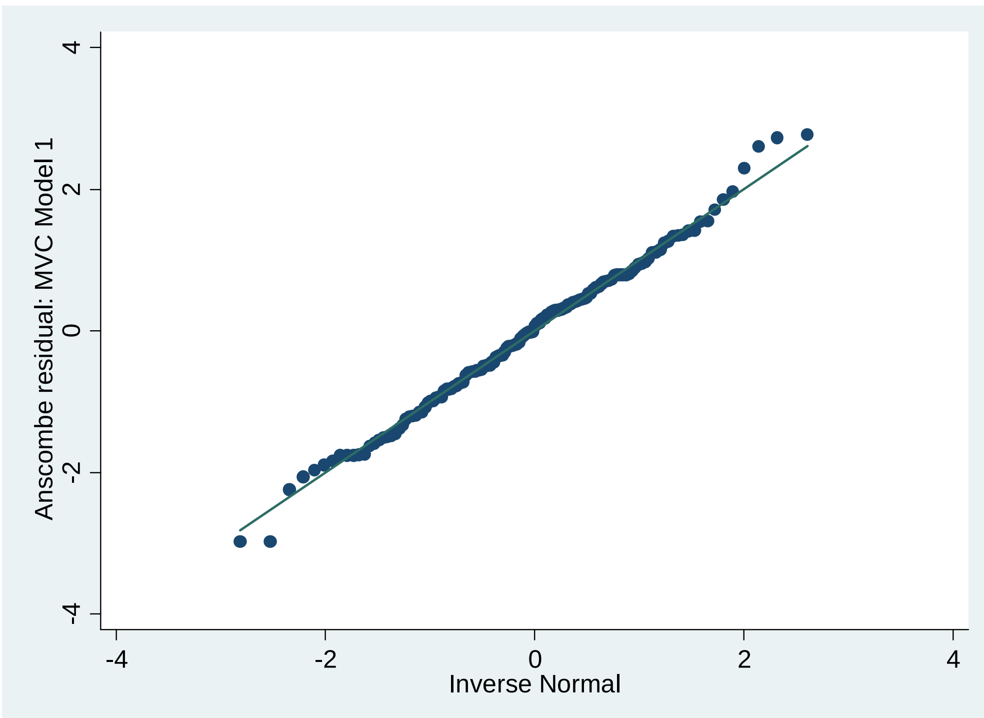


Figure 6.3 Quantiles of the Anscombe residual plotted against quantiles of the normal distribution for the multi-vehicle crash model with traffic volume

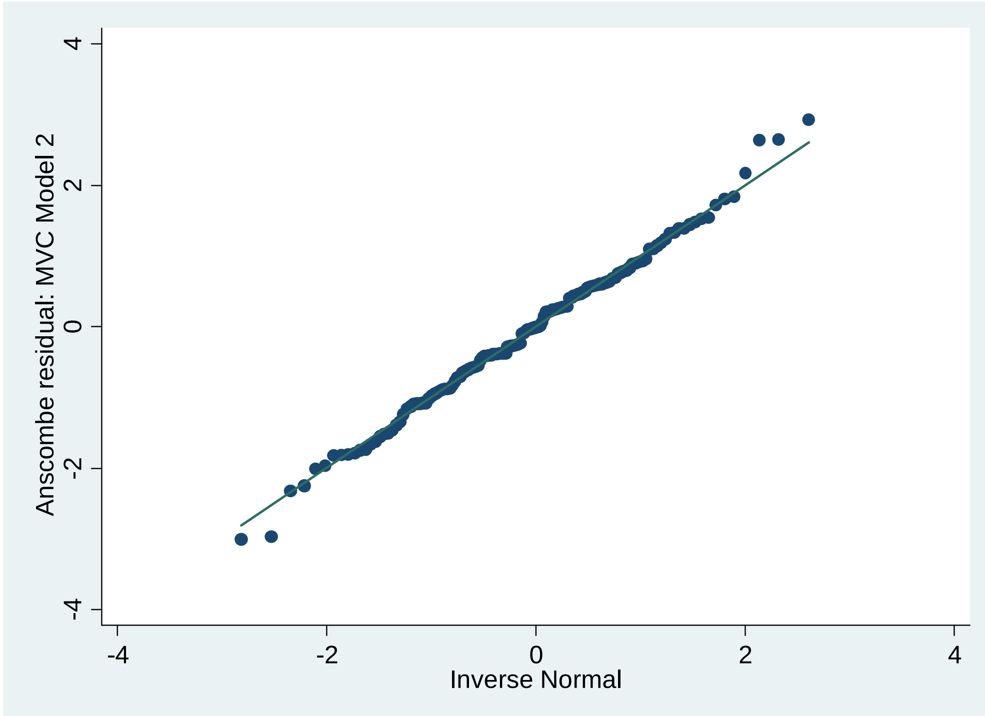


Figure 6.4 Quantiles of the Anscombe residual plotted against quantiles of the normal distribution for the multi-vehicle crash model with traffic density

Figure 6.5 depicts scatter plots of the Anscombe residual against the predicted means from MVC Model 1 (left panel) and MVC Model 2 (right panel), with locally weighted scatterplot smoothing. In general, the Anscombe residuals clustered around zero and the lowess line was relatively flat, apart from one road segment (which had the maximum observed number of MVC) that did have an influence on the line at high values of MVC frequency. Overall, this indicated there was no strong evidence of a relationship between the residuals and the predicted values.

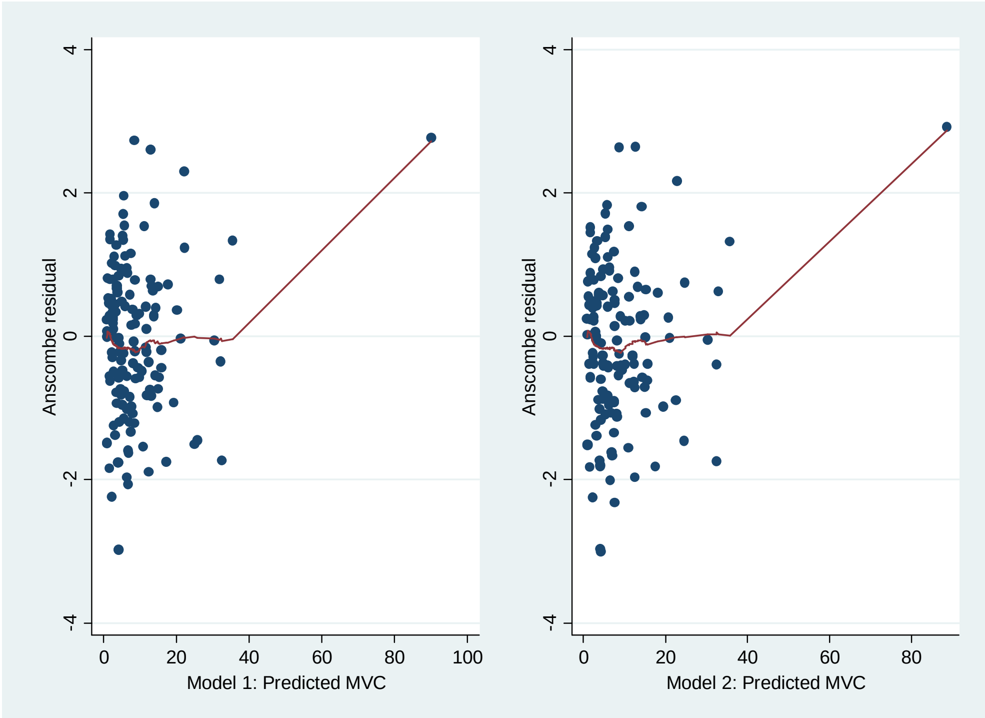


Figure 6.5 Locally weighted smoothed scatterplot of the Anscombe residual against the predicted mean number of multi-vehicle crashes (MVC) for MVC Model 1 (with traffic volume) and MVC Model 2 (with traffic density)

There were four road segments with potentially high leverage in MVC Model 1 and three of the same road segments also displayed potentially high leverage in MVC Model 2. Removing these potentially influential segments had a negligible impact on the magnitude or statistical significance of the IRR estimates. The one exception was that the IRR for the relationship between parking on "three sides" of the road (that is, parking on both sides and in the centre) and MVC was no longer statistically significant in either model ($p>0.15$), even though the IRR did not change. Given that there were only three road segments with this parking configuration in the sample, and that one of them was removed as a potentially influential observation, it is not surprising that there was no longer enough power to detect an effect. As such, the potentially influential observations were not found to significantly alter the outcomes of the MVC models.

6.2.2.6 Overfitting

Both of the MVC models were re-fitted twice using bootstrap resampling with 5,000 replications to determine the effect on the standard error estimates. For both MVC Model 1 and MVC Model 2, all variables remained significantly associated with MVC frequency ($p<0.05$) when bootstrapped standard error estimates were used. This suggests that the models were not overfitted when the p-value for inclusion in the model was set at $p<0.01$.

As mentioned previously, the p-value for inclusion in the models was set at $p<0.01$ at the end of the model development phase to avoid potential overfitting. Eight variables that were associated with MVC frequency with a p-value of between 0.01 and 0.10 were therefore excluded (refer to Table 6.6; seven of the eight had a p-value of less than 0.05 for Model 1 and six of eight had a p-value<0.05 for Model 2). If a less stringent criterion of $p<0.05$ was applied for inclusion of variables in the models, the bootstrapped standard error estimates resulted in p-values of >0.05 for six of the extra seven variables in MVC Model 1 and three of the extra six variables in MVC Model 2. This suggests that restricting the p-value for inclusion into the models to $p<0.01$ avoided potential problems with overfitting.

6.2.2.7 Sensitivity analysis of fixed vs. random parameters

The MVC model with traffic volume as a predictor was re-fitted with a random intercept to determine if there was any evidence that the baseline risk differed across sites. The MVC model with the random intercept did not fit the data significantly better than the MVC model with the fixed intercept (likelihood ratio test, $p=0.38$).

Attempts were made to fit the MVC model with all random coefficients, however, the model would not converge within a reasonable time. As a compromise, the MVC model was re-fitted 14 times; each time the intercept and one of the 14 variables were random and the other 13 variables were fixed. There was no evidence of better MVC model fit when any of the coefficients were random compared to the equivalent fixed effects MVC model (that is, the relationship between the risk factors and MVC frequency did not vary significantly across sites). Therefore, there was no evidence that the assumption of fixed effects was unreasonable.

6.2.2.8 Summary of model fit and diagnostic tests

To summarise, tests revealed that for both MVC models, the link function was properly specified and the observed probabilities were a reasonable match to the predicted probabilities (despite some noise). The relationships between the continuous variables in the linear predictor and MVC frequency were linear and there was no serious multicollinearity between predictor variables. Identification and removal of observations with high leverage had no appreciable influence on the association between the predictor variables and MVC frequency. Restricting the criterion for inclusion in the models to $p<0.01$ (i.e. those variables that entered the model during the initial development phase and the first iteration of post-hoc testing) appears to have avoided overfitting the models. There was no evidence to suggest that the assumption of fixed effects was unreasonable.

Finally, a comparison of the fit statistics for the two models indicates that the model that used ln(traffic volume) as the traffic exposure variable showed slightly better fit for the data than the model that used ln(vehicles per lane) and the number of lanes. For this reason, and because the models were so similar, the interpretation of model results presented in the next section will focus mainly on MVC Model 1. An exception will be made where the models differed appreciably: that is, in the relationship between the traffic exposure variable and MVC, and the number of lanes and MVC, where the results of both models will be discussed. Each of the broad categories of variables describing different characteristics of the road segments will be presented in turn, first identifying those variables that were significantly associated with MVC frequency at the $p<0.01$ level, then those that were significantly associated with MVC frequency when the p-value for inclusion was relaxed to $p<0.05$, and finally, those variables that were not significantly associated with MVC.

6.2.3 Model results and interpretation

Table 6.5 presents the characteristics of the strip shopping road segments in metropolitan Melbourne that were associated with MVC frequency in MVC Model 1 (where the natural log of traffic volume per day was used as the traffic exposure variable) and MVC Model 2 (in which the natural log of the number of vehicles per lane per day was used as the traffic exposure variable) when the p-value for inclusion in the models was set at $p<0.01$. The results are tabulated by broad category of built environment characteristic. Those variables that were associated with MVC frequency in each model when the p-value was set at a more lenient $p<0.05$ or $p<0.10$ are shown in Table 6.6.

Table 6.5 Risk factors for multi-vehicle crashes (MVC) for models developed with traffic volume and traffic density (p<0.01)

.		MVC Model 1 – Traffic volume	MVC Model 2 – Traffic density
Variable	N (%) or Median (Interquartile range)	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)
ROAD USER EXPOSURE			
ln(traffic volume/1000 per day)	3.18 (0.67)	2.02 (1.61–2.53)	N/A
ln(vehicles per lane/1000 per day)	1.86 (0.46)	N/A	1.99 (1.51–2.61)
CROSS SECTION			
Minimum carriageway width			
<=10m	22 (15.5)	1.39 (1.14–1.71)	1.36 (1.11–1.67)
>10m	120 (84.5)	Reference	Reference
Number of lanes			
<4	23 (16.2)	N/A	0.58 (0.45–0.74)
4	86 (60.6)		Reference
>4	33 (23.2)		1.35 (1.06–1.73)
ACCESSES & INTERSECTIONS			
Number of signalised intersections/km	2.74 (2.80)	*0.96 (0.93–1.00)**	*0.96 (0.92–1.00)***
Number of unsignalised intersections/km	8.47 (6.75)	1.04 (1.02–1.06)	1.04 (1.02–1.06)
Number of driveways or laneways/km	16.97 (20.8)	0.992 (0.987–0.997)	0.992 (0.986–0.997)
Number of roundabouts/km	0 (0)	1.25 (1.13–1.38)	1.25 (1.13–1.39)
Service road present			
No	115 (81.0)	Reference	Reference
Yes	27 (19.0)	1.53 (1.16–2.03)	1.50 (1.13–1.999)
MEDIANS			
Maximum median width			
No median	40 (28.2)	Reference	Reference
<1.2m	14 (9.9)	*1.15 (0.88–1.50)*	*1.13 (0.87–1.48)*
1.2–3m	62 (43.7)	*0.87 (0.71–1.07)*	*0.86 (0.69–1.06)*
>3m	26 (18.3)	0.66 (0.50–0.86)	0.66 (0.50–0.86)

.		MVC Model 1 – Traffic volume	MVC Model 2 – Traffic density
Variable	N (%) or Median (Interquartile range)	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)
ROAD TYPE AND TRAFFIC MIX			
Over-dimensional route No Yes	 135 (95.1) 7 (4.9)	 Reference 1.56 (1.16–2.09)	 Reference 1.53 (1.13–2.06)
ROADSIDE PARKING			
Roadside parking None 1 side 2 sides 2 sides and centre	 18 (12.7) 16 (11.3) 105 (73.9) 3 (2.1)	 Reference *0.99 (0.69–1.41)* 1.47 (1.08–1.98) 0.46 (0.22–0.95)	 Reference *0.94 (0.66–1.33)* 1.45 (1.08–1.96) 0.46 (0.22–0.94)
PUBLIC TRANSPORT AND BICYCLE FACILITIES			
No variables significantly associated with MVC frequency			
PEDESTRIAN FACILITIES			
No variables significantly associated with MVC frequency			
ROADSIDE			
Nature strip present No 1 side 2 sides	 105 (73.9) 16 (11.3) 21 (14.8)	 Reference 0.60 (0.45–0.80) 0.52 (0.39–0.68)	 Reference 0.61 (0.46–0.81) 0.51 (0.38–0.68)
Predominantly 2 storey development No Yes	 88 (62.0) 54 (38.0)	 Reference 0.76 (0.64–0.91)	 Reference 0.74 (0.62–0.89)
PAVEMENT CONDITION			
No variables significantly associated with MVC frequency			
HEIGHT CLEARANCE			
No variables significantly associated with MVC frequency			
STATIC ENFORCEMENT CAMERAS			
No variables significantly associated with MVC frequency			

.		MVC Model 1 – Traffic volume	MVC Model 2 – Traffic density
Variable	N (%) or Median (Interquartile range)	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)
AMENITIES AND FACILITIES			
Number of off-street parking facilities/km	2.07 (4.98)	0.95 (0.92–0.97)	0.95 (0.92–0.97)
Number of late night liquor licences/km	0 (1.78)	1.08 (1.05–1.12)	1.08 (1.05–1.12)
SPEED LIMIT			
No variables significantly associated with MVC frequency			

All variables were associated with MVC frequency at p<0.01. For categorical variables, the overall association (all categories considered together) with MVC frequency was <0.01, but associations for individual categories may not have reached the level for inclusion; as such, p-values >0.05 (where 95% CI includes one) shown in italics.
*p=0.08 & **p=0.06 but the rate of signalised intersections was kept in all models to adjust for exclusion of crashes at signalised intersections.

Table 6.6 Additional risk factors for multi-vehicle crash (MVC) frequency (p<0.05 or p<0.10)

		MVC Model 1 – Traffic volume	MVC Model 2 – Traffic density
Variable	N (%) or Median (Interquartile range)	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)
ROAD USER EXPOSURE			
Pedestrian exposure	0.62 (1.04)	1.03 (1.01–1.04)	1.02 (1.01–104)
CROSS SECTION			
Presence of curve/s			
No	125 (88.0)	Reference	Reference
Yes	17 (12.0)	0.77 (0.62–0.94)	0.73 (0.59–0.90)
ROAD TYPE AND TRAFFIC MIX			
Primary state arterial			
No			
Yes	28 (19.7)	Reference	Reference
	114 (80.3)	1.22 (1.02–1.47)	*1.18 (0.98–1.44)*
PUBLIC TRANSPORT AND BICYCLE FACILITIES			
Shared lane with trams			
No	99 (69.7)	Reference	Reference
Yes	43 (30.3)	1.24 (1.06–1.46)	1.26 (1.07–1.50)
ROADSIDE			
Number of non-frangible poles on the roadside/km	69.65 (36.41)	0.997 (0.995–0.999)	0.997 (0.995–0.999)
AMENITIES			
Presence of petrol station			
No	81 (57.0)	Reference	Reference
Yes	61 (43.0)	0.86 (0.76–0.98)	0.87 (0.76–0.99)
Presence child-care, kindergarten or maternal child health centre			
No	117 (82.4)	Reference	Reference
Yes	25 (17.6)	1.39 (1.15–1.68)	1.37 (1.13–1.66)
Presence of place of worship			
No	109 (76.8)	Reference	Reference
Yes	33 (23.2)	*0.87 (0.75–1.00)*	*0.87 (0.76–1.01)*

Variables associated with MVC frequency with a p-value between 0.01 and 0.05 shown in normal font; variables associated with MVC frequency with a p-value between 0.05 and 0.10 shown in italics.

The remainder of this section presents the results of the MVC model 1 which included the natural log of traffic volume as the exposure variable summarised by category and interpreted in detail (previously presented in Table 6.5). Variables that were not found to be associated with MVC will also be identified. Those results that were specific to MVC Model 2 will also be presented (traffic density and number of lanes).

6.2.3.1 Road user exposure

Measures of traffic exposure were available for the road segment of interest and estimates of pedestrian activity were derived from aggregate data from travel surveys at the level of the local area.

The natural log of traffic volume was significantly associated with MVC frequency which was proportional to traffic volume (thousand vehicles per day) raised to the power of 0.70 (95% CI 0.48–0.93). Thus, including traffic volume as part of the offset term (which effectively sets the coefficient to be equal to one) would have been a mistake, because the 95% CI for the coefficient did not include one.

When the natural log of the number of vehicles per lane (traffic density) was included in the model instead of traffic volume, there was also a significant association with MVC frequency. MVC frequency was proportion to traffic density raised to the power of 0.69 (95% CI 0.41–0.96). Again, it would have been a mistake to include this variable as part of the offset term, as the coefficient was significantly different to one.

The measure of pedestrian activity (distance walked per square km of statistical local area x segment length) was not included in the final MVC model 1 ($p>0.01$), however, it was one of the eight variables that were associated with MVC frequency when using an inclusion criterion of $p<0.05$. Therefore, there was some evidence that MVC frequency increased as pedestrian activity increased, however this association disappeared when bootstrapped standard error estimates were used ($p=0.19$). Nevertheless it may be worth considering in future research.

6.2.3.2 Road cross section

Aspects of the road cross section that were investigated were the carriageway width, the lane width, the number of lanes and the presence of a curve (or curves) on the road segment.

Minimum carriageway width (categorised as <=10m and >10m) was significantly associated with MVC frequency. Roads with a minimum carriageway width narrower than 10m were associated with a 39% (95% CI 14%–71%) increase in MVC frequency compared to those with a minimum carriageway width of greater than 10m.

The number of lanes entered into the model when the traffic density was included as the traffic exposure variable (MVC Model 2) but not when traffic volume (the number of vehicles) was included (MVC Model 1). This is to be expected, given that traffic density is equal to the traffic volume divided by the number of lanes. In MVC Model 2, MVC frequency increased as the number of lanes increased. Compared to roads with four lanes, road with one or two lanes were associated with a statistically significant 42% (95% CI 26%–55%) reduction in MVC frequency, while roads with more than four lanes were associated with a statistically significant 35% (95% CI 6%-73%) increase in MVC.

The presence of a curve on the road segment was not included in the final model, however, this variable was associated with MVC frequency when the less stringent criterion for inclusion was applied ($p<0.05$). If the road segment was not completely straight, MVC frequency was 20% (95% CI 2%–35%) lower compared to when the road segment was completely straight. This association was no longer statistically significant when bootstrapped standard error estimates were used ($p=0.08$), nonetheless, it may be worth considering in future research.

Lane width (mode, minimum or maximum) was not significantly associated with MVC frequency.

6.2.3.3 Accesses and intersections

Accesses and intersections were measured in terms of the number of signalised and unsignalised intersections per km, the number of driveways and laneways per km, the number of roundabouts per km, the number of keep clear zones per km, the number of dedicated turning lanes per km and the presence and type of service roads. Most of these aspects were significantly associated with MVC frequency.

The number of signalised intersections per km was included in all models to adjust for the exclusion of crashes that occurred at signalised intersections (that is, this variable was included in the models even if not significantly associated with MVC frequency). It would be expected, therefore, that the MVC frequency would decrease as the number of signalised intersections per km increased. This expectation was borne out: for each extra signalised intersection per km, the MVC frequency decreased, however, this association was not statistically significant.

MVC frequency significantly increased as the number of unsignalised intersections increased. For every extra unsignalised intersection per km, MVC frequency increased by 4% (95% CI 2%–6%). In comparison, driveways and laneways had the opposite effect on MVC frequency. For every extra 10 driveways or laneways per km, MVC frequency decreased by 8% (95% CI 3%–13%).

Roundabouts were associated with increased MVC frequency. For every extra roundabout per km, the MVC frequency increased by 25% (95% CI 13%–38%). Service roads were also related to increases in MVC. If there was a service road present on either or both sides of the road, MVC frequency increased by 53% (95% CI 16%–103%).

Keep clear zones and dedicated turning lanes were not significantly associated with MVC frequency.

6.2.3.4 Medians

The presence, type and width of medians, as well as the proportion of the road that had a median of various types and widths were measured. Of these variables, maximum median width had the strongest relationship with MVC frequency.

Maximum median width was categorised as follows: no median, less than 1.2m, between 1.2m and 3m and greater than 3m. Compared to all other maximum median width categories, roads with a maximum median width of greater than three metres had significantly fewer MVC: 34% (95% CI 14%–50%) fewer compared to roads with no median, 25% (95% CI 7%–39%) fewer compared to roads with a maximum median width between 1.2m and 3m and 43% (95% CI 23%–57%) fewer compared to roads with a maximum median width of less than 1.2m. Roads with a maximum median width of between 1.2m and 3m had significantly fewer MVC than those with a maximum median width of <1.2m (IRR=0.76, 95% CI 0.60–0.97). There was no significant difference in MVC frequency between roads with no medians and those with a maximum median width less than 1.2m or those with a maximum median width of between 1.2m to 3m. It is worth emphasising that undivided roads with even one median island will have a maximum median width, so this variable essentially describes the presence and width of both median islands and longer medians.

6.2.3.5 Road type and traffic mix

All road segments in the sample were on arterial roads, however, some were primary state arterials while others were secondary state arterials. The traffic distribution on each road segment was measured, in terms of the percentage of the traffic that were heavy vehicles. Road segments were also classified according to whether or not heavy vehicles or over-dimensional vehicles were permitted to use the road.

Over-dimensional vehicle routes were associated with a 56% (95% CI 16%–109%) increase in MVC frequency compared to routes that did not allow over-dimensional vehicles.

Road type was not included in the final model, however, when the p-value for inclusion was set at $p<0.05$ (rather than the more stringent $p<0.01$), primary state arterials were associated with a 21% (95% CI 1%–46%) increase in MVC frequency compared to secondary state arterials. This variable, however, was no longer significantly associated with MVC frequency when bootstrapped standard error estimates were used ($p=0.12$).

The remaining variables related to traffic distribution were not significantly associated with MVC frequency. Specifically, these were the percentage of traffic that were heavy vehicles, heavy vehicle approved routes and whether or not an over-dimensional route intersected with the road segment.

6.2.3.6 Roadside parking

The presence and type of roadside parking and whether or not there was a parking clearway was determined for each road segment. The number of loading zones per km was also measured.

Roadside parking was significantly associated with MVC frequency. There were four categories of roadside parking: no parking, parking on one side of the road, parking on two sides of the road, and parking on two sides and the centre of the road. MVC frequency was not significantly different between roads with no parking and roads with parking on one side of the road. When parking was allowed on both sides of the road, MVC frequency was significantly higher than all other parking combinations: 47% (95% CI 8%–98%) higher than when no roadside parking was allowed, 48% (95% CI 8%–105%) higher than when parking was on one side only and 217% (95% CI 57%–542%) higher than when parking was on both sides and in the centre of the road. When parking was allowed on both sides of the road and parking in the centre median was also available, MVC frequency was significantly lower than all other parking combinations: 54% (95% CI 5%–78%) decrease in MVC frequency compared to when no parking was permitted, 53% (95% CI 1%–78%) decrease compared to when parking was allowed on one side of the road and 68% (95% CI 36%–84%) decrease compared to when parking was on both sides but there was no centre parking.

The presence of parking clearways and the number of loading zones per km were not significantly associated with MVC frequency.

6.2.3.7 Public transport and bicycle facilities

Public transport facilities included railway level crossings, the presence of trams and buses, tram lane type and the number of stops per km for both buses and trams. Bicycle facilities included the presence, type and width of bicycle lanes

None of the variables relating to public transport and bicycle facilities were significantly associated with MVC frequency except when the less stringent criterion ($p<0.05$) for inclusion in the model was used. On road segments where vehicles were required to share a travel lane with trams, the MVC frequency was 22% (95% CI 4%–44%) higher than on road segments where vehicles did not share a lane with trams (that is, where there were no trams, or where trams had a separated lane). This association became less strong, however, when bootstrapped standard error estimates were used (p-value increased to approximately 0.08).

6.2.3.8 Pedestrian facilities

The number of midblock pedestrian crossings per km was not significantly associated with MVC frequency, regardless of type (signalised or signed). The presence of pedestrian fencing at pedestrian crossings was also not significantly associated with MVC frequency.

6.2.3.9 Roadside

Roadside characteristics included development height, offset between roadside and development, the presence of nature strips, whether or not there were shops on one or both sides of the road, and the number of poles (frangible and/or non-frangible) on the roadside and median.

The presence of a nature strip on one or both sides of the road was associated with a reduction in MVC. When there was a nature strip on one side of the road, MVC frequency was 40% (95% CI 20%–55%) lower than when there was no nature strip present. The presence of a nature strip on both sides of the road was associated with a 48% (95% CI 32%–61%) reduction in MVC compared to when there was no nature strip.

Roadside development height was also significantly associated with MVC frequency. When the roadside development was predominantly two-storey, MVC frequency was 24% (95% CI 9%–36%) lower than when roadside development was predominantly single storey or a mix of different heights.

When a less severe threshold was used for inclusion into the model ($p<0.05$ instead of $p<0.01$), the number of non-frangible poles or trees per km of roadside was associated with a reduction in MVC frequency. For every extra 10 non-frangible poles or trees per km, MVC frequency decreased by 3% (95% CI 1%–5%). When bootstrapped standard errors were used, however, this association was no longer statistically significant ($p=0.06$).

The remaining variables describing the roadside environment were not associated with MVC frequency. These were: the presence of shops on one or both sides of the road, the offset distance between the road and roadside development, the height of the highest building, the number of frangible poles and trees on the roadside and the number of poles and trees (frangible or non-frangible) on the median.

6.2.3.10 Road pavement condition

Road pavement condition (the presence of pavement distress or roughness) was not significantly related to MVC frequency.

6.2.3.11 Height clearance

Height clearance was not associated with MVC frequency. This included the presence of low tram wires (which affect the whole road segment) and low clearance due to bridges (which affect only part of the road segment).

6.2.3.12 Static enforcement cameras

The presence of static red light and speed enforcement cameras at signalised intersections on the road segments was not significantly associated with MVC frequency on the road segment midblocks.

6.2.3.13 Amenities and facilities

Amenities and facilities included the rate of establishments with different types of liquor licences, the rate of off-street parking facilities and the presence of amenities and facilities likely to affect the number of type of road users in the area.

Off-street parking facilities were associated with a reduction in MVC frequency. For each extra off-street parking facility per km, MVC frequency was 5% (95% CI 3%–8%) lower.

Liquor establishments with a late night liquor licence were also associated with MVC frequency. For every establishment with a late night liquor licence per km, MVC frequency increased by 8% (95% CI 5%–12%). Other types of liquor licence (non-late night licences and licences to sell alcohol to consume off-premises) were not significantly associated with MVC frequency.

None of the other facilities or amenities were significantly associated with MVC frequency. These included liquor licences (apart from late-night liquor licences), the presence of primary, secondary or tertiary education institutions, various types of community centres, parks and sporting fields, emergency services, hospitals, places of worship, railway stations and indoor shopping centres. When the p-value for inclusion in the model was set at a less strict $p<0.05$, two facilities/amenities were associated with MVC frequency. The presence of one or more petrol stations was associated with a 14% (95% CI 2%–25%) reduction in MVC while the presence of kindergartens, child-care or maternal and child health facilities was associated with a 31% (95% CI 9%–56%) increase in MVC frequency. When bootstrapped standard error estimates were used, however, the presence of petrol stations was no longer significantly associated with MVC ($p=0.07$).

6.2.3.14 Speed limit

The speed limit of the road segment was not significantly associated with MVC frequency.

6.2.3.15 Summary of variables associated with multi-vehicle crash frequency

In summary, a number of factors were associated with significant increases in MVC when the criterion for inclusion in the model was set at $p<0.01$: traffic volume (or traffic density and the number of lanes), carriageway widths narrower than 10m, the number of unsignalised intersections per km, the number of roundabouts per km, the presence of a service road, the presence of parking on both sides of the road, over-dimensional vehicle routes and establishments with late night liquor licences. With a less severe threshold for inclusion of $p<0.05$, pedestrian activity, the requirement to share a lane with trams, road type (primary state arterials) and the presence of a kindergarten, childcare or material and child health facility were also associated with increases in MVC.

The following aspects were associated with significant decreases in MVC frequency ($p<0.01$): the number of driveways or laneways per km, medians or medians islands with a maximum width of greater than 3m, parking on both sides and in the centre median, double-storey roadside development, nature strips and off-street parking facilities. With a less severe criterion for inclusion ($p<0.05$), the presence of a curve (or curves), non-frangible roadside poles or trees and the presence of petrol stations were also associated with decreased MVC frequency.

6.3 Single-vehicle crashes

6.3.1 Selecting the optimal method for specifying exposure

A series of negative binomial regression models were fitted to determine the optimal way to specify segment length and traffic exposure. An offset term was included to account for the different number of days for the collection of crash data across road segments. The number of signalised intersections per km was also included as a predictor in the model, to adjust for the exclusion of crashes at signalised intersections.

Table 6.7 displays the different ways that were considered to specify traffic volume and length, and the AIC value for each SVC model. The best model fit (lowest AIC value) resulted from the model that included the natural log of traffic volume and the natural log of segment length. This model showed that the natural log of the traffic volume was significantly related to SVC (β=0.46, 95%CI 0.08–0.85). The 95% CI for the coefficient did not include one, which indicated that the relationship between traffic volume and SVC was not directly proportional so traffic volume was retained in the linear predictor. The natural log of segment length was also significant related to SVC (β=0.98, 95%CI 0.72–1.24) and the coefficient was not significantly different to one. Hence, the segment length was included as part of the offset term during subsequent development of SVC model 1, using the natural log of traffic volume as the traffic exposure variable.

Table 6.7 Comparison of different ways to specify traffic volume and segment length for modelling single-vehicle crashes

Traffic volume	Segment length	AIC
AADT/1000	length	2.738
AADT/1000	ln(length)	2.731
ln(AADT/1000)	length	2.727
ln(AADT/1000)	ln(length)	2.721 (lowest)
thousand vehicle km per day		*
ln(thousand vehicle km per day)		*

* Coefficients for traffic volume and length were significantly different when entered as separate variables, so there was no justification for pooling as one combined measure of exposure.

Table 6.8 shows the different ways that were considered to specify traffic density and segment length. None of the traffic density variables were significantly associated with SVC frequency. Therefore, no model was developed for SVC frequency with traffic density as the traffic exposure variable.

Table 6.8 Comparison of different ways to specify traffic density and segment length for modelling single-vehicle crashes

Traffic density	Segment length	AIC
AADT/1000 per lane*	length	2.774
AADT/1000 per lane*	ln(length)	2.755
ln(AADT/1000 per lane)*	length	2.775
ln(AADT/1000 per lane)*	ln(length)	2.757
thousand vehicle km per lane per day		**
ln(thousand vehicle km per lane per day)		**

* These variables were not significantly associated with SVC frequency.
** Coefficients for traffic density and length were significantly different when entered as separate variables, so there was no justification for pooling as one combined measure of exposure

6.3.2 Model development and diagnostic tests

6.3.2.1 Model development

Only one model was developed for SVC frequency; this model included the natural log of traffic volume as the traffic exposure variable, and segment length was included in the offset term. Two iterations of post-hoc testing were required to finalise the model. At the end of the model development process, there were 13 predictor variables. When the model was finalised and the p-value for inclusion in the model was reduced from $p<0.10$ to $p<0.05$, all of the 13 variables were retained in the model. The final model showed no evidence of overdispersion ($\alpha = 5.86e^{-8}$) and the likelihood ratio test of alpha being equal to zero was non-significant ($\chi^2(1)=0.00$, $p=0.500$). Therefore, Poisson regression was used to fit the final model. Because there were only 170 SVC over the five year observation period, there was not enough power to reduce the criterion for inclusion to less than $p<0.05$ (compared to the MVC models where there were 1294 crashes).

6.3.2.2 Model fit

The deviance statistic divided by the number of degrees of freedom (0.89) and the Pearson statistic divided by the number of degrees of freedom (0.82) were less than one, indicating the data may be slightly underdispersed.

Figure 6.6 displays the observed probabilities that the number of SVC was equal to x (where x=0 to 16) from the raw data, and the predicted probabilities from the Poisson regression model. Visual inspection revealed that the model fitted the data well. There was no evidence for an excess number of zero observations beyond that predicted by the Poisson model, therefore, there was no evidence that a zero-inflated model was required.

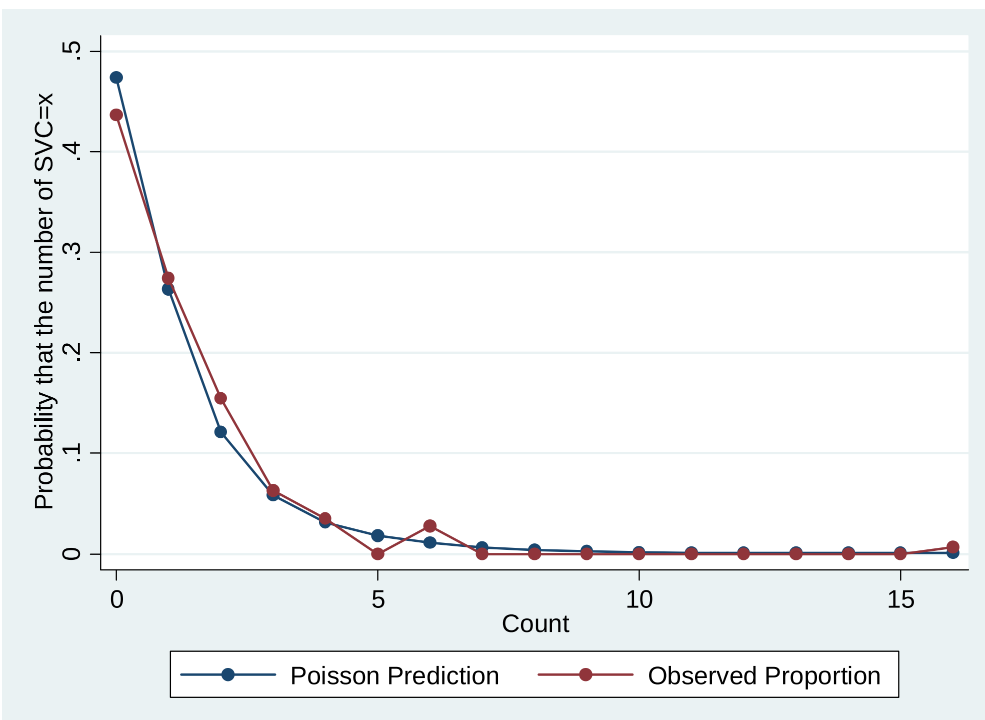


Figure 6.6 Probability that the number of single-vehicle crashes (SVC)=x: observed proportion and predicted proportion

There was no evidence that the link function was misspecified (p-value=0.70), nor that any relevant variables were excluded so interactions between variables were not explored. Neither the deviance statistic ($\chi^2(125)=111.93$, $p=0.81$) nor the Pearson statistic ($\chi^2(125)=103.22$, $p=0.93$) were statistically significant, indicating that the model form was appropriate.

6.3.2.3 Tests of linearity

The Box Tidwell power and exponential transformation regression models showed no evidence that the relationship between any of the continuous variables and SVC frequency was non-linear (all p-values greater than 0.23).

6.3.2.4 Tests for multicollinearity

The predictor variables were not found to be multicollinear. All VIFs were lower than 10, which is the threshold for concern regarding multicollinearity between variables (range 1.07 to 1.67). The CN was 30.66, which is a little high given that 30 is the threshold for concern. There was only one principal component where more than one variable loaded highly. The constant and the natural log of traffic volume loaded highly on the last principal component. The constant is essentially a nuisance variable, therefore because more than one predictor variable did not load highly on any principal component, there was no evidence that multicollinearity was of concern.

6.3.2.5 Analysis of residuals

Inspection of the distribution of the residuals showed that the adjusted deviance residual showed the least skew and kurtosis compared to other residuals so this residual was chosen for further diagnostic testing. The quantiles of the adjusted deviance residual plotted against quantiles of the normal distribution showed no obvious departures from normality (Figure 6.7).

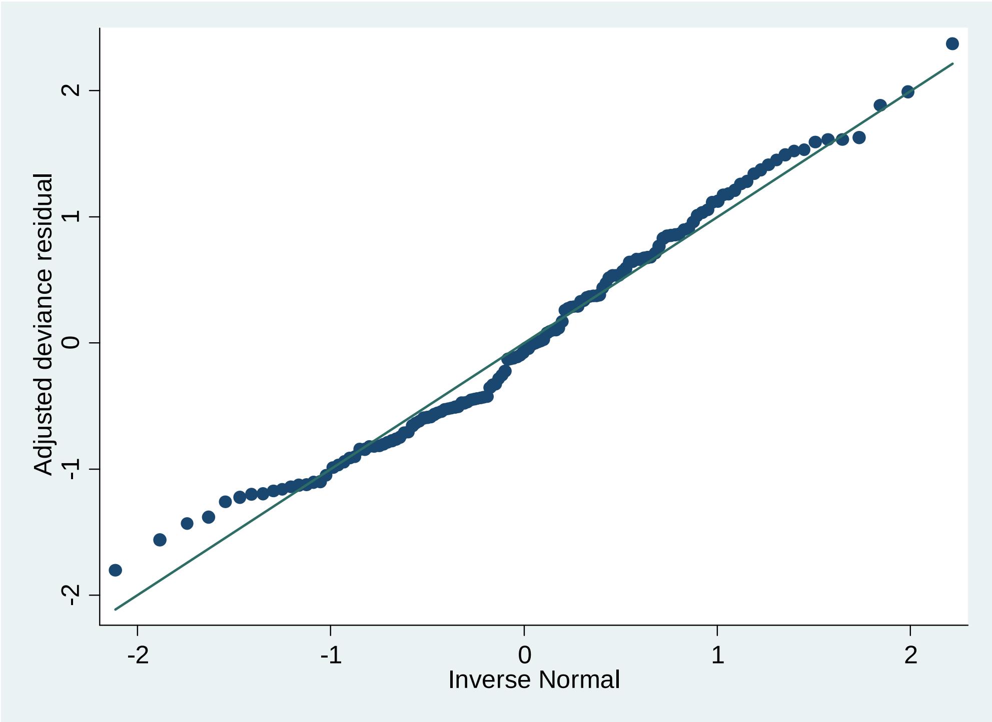


Figure 6.7 Quantiles of the adjusted deviance residual plotted against quantiles of the normal distribution

Figure 6.8 shows a scatter plot of the adjusted deviance residual against the predicted mean SVC with locally weighted scatterplot smoothing. The residuals clustered around zero and the lowess line was reasonably flat. As such, there was no evidence of a relationship between the residuals and the predicted values.

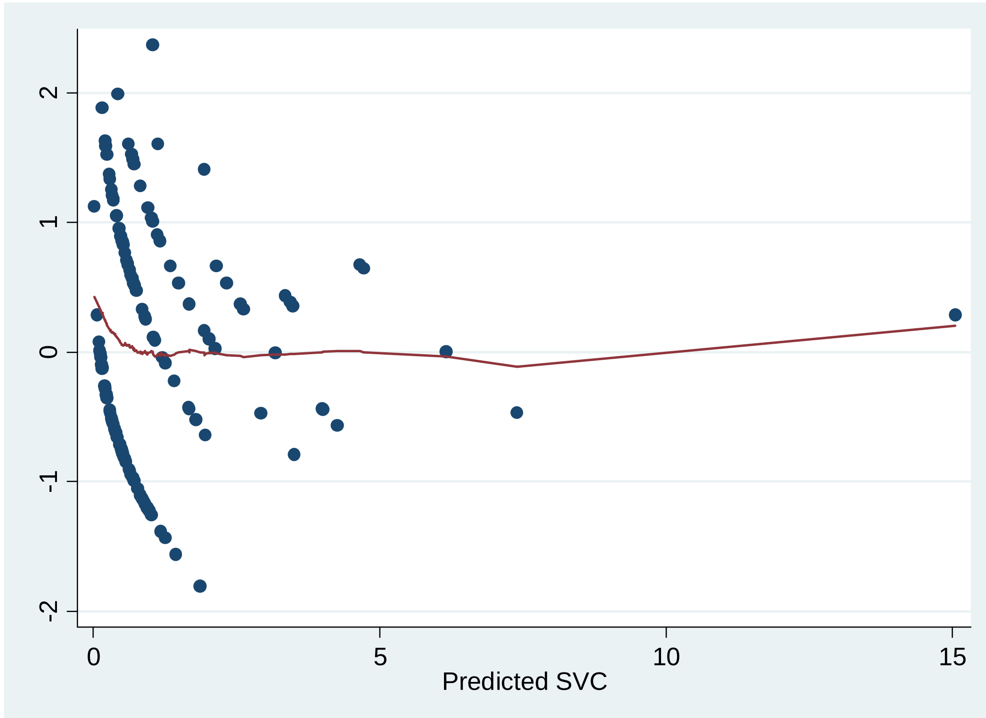


Figure 6.8 Locally weighted smoothed scatterplot of the adjusted deviance residual against the predicted mean number of single vehicle crashes (SVC)

There were 11 road segments with potentially high leverage. Once the potentially influential road segments were removed, two characteristics of the road segments were no longer significantly associated with SVC frequency (the presence of low tram wires ($p=0.79$) and the presence of bicycle lanes > 1.2m wide ($p=0.15$)). For two other variables, the p-value increased to lie between the range 0.05 and 0.09 (the presence of parks and sporting fields and the rate of non-frangible poles and trees on the roadside). While the low number of SVC means there was less power to detect true effects, however, there was still reasonable evidence for these latter two variables to be related to SVC frequency when potentially influential observations were removed.

6.3.2.6 Overfitting

The SVC model was re-fitted using bootstrap resampling with 5,000 replications to determine the effect on the standard error estimates as an aid to investigate whether the model may have been overfitted. With bootstrapped standard error estimates, two aspects of the road environment were no longer significantly associated with SVC frequency: the presence of low tram wires ($p>0.37$) and bicycle lanes wider than 1.2m ($p>0.72$). These were the same two variables that were no longer significantly associated with SVC frequency when observations with high leverage were removed. Thus it appears that these associations may not be generalisable to other samples of urban road segments.

6.3.2.7 Sensitivity analysis of fixed vs. random parameters

The SVC model re-fitted with a random intercept to determine if there was any evidence that the baseline risk differed across sites. The SVC model with the random intercept did not fit the data significantly better than the SVC model with the fixed intercept. Hence the baseline risk of SVC was not significantly different across sites.

The SVC model would not converge with all random coefficients, so as a compromise, attempts were made to re-fit the SVC model a number of times. Each time the intercept and one of the variables were random and the other variables were fixed. Unfortunately, none of these models would converge. This is likely because the models were too complex for the data available (only 170 SVC were observed).

6.3.2.8 Summary of model fit and diagnostic tests

In summary, tests of model fit and diagnostics revealed that the link function was properly specified and that the model form was appropriate, with observed probabilities matching the predicted probabilities. There was no evidence to suggest that the relationship between the continuous predictor variables and the SVC frequency was non-linear, nor that there was significant multicollinearity between predictors. Several observations with high leverage were identified, however, the removal of these observations did little to change the regression estimates for all but two of the variables: the presence of low tram wires and bicycle lanes wider than 1.2m. When bootstrapped standard error estimates were used, the same two variables were no longer significantly associated with SVC frequency. Overall, model testing revealed no reason to reject the fitted model, notwithstanding the potential for generalisability of the risk associated with low tram wires and wide bicycle lanes. There was no evidence that the baseline risk differed across sites so there was no need to include a random intercept.

6.3.3 Model results and interpretation

Table 6.9 presents the characteristics of the strip shopping road segments in metropolitan Melbourne that were associated with SVC frequency ($p<0.05$). The results are tabulated by the broad categories of built environment characteristics. Following Table 6.9, the results are interpreted in detail, by category. Variables that were not found to be associated with SVC will also be identified.

Table 6.9 Risk factors for single-vehicle crashes

Variable	N (%) or Median (Interquartile range)	Incidence rate ratio (95% Confidence Interval)
EXPOSURE		
ln(traffic volume/1000 per day)	3.18 (0.67)	2.43 (1.58–3.73)
CROSS SECTION		
Minimum carriageway width		
5.5–8.5m	16 (11.3)	1.97 (1.25–3.12)
8.5–10m	6 (4.2)	*0.64 (0.15–2.68)*
>10m	120 (84.5)	Reference
Presence of curve/s		
No	125 (88.0)	Reference
Yes	17 (12.0)	2.33 (1.49–3.64)
ACCESSES AND INTERSECTIONS		
Number of signalised intersections/km	2.74 (2.80)	*0.96 (0.88–1.06)**
MEDIANS		
No variables significantly associated with SVC frequency		
ROAD TYPE AND TRAFFIC MIX		
No variables significantly associated with SVC frequency		
ROADSIDE PARKING		
No variables significantly associated with SVC frequency		
PUBLIC TRANSPORT AND BICYCLE FACILITIES		
Bicycle lane width		
None	118 (83.1)	Reference
<1.2m	14 (9.9)	2.19 (1.33–3.61)
>1.2m	10 (7.0)	0.50 (0.25–0.97)
Shared lane with trams		
No	99 (69.7)	Reference
Yes	43 (30.3)	2.04 (1.37–3.05)
Number of bus stops per km	3.55 (6.92)	1.06 (1.01–1.11)
PEDESTRIAN FACILITIES		
No variables significantly associated with SVC frequency		
ROADSIDE		
Number of non-frangible poles on the roadside/km	69.65 (36.41)	1.006 (1.001–1.010)
PAVEMENT CONDITION		
Presence of distress		
No	66 (46.5)	Reference
Yes	76 (53.5)	0.69 (0.50–0.95)
HEIGHT CLEARANCE		
Presence of low tram wires		
No	137 (96.5)	Reference
Yes	5 (3.5)	0.31 (0.11–0.90)
STATIC ENFORCEMENT CAMERAS		
No variables significantly associated with SVC frequency		
AMENITIES AND FACILITIES		

Variable	N (%) or Median (Interquartile range)	Incidence rate ratio (95% Confidence Interval)
Number of off-street parking facilities/km	2.07 (4.98)	0.91 (0.85–0.97)
Number of bring your own liquor licences/km	2.03 (4.50)	1.08 (1.03–1.13)
Presence of sports centres, fields, parks		
No	110 (77.5)	Reference
Yes	32 (22.5)	0.61 (0.39–0.94)
SPEED LIMIT		
No variables significantly associated with SVC frequency		

All variables were associated with SVC frequency at $p<0.05$. For categorical variables, the overall association (all categories considered together) with SVC frequency was <0.05, but associations for individual categories may not have reached the level for inclusion; p-values >0.05 (where 95% CI includes one) shown in italics.
*p=0.44 but the rate of signalised intersections was kept in all models to adjust for exclusion of crashes at signalised intersections.

6.3.3.1 Road user exposure

SVC frequency is related to the traffic volume (thousand vehicles per day) raised to the power of 0.89 (95% CI for β 0.46–1.32). In the final model, the CI for the estimate of β did include one so it would be acceptable to include traffic volume in the offset term alongside segment length and the number of days (which would effectively model crashes per thousand vehicle km per day).

Pedestrian activity, estimated using aggregate area-level data, was not significantly associated with SVC frequency.

6.3.3.2 Road cross section

The incidence of SVC when a curve or curves were present on the road segment was approximately 2.3 times the incidence of SVC when the segment was straight (IRR=2.33, 95% CI 1.49–3.64).

Minimum carriageway width (categorised as 5.5 to 8.5m, 8.5 to 10m and >10m) was also associated with SVC frequency. Roads with a minimum carriageway width of between 5.5 to 8.5m had almost double the SVC as roads with a minimum carriageway wider than 10m. However, none of the other pairwise comparisons between categories revealed any statistically significant differences.

Lane width and the number of lanes were not significantly associated with SVC frequency.

6.3.3.3 Accesses and intersections

None of the variables relating to accesses and intersections were significantly associated with SVC frequency. This included the number of unsignalised intersections per km, the number of driveways or laneways per km, the number of roundabouts per km, the number of keep clear zones per km, the number of dedicated turning lanes per km and the presence of service roads. The number of signalised intersections per km was included to adjust for the exclusion of crashes that occurred at signalised intersections even though it was not significantly associated with SVC.

6.3.3.4 Medians

The presence, type, width of medians and the proportion of the road segment that had a median of various types and widths were not significantly related to SVC frequency.

6.3.3.5 Road type and traffic mix

The type of road (primary or secondary arterial), heavy vehicles access, over-dimensional routes and the proportion of traffic that were heavy vehicles were not significantly associated with SVC frequency.

6.3.3.6 Roadside parking

The presence and type of roadside parking, the presence of clearways and the number of loading zones per km were not significantly associated with the incidence of SVC.

6.3.3.7 Public transport and bicycle facilities

When vehicles and trams shared a travel lane on a road segment, the incidence rate of SVC was more than double when vehicles and trams did not share a lane. SVC frequency was positively related to the number of bus stops. For every extra bus stop per km, SVC incidence increased by 6% (95% CI 1%–11%).

There was no significant association discovered between the presence of a railway level crossing and the frequency of SVC.

The presence and width of on-road bicycle lanes was significantly associated with SVC frequency. When there was a bicycle lane present of width less than 1.2m, the incidence of SVC was more than double that when there was no bicycle lane present. In contrast, if there was a wider bicycle lane present (>1.2m), there frequency of SVC was halved compared to when there was no bicycle lane. Wide bicycle lanes, however, were no longer significantly associated with SVC frequency when potentially influential observations were removed, nor when bootstrapped standard error estimates were used, therefore this result may not be generalisable to other samples of urban road segments.

6.3.3.8 Pedestrian facilities

The incidence of SVC was not significantly related to the number of pedestrian crossings per km nor to the presence of roadside fencing at pedestrian crossings.

6.3.3.9 Roadside

There was a positive relationship between roadside non-frangible poles and trees and SVC occurrence. For every extra 10 non-frangible poles or trees per km of roadside, SVC frequency increased by 6% (95% CI 1%–10%). In contrast, frangible poles and trees on the roadside were not significantly associated with SVC incidence and nor were poles and trees (either frangible or non-frangible) on the median.

None of the other roadside variables were significantly associated with SVC frequency. This included the presence of nature strips, the presence of shops on one or both sides of the road, development height or the offset between the road and roadside buildings.

6.3.3.10 Pavement condition

The presence of distress (lane rutting and cracking) on the road pavement was associated with 31% (95% CI 5%–50%) decrease in SVC frequency compared to road segments with no distress. In contrast, the presence of rough pavement on a road segment was not significantly related to SVC frequency.

6.3.3.11 Height clearance

On road segments with low tram wires (<4.6m), the SVC frequency was 69% (95% CI 10%–89%) lower than when there were no low tram wires. This result was not statistically significant, however, when influential observations were removed nor when bootstrapped standard error estimates were used and thus must be viewed with caution. Low clearance due to the presence of bridges was not significantly associated with SVC frequency.

6.3.3.12 Static enforcement cameras

The presence of static red light and speed enforcement cameras at signalised intersections on the road segments was not significantly associated with the frequency of SVC frequency on midblock segments.

6.3.3.13 Amenities and facilities

The number of establishments with a BYO liquor licence was associated with significant increases in SVC frequency. For every extra BYO licence per km, there was an 8% (95% CI 3%–13%) increase in SVC. Other types of liquor licence, however, were not associated with SVC frequency.

Two facilities in the environment were associated with reductions in SVC. For each extra off-street parking facility per km, the SVC incidence was reduced by 9% (95% CI 3%–15%). The presence of sports centres, sports fields or parks on the road segment was associated with a 39% (95% CI 6%–61%) reduction in SVC frequency compared to road segments with no such facilities.

None of the other amenities and facilities were significantly associated with SVC frequency including liquor licences (apart from BYO licences), the presence of primary, secondary or tertiary education institutions, various types of community centres, emergency services, hospitals, places of worship, railway stations, indoor shopping centres, petrol stations and kindergartens, child-care and maternal and child health facilities.

6.3.3.14 Speed limit

Speed limit was not significantly associated with SVC frequency on these strip shopping road segments in metropolitan Melbourne.

6.3.3.15 Summary of variables associated with single-vehicle crash frequency

The following characteristics of the strip shopping road segments were associated with increases in SVC: minimum carriageway width of less than 8.5m (compared to >10m), the presence of a curve or curves on the road segment, the presence of narrow bicycle lanes, the requirement to share a travel lane with trams, non-frangible trees and poles on the roadside and the number of establishments with a BYO liquor licence.

Decreases in SVC frequency were associated with distressed road pavement conditions, off-street parking facilities, and the presence of sports centres, sports fields or parks. The presence of bike lanes wider than 1.2m and the presence of low tram wires were also associated with fewer SVC, however, this result may not be generalisable to other samples because the association disappeared when influential observations were removed and also when standard error estimates were bootstrapped.

6.4 Pedestrian vehicle crashes

6.4.1 Selecting the optimal method for specifying exposure

The optimal way to specify segment length and traffic exposure was assessed by fitting a series of negative binomial regression models with an offset term to account for the different number of days for the collection of crash data across road segments. The number of signalised intersections per km was also included, to adjust for the exclusion of crashes at signalised intersections. Table 6.10 and Table 6.11, respectively, show the different ways to specify traffic volume and length, and traffic density and length, and the AIC value for each PVC model. None of the traffic volume or traffic density variables were significantly associated with the frequency of PVC. Models that included the natural log of segment length rather than untransformed segment length had lower AIC values. The coefficient for the relationship between the natural log of segment length and PVC frequency was not significantly different from one (β=1.07, 95% CI 0.86–1.28). Therefore, the PVC model was developed with length as part of the offset term and without any traffic exposure variables (although these were included in post-hoc testing to determine if they were significantly related to PVC when other factors were adjusted for).

Table 6.10 Comparison of different ways to specify traffic volume and segment length for modelling pedestrian-vehicle crashes

Traffic volume	Segment length	AIC
Traffic volume per day*	length	4.491
Traffic volume per day*	ln(length)	4.331
ln(traffic volume per day)*	length	4.497
ln(traffic volume per day)*	ln(length)	4.340
thousand vehicle km per day		**
ln(thousand vehicle km per day)		**

* These variables were not significantly associated with SVC frequency.
** Coefficients for traffic volume and length were significantly different when entered as separate variables, so there was no justification for pooling as one combined measure of exposure.

Table 6.11 Comparison of different ways to specify traffic density and segment length for modelling pedestrian-vehicle crashes

Traffic density	Segment length	AIC
Vehicles per lane per day*	length	4.496
Vehicles per lane per day*	ln(length)	4.341
ln(vehicles per lane per day)*	length	4.497
ln(vehicles per lane per day)*	ln(length)	4.341
thousand vehicle km per lane per day		**
ln(thousand vehicle km per lane per day)		**

* These variables were not significantly associated with SVC frequency.
** Coefficients for traffic density and length were significantly different when entered as separate variables, so there was no justification for pooling as one combined measure of exposure.

6.4.2 *Model development and diagnostic tests*

6.4.2.1 Model development

Only one model was developed for PVC frequency and exposure was accounted for by including segment length in the offset term because neither traffic volume nor traffic density were significantly associated with PVC frequency. Two iterations of post-hoc testing were required to finalise the model. There were 15 variables at the end of model development stage, and when the model was finalised and the p-value was reduced to $p<0.05$, two of these were removed. Once the model was developed, the overdispersion parameter was extremely small ($\alpha=3.11e^{-08}$, likelihood ratio test $\alpha=0$, $\chi^2(1)=0.00$, $p=0.500$), so the model was refitted using Poisson regression.

6.4.2.2 Model fit

The deviance statistic divided by the number of degrees of freedom (1.06) and the Pearson statistic divided by the number of degrees of freedom (0.97) were both close to one, indicating the data were neither underdispersed nor overdispersed.

Figure 6.9 presents the observed probabilities that the number of PVC was equal to x (where x=0 to 43) from the raw data, and the predicted probabilities from the Poisson regression model. The model fitted the data well and there was no evidence for an excess number of zero observations beyond that predicted, therefore, there was no evidence that a zero-inflated model was required.

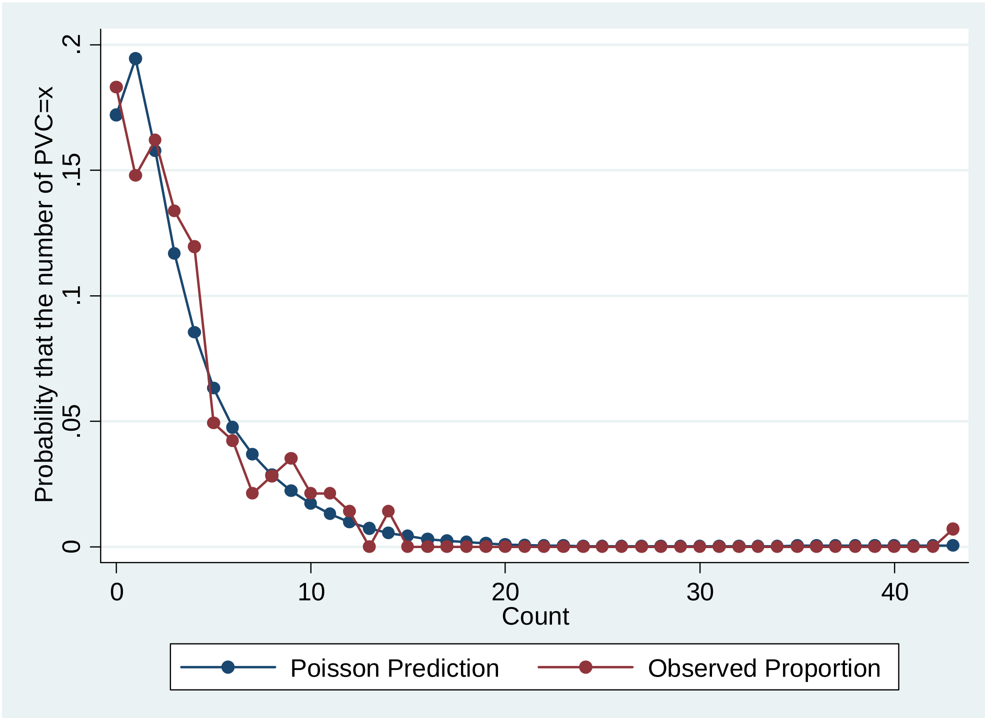


Figure 6.9 Probability that the number of pedestrian-vehicle crashes (PVC)=x: observed proportion and predicted proportion

The non-significant link test (p=0.988) provides no evidence that the link function was misspecified nor that there were any omitted variables. Hence, potential interactions were not investigated. The deviance statistic ($\chi^2(125)=131.92$, $p=0.32$) and the Pearson statistic ($\chi^2(125)=120.90$, $p=0.59$) were not statistically significant, which supported the appropriateness of the choice of model form.

6.4.2.3 Test of linearity

The Box Tidwell power and exponential transformation regression models showed no evidence that the relationship between any of the continuous variables and PVC frequency was non-linear (all p-values >0.43).

6.4.2.4 Tests for multicollinearity

There was no evidence of multicollinearity between the predictor variables. The VIFs ranged from 1.09 to 1.55 which were below the threshold for concern (10). The CN was 19.74, which is below 30, the threshold for potential problems with collinearity. In addition, no more than one variable that loaded highly on any principal component, therefore, there was no evidence for multicollinearity between variables.

6.4.2.5 Analysis of residuals

The Anscombe residual was chosen for further diagnostic tests because it had the least skew and kurtosis compared to other residuals. When the quantiles of the Anscombe residuals were plotted against quantiles of the normal distribution, there were no obvious departures from normality (refer to Figure 6.10).

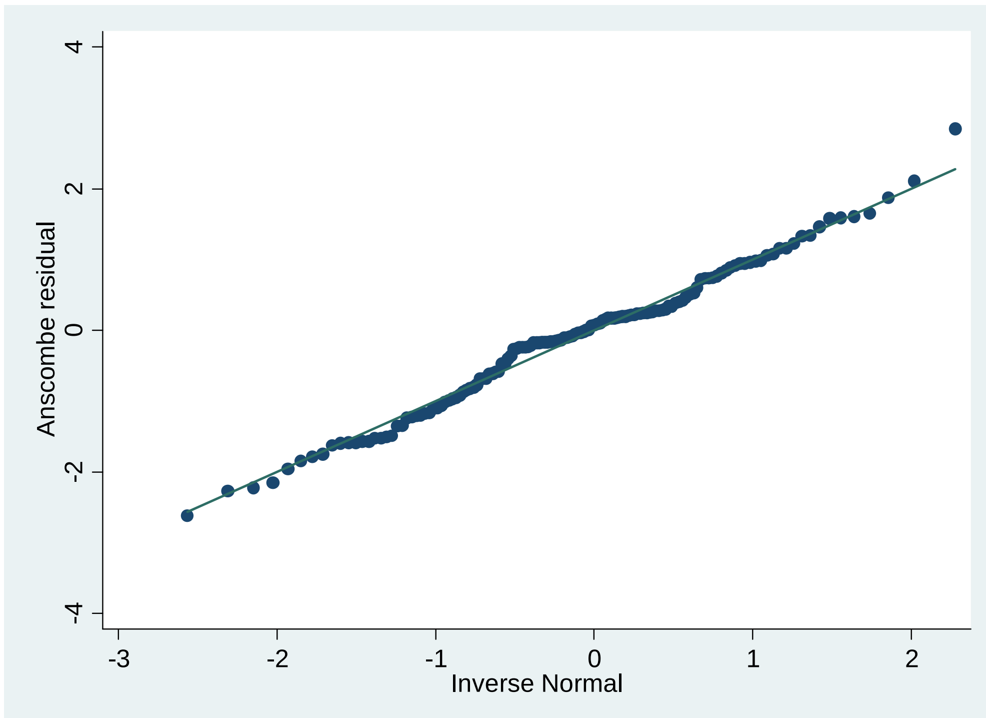


Figure 6.10 Quantiles of the Anscombe residual plotted against quantiles of the normal distribution

A scatter plot of the Anscombe residual against the predicted mean PVC with locally weighted scatterplot smoothing showed a lowess line that was flat and residuals that clustered around zero (refer to Figure 6.11) indicating no evidence of a relationship between the residuals and the predicted values.

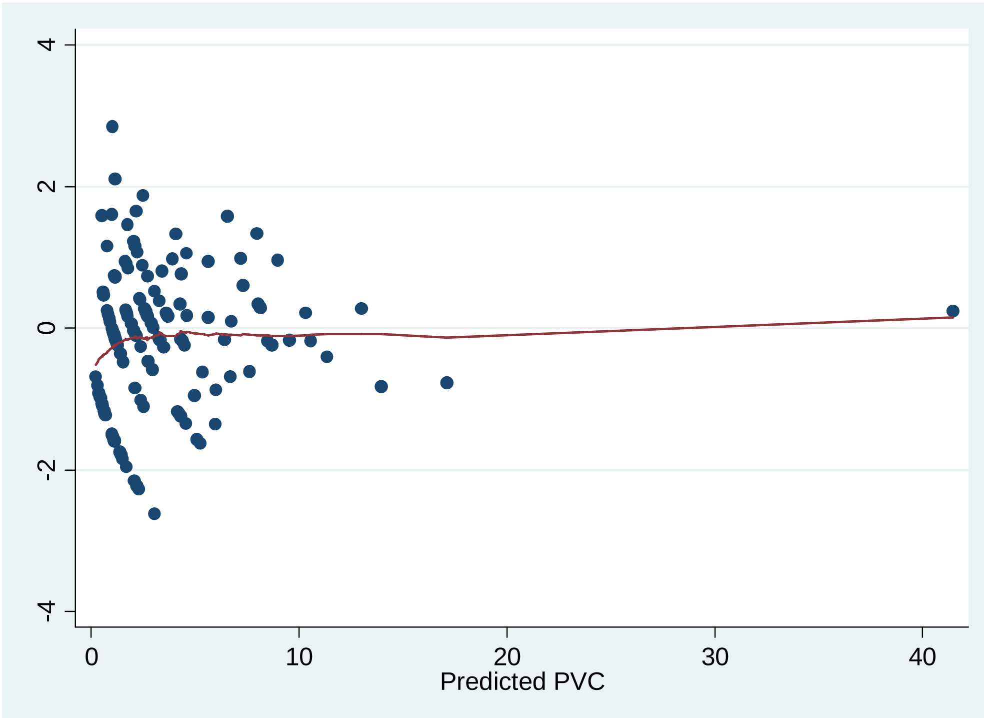


Figure 6.11 Locally weighted smoothed scatterplot of the Anscombe residual against the predicted mean number of pedestrian-vehicle crashes (PVC)

Five road segments with potentially high leverage were identified. When the Poisson regression model was re-fitted without these observations the IRRs did not change appreciably, however, the p-value for the number of late night liquor establishments per km increased from 0.0001 to 0.086 and that for comparing roads with speed zones of 40 or 50 km/h to zones of 60 km/h increased from 0.017 to 0.085. Considering these p-values still remained below 0.10, it was deemed that these two associations were still of reasonable interest even with the influential observations removed.

6.4.2.6 Overfitting

The PVC model was re-fitted using bootstrapped standard error estimates with 15,000 replications and the effect on the standard error estimates and p-values was determined. The only major change was that the PVC frequency in 80 km/h speed zones was no longer significantly different to those of other speed zones. For example, in the original model, the PVC frequency in 80 km/h zones was 70% lower than in 60 km/h zones (p=0.003), however, with bootstrapped standard error estimates, the p-value increased to p=0.70. Therefore, this result may not be generalisable to other samples of urban road segments. The p-value for the association between the number of signalised intersections per km and PVC frequency also increased from 0.08 to 0.13, however, this variable was only included in the model to adjust for the exclusion of crashes at signalised intersections, so this was not an issue of overfitting.

6.4.2.7 Sensitivity analysis of fixed vs. random parameters

The PVC model re-fitted with a random intercept to determine if there was any evidence that the baseline risk differed across sites. The PVC model with the random intercept did not fit the data significantly better than the PVC model with the fixed intercept. Hence the baseline risk of PVC was not significantly different across sites.

The PVC model would not converge with all random coefficients, so as a compromise, attempts were made to re-fit the PVC model a number of times. Each time the intercept and one of the variables were random and the other variables were fixed. Unfortunately, none of these models would converge. This is likely because the models were too complex for the data available.

6.4.2.8 Summary of model fit and diagnostic tests

To summarise, tests of model fit and diagnostics revealed that the model form was appropriate and the link function was properly specified. Observed probabilities matched the predicted probabilities well. There was no evidence for non-linearity of the relationship between the continuous predictor variables and the PVC frequency, nor was there significant multicollinearity between predictors. Five observations with high leverage were identified, however, the removal of these observations did little to change the regression estimates. The p-values for two variables did increase, however, it was felt that there was still reasonable evidence for an association even with the influential observations removed. When bootstrapped standard error estimates were used, 80 km/h zones no longer had significantly fewer PVC than 60 km/h zones: this association may not be generalisable to other samples of urban roads. There was no evidence that the baseline risk differed across sites so there was no need to include a random intercept.

6.4.3 Model results and interpretation

The characteristics of the strip shopping road segments in metropolitan Melbourne that were associated with PVC frequency ($p<0.05$) are presented in

Table 6.12, tabulated by broad category.

Table 6.12 Risk factors for pedestrian-vehicle crashes

Variable	N (%) or Median (Interquartile range)	Incidence Rate Ratio (95% Confidence Interval)
EXPOSURE		
No variables significantly associated with PVC frequency		
CROSS SECTION		
Carriageway width <=10m >10m	 10 (7.0) 132 (93.0)	 Reference 1.70 (1.04–2.78)
Lane width <=3.3m >3.3m	 93 (65.5) 49 (34.5)	 Reference 0.74 (0.59–0.93)
ACCESSES AND INTERSECTIONS		
Number of signalised intersections/km	2.74 (2.80)	*1.04 (0.99–1.10)**
Number of unsignalised intersections/km	8.47 (6.75)	1.04 (1.01–1.06)
MEDIANS		
Presence of median or median island (maximum median width>0m) No Yes	 40 (28.2) 102 (71.8)	 Reference 1.56 (1.24–1.97)
ROAD TYPE AND TRAFFIC MIX		
No variables significantly associated with PVC frequency		
ROADSIDE PARKING		
Presence of parking clearway No Yes	 86 (60.6) 56 (39.4)	 Reference 0.66 (0.53–0.82)
PUBLIC TRANSPORT AND BICYCLE FACILITIES		
No variables significantly associated with PVC frequency		
PEDESTRIAN FACILITIES		
Number of midblock pedestrian crossings per km	1.19 (2.77)	1.12 (1.05–1.19)
ROADSIDE		
Height of highest building <=2 storey > 2 storey	 77 (54.2) 65 (45.8)	 Reference 1.30 (1.06–1.58)
PAVEMENT CONDITION		
No variables significantly associated with MVC frequency		
HEIGHT CLEARANCE		
No variables significantly associated with MVC frequency		
STATIC ENFORCEMENT CAMERAS		
No variables significantly associated with MVC frequency		
AMENITIES AND FACILITIES		
Number of off-street parking facilities/km	2.07 (4.98)	1.04 (1.01–1.08)
Number of late night liquor licences/km	0 (1.78)	1.07 (1.04–1.11)

Variable	N (%) or Median (Interquartile range)	Incidence Rate Ratio (95% Confidence Interval)
Presence of railway station		
no	92 (64.8)	Reference
yes	50 (35.2)	1.44 (1.19–1.75)
Presence of tertiary education institution		
no	136 (95.8)	Reference
yes	6 (4.2)	2.12 (1.56–2.87)
SPEED LIMIT		
40 or 50 km/h	4 (2.8)	0.56 (0.35–0.90)
Variable 40/60 km/h	14 (9.9)	*0.88 (0.68–1.14)*
60 km/h	100 (70.4)	Reference
70 km/h	17 (12.0)	*0.81 (0.57–1.14)*
80 km/h	7 (4.9)	0.30 (0.14–0.66)

All variables were associated with PVC frequency at p<0.05. For categorical variables, the overall association (all categories considered together) with PVC frequency was <0.05, but associations for individual categories may not have reached the level for inclusion; as such, p-values >0.05 (where 95% CI includes one) shown in italics.
*p=0.09 but the rate of signalised intersections was kept in all models to adjust for exclusion of crashes at signalised intersections.

In the next sections, the results are interpreted in detail, by category. Variables that were not significantly associated with PVC in each category will also be identified.

6.4.3.1 Road user exposure

None of the variables measuring road user exposure were significantly related to PVC frequency. This included the estimated pedestrian activity, traffic volume per day or traffic density (vehicles per lane per day).

6.4.3.2 Road cross section

Road segments where the carriageway was more than 10m wide had 70% (95% CI 4%–178%) more PVC than roads with narrower carriageways. In contrast, wider lanes were associated with fewer PVC: roads with lanes wider than 3.3m had 26% (95% CI 7%–41%) fewer PVC than roads with narrower lanes.

The number of lanes and the presence of a curve, or curves, on the road segment were not significantly associated with the frequency of PVC.

6.4.3.3 Accesses and intersections

For every extra unsignalised intersection per km, the frequency of PVC increased by 4% (95% CI 1%–6%). The number of signalised intersections per km was included in the model to adjust for the exclusion of crashes that occurred at signalised intersections on the segment. Thus, the number of crashes would be expected to be negatively associated with the rate of signalised intersections. Yet there was a non-significant increase in PVC on midblocks as the rate of signalised intersections increased (IRR=1.04, 95% CI 0.99–1.10, p=0.09).

No other variables describing accesses and intersections were significantly associated with PVC (the number of driveways or laneways per km, the number of roundabouts per km, the number of keep clear zones per km, the number of dedicated turning lanes per km and the presence of service roads).

6.4.3.4 Medians

Medians were associated with increases in PVC frequency. The variable which led to the best fit was a binary variable describing maximum median width (where zero indicates no maximum median width and one indicates a maximum median width of greater than zero metres). Therefore, the variable essentially identifies road segments where there was a median island (e.g. a pedestrian refuge) or longer median present compared to road segments with no refuges or medians. The incidence of PVC was 56% (95% CI 24%–97%) higher when there was a median island or median present compared to when there was no median island or median.

6.4.3.5 Road type and traffic mix

The type of road (primary or secondary arterial), heavy vehicle access, over-dimensional routes and the proportion of traffic that was heavy vehicles were not significantly associated with PVC frequency.

6.4.3.6 Roadside parking

If there was a parking clearway present on one or both sides of the road, the frequency of PVC was reduced by 34% (95% CI 18%–47%). The presence and type of roadside parking and the number of loading zones per km were not significantly associated with the incidence of PVC.

6.4.3.7 Public transport and bicycle facilities

None of the variables relating to public transport and bicycle facilities were significantly associated with PVC frequency, including the presence of railway level crossings, the presence of bus routes or the number of bus stops, the presence of trams or the number of tram stops nor the presence and type of bicycle lanes.

6.4.3.8 Pedestrian facilities

Midblock pedestrian crossings were positively related to PVC. For every extra midblock pedestrian crossing per km, the incidence of PVC increased significantly by 12% (95% CI 5%–19%). There was, however, no association between the presence of pedestrian fencing at pedestrian crossings and PVC frequency.

6.4.3.9 Roadside

Development height was significantly related to the incidence of PVC. On road segments where the highest building was three storeys or more, PVC frequency was 30% (95% CI 6%–58%) higher than on road segments where the highest building was two storeys or less.

None of the other variables relating to the roadside were significantly associated with PVC frequency (presence of nature strips, the presence of shops on one or both sides of the road, predominant development height, the offset distance between the road and roadside buildings or the number of poles and tress on the roadside or median).

6.4.3.10 Pavement condition

The presence of pavement distress or roughness were not significantly associated with PVC frequency.

6.4.3.11 Height clearance

Height clearance (the presence of low tram wires or low clearance because of bridges) were not associated with PVC frequency.

6.4.3.12 Static enforcement cameras

The presence of static red light and speed enforcement cameras at signalised intersections on the road segments was not significantly associated with midblock PVC frequency.

6.4.3.13 Amenities and facilities

Off-street parking facilities were positively associated with PVC frequency. For each extra off-street parking facility per km, the incidence of PVC rose by 4% (95% CI 1%–8%).

Certain types of liquor licences were associated with PVC. For every extra establishment per km with a late-night liquor licence, there was a 7% (95% CI 4%–11%) increase in PVC frequency.

If there was a railway station on the road segment, or on a small access road that was only accessible from the road segment, the frequency of PVC was 44% (95% CI 19%–75%) higher than when there was no railway station present.

The incidence of PVC more than doubled if there was a tertiary education institution on the road segment compared to segments where there was no tertiary education institution (IRR=2.12, 95% CI 1.56–2.87).

None of the other amenities and facilities were significantly associated with PVC frequency, including liquor licences (apart from late-night liquor licences), the presence of kindergartens, child-care and maternal and child health facilities, schools (primary or secondary), various types of community centres, emergency services, hospitals, places of worship, indoor shopping centres, petrol stations and sports fields, sports centres or parks.

6.4.3.14 Speed limit

The speed limit of the road segment was significantly associated with PVC frequency. PVC frequency was highest in 60 km/h zones and reduced as the speed zone increased or decreased. PVC frequency in 80 km/h zones was significantly lower than all other speed zones except for 40 or 50 km/h zones: 63% (95% CI 15%–85%) lower than 70 km/h zones, 70% (95% CI 34%–86%) lower than 60 km/h zones, and 66% (95% CI 12%–85%) lower than in variable 60/40 km/h speed zones. When bootstrapped standard error estimates were used, however, the PVC frequency in 80 km/h zones was not significantly different to any other speed zone, so this result may not be generalisable to other samples of urban roads. The only other pairwise comparison that was statistically significant was that the PVC frequency in 40 or 50 km/h zones was 44% (95% CI 10%–65%) lower than in 60 km/h zones.

6.4.3.15 Summary of variables associated with pedestrian-vehicle crash frequency

PVC frequency was significantly higher when carriageways were more than 10m wide, when there was a median island or median present, when the maximum development height was 3 storeys or more, when there was a railway station present and when there was a tertiary education institution on the road segment. PVC incidence also increased with increases (per km) in the number of unsignalised intersections, the number of midblock pedestrian crossings, the number of off-street parking facilities, and the number of late night liquor licences.

The following characteristics of the road segment were associated with decreases in PVC frequency: lane width greater than 3.3m, the presence of a parking clearway, and low (40 or 50 km/h) or high (80km/h) speed zones.

6.5 Overall summary: characteristics of the built environment associated with multi-vehicle crashes, single-vehicle crashes and pedestrian-vehicle crashes

All three crash types were directly proportional to the number of days and segment length so these variables were included in the offset term, therefore the risk being measured was the number of casualty crashes per km per day (transport injury risk). The phased approach to model building worked well for all crash types; the models converged and a solution was reached with relatively few iterations of post-hoc testing. The model forms were appropriate and the models fitted the data well. There was no evidence of multicollinearity between variables in the models or that the models were overfitted. There was also no evidence that random parameters provided better model fit than fixed parameters in any of the models. An overall summary of the road segment characteristics associated with each of the three crash types are presented in Table 6.13 to enable comparison between the risk factors for different crash types. Recall that the p-value for inclusion in the MVC model was set at $p<0.01$ to avoid overfitting, whereas the p-value for inclusion in the SVC and PVC models was $p<0.05$. For comparison, the variables that were associated with MVC frequency at $p<0.05$ are also displayed, shaded in light grey. The results, including differences between the models will be discuss in Chapter 7.

Table 6.13 Risk factors for multi-vehicle crashes (MVC), single-vehicle crashes (SVC) and pedestrian-vehicle crashes (PVC)

	MVC	SVC	PVC
Model form	Negative binomial	Poisson	Poisson
Variable	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)
EXPOSURE			
ln(traffic volume/1000 per day)	2.02 (1.61–2.53)	2.43 (1.58–3.73)	
Pedestrian activity	1.03 (1.01–1.04)		
CROSS SECTION			
Minimum carriageway width <=10m >10m	 1.39 (1.14–1.71) Reference		
Minimum carriageway width 5.5–8.5m 8.5–10m >10m		 1.97 (1.25–3.12) *0.64 (0.15–2.68)* Reference	
Mode carriageway width <=10m >10m			 Reference 1.70 (1.04–2.78)
Lane width <=3.3m >3.3m			 Reference 0.74 (0.59–0.93)
Presence of curve/s No Yes	 Reference 0.77 (0.62–0.94)	 Reference 2.33 (1.49–3.64)	
ACCESSES AND INTERSECTIONS			
Number of signalised intersections/km	*0.96 (0.93–1.00)**	*0.96 (0.88–1.06)**	*1.04 (0.99–1.10)**
Number of unsignalised intersections/km	1.04 (1.02–1.06)		1.04 (1.01–1.06)
Number of driveways or laneways/km	0.992 (0.987–0.997)		
Number of roundabouts/km	1.25 (1.13–1.38)		

	MVC	SVC	PVC
Model form	Negative binomial	Poisson	Poisson
Variable	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)
Service road present			
No	Reference		
Yes	1.53 (1.16–2.03)		
MEDIANS			
Maximum median width			
No median	Reference		
<1.2m	*1.15 (0.88–1.50)*		
1.2–3m	*0.87 (0.71–1.07)*		
>3m	0.66 (0.50–0.86)		
Presence of median or median island (maximum median width>0m)			
No			Reference
Yes			1.56 (1.24–1.97)
ROAD TYPE AND TRAFFIC MIX			
Over-dimensional route			
No	Reference		
Yes	1.56 (1.16–2.09)		
Primary state arterial			
No	Reference		
Yes	1.22 (1.02–1.47)		
ROADSIDE PARKING			
Roadside parking			
None	Reference		
1 side	*0.99 (0.69–1.41)*		
2 sides	1.47 (1.08–1.98)		
2 sides and centre	0.46 (0.22–0.95)		
Presence of parking clearway			
No			Reference
Yes			0.66 (0.53–0.82)

	MVC	SVC	PVC
Model form	Negative binomial	Poisson	Poisson
Variable	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)
PUBLIC TRANSPORT AND BICYCLE FACILITIES			
Bicycle lane width None <1.2m >1.2m		 Reference 2.19 (1.33–3.61) 0.50 (0.25–0.97)	
Shared lane with trams No Yes	 Reference 1.24 (1.06–1.46)	 Reference 2.04 (1.37–3.05)	
Number of bus stops per km		1.06 (1.01–1.11)	
PEDESTRIAN FACILITIES			
Number of midblock pedestrian crossings per km			1.12 (1.05–1.19)
ROADSIDE			
Nature strip present No 1 side 2 sides	 Reference 0.60 (0.45–0.80) 0.52 (0.39–0.68)		
Predominantly 2 storey development No Yes	 Reference 0.76 (0.64–0.91)		
Height of highest building <=2 storey > 2 storey			 Reference 1.30 (1.06–1.58)
Number of non-frangible poles on the roadside/km	0.997 (0.995–0.999)	1.006 (1.001–1.010)	
PAVEMENT CONDITION			
Presence of distress No Yes		 Reference 0.69 (0.50–0.95)	
HEIGHT CLEARANCE			

	MVC	SVC	PVC
Model form	Negative binomial	Poisson	Poisson
Variable	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)
Presence of low tram wires No Yes		Reference 0.31 (0.11–0.90)	
STATIC ENFORCEMENT CAMERAS			
Not significantly associated with MVC, SVC or PVC frequency			
AMENITIES			
Number of off-street parking facilities/km	0.95 (0.92–0.97)	0.91 (0.85–0.97)	1.04 (1.01–1.08)
Number of late night liquor licences/km	1.08 (1.05–1.12)		1.07 (1.04–1.11)
Number of bring your own liquor licences/km		1.08 (1.03–1.13)	
Presence of sports centres, fields, parks No Yes		Reference 0.61 (0.39–0.94)	
Presence of railway station no yes			Reference 1.44 (1.19–1.75)
Presence of tertiary education institution no yes			Reference 2.12 (1.56–2.87)
Presence of petrol station No Yes	Reference 0.86 (0.76–0.98)		
Presence child-care, kindergarten or maternal child health centre No Yes	Reference 1.39 (1.15–1.68)		
SPEED LIMIT			

	MVC	SVC	PVC
Model form	Negative binomial	Poisson	Poisson
Variable	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)	Incidence Rate Ratio (95% Confidence Interval)
40 or 50 km/h			0.56 (0.35–0.90)
Variable 40/60 km/h			*0.88 (0.68–1.14)*
60 km/h			Reference
70 km/h			*0.81 (0.57–1.14)*
80 km/h			0.30 (0.14–0.66)

*p>0.05 but the rate of signalised intersections was kept in all models to adjust for exclusion of crashes at signalised intersections. Additional variables associated with MVC frequency when the p-value was set at p<0.05 (rather than p<0.01) are shaded light grey.

CHAPTER 7. DISCUSSION: IDENTIFYING RISK FACTORS

In this chapter, the results of Component 1 will be summarised and interpreted with reference to the aims of the study and compared to the findings of relevant previous research. Methodological issues, including the strengths and limitations of the current study, will be discussed. Finally, the implications for future research (including the second research component of this thesis) will be presented.

7.1 Achievement of primary aims

The review of the literature in Chapter 3 established that, beyond the effect of traffic volume and accesses and intersections, there was a lack of strong evidence about the influence of the built urban environment on crash risk. Moreover, the possible influence of the surrounding built environment external to the roadway was previously overlooked. This thesis was therefore novel in the development of a list of characteristics of the built environment, including the road, roadside and human activity, with the potential to influence crashes (which addressed the first aim of the study). The characteristics were selected because of their potential to influence risk through their effect on the field of safe travel, safety margins, complexity of the environment and the number and movement of road users. The list was used to define the data requirements for this study. The second aim of the study was also achieved, as sources of data were found for almost all of the data items in the list. Hence the first two aims of the study, which were methodological, were achieved.

With reference to the third aim, the results of the current study revealed that, on strip shopping centre road segments, casualty crash frequency was related not just to the traditionally studied risk factors such as traffic volume and road design features, but also the physical characteristics of the roadside and amenities and features in the area that impact the number and type of road users and their behaviour. In addition, different risk factors were found to be associated with different crash types (re aim 4). Unfortunately, however, no data could be found to estimate bicycle exposure on the study road segments so a separate model was not developed for BVC.

7.2 Summary and interpretation of results

The characteristics of the built environment that were found to be associated with different types of crashes on strip shopping centre road segments in metropolitan Melbourne will be discussed compared to the results of relevant previous studies. While it is tempting to offer explanations for why the identified risk factors affect crashes, as emphasised throughout the thesis, cross-sectional studies can only establish associations and other research methods are required to investigate causative mechanisms. As such, possible mechanisms for the observed associations are proposed but only in terms of hypotheses to be tested using other research methods rather than as stand-alone post-hoc explanations of the results.

When discussing the results, it must be kept in mind that the outcome in this study was the frequency of casualty crashes (obtained from the police-reported crash data held by the Victorian state road authority). This means that the risk being measured in this study was the risk of *injurious* crash occurrence (transport injury risk) rather than the risk of crash occurrence (transport risk), per se. Therefore the risk factors that were identified may be related to the probability of crash occurrence, the probability of an injury, given a crash occurred, or a combination of the two.

7.2.1 Road user exposure and traffic distribution

As expected, and similar to what has been found in previous research (Abdel-Aty et al., 2009; Avelar et al., 2013; Bonneson & McCoy, 1997; El-Basyouny & Sayed, 2009; Greibe, 2003; Haynes et al., 2008; Jonsson, 2005; Manuel et al., 2014; Potts et al., 2007; Sawalha & Sayed, 2001), an increase in the AADT led to a non-linear increase in the frequency of MVC and SVC.

In addressing the secondary aim of this research component, it was discovered that the model for MVC developed using AADT resulted in the same set of predictors as the model developed using AADT per lane, except that the number of lanes was also significantly associated with MVC in the latter model. Changing the exposure measure from AADT to AADT per lane and the number of lanes had no substantive effect on the other parameter estimates, although the fit of the model was slightly poorer.

No model was developed for SVC with traffic density as the exposure variable because it was not significantly related to SVC during phase 2 of the model development process when the best exposure variables were chosen. If, however, AADT per lane and the number of lanes were substituted into the SVC model that was developed using AADT during Phase 3 (the final phase), both were significantly associated with SVC and there was negligible change in the parameter estimates for the other variables in the model. Thus, AADT per lane and the number of lanes can be substituted into a model for AADT with no impact on other estimates. The choice of which to use in future studies will depend on whether the effect of total traffic exposure, or the independent contribution of traffic per lane and the number of lanes, are of interest.

No significant relationship was found between PVC and traffic volume (AADT), traffic density (AADT/lane) or pedestrian activity. In contrast, PVC were related to both AADT and pedestrian volumes in previous research. On road links in Sweden, PVC increased in a non-linear fashion with increases in AADT and pedestrian volumes (Jonsson, 2005), while PVC that occurred during the day in the Melbourne CBD were related to the product of AADT and pedestrian crossing volumes (Alavi, 2013). It is highly likely that the measure of pedestrian activity used in this study (estimated from travel survey data measured at the area-level) did not accurately reflect pedestrian volumes on specific strip shopping centre road segments located in that area. This may explain the lack of relationship between estimated pedestrian activity and PVC in this study. It is unclear, however, why AADT was not related to PVC frequency in this study. Future research that includes pedestrian crossing volumes measured for the specific road segments should shed light on whether the interactive relationship between AADT and pedestrian volumes found in the Melbourne CBD is also observed for complex urban road segments in metropolitan Melbourne.

7.2.2 Roadway

7.2.2.1 Number of lanes

Previous research has been equivocal regarding the association between the number of lanes and urban crashes (controlling for traffic volume) with some studies finding a positive association (Abdel-Aty et al., 2009; El-Basyouny & Sayed, 2009; Sawalha & Sayed, 2001) and some finding no association (Alavi, 2013; Greibe, 2003;

Jackett, 1993; Jonsson, 2005). In this study, the number of lanes was not significantly associated with MVC, SVC or PVC, controlling for AADT. When traffic density (AADT per lane) was included as the traffic exposure variable instead of traffic volume (AADT), however, the number of lanes was positively associated with MVC and SVC. This is not surprising, given that the total number of vehicles per day (AADT) is equal to the number of vehicles per lane multiplied by the number of lanes. Hence, if traffic density is included in the model instead of traffic volume, the number of lanes becomes a surrogate exposure variable.

7.2.2.2 Carriageway width and lane width

The carriageway width is defined as the distance from kerb to kerb and is therefore a combined measure of the number of lanes, the lane width and whether or not the road has a concrete median. The results for carriageway width and lane width will therefore be discussed together. Results were similar for MVC and SVC but differed for PVC, so they will discussed in turn.

Roads with a minimum carriageway width of more than 10m had fewer MVC and SVC than roads with a narrower minimum carriageway width. For most roads (123; 87%), the width did not change throughout the road segment. For 19 of the 22 roads with a minimum carriageway width of 10m or less, however, the width changed throughout the segment. A prior study found that a change in roadway width was associated with an increase in crashes (aggregated by type) on two lane urban residential collector roads (Manuel et al., 2014). It is possible that it was also the change in carriageway width that influenced risk in this study. Post-hoc analyses revealed that although the minimum carriageway width variable could be replaced in the MVC and SVC models with a variable that indicated the presence of a change in width somewhere in the segment, the model fit was better with minimum carriageway width. Wider roads were associated with reduced frequency of vehicle crashes which may have been due to the resultant wider field of safe travel, and therefore increased safety margins in terms of lateral clearance to potential obstacles including other vehicles.

For a given carriageway width, lane width was not independently associated with MVC or SVC. While narrower lanes (even for a given carriageway width) would be expected to impact the field of safe travel by reducing safety margins in terms of the lateral clearance to other vehicles, it is possible that the drivers on the road segments in this study felt that risk was higher when lanes were narrow and adapted their behaviour by reducing travel speed (as predicted by the multiple comfort zone model (Summala, 2007)). Lewis-Evans and Charlton (2006) showed that drivers in a simulator reduced their speed and moved further away from the kerb when lane width (and road width) was reduced in a simple rural environment—further investigations should be performed to determine whether this occurs in more complex environments, as proposed from the results of this thesis.

In contrast, when the mode carriageway width was greater than 10m, the frequency of PVC increased by 70% compared to narrower carriageways. Hence, wider roads were associated with increased frequency of PVC. This may be related to exposure as the greater the distance between kerbs, the greater the distance (and time) that a crossing pedestrian is exposed to the risk of being hit by vehicles travelling along the road (Lassarre, Papadimitriou, Yannis, & Golias, 2007).

Lane width was also independently associated with PVC, controlling for carriageway width. For a given carriageway width, roads with lanes wider than 3.3m had 26% fewer PVC than roads with narrower lanes. This is a novel finding, as the previous studies of PVC in urban areas did not assess the influence of lane width. Thus, wider lanes (given a particular road width) reduce the risk of PVC. It is possible that wider lanes provide more space for both vehicles and pedestrians to negotiate a path of safe travel to avoid a collision when pedestrians are on the roadway attempting to cross. This hypothesis could be tested in observational studies. It is also possible that pedestrians are more likely to avoid crossing at a midblock location on roads with wide lanes. Accurate data on pedestrian crossing volumes is required to test this hypothesis.

Even though carriageway width must, to some extent, be related to the lane width, the two were not correlated (at the level of measurement used in this study) which is why they were both permitted to be tested for entry into the same model. Future research on the relationship between road and lane width and urban crashes should try to tease apart the separate influences of the number of lanes, road width, lane width, and whether the road has a concrete median. Measurement of lane and road width at a finer resolution than the categories used in this study may be required.

7.2.2.3 Horizontal curves

Surprisingly, only one previous multivariable study investigated curves and crashes on urban road segments and failed to find an association with crashes, aggregated by type (Manuel et al., 2014). In the present study, the frequency of SVC more than doubled when the road segment was not straight, indicating that curved or crooked road segments may contribute to a loss of control. There was, however, no relationship with MVC or PVC even though curves have the potential to obstruct vision of pedestrians or turning cars (Gibson & Crooks, 1938). Only 17 (12%) of the road segments in this study were not straight and the type of curvature varied from a slight curve to a 90° change in direction. This variation precluded the investigation of whether the type and degree of curvature affected SVC risk, as has been found in previous studies in rural areas (e.g. Miaou, 1994; Milton & Mannering, 1998; Shankar et al., 1995). Future research should be conducted to further explore the relationship between the degree and type of road curvature and SVC risk on urban roads.

7.2.2.4 Medians

There were several variables that described medians; presence, type, width and the proportion of the road with a median. Most of the previous multivariable studies assessed the influence of median type and presence on crashes in urban areas and the results have been equivocal (Alavi, 2013; Avelar et al., 2013; Bonneson & McCoy, 1997; Brown & Tarko, 1999; Greibe, 2003; Jackett, 1993; Jonsson, 2005; Sawalha & Sayed, 2001). In this study, maximum median width was the median-related variable that led to the best fitting models. Maximum median width was significantly associated with MVC and PVC, but not SVC.

On roads with a median (or traffic island) with a maximum width greater than 3m, MVC frequency was significantly lower than if the maximum median width was less than 3m. Wide medians are usually made from raised concrete, or on some road segments with wide medians, the median contains a tram lane separated from the traffic. These road features decrease potential for conflict between traffic moving in opposite directions and also reduce U-turn manoeuvres, which were a common MVC type.

In contrast, if any median or median island was present the frequency of PVC increased by 56% (adjusted for the other variables in the model, e.g. carriageway width). This result seems counterintuitive as medians and islands provide a refuge for crossing pedestrians and therefore might be expected to improve pedestrian safety. It is possible, however, that this result may be related to pedestrian crossing volumes. On roads with no medians or median islands, pedestrians may choose to cross at signalised intersections, hence reducing exposure at midblock locations. Pedestrians may be more likely to cross road segment midblocks with medians and median islands than those without because they provide a refuge, thus increasing exposure to risk while crossing.

Yet, even though medians provide a refuge, they provide no protection for a pedestrian if a vehicle hits, or mounts, the refuge. It is also possible that drivers are less cautious of crossing pedestrians when there is a refuge present. It is impossible to measure pedestrian crash risk at the level of exposure to hazards (as per the conceptual framework of the thesis) without data on the number of pedestrians who cross the road. These hypotheses could form the basis for future research into the relationship between medians, median islands and pedestrian and driver behaviour, as long as data are collected for pedestrian exposure.

7.2.2.5 Accesses and intersections

Accesses and intersections increase the opportunity for conflict between road users travelling in different directions, or those vehicles attempting to join, leave or cross the traffic stream. This impacts the field of safe travel and safety margins, in addition to increasing task demands in terms of the requirement to monitor a number of moving entities from different directions. Crashes involving cross traffic are also likely to be more severe and therefore more likely to result in a casualty that is reported in the Victorian crash database. Previous research found a strong positive relationship between the number of accesses and intersections and the frequency of crashes aggregated by type and MVC in urban areas (Bonneson & McCoy, 1997; Brown & Tarko, 1999; El-Basyouny & Sayed, 2009; Greibe, 2003; Jackett, 1993; Jonsson, 2005; Manuel et al., 2014; Sawalha & Sayed, 2001; Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013).

In the current study, SVC were not associated with any of the variables relating to accesses and intersections. This makes intuitive sense, as the frequency of SVCs should not be influenced by road features that increase conflict between moving road users.

In contrast, several of the variables describing accesses and intersections were significantly related to MVC. As the number of unsignalised intersections and roundabouts on a road segment increased, so did the frequency of MVC. This may be explained by the increase in potential conflict between road users. A positive relationship between unsignalised intersections and MVC has been reported previously (Jonsson, 2005), however none of the previous studies investigated the effect of roundabouts on urban crash frequency, so this finding is new.

Another new finding emanating from this study is that the presence of service roads was also associated with increased MVCs (on the through road—crashes that occurred on the service road were not included in this study). Although service roads remove vehicles accessing businesses from the traffic stream (and thus reduce the risk associated with manoeuvring and parking vehicles), service roads are often located on higher speed roads. Because manoeuvring traffic is removed from the traffic stream, the speed of traffic on the through road is likely to be higher. Hence, those vehicles slowing to enter the service road, or those trying to re-join the main thoroughfare may have to contend with high speed traffic, which may increase the frequency of injurious MVC.

Prior studies found that as the number of driveways increased (particularly in commercial areas), the number of crashes (aggregated by type) increased (Avelar et al., 2013; Sawalha & Sayed, 2001; Xu, Kouhpanejade, et al., 2013; Xu, Kwigizile, et al., 2013). In this study, unexpectedly, as the number of driveways/laneways increased, the MVC frequency decreased, even though these are locations at which road users may come into conflict with each other. When entering the road from a driveway or laneway, visibility of the through traffic is often obstructed, and likewise, visibility of the entering vehicle for through traffic is often reduced. It is possible, however, that the reduction in crash risk with driveways and laneways could be because drivers were more cautious when there were frequent driveways and lane ways—this issue warrants further investigation.

The number of PVC increased significantly as the number of unsignalised intersections per km increased which was in conflict with earlier studies that found no association (Alavi, 2013; Jonsson, 2005). This positive relationship is not unexpected, however, as pedestrians crossing a minor road are exposed to traffic turning into, and out of, the minor road. Drivers turning at unsignalised intersections may be concentrating on the traffic rather than crossing pedestrians. PVC were not associated with roundabouts, service roads or driveways and laneways. The latter result is unexpected, as pedestrians crossing laneways and driveways would also be exposed to turning traffic and the frequency of PVC that occurred during the day in the Melbourne CBD was positively related to driveway density (Alavi, 2013). It is possible, however, that the traffic on the strip shopping centre road segments in this study was driving more slowly when entering and exiting narrow openings, which could reduce the risk of injurious PVC compared to other types of unsignalised intersections. This hypothesis could be investigated in an observational study.

Finally, the number (per km) of keep clear zones, dedicated right turn lanes or dedicated left turn lanes were not significantly associated with any crash type.

7.2.2.6 Road type and traffic mix

The distribution of traffic, in terms of the proportion of heavy vehicles, has only been included in one previous study in which it was not found to be associated with PVC in the Melbourne CBD (Alavi, 2013). The current study adds to the meagre evidence that the percentage of heavy vehicles does not influence crashes in urban areas, as none of the three crash types (MVC, SVC or PVC) were significantly associated with the proportion of traffic that were heavy vehicles.

The type of arterial (primary or secondary) and whether or not access to the road segment was approved for heavy vehicles were also not associated with any of the crash types. If over-dimensional vehicles were permitted on the road segment, however, the frequency of MVC was significantly higher. Over-dimensional routes are wide, busy roads that are capable of accommodating over-size vehicles, however carriageway width and traffic volumes were controlled for in the MVC model, so there must have been additional factors driving the increased risk (for example, different aspects of road design or traffic distribution). There were only seven road segments that were classified as over-dimensional routes in this study which limits the opportunity to investigate if they differ systematically from other road segments. A larger study of crashes on urban arterials (not restricted to strip shopping segments) would be needed to investigate this issue more thoroughly.

7.2.2.7 Roadside parking

Roadside parking may contribute to risk in a number of ways. Parked (unattended) vehicles are stationary objects located close to the traffic stream that moving vehicles may collide with. Drivers must monitor and adjust their speed or travel path when vehicles enter and exit parking bays, or when car doors are opened. Parked vehicles can obstruct a motorist's view of other road users, for example, pedestrians who are about to cross the road, and vehicles turning into the road. The view of oncoming traffic for vehicles waiting to turn into the road can also be obstructed by parked vehicles. Hence, roadside parking could potentially contribute to all three crash types: MVC, SVC and PVC. The evidence from past studies was equivocal with regard to the relationship between roadside parking and crash frequency (Alavi, 2013; Bonneson & McCoy, 1997; Greibe, 2003; Jonsson, 2005; Potts et al., 2007; Sawalha & Sayed, 2001).

In this study, MVC frequency increased when roadside parking was present on both sides of the road, however, when there was also parking in the centre median, MVC risk was reduced. Roadside parking was not significantly associated with SVC frequency. Specific crash types were inspected in an attempt to shed light on these findings. A relatively large proportion (42%) of SVC were crashes that involved a collision into an object or parked vehicle, and although it was impossible to determine whether it was a parked vehicle or another object, the presence of roadside parking on a road segment apparently did not influence SVC frequency. In contrast, approximately 5% of MVC involved stationary objects or parked vehicles, 5% of MVC involved road users entering or exiting a vehicle and only 3% of MVC involved a collision with a vehicle manoeuvring into or out of a parking space. The discovery of a significant positive relationship between roadside parking and MVC, even though only a small proportion of MVC were directly related to parked or parking vehicles, suggests that roadside parking may affect other types of MVC too. These results warrant further investigation using behavioural research methods similar to that demonstrated in the second part of this thesis.

While roadside parking, per se, was not associated with PVC, the frequency of PVC was significantly lower on road segments that had a parking clearway (which usually means that parking is prohibited during peak travel times). It is possible that pedestrians are easier to see, and pedestrians have a clearer view of oncoming traffic, when there are no parked vehicles and hence PVC risk is reduced when there are parking clearways. This hypothesis could be tested using behavioural research methods (observational or experimental). It could also be that pedestrians avoid crossing arterial strip shopping centre midblocks that are busy enough to have a parking clearway during peak traffic times (high traffic volumes dissuade pedestrians from crossing at midblock locations (Papadimitriou, Yannis, & Golias, 2014)). Measurement of pedestrian crossing volumes would be required to establish whether this is the case for these strip shopping centre road segments.

7.2.2.8 Public transport and bicycle facilities

It might be expected that the presence of buses and trams (and the number of stops) would be related to MVC and PVC. Public transport vehicles frequently stop to pick up and drop off transport users and drivers must monitor frequently stopping buses and trams. Drivers must also monitor pedestrians boarding or alighting public transport vehicles, or crossing the road (perhaps dangerously when in a rush) to catch public transport. Railway level crossings are also locations of potential conflict; both between vehicles and trains, and rear-end crashes involving vehicles stopped at a level crossing. It was expected that SVC would be relatively uninfluenced by buses, trams and level crossings.

In this study, however, public transport facilities were not significantly associated with MVC or PVC (apart from the presence of train stations, which will be discussed in Section 7.2.4). Surprisingly, SVC frequency was positively associated with the number of bus stops and the presence of a lane shared by both trams and other road vehicles. The first result is puzzling and requires replication and further investigation. There are several possible explanations for the second result that could be explored in future research. On roads with shared tram lanes, tram stops are located in the middle of the road, some of which have barriers and other roadside furniture that an out of control vehicle may crash into, which could contribute to SVC resulting in injury. Tram tracks can be slippery and affect braking, especially in wet weather. For this reason, drivers with experience driving on roads with tram tracks choose a lane position so that the wheels of the vehicle are not on the tram tracks. Drivers without experience driving on roads with tram tracks may not realise the importance of lane position and hence may be at risk of loss of control when braking suddenly or in wet weather. In addition, shared tram lanes may increase the risk of moving vehicles colliding with parked vehicles. On four lane undivided roads with a shared tram lane where roadside parking is permitted in the left lane, traffic is often restricted to driving in the shared tram (right) lane which can result in a line of traffic forming behind the frequently stopping tram (see Figure 7.1). It is illegal to pass a stationary tram with the doors open for boarding or alighting passengers. Hence, there are limited opportunities to pass a tram on these types of roads. Drivers following a tram must wait for a stretch of road with no parked vehicles before they can attempt to pass the tram on the left whilst the tram is moving. They may, however, misjudge how much time they have to pass the tram because modern trams accelerate much faster than older trams, which reduces the time available to pass. A frustrated driver who attempts to pass a tram in the outer lane when there is a parked vehicle ahead is therefore at risk of colliding with the parked vehicle if they misjudge the distance available and the speed of the tram. Hence, shared tram lanes may contribute to SVC in a number of ways—these hypotheses could be investigated in future research.

Figure 7.1 Four-lane urban arterial in Melbourne with shared tram lane and roadside parking, photograph used with permission from Edquist, J.

The association between bicycle facilities and crash frequency on urban road segments has rarely been investigated in multivariable studies. The few previous studies that included this risk factor found no association between the presence of bicycle lanes and crash frequency (aggregate, MVC, SVC or PVC) (Alavi, 2013; Greibe, 2003; Jonsson, 2005). The current study also found no association between bicycle lanes and MVC (which included those where bicycles were involved) or PVC. In contrast, the presence and width of bicycle lanes were significantly associated with SVC. Roads with a narrow (<1.2m) bicycle lane had more than double the frequency of SVC compared to roads with no bicycle lane, while roads with a wide (>1.2m) bicycle lane had approximately half the SVC risk of roads with no bicycle lane. It should be noted that the latter result was no longer significant once observations with high leverage were removed or bootstrapped standard error estimates were used.

Further investigation was conducted to determine why bicycle lane presence and width might impact SVC frequency. The roads in this study with narrow bicycle lanes were all four-lane roads with parking permitted on both sides and the majority had parking clearways during peak travel periods. For all except one road, the bicycle lane was located in the same space as the parking bays (refer to Figure 7.2). The majority of these roads also required vehicles and trams to share the middle lane. Thus these were complex road environments with little space to manoeuvre when circumstances went awry. In comparison, almost all of the wide bicycle lanes observed in this study were for bicycle use only (that is, they were not shared with parked vehicles). On roads with wide bicycle lanes, trams were rare and although most roads had roadside parking, not many had clearways. It is possible that the wide bicycle lane provided a buffer zone for out of control vehicles to recover before potentially hitting a parked vehicle or object (hence the association with reduced SVC frequency). A study of the influence of the type of bicycle lanes on crashes on urban segments in the greater Melbourne metropolitan area (not just strip shopping centres) could be conducted to try to replicate this finding and, if replicated, to investigate potential reasons.

Figure 7.2 Four-lane urban arterial in Melbourne with bicycle lane that shares space with roadside parking, photograph used with permission from Smith, S.

7.2.2.9 Pedestrian facilities

Previous multivariable studies provided weak to medium strength evidence that pedestrian crossings are associated with increases in crash frequency (crashes aggregated by type)—these studies did not control for pedestrian crossing volumes (El-Basyouny & Sayed, 2009; Sawalha & Sayed, 2001). Two previous studies focused exclusively on PVC and did control for pedestrian crossing volumes: one found that PVC increased along with the number of pedestrian crossings (Jonsson, 2005), while another found no relationship (Alavi, 2013). In the current study, there was no significant association between the number of midblock pedestrian crossings per km and MVC or SVC frequency, however, an increase in the number of midblock pedestrian crossings was associated with a significant increase in PVC. It is possible that the increase in PVC risk was due to increased exposure, as pedestrian crossing volumes would be expected to be higher at midblock pedestrian crossings than at other midblock locations. Given that previous research (Jonsson, 2005) has also found that pedestrian crossings increase PVC even when pedestrian crossing volumes are controlled for, however, there may also be a residual increase in risk not accounted for by pedestrian crossing exposure . Endogeneity is a possible explanation—this is discussed in Section 7.3.2.3. These issues should be explored in future research.

This study also explored whether the presence of pedestrian fencing adjacent to the pedestrian crossings was related to crash frequency, a feature that has not previously been studied in this manner. There was, however, no significant association between pedestrian fencing at pedestrian crossings and the frequency of MVC, SVC or PVC.

7.2.2.10 Pavement condition, height clearance and enforcement cameras

Rough or distressed pavement may increase crash risk through reduced vehicle braking capabilities or compromised vehicle handling. Alternatively, drivers may travel at a slower speed over rough or distressed pavement in order to provide a more pleasant travel experience (that is, to remain in their comfort zone) which may result in a reduction in crash risk. None of the prior multivariable studies investigated the relationship between pavement condition and urban crashes. In this study, the presence of rough pavement on the road segment was not significantly associated with any crash type. The presence of pavement distress was not related to MVC or PVC, however, there were 31% fewer SVC on roads with distressed pavement than roads without distressed pavement providing some support for the hypothesis of reduced travel speeds to maintain comfort.

Low height clearance limits the type of vehicles that can use a road. In addition, the presence of bridges or low tram wires could make the environment appear more cluttered, or complex. The presence of low clearance due to bridges was not significantly associated with any of the three crash types. The presence of low tram wires was significantly associated with reductions in SVC but not MVC or PVC. The association with SVC was no longer statistically significant when observations with high leverage were removed, or bootstrapped standard error estimates were used, thus, this result must be viewed with caution and requires replication.

Static red light and speed enforcement cameras were located at 25 intersections within the 142 road segments in this study. All were signed. Formal evaluation has found they are effective at reducing crashes at Victorian intersections (Budd, Scully, & Newstead, 2011). Though crashes at signalised intersections were excluded from this study, it is possible that enforcement cameras at intersections may affect crashes on nearby midblocks through their effect on driver behaviour and speed choice. It seems, however, that any effects are localised to the intersection because there was no association found between enforcement cameras at signalised intersections and midblock crashes in this study.

7.2.2.11 Speed limit

Speed zones are set to manage risk on the road network and are chosen with reference to the characteristics of the road, the extent and nature of abutting development and road user movements and potential for conflict (VicRoads, 2010). Therefore, the statistical models in the current study already included many of the features that influence speed limits so speed limit would not be expected to be independently related to crash risk. Assessing the effect of speed limits is not possible if the other variables are proxies for speed limit. Previous multivariable studies of crashes in urban areas were equivocal as to the role of speed limits, possibly due to inadequate control of potential confounders. The current study found no significant association between speed limit and MVC or SVC. There was, however, a significant association with PVC which may suggest that speed limit setting practices do not fully take into account the factors that affect the risk of injurious pedestrian crashes occurring in strip shopping zones.

PVC frequency was significantly lower on 80 km/h roads than for all other speed zones. This result, however, was no longer statistically significant when bootstrapped standard error estimates were used, indicating it may not be generalisable to other samples. It is also possible that this result was due to confounding with pedestrian exposure. None of the 80 km/h roads had midblock pedestrian crossings and pedestrians are unlikely to choose to cross at uncontrolled midblock locations on high speed roads (Papadimitriou et al., 2014). The frequency of PVC involving injury was also 44% lower on roads with either a 40 km/h or 50 km/h speed limit compared to roads with a 60 km/h speed limit. This result is unlikely to be due to confounding with pedestrian exposure because lower speed limits are applied in areas of high pedestrian activity. It is possible that this reduction is due to both the reduced probability of a crash occurring and the lower probability of the pedestrian being injured if a crash occurs in these lower speed zones.

7.2.3 Roadside

The design and layout of the urban roadside can affect road safety in a number of ways yet few previous studies have investigated the association between roadside characteristics and crashes in urban areas. Roadside design affects the separation of pedestrian and vehicular traffic. On some road segments in metropolitan Melbourne, the footpath directly abuts the roadway, while on others, the footpath and road are separated by nature strips. The presence of nature strips provides a buffer that may increase the time available for drivers to recognise that a pedestrian is about to cross the road. The roadside (including nature strips) can also become the path of safest travel when the roadway suddenly becomes obstructed (Gibson & Crooks, 1938)—that is, drivers may choose to drive or ride onto the roadside to avoid a collision. Vehicles may also end up on the roadside if a driver or rider loses control. Both situations (leaving the roadway intentionally or unintentionally) can result in a collision with roadside objects (e.g. poles, trees). The roadside can also affect the complexity of the visual environment, especially in urban areas, which may affect a road user's workload and their ability to detect and respond to hazards.

In the current study, a nature strip on one side of the road was associated with a 40% reduction in MVC, while nature strips on both sides were associated with a 48% reduction in MVC, compared to when there was no nature strip. It is possible, therefore, that nature strips provided an alternative pathway for vehicles to avoid on-road obstructions (thus reducing MVC). It is also possible that the environment appeared more complex when there were no nature strips, and the increase in workload associated with monitoring objects (and pedestrians) close to the road may adversely affect responses to moving vehicles (hence increasing MVC). There was no association between nature strips and SVC or PVC. Therefore, the hypothesis that nature strips may reduce PVC because they form a safety buffer between pedestrians on the footpath and moving vehicles on the roadway was not supported.

The offset distance between the roadway and roadside buildings also affects the distance and time available for drivers to react and avoid a collision if the vehicle leaves the roadway. There was no independent association between offset distance and crashes in this study, however, offset distance to buildings was positively correlated with nature strips: for 60% of roads with a nature strip on both sides of the road, the average offset distance was wider than 5m, compared to 37.5% of road with a nature strip on one side of the road and only 7.6% of roads with no nature strip.

The frequency of SVC increased as the number of non-frangible poles per km of roadside increased. It makes intuitive sense that when a vehicle leaves the road, either intentionally to avoid a collision or unintentionally because of a loss of control, that an increased number of large trees and poles on the roadside increases the chance of an injurious SVC. MVC and PVC were not associated with roadside poles and objects. The distance between the roadway and roadside objects may also influence crash risk, however, in this study, there was little variation between road segments in terms of the distance to roadside objects, so the association with crashes could not be tested.

Roadside development height was significantly associated with MVC and PVC in this study. MVC frequency was significantly lower when the predominant development height was two storeys, compared to roads with predominantly single storey buildings or a mix of development heights. It is not immediately clear why this was the case. PVC frequency was significantly higher when the highest building was three storeys or higher. This may be related to the more visually complex environment making it more difficult to detect hazards (e.g. crossing pedestrians). It may also be because of the increased floor area in multiple storey buildings compared to buildings of lower height. The floor area may be used by business and/or accommodation. Thus, the extra businesses and/or accommodation may attract more pedestrians, however, without pedestrian volume data, this hypothesis cannot be tested. Previous research, controlling for pedestrian volumes, however, found that PVC frequency on weekdays in the Melbourne CBD was positively associated with office floor space density and shop density (Alavi, 2013)—which would be greater in areas with higher development heights.

The final roadside-related factor that was investigated in this study was the relationship between crashes and whether there were shops and businesses on one or both sides of the road. It was hypothesised that PVC may be more frequent on road segments with shops on both sides of the road, because of the potential increased exposure of pedestrians crossing the road to access businesses. This hypothesis was not supported since there was no association with PVC, or other crash types.

7.2.4 Amenities and facilities

The amenities and facilities in an area can influence the number and type of road users and their behaviour. Few of the previous multivariable studies of urban crashes have included variables describing amenities and facilities.

Off-street parking facilities were significantly associated with all three types of crashes. An increase in off-street parking facilities was associated with a significant reduction in both MVC and SVC. This may be related to a reduction in the number of vehicles parking, or searching for a parking space on the side of the roadway, which, as discussed previously, can impact crashes in a number of ways. In contrast, as the number of off-street parking facilities per km increased, there was a significant increase in the frequency of PVC. It is possible that pedestrians walking along the footpath are at risk of being hit by vehicles entering and exiting the off-street parking facility, however, further investigation is needed to confirm this hypothesis.

The presence of parks, sports centres or sports fields on the road segment was associated with a 39% reduction in the frequency of SVC. This may be because the environment appears less complex (the presence of parks implies there are fewer buildings) and there may be fewer roadside objects to collide with. Alternatively, drivers may drive in a safer manner in such areas. This warrants further investigation.

There was an association between the number of establishments per km with particular types of liquor licences and urban crashes in this study. An increase in the number of establishments (or restaurants) per km with a BYO liquor licence was associated with a significant increase in SVC. Previous research also found an increase in night-time SVC as the density of restaurants in an area increased (Gruenewald et al., 1996). As the number of establishments per km with a licence to serve liquor after 1am increased, so did the frequency of MVC on strip shopping centre road segments in metropolitan Melbourne. It is possible that there is an association between the number of impaired drivers and the number of establishments that serve alcohol. Unfortunately it was not possible to pursue this issue further as the data on road user impairment (blood alcohol content) was incomplete in the Victoria crash data. It is unclear though, why different types of liquor licence would affect MVC and SVC differently.

The number of late night liquor licences per km was also associated with PVC on strip shopping centre road segments in metropolitan Melbourne. Again, this may be due to an increase in impaired road users (both drivers/riders and pedestrians). A study of pedestrian injury hot-spots in Vancouver, Canada, found that almost two-thirds had bars nearby and over a third had a high density of alcohol outlets (Schuurman, Cinnamon, Crooks, & Hameed, 2009). It is also possible that road segments with late night liquor establishments attract more pedestrians than other roads. Previous research into risk factors for urban crashes at the area level found that the density of bars was positively associated with pedestrian crashes where the pedestrian had been drinking but not pedestrian crashes where the person had not been drinking, indicating that the alcohol establishments were directly implicated, however, there was no measure of pedestrian exposure in those studies (LaScala et al., 2000; LaScala et al., 2001). In a study of crashes in the Melbourne CBD that did control for pedestrian crossing volumes, however, alcohol establishments were still associated with an increase in PVC that occurred at night-time, indicating that there is an increase in PVC risk associated with alcohol establishments independent of the number of pedestrians (Alavi, 2013).

In addition, more PVCs occurred when there were railway stations or tertiary education institutions present on a road segment. This may be due to a higher number of pedestrians crossing the road in these road segments, but it is also likely due to the behaviour of those road users. Public transport users (pedestrians and those driving to the station) may take risks on the road in order to arrive at the station on time to catch a train. A high proportion of students at tertiary education institutions are young adults, a group known to display risky behaviours. Though pedestrian exposure data were not available, the finding that pedestrian crash frequency is higher in the presence of railway stations and tertiary institutions suggests an observational study of the behaviour of pedestrians and drivers in such locations is warranted. It would also be interesting to determine if stations and tertiary education institutions influence PVC risk on other urban road segments, not just strip shopping centres.

None of the other facilities or amenities considered in this study were significantly associated with crashes (MVC, SVC or PVC). It is interesting that there were more PVC on road segments with tertiary education institutions but not on road segments with primary or secondary education institutions, despite the presence of young vulnerable pedestrians around schools. In the state of Victoria, Australia, speed limits around primary and secondary schools are lower during hours when students are arriving at, or departing, the school. This study provides indirect evidence that this policy is effective in controlling PVC risk. Other facilities and amenities considered in this study that were not related to urban crash frequency were: take-away liquor licences, hospitals and nursing homes, community facilities (including community centres, libraries, neighbourhood houses, senior citizens centres, town halls), kindergartens and maternal child health facilities, places of worship, emergency services, petrol stations and indoor shopping centres.

7.3 Methodological issues

7.3.1 Strengths

7.3.1.1 Scope of study

One of the main strengths of this research was the collection of a large amount of data on a range of different potential risk factors for crashes in urban areas, many of which have not been considered in previous research. In particular, the inclusion of a range of variables describing the roadside environment and amenities and facilities on the road segment was novel. In addition, this study included a large range of variables describing the roadway whereas previous studies were often limited to only a few risk factors. Hence, this study is likely to suffer less from confounding and omitted variable bias than previous research.

7.3.1.2 Modelling approach

The statistical models of the relationship between risk factors and urban crash frequency were developed using an innovative phased modelling approach developed for this thesis that enabled the identification of risk factors significantly associated with crashes while taking into account potential correlations between variables. That the models were developed with relatively few iterations and that diagnostic tests showed no problem with inter-correlation between variables in the model both support the utility of the modelling approach.

Researchers often apply very complex statistical methods (e.g. zero-inflated models, random parameters models) to simplistic models that are not fully specified (e.g. missing important predictors) in order to improve model fit but the results may be compromised by omitted variable bias (Mannering & Bhat, 2014). For example, Mitra and Washington (2007) demonstrated a well specified model that has no omitted variables has no need for a random dispersion parameter. The present research differs from previous studies because of the large range of potential risk factors that were identified and because separate models were derived for different crash types. The final SVC and PVC models displayed no significant overdispersion and diagnostic tests showed Poisson models fitted the data well. The final MVC model was overdispersed and was therefore fitted using negative binomial regression; likewise, model fit was good. There was no evidence that the use of random parameters led to better model fit than fixed parameters, that is, neither the baseline risk of MVC, SVC or PVC nor the relationship between the risk factors and MVC frequency varied significantly across sites. This indicates that the models were well specified and that was no need to use more complex statistical modelling methods to improve model fit. For example, there was no evidence that zero-inflated models were required, even for SVC (43% of road segments had no SVC over the five year period between 2005 and 2009). This is advantageous because Poisson regression and negative binomial regression require fewer assumptions than zero-inflated and random parameters models and interpretation is more straightforward.

Investigations were conducted to explore why the MVC model was overdispersed but the SVC and PVC models were not. The distribution of MVC, SVC and PVC across the broad definitions for classifying accidents (VicRoads, 2008a) and specific crash types were inspected. Table 7.1 shows the three most commonly occurring crash types for MVC, SVC and PVC. The MVC were not homogeneous: between them, the three most common crash types accounted for less than half of the MVC. These were rear end crashes (24%), crashes where a vehicle turned right in front of an oncoming vehicle (15%) and U-turn crashes (8%). In comparison, SVC were relatively homogeneous: 91% of SVC (including the three most common types) involved the vehicle going “off path on straight”. PVC were also more homogeneous than MVC, with the two most common crash types being pedestrians crossing the road from the near side (36%) or the far side (27%). It is therefore possible that SVC and PVC models were Poisson distributed because they were modelling relatively homogeneous groups of crashes, whereas the MVC model was overdispersed because there were different crash types each with their own (perhaps Poisson) distribution. To test this hypothesis, a model could be developed for the most common MVC type (rear end crashes) to see if it is Poisson distributed.

Table 7.1 The three most commonly occurring types of crashes for multi-vehicle crashes (MVC), single-vehicle crashes (SVC) and pedestrian-vehicle crashes (PVC)

Definition for Classifying Accidents	Specific crash type	Number (%)	%	Cumulative %
		MVC		
Vehicles from same direction	Rear end (vehicles in same lanes)	312	24.2	24.2
Vehicles from opposing directions	Right through	196	15.2	39.4
Manoeuvring	U-turn	101	7.8	47.2
		SVC		
Off path on straight	Out of control on carriageway	74	44.7	44.7
Off path on straight	Left off carriageway into object/ parked vehicle	39	22.9	67.7
Off path on straight	Right off carriageway into object/ parked vehicle	30	17.7	85.3
		PVC		
Pedestrian on foot in toy/pram	Crossing from near side of vehicle	188	36.2	36.2
Pedestrian on foot in toy/pram	Crossing from far side of vehicle	138	26.6	62.8
Pedestrian on foot in toy/pram	Other (not descriptive)	53	10.2	73.0

7.3.2 Limitations

The limitations of this study related to the available data and the use of cross-sectional studies for analysing crash frequencies (for reviews of the latter issue, see Lord & Mannering, 2010; Mannering & Bhat, 2014). These are discussed below, including consideration of the potential influence on the modelling results.

7.3.2.1 Data issues

An issue with the use of police-reported data is that fatalities are more likely to be reported than serious injury crashes which in turn are more likely to be reported than minor injury crashes (Cercarelli, Rosman, & Ryan, 1996). Capture of PVC in police-reported data is likely to be more complete than for MVC and SVC given that even low speed crashes involving pedestrians are likely to cause serious injuries. Under-reporting, however, is not likely to bias the estimates of a cross-sectional study if the probability of under-reporting is equally likely across road segments (which is a fair assumption) and if the risk factors for injury crash occurrence do not differ by crash severity. The second assumption could be tested by using logistic regression to determine the risk factors associated with outcome severity, given a crash has occurred. Separate models to predict injury severity were not developed in this thesis, however, because issues have been raised about the accuracy of the reporting of injury severity in the Victorian police-reported crash data (D'Elia & Newstead, 2015). Crashes are classified as severe if the police officer reports that the person was taken to hospital, which has been shown to over-estimate serious injury. Linkage of the Victorian crash data to the injury insurance data held by the Transport Accident Commission (the Government owned compulsory third-party insurer) results in more accurate measures of severity (D'Elia & Newstead, 2015), however, the number of true serious injury crashes is low and it is doubtful there would be enough power to conduct the analysis. Arguably, the identification of risk factors for the frequency of reportable injury crashes (as per this thesis) is of more importance for preventing road trauma than focusing on minor injury crashes because of the potential for more serious long-term outcomes. Statistical methods have been used to deal with under-reporting (e.g. Kumara & Chin, 2005), however these are complex and beyond the scope of this thesis

Despite data being collected for a large number of variables, it was not possible to find or collect data for some potentially important risk factors. Arguably the most important potential risk factors for which data were unavailable were vulnerable road user exposure (e.g. pedestrian crossing volumes, cyclist volumes). No models were developed for BVC and the lack of bicycle exposure data is not likely to have affected the models for other crash types. The lack of pedestrian exposure data was most likely to have affected the PVC model. Hence, the relationship between some of the identified risk factors and PVC may be confounded by pedestrian exposure. Diagnostic tests, however, revealed no significant collinearity between the variables in the PVC model. If more than one of the risk factors in the model were simply related to crashes through their association with pedestrian volumes (that is, complete confounding), then those variables should also be correlated. This provides evidence that the risk factors in the PVC model do have an independent relationship with PVC frequency and suggests any confounding due to the omission of pedestrian exposure data must only be partial.

Some of the variables with an underlying continuous distribution were coded as categorical variables in this study. This decision was made to enhance data quality, as the data coders were experienced at coding the data categorically, not by entering precise measurements. The categories used were those commonly used in other similar applications, for example, AusRAP (Australian Automobile Association, 2006). It is possible, however, that the choice of categories may have obscured a statistically significant association—for example, the offset distance between the road edge and roadside objects was coded as between zero to three metres for all road segments in this study. While this level of resolution may be appropriate for some environments (e.g. rural roads), it may not be detailed enough for urban roads.

Variables that were measured at the area-level rather than at the road segment level did not prove to be useful predictors of crash frequency in this study. Pedestrian activity at the local area level may not be representative of pedestrian exposure on strip shopping centre road segments within that area. The sociodemographic and socioeconomic variables measured at the area-level were also problematic when included in the models. In particular, there appeared to be some relationship between the socioeconomic ranking of an area and road design and infrastructure. While the correlations were not strong in a statistical sense, they were strong enough to cause instability during model development and it was found that the models converged more readily when socioeconomic variables were excluded. This was seen as a reasonable compromise as the residents of an area may not be representative of the road users (particularly drivers) driving on arterial roads in that area. Considering the road safety implications, it was considered more important to identify of aspects of the built environment with the potential for countermeasures to be developed, than to identify characteristics of the population of the local area that were associated with crashes. A thorough investigation of the relationship between socioeconomic indicators, road infrastructure and roadside design was beyond the scope of this thesis but could prove to be an interesting avenue for further research.

7.3.2.2 Modelling issues

When a small sample of count data has a low mean value, the data can become skewed towards zero and fail to exhibit large sample properties so maximum likelihood techniques may not be appropriate (Lord & Mannering, 2010). The current study included all arterial strip shopping centre road segments in metropolitan Melbourne (n=142) and a five-year time period was chosen for measuring crash frequency in order to maximise the number of crashes while minimising the possibility that the road or surrounding environment had changed over that time period. This provided a reasonable sample size of MVC (1,291) and PVC (519). There were, however, only 170 SVC observed on the 142 road segments over the five year period (mean=1.20) with 62 road segments having no SVC over that time. The SVC model fit, however, was good and the predicted probabilities matched the observed probabilities well. In particular, the observed proportion of road segments with no crashes matched that predicted from the SVC Poisson model. This provides evidence that the low mean value of SVC did not adversely affect the success of the modelling.

When there are a large number of potential predictors and relatively few units of analysis, the resulting model can be overfitted. Overfitted models are complex and too dependent on the data used for development which means results cannot be generalised to other samples of interest. There was a risk of overfitting the models in this study, given the large number of potential predictors and limited sample size. Diagnostic tests, however, indicated that overfitting was not likely to be a problem with the models that were developed. The ultimate test, however, would be to determine how well the models fit a new sample of urban road segments, which was beyond the scope of the current thesis.

7.3.2.3 Interpretation issues

An issue with cross-sectional studies is that there may be an endogenous relationship between a risk factor and the outcome as there is no information on the temporal relationship between the two (Mannering & Bhat, 2014). Road safety countermeasures are usually installed where there is an existing crash problem. Even if a road safety countermeasure effectively reduces crashes at the locations where it is installed (best measured using an experimental or quasi-experimental design), those locations may still have more crashes than other similar locations without the countermeasure. If a cross-sectional study is then conducted, a positive association between the road safety countermeasure and crash frequency may be observed. Thus, even though the countermeasure may be effective at reducing crashes, it may appear to be associated with an increased frequency of crashes in a study that does not take into account the endogenous relationship. As a very basic example, imagine there were 20 PVC per year at one midblock location and 5 PVC per year at another location. A pedestrian crossing is installed at the first location only, and the number of crashes fell to 10 PVC per year (the number of crashes at the other location remains constant). Even though the pedestrian crossing was effective in halving the number of PVC, a cross-sectional study performed after installation of the crossing (that is, a snapshot in time) would make it appear that pedestrian crossings were associated with increased PVC, because the number of crashes on the segment with the pedestrian crossing is still double that on the segment without the crossing.

In the current study, the number of midblock pedestrian crossings per km was significantly associated with increases in PVC frequency—this may have been partially confounded by pedestrian volumes (as discussed previously) or there may be an endogenous relationship between pedestrian crossings and PVC. That is, pedestrian crossings may have been installed where there was an existing pedestrian crash problem. Some researchers have successfully dealt with endogeneity using statistical methods but this is more difficult for count data models than linear models (Lord & Mannering, 2010). Kim and Washington (2006) used a method for dealing with endogeneity in count data models using a structural equation modelling-type approach, however, the method was complex and is not currently incorporated in statistical modelling software. Explicitly modelling endogeneity was beyond the scope of the current research. It was, however, important to identify where this might have occurred. The most obvious example is that given above, of the positive relationship between pedestrian crossings and PVC.

Some variables described the characteristics of the entire road segment (e.g. lane width, number of lanes, presence of trams, etc.) whereas others were only relevant to one or more locations on that segment (e.g. presence of amenities and facilities). For the latter class of variables, an observed association with crash frequency on the road segment simply means that the presence of that risk factor is associated with risk on the road segment overall. It does not imply that the crashes happened at that particular location. For example, the increased frequency of PVC on road segments with a railway station does not mean that the PVC all occurred near the railway station. It could be hypothesised that the presence of the station may affect risk and that the risk may diminish with distance from the station—this could not be assessed in the current study. Similarly, the observed increase in crashes as the number of alcohol establishments increased does not necessarily indicate that the crashes were alcohol-related, although previous research indicates that is probably the case (LaScala et al., 2001). It may be that these areas were busy and popular and that overall exposure was increased, or it may be a surrogate for the number of impaired road users. Further research is needed to investigate these issues.

This study focused on strip shopping centre arterial road segments in metropolitan Melbourne. It is unknown whether the results of this study can be generalised to other urban areas. Future research should therefore focus on a broader range of urban road segments. It is important to recognise, however, the methodological contribution of this research. One of the reasons strip shopping centre arterial road segments were chosen as the focus for this study was because they represent a highly complex environment that is challenging to characterise in order to measure the influence of the built environment on crashes. The current research successfully demonstrated that these challenges can be overcome for a sample of highly complex urban road segments. This means that it is eminently feasible to conduct similar studies of other, perhaps less complex, urban road segments (e.g. local roads, residential areas, other arterial road segments).

7.4 Implications for future research and practice

This research has established that, in addition to traffic exposure and road design, roadside features and the types of facilities and amenities on a road segment are also associated with crash frequency in complex urban areas (strip shopping centre road segments). In future, researchers aiming to develop well-fitting statistical models of crash risk in urban areas need to incorporate, or control for, a larger range of risk factors than simply traffic exposure and road geometry. Focusing only on road geometry and traffic, especially in urban areas, fails to recognise that the road system exists within, and is affected by, the broader built urban environment. Therefore, programs to improve road safety also need to look beyond the design of the road and consider the broader context of the built environment in which the road exists.

It could be argued that the crash risks associated with non-road aspects of the built environment are not relevant to road safety practitioners because these aspects lie outside the realm of road infrastructure and are therefore less amenable to direct intervention. For example, road safety practitioners have no power over the liquor licensing process or the presence of facilities like railway stations and tertiary education institutions. While this is true, road safety interventions can still be used to successfully manage risk in areas with these amenities and facilities. Examples include the reduction of speed limits around schools and the use of modified signal phasing to reduce traffic speeds around licensed premises (Lenné , Corben, & Stephan, 2007). Hence, although the risk factor itself may not be amenable to change, effective countermeasures to manage risk may still be available. This thesis also highlights the need for agencies beyond those directly responsible for road safety to consider the road safety (and perhaps broader health system) implications of urban planning decisions: for example, by restricting the number of late night liquor licences allowed within a certain distance of each other or by working with road authorities to install appropriate countermeasures around late-night liquor establishments.

Most of the data items required for measuring the influence of the built environment on crash risk were available from existing sources however only the data for crashes, traffic volumes, speed limits and pavement condition were accessible from databases held by VicRoads. Important data were also sourced from the databases of other government departments—for example, liquor licensing information was obtained from the Victorian Department of Justice. Yet most of the data describing the road, roadside design and facilities/amenities on the road segment had to be coded specifically for the project from video images and maps. This was a time-consuming process and is a barrier to conducting thorough investigations of crash risk in urban areas. Technological advancements should solve this problem in future. For example, the ARRB group have developed a side-scanning laser system for their Network Survey Vehicles that can automatically and accurately measure offset distance (Roberts, Cammack, & Rodwell, 2010). Geo-spatially coded data are becoming increasingly available which introduces the potential to link different sources to obtain detailed information on the risk factors present at geographic locations.

Other data were not available from existing sources and to collect the data and site visits would have been required, which was beyond the scope and budget of the current project. In particular, there were no exposure data available for bicycles or pedestrians. Alavi (2013) discovered that collection of pedestrian activity data in the Melbourne CBD was irregular and non-systematic and that the data that were collected were not sufficient for road safety purposes. The greater Melbourne metropolitan area is no different in this respect. Cycling and walking are promoted as means of transport to promote health and environmental sustainability yet research into risk factors for crashes involving vulnerable road users in urban areas is severely hampered without access to accurate exposure data. Identifying risk factors is an essential step toward the goal of preventing injuries—as such, it is vitally important for governments to invest in the systematic collection of exposure data for vulnerable road users. Large-scale evaluations of the health effects of programs to encourage cycling and walking should consider injury as a potential adverse outcome so these also require accurate data on exposure of vulnerable road users. In addition, this study was restricted to roads managed by the state road authority, because traffic volume data were not consistently available for local roads. The urban road safety problem cannot be adequately addressed without also considering local roads, yet it is currently impossible to measure risk on local roads, as no exposure data are available.

As emphasised throughout the thesis, cross-sectional studies conducted to measure the association between characteristics of the built environment and crash frequency are useful for identifying potential risk factors but do not provide any information regarding the reasons for the change in risk. Association does not imply causation. Other research methods are therefore required to establish whether the relationship is causal and the causative mechanisms. The results of cross-sectional studies, when considered in conjunction with theories of driver behaviour, can lead to hypotheses to be tested using behavioural research methods. This forms the focus of the second research component of the thesis, which is presented in the following chapters.

CHAPTER 8. THE BUILT ENVIRONMENT AND DRIVER BEHAVIOUR

The first research component of this thesis discovered a range of characteristics of the built environment that were significantly associated with crash incidence (MVC, SVC and PVC) on strip shopping centre road segments in metropolitan Melbourne. As discussed in Chapter 1, identifying aspects of the built environment that are associated with crashes is only one step in the process of improving road safety. To prevent crashes, effective countermeasures must be developed and applied which requires an understanding of the underlying factors behind the increase in crash risk, including how road users behave in different environments. Understanding the effect of the environment on the performance and behaviour of road users is essential for understanding how and why aspects of the built environment may influence crash risk.

The second research component of this thesis therefore sought to demonstrate how behavioural research methods can be used to improve the understanding of why aspects of the road and/or roadside increase the risk of a crash occurring. One of the risk factors identified during Component 1 of the thesis was chosen to conduct a case study of the influence of the risk factor on driver behaviour. This research component also serves to demonstrate how the process could be applied to study other risk factors as well.

This chapter reviews the range of behavioural research methods that can be used to investigate the influence of the built environment on road user behaviour. Strengths and limitations of the various methods will be identified. A rationale for the choice of driving simulation as the research method to be used in this thesis is provided. A set of criteria were developed and applied to select a risk factor, identified during research Component 1, for the case study. The risk factor that was chosen for further investigation was the presence of roadside parking. The chapter concludes with a review of previous research into roadside parking and driver behaviour that identifies the gaps in current knowledge and leads to the aims of the research conducted in Component 2.

8.1 Behavioural research methods

The choice of the most appropriate research method to use to investigate the influence of a risk factor on road user behaviour will depend on the risk factor being investigated. Although driving simulation was chosen as the most appropriate method to investigate the effect of roadside parking on driver behaviour for the case study in this thesis (as discussed below), driving simulation will not be the most appropriate method for investigating all risk factors. A wide range of risk factors for crashes on strip shopping centre road segment were identified in research Component 1 of this thesis. Hence, this review will cover the range of behavioural methods that can be used to investigate the effect of the environment on road user behaviour. These range from methods that involve collecting data on self-reported behaviour to those that involve measurement of road user behaviour in the real world. Most of the methods are better developed for studying the behaviour of drivers than other road users, although those methods that have been used for studying the behaviour of other road users will be identified.

8.1.1 Self-reported behaviour

There are a range of research methods in which data are collected via self-report, for example, surveys, interviews and focus groups. In these studies participants are asked to report how they behave or what they would do in particular situations. Prompts such as photographs or videos of road segments that show different roads and roadside environments are often used. These studies can include participants from different road user groups.

Although these methods could be applied to investigate how road users believe they behave in different road environments, self-report is not always representative of actual behaviour. Participants may not explicitly know how they would behave in a given situation or a participant's responses may be biased toward reporting more socially acceptable behaviour. For measuring how behaviour actually changes in different environments, objective measurement methods are necessary. Hence, research methods that solely involve self-report were not considered for use in this study and they are therefore not reviewed in further detail here.

8.1.2 Objective measurement of behaviour

There are a number of methods that can be used to objectively measure how road users behave within the road environment and how aspects of the environment affect behaviour. Sometimes self-report data are used in conjunction with measurements of behaviour in order to measure determinants of behaviour that are not amenable to being objectively measured. Some examples are self-report scales to measure constructs like mental workload or situation awareness, or techniques like verbal protocol analysis in which road users are asked to report what they are doing and thinking to elicit information about cognitive processes.

Broadly, study designs range from the purely experimental to the purely observational and vary according to the fidelity (or face validity) of the method and the level of control that the researcher has. Figure 8.1 presents the five main contexts in which road user behaviour can be studied and ranks them (from left to right) in terms of increasing face validity (or fidelity) and decreasing researcher control. Practically, the methods are also ranked in terms of cost with laboratory studies being the cheapest and observational studies (e.g. naturalistic driving studies) being most costly to conduct. Each context will be discussed in turn.

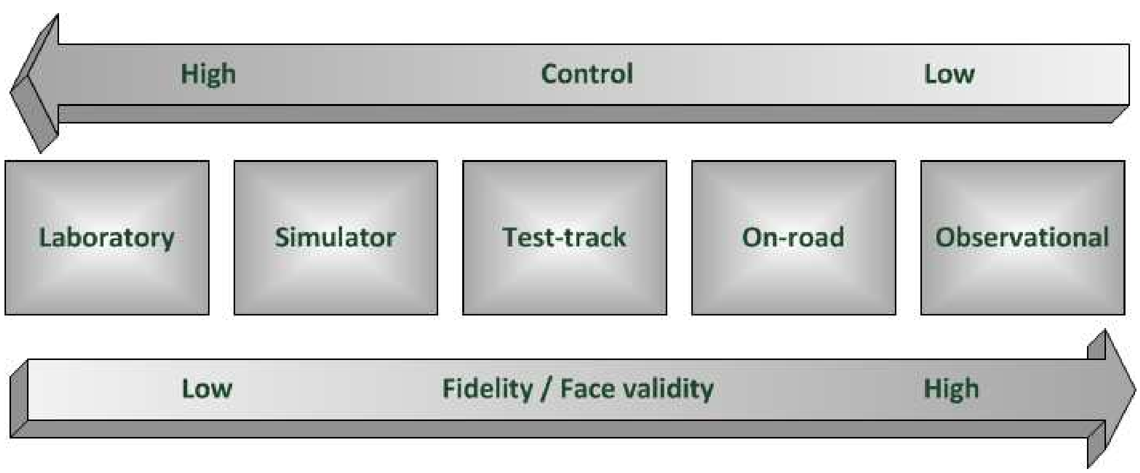


Figure 8.1 Methods for investigating the influence of the environment on road user behaviour

8.1.2.1 Laboratory experiments

Well-designed experiments conducted in a laboratory setting allow a high degree of experimental control over the independent variables and confounding factors. Measurement techniques can be used that are impractical in other settings, for example, physiological measurements like electro-encephalogram recordings. Well-controlled experiments provide strong evidence for a causal relationship between independent and dependent variables. For applied problems however, laboratory settings are often so tightly controlled that the results may not be generalised to the real world. In addition, the tasks that participants are asked to perform in the laboratory are often different to those performed in real-world situations.

Laboratory studies are useful for investigating the effect of one or a small number of factors on fundamental physical, perceptual or psychological processes; for example, the effect of alcohol consumption on motor function (Ando, Iwata, Ishikawa, Dakeishi, & Murata, 2008) or working memory (Boha et al., 2009). In terms of investigating the effects of the built environment on road user behaviour, laboratory experiments are best suited to detailed investigations of the effects of perceptually salient road features (e.g. road width, curves) on driving related visuo-motor tasks (e.g. tracking performance) or hazard perception.

Thus, while there is an important role for laboratory based research in road safety, investigation of the effect of the environment on road user behaviour often needs to be conducted within a more operationally valid context. For this reason, laboratory-based studies were not considered for the current study.

8.1.2.2 Simulator

Simulators provide a controlled environment in which to conduct experiments to safely measure the effect of various factors on tasks relevant for driving. They are also an important tool for developing and testing countermeasures in a safe and controlled context. Scenarios can be designed to closely represent the real-world road environment. The scenario can be designed so events and conditions are the same for each driver. Similar to laboratory studies, measurement techniques can be used in simulators that cannot be used in the real world, for example, complex physiological measurements and intrusive secondary task techniques. Although the level of researcher control is not as high as for laboratory studies, simulation studies have greater face validity for investigating the effect of different aspects of the environment on behaviour. The vast majority of simulation studies in road safety and traffic psychology involve driving simulators, although there are also examples of bicycle simulators (e.g. Plumert, Kearney, & Cremer, 2004) and immersive environments for studying pedestrian behaviour (e.g. Clancy, Rucklidge, & Owen, 2006). The discussion of the strengths and limitations of the method will therefore focus on driving simulation as this is the most advanced application.

Although they have greater face validity than laboratory studies, studies conducted in driving simulators are sometimes criticised because the road environment does not fully replicate the real world. In addition, participants have different motivators within a driving simulator experiment than they do in real life which can affect the behaviour they display. The tasks may also be somewhat artificial. Nevertheless, driving simulation holds an important place as a valid method for conducting rigorously controlled experiments to investigate the effect of a small number of experimental manipulations on driver behaviour and in the first stages of evaluating countermeasures prior to implementation on road.

8.1.2.3 Test-tracks

Experiments involving drivers in control of real vehicles are sometimes conducted on test-tracks in order to provide a more realistic driving environment. The added realism, however, comes at the cost of experimental control. Although it is possible to manipulate the road environment, it may be difficult to accurately reproduce experimental manipulations (e.g. placement of road signs, cones, etc.) and light and weather and conditions may vary between participants. Investigation of complex driving scenarios that include many other road users and infrastructure is difficult. Collecting good quality data is more challenging in less controlled environments and it may be necessary to rely on observations by the researcher, which are prone to observer bias. The added level of realism also brings an added level of risk which therefore raises insurance issues and ethical and occupational health and safety concerns. In practice, there are very few test-tracks available for use by researchers in Victoria, Australia. The relative lack of control, concerns regarding safety and insurance and the lack of facilities meant that conducting this study in a test-track environment was not a viable option.

8.1.2.4 On-road studies

Vehicles instrumented to record data from the vehicle (e.g. location, dynamics) and the driver (e.g. physiological measurements, observable behaviour) are a valuable tool for measuring driver behaviour and consequences in the real world. On-road studies of drivers are the most common so these are the main focus here, however, on-road studies have also been conducted of the behaviour of motorcyclists, cyclists and pedestrians (Salmon, Young, & Cornelissen, 2011).

In on-road studies using instrumented vehicles, the researcher defines the route to be driven and the type and, as far as practicable, timing of tasks to be performed. The researcher is sometimes also present in the vehicle with the participant (e.g. Young, Salmon, & Lenné 2013) but not always (e.g. Young, Lenné , Beanland, Salmon, & Stanton, 2015).

If good quality data can be captured on the environment the driver is travelling through, then studies using instrumented vehicles can measure the influence of the environment on driving behaviour. Data regarding the road and environment could be captured by strategically placed cameras, global positioning system (GPS) technologies or even experimenters in (or behind) the vehicle manually recording data.

On-road studies are useful for in-depth investigation of behaviour in real-world environments (e.g. at intersections, rail level crossings). Disadvantages of on-road studies relate to the lack of control over aspects such as weather, traffic volumes and actions of other road users. Sample sizes are often limited due to practical considerations of time and cost which raises questions of statistical power. Data extraction, manipulation and analyses of in-depth studies can also be complex. Evaluation of countermeasures can be conducted using on-road studies, however, other methods are preferred for initial investigations because of the potential for unintended effects that may influence safety.

8.1.2.5 Observational studies

In observational studies, road users are observed without any intervention on behalf of the researcher. At the population level, an observational study can focus on the behaviour of all road users at a particular location, or type of location. For example, Read, Salmon, Lenné and Grey (2014) observed how pedestrians and cyclists interacted with infrastructure at railway level crossings in Melbourne, Victoria and found that observed behaviour did not always match that expected by designers of the system, which has implications for the design of level crossings. Such observational studies could be used to record how road users actually interact with particular environmental features. Observational studies of behaviour at a location are often used for evaluating the effects of countermeasures using before-after studies or quasi-experiments (before-after studies with comparison locations).

Observational studies can also involve more detailed observation of individuals as they travel through the road system. Researchers can collect data while surreptitiously following the unaware road user (Papadimitriou et al., 2014) or instrumented vehicles can be used to conduct naturalistic driving studies (NDS). NDS are similar to on-road studies in that they involve in-depth recording of behaviour in the real world, however, unlike on-road studies, once the driver takes charge of the instrumented vehicle in a NDS, the researchers have no control over where or when the driver travels or how the driver behaves. The researchers merely observe (and record) the driver's travel behaviour. Examples of NDS are the 100 Car study conducted by the Virginia Tech Transportation Institute in the USA (Neale, Dingus, Klauer, Sudweeks, & Goodman, 2005) and a naturalistic driving study of cyclist behaviour conducted in Melbourne, Australia (Johnson, Charlton, & Oxley, 2010; Johnson, Charlton, Oxley, & Newstead, 2010).

Observational studies have high face validity as they involve observing road users interact with the environment in real-world situations. However, the researcher has no control over any aspect of the situation (the trip, the environment, driver behaviour). In NDS in particular, a huge amount of data are recorded and researchers are still grappling with the most appropriate study designs and analysis methods to answer research questions using the data. Some of the major methodological issues relate to how to effectively deal with repeated in-depth measurements on individuals and the huge potential for bias and confounding. Therefore observational studies are appropriate for observing behaviour in a real-world context but may not be the most appropriate method for controlled investigation of the effect of a particular risk factor on driver behaviour.

8.1.3 Summary and implications for this study

Drivers, cyclists and pedestrians are influenced by the environment. Hence, to reduce crashes and injuries, it is necessary to investigate the effect of the environment on all road users. However, methods for measuring the effect of the environment on driver behaviour are generally more advanced and therefore more feasible than those for measuring cyclist and pedestrian behaviour. It is more difficult to study the effect of the environment on pedestrians and cyclists, apart from perhaps using observational studies. Therefore, the behaviour of drivers was the focus of Component 2 of this thesis.

It is, however, possible to investigate the response of drivers to other road users (e.g. pedestrians, cyclists) in different environments using the different methods, particularly driving simulation. It must be recognised, however, that this only allows investigation of one side of the interaction—the response of the driver to the other road user (ignoring the response of the other road user to the driver). The use of linked simulators to investigate the interactions between road users has been reported in the literature (Cai, Lin, & Mourant, 2007), however this approach is quite rare.

Of the research methods reviewed here, driving simulation studies, on-road studies and observational studies (at the level of the individual: NDS) are the most promising methods for investigating the effect of characteristics of the built environment on individual driver behaviour. Driving simulation affords the experimenter the highest level of control, while observational studies offer the opportunity to observe the driver behaving naturally in real-world environments.

Table 8.1 compares the driving simulation, on-road and observational (NDS) studies in terms of whether the method is appropriate for investigating the effect on driver behaviour of the various categories of risk factors that were considered in Component 1 of this thesis. Any of the three methods could be used to investigate most of the types of risk factors, with the exception of perhaps road pavement condition, enforcement cameras and amenities and facilities, which would be difficult to investigate using a driving simulator. Pavement condition could only be investigated in on-road studies and NDS if the instrumented vehicle was capable of collecting pavement condition data, or if the GPS location could be linked to existing databases containing such data. Obviously the effect of many of the environment-based risk factors on behaviour could only be investigated in on-road studies and NDS if the cameras adequately captured the external environment or if GPS data were available to link the vehicle location to the locations of risk factors (e.g. intersections, road type, public transport facilities, pavement condition, enforcement cameras, amenities and facilities). Alternatively, the route for on-road studies could be chosen so drivers travelled on road segments where these factors were present.

Table 8.1 Applicable driver behaviour research methods for measuring the effect of built environment on observable driver behaviour: by category of risk factor

	Driving simulator	On-road	Observational (NDS)
Road user volumes	✓ Difficult to sufficiently simulate all effects	✓ From the perspective of the driver	✓ From the perspective of the driver
Road cross section	✓	✓	✓
Accesses and intersections	✓	✓	✓
Medians	✓	✓	✓
Road type	✓	✓	✓
Traffic mix	✓	✓ Interactions with other road users	✓ Interactions with other road users
Roadside parking	✓	✓	✓
Public transport facilities	✓	✓	✓
Bicycle facilities	✓	✓	✓
Pedestrian facilities	✓	✓	✓
Roadside development	✓	✓	✓
Pavement condition	×	✓	✓
Height clearance	✓	✓	✓
Enforcement cameras	×	✓	✓ If GPS location of cameras known
Amenities and facilities	×	✓	✓
Speed limit	✓	✓	✓

The choice of the most appropriate method therefore depends on the research question to be addressed. Driving simulation is the most appropriate method if the effect of a small number of environmental factors on a specified range of driving behaviours is of interest and a high level of experimental control of the environment and other road user activity is desired. This high level of control provides more evidence for a causal relationship between the risk factor and the behaviour. Driving simulation is also more appropriate for investigating the development of countermeasures and initial evaluations of countermeasures, particularly to measure unintended, potentially unsafe, effects.

On-road studies and NDS are appropriate for measuring behaviour in the natural environment where control is less essential. These studies are valuable for validating the results of driving simulation studies (providing the combination of factors of interest exists in the real world) to determine if drivers do act as predicted from the results of experimental studies. The results of driving simulation studies can therefore inform the types of activities that could be the focus of on-road studies or NDS and, in turn, situations in which expected behaviour does not occur in the real world could be identified for future investigation in the controlled environment of the driving simulator. Observational studies are also useful for evaluating the effect of countermeasures that have been implemented (preferably after testing in the safe environment of the driving simulator).

The research methods therefore can be considered as to how they fit into a cycle for investigating the effect of the road and roadside on road user behaviour. Laboratory studies should be used for the investigation of underlying fundamental physical, physiological and perceptual mechanisms and how they vary in response to specific features relevant to the driving environment. Driving simulation studies are appropriate for the initial controlled investigation of driver behaviour in response to changes in the environment and for the development and initial experimental evaluation of countermeasures. Driving simulation is particularly valuable for safely evaluating countermeasures that have not yet been implemented. On-road studies and NDS can be employed to investigate behaviour in the real world. These can lead to hypotheses about fundamental perceptual processes that could be addressed in laboratory studies, or about specific effects of road and roadside features that could be tested in driving simulation. Simulator results could then be validated using real-world data. The research process for investigating driver behaviour is therefore cyclic and can be mapped onto the research and countermeasure development cycle for injury prevention that was presented in Chapter 1 (Figure 8.2).

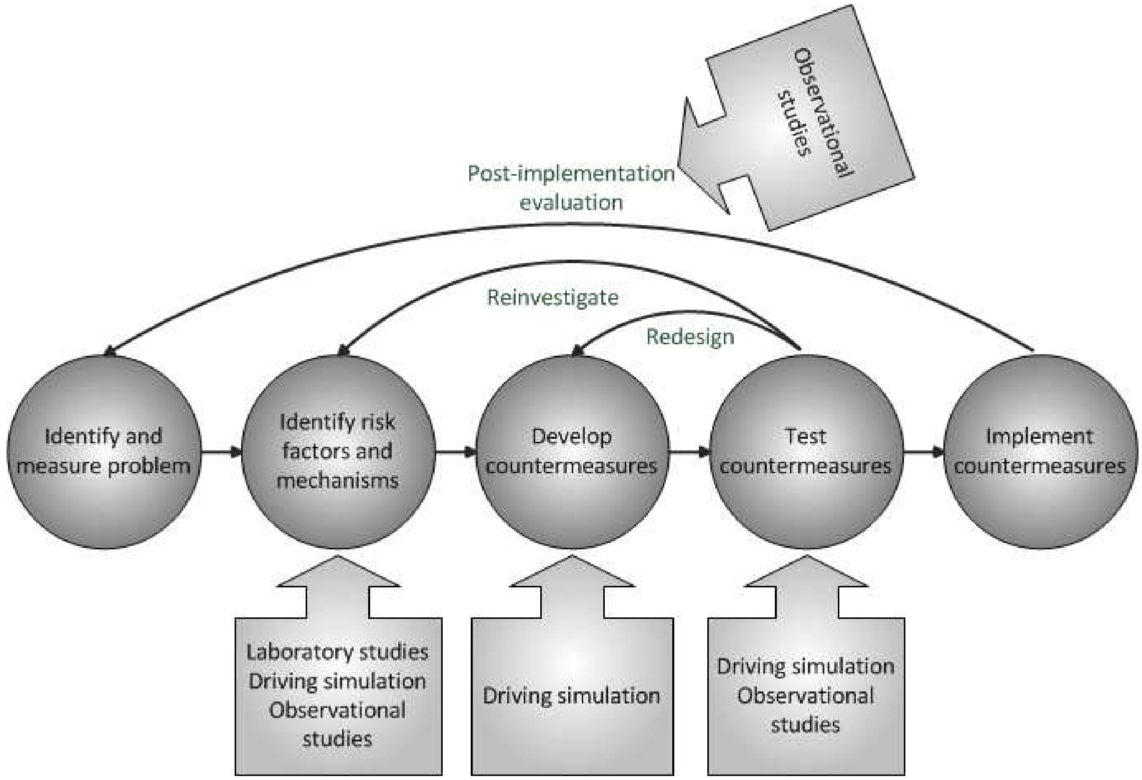


Figure 8.2 Behavioural research methods applicable for investigating driver behaviour in different stages of the research and countermeasure development cycle for injury prevention

For the purposes of this thesis, a controlled experimental investigation was conducted to investigate the effect on driver behaviour of one particular characteristic of the road and roadside that was found to affect crash risk in Component 1. In addition, the effect of a simple countermeasure was also of interest. For this reason, driving simulation was chosen as the most appropriate research method, because of the high level of experimental control it affords to safely manipulate aspects of the environment and measure the resulting changes in behaviour within a realistic context. It is the most appropriate method for the initial investigation of driver behaviour in response to changes in the environment and for the development and initial evaluation of countermeasures.

8.1.4 Measurement of behaviour in the driving simulator

Since driving simulation was chosen as the research method for the research in Component 2 of this thesis, it was necessary to outline how the aspects of behaviour that are hypothesised to change with changes in the environment can be measured in the driving simulator. The multiple comfort zone model predicts that when the threshold level of risk is violated (e.g. by changes in the environment), the driver feels uncomfortable and will act to restore safety margins to an acceptable level. Perceptions of task difficulty and mental workload will also be affected.

A driver's behavioural response to the change in safety margins can be measured in the driving simulator through their control of the vehicle (lateral and longitudinal) and their tactical response to hazards. Frequently used measures can be categorised according to whether they measure driver behaviour, or the result of that behaviour in terms of the consequent effect on the vehicle position/speed within the environment. For example, lateral vehicle control can be measured through measuring driver actions (e.g. steering wheel movements) or the results of those actions (lane excursions, lane position and the standard deviation of lane position, which is a measure of weaving) (Knappe, Keinath, Bengler, & Meinecke, 2007). Likewise, measures of longitudinal control can include driver's actions such as the pressure placed on the accelerator, the variability in accelerator pressure, and braking (e.g. number and duration of braking episodes, brake pressure, variability in brake pressure) or the results of those actions, in terms of travel speed and speed variability. The response to moving hazards that encroach onto the field of safe travel can also be measured as the time to respond (e.g. time to accelerator release, time to brake, or time to steer away from the hazard) and the result of the response (successful or unsuccessful hazard avoidance) (e.g. Edquist et al., 2012).

The effect of manipulations of the risk factor on perceived risk, discomfort and task difficulty can be measured via self-report techniques, for example, rating scales (e.g. Lewis-Evans & Rothengatter, 2009). Mental workload is a frequently studied construct within the broader human factors field. In order to highlight the necessity to use several methods to measure mental workload, it is first necessary to review the relationship between task demands, performance and mental workload, shown in Figure 8.3 (de Waard, 1996). At very low levels of task demand (region D in Figure 8.3), the task is monotonous and it takes an operator a great deal of effort to perform the task, hence mental workload is high and performance is poor. In region A1, task demands have increased a little; the task is not as monotonous as in region D and performance is good, however, due to relatively low task demands the operator still has to expend effort (operator state-related effort) to perform well. In region A2, the task is demanding enough but not too demanding, hence performance is optimal and operator mental workload is low. This is the ideal situation and designers should aim to design systems (e.g. roads) so that task demands are within this region. As the tasks become more demanding, operators have to expend more effort (that is, use more of their processing capabilities) to maintain good performance; this is called task-related effort (region A3). It is important to be able to identify when operators are working in this region, because they will have fewer spare resources to respond to any additional tasks, for example, a hazard on the roadway. Finally, as task demands increase beyond the capabilities of the operator, mental workload increases and performance diminishes (region B) until the operator finds the task so difficult that mental workload is at maximum yet task performance degrades (region C). Although both mental underload (regions D and A1) and mental overload (regions A3, B and C) are dangerous for road safety, mental overload is more likely to be a problem than underload in complex urban areas whereas the converse is true for rural areas.

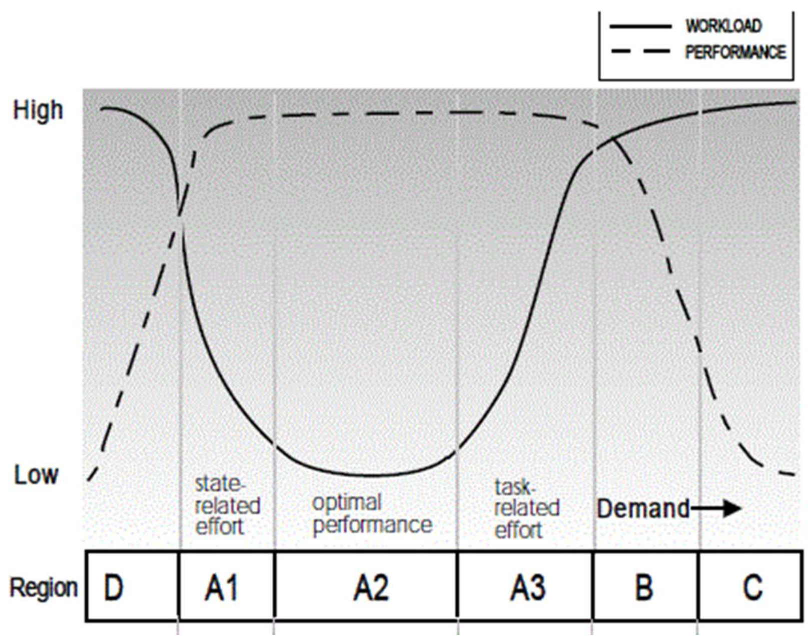


Figure 8.3 Relationship between task demand, mental workload and performance. Reprinted with permission from *The Measurement of Drivers' Mental Workload* (p. 24), by de Waard, 1996

Several types of measures have been used to measure mental workload. The first is primary task performance (e.g. lane-keeping and speed control). Primary task performance (represented as the dotted line in Figure 8.3) cannot distinguish between regions A1 to A3. System design should aim for task demands to be within region A2. If the operator is operating within region A1, the task is too easy and effort is required to maintain performance and if the operator is working within region A3, then the task is becoming more difficult and they are using more spare capacity to perform the task. It is important to know if changing the road and roadside environment places the driver into this region, in particular. Hence, on their own, primary performance measures are inadequate to measure mental workload.

Self-report measures are a common method of assessing mental workload that require operators to assess the level of mental workload they feel when performing a task. Considering mental workload is a subjective state, self-report measures are considered to be a valid method of measuring mental workload (Annett, 2002). The benefit of subjective mental workload ratings is that, unlike primary task performance, they are sensitive to changes in task demand within the A1 and A3 regions. Another advantage is that subjective ratings scales are relatively non-intrusive, especially if they are administered after the primary task (e.g. driving) has been performed.

Rating scales for measuring mental workload can be classified as multidimensional or unidimensional. Multidimensional scales involve rating mental workload along several dimensions. Unidimensional scales require operators to rate only overall feelings of mental workload.

The NASA task load index (NASA-TLX) is an often used, validated, multidimensional workload rating scale that was developed in the aviation context (Hart & Staveland, 1988). The NASA-TLX has six dimensions: mental demand, physical demand, temporal demand, effort, performance and frustration. Operators rate the workload associated with a task on each dimension, on a scale from very low to very high. They are also asked to judge which dimensions are most important for workload on that task using a series of pairwise comparisons, from which a weighting is derived for each dimension. An overall workload score is calculated based on combining the ratings and the weights of all dimensions. Often, the weighting component is omitted, in which case the technique is known as the Raw-TLX (Hart, 2006). The NASA-TLX has frequently been used to measure driving-related workload. Workload, however, is related to task demands and therefore the dimensions applicable to one domain may not be relevant to another. In road safety research, researchers often modify one or more of the NASA-TLX dimensions to make them more relevant to the driving task (e.g. Edquist et al., 2012), which removes the benefits of using a validated scale. The Driving Activity Load Index (DALI), based on the NASA-TLX, was developed to measure driving-related workload for evaluating new human-machine interfaces (Pauzié, Manzano, & Dapzol, 2007), however it has not been validated for other purposes.

An example of a validated, unidimensional mental workload scale is the Rating Scale Mental Effort (RSME, also known as the BSMI in Dutch; Zijlstra, 1993). Operators are asked to rate their mental effort on a 15 cm scale with 1 cm markings. In contrast to most scales which only have anchors at each end (e.g. very low to very high as in the NASA-TLX), the RSME has nine labels along the scale at irregular (but validated) intervals; from absolutely no effort (at a rating of just above zero) through rather much effort (at a rating of about 58) to extreme effort (at a rating of around 112). The RSME has been used successfully for measuring overall mental workload, in terms of the effort expended, in driving research (de Waard, 1996; Lewis-Evans, 2012).

The choice between using a multidimensional scale or a unidimensional scale depends on the research question. If it is important to determine which aspect of the task is contributing the most to overall workload, a multidimensional scale is necessary. Of course, it must be demonstrated that the multidimensional scale is measuring dimensions of workload relevant for the task being performed. If, however, an overall measure of workload is desired, then a unidimensional scale is to be preferred (de Waard, 1996). One reason is that it is requires less of the participants because they only need to provide one rating rather than multiple ratings for different dimensions. In addition, unidimensional scales are more sensitive measures of overall mental workload than multidimensional scales; in particular, the RSME is more sensitive to task demands than the NASA-TLX (Veltman & Gaillard, 1996).

Another technique for measuring spare capacity, or mental workload, is the secondary task technique. When someone is operating with region A2 (the region of optimal performance), their mental workload is at a level where they have spare resources to effectively perform tasks other than the primary task. As task demands increase and they enter region A3, however, they have fewer spare resources available to perform a secondary task. Hence, as mental workload increases in region A3, performance on a secondary task will become poorer even though primary task performance is still high. Secondary task techniques are appropriate for use in driving simulation experiments because the degree of control available means they can be presented at set intervals and responses measured accurately.

There are many options for choosing a secondary task and it is necessary to choose a task that will not affect performance on the primary task (Gawron, 2008). For example, requiring people to perform a secondary visuo-motor tracking task while driving in a simulator would be inappropriate. Several types of secondary tasks have been used successfully in driving simulator studies, for example, simple reaction time tasks and choice reaction time tasks. The peripheral detection task (PDT) is a simple reaction time task in which a light is displayed in the periphery of the participant's field of view and the participant is required to respond by pressing a button (Martens & van Winsum, 2000). The PDT is sensitive to changes in driving task demands: as primary task demands increase, the response rate to the secondary task decreases and the response time is slower (Jahn, Oehme, Krems, & Gelau, 2005; Martens & van Winsum, 2000). Experienced drivers perform better on the PDT than inexperienced drivers, which supports the view that mental workload is a function of task demands and the abilities and state of the driver (Patten, Kircher, Ostlund, Nilsson, & Svenson, 2006). Some researchers used a more complex choice reaction time task where a stimulus is presented peripherally, either to the right or to the left, and participants are asked to respond by indicating which side the stimulus was on, for example, by using the indicator stalk. The correct response rate (Edquist et al., 2012) and response times for correct responses have been shown to be sensitive to manipulations of the complexity of the environment (Edquist et al., 2012; Stinchcombe & Gagnon, 2010).

Finally, physiological measurements have also been used to measure workload (for reviews, see Brookhuis & de Waard, 2010; Kramer, 1991). Cardiac measures are the most commonly used physiological measure of workload. Heart rate generally increases with workload (HR; beats per minute) while HR variability generally decreases (inter-beat interval; IBI), although IBI is has been shown to be more sensitive than HR for measuring mental workload (Brookhuis & de Waard, 2010; Kramer, 1991). Other less frequently used measures are skin conductance, eye movements and brain activity (measured using EEG recordings). Advantages of physiological measures are that they can be continuously measured and do not require an active response from the participant (de Waard, 1996). There are, however, some disadvantages of physiological measurements: they can be affected by factors other than mental workload (for example, physical movements), they can be more physically invasive than other techniques, specialised equipment is usually required, and extracting, analysing and interpreting data is complex (Kramer, 1991).

Thus, the aspects of objective behaviour and subjective perceptions that are hypothesised to be affected by manipulations of behaviour, as predicted by the multiple comfort zone model, can be measured in the driving simulator. In the research conducted for Component 2 of this thesis, objective measures of lane-keeping, speed control and hazard perception were used alongside a secondary task technique for workload measurement and self-report measures of perceived risk, discomfort, task difficulty and workload. Physiological measures of workload were not used to measure workload in this study.

8.2 Selection of a risk factor for further investigation

Due to time and resource constraints, only one of the many risk factors that were identified during Component 1 was chosen to investigate further in terms of its effect on driver behaviour. For the purposes of this thesis, a set of criteria were specifically developed to select a risk factor for further investigation. The criteria related to the feasibility of investigating the risk factor using the chosen method (driving simulation), theoretical considerations and gaps in current knowledge. Seven criteria were devised to inform the final choice of risk factor to be investigated in the driving simulation research component. Four criteria were considered to be essential while three were classed as desirable (in that the research question would be more interesting if these criteria were met). Each criterion will be presented in turn, and risk factors identified in Component 1 that did not meet each criterion will be ruled out before moving onto discuss the next criterion.

8.2.1 Essential criteria

Criterion 1. The risk factor must be amenable to manipulation in the driving simulator (Essential).

It must be possible to systematically vary aspects of the risk factor in the driving simulator in order to measure the effect on driver behaviour. The risk factors identified during Component 1 that could be systematically manipulated in the driving simulator are the following aspects of the road and roadside environment: lane width, curves, different types of accesses and intersections (unsignalised, driveways/laneways, roundabouts, service roads), roadside parking (and clearways), bicycle lane width, shared tram lanes, bus stops, midblock pedestrian crossings, nature strips, development height, roadside poles/trees and speed limit.

Several risk factors were ruled out for consideration using this criterion. The effect of amenities and facilities on crash risk is likely due to a combination of factors relating to the number and type of road users and their behaviour that would be difficult to reproduce in a simulator and therefore best explored in the real world. It is difficult to comprehensively measure the effect of traffic volume on driving behaviour in the driving simulator because it is not just the raw number of vehicles but also the interaction between them that is important. Although the number of vehicles can be systematically varied in the driving simulator, it is difficult to develop scenarios in which all other vehicles behave as if driven by intelligent drivers (rather than vehicles programmed to drive at a certain speed with a limited range of responses to the driver). It may be possible to simulate pavement condition in high fidelity simulators with a highly realistic motion base, however, there are no such high end driving simulators in Australia. The presence of low tram wires as a risk factor doesn't lend itself to investigation in a driving simulator as the change in crash risk may be related to the reduction in heavy vehicle traffic rather than the low tram wires themselves. In addition, the higher risk associated with over-dimensional routes is likely due to multiple reasons relating to road design and traffic mix that were not captured in the model that would need further investigation before being considered for incorporation into a driving simulation scenario.

Though most physical characteristics of the road and roadside are amenable to systematic manipulation in the driving simulator, some risk factors would be more complex to manipulate than others. Risk associated with carriageway width, (the width of the road from kerb to kerb) may reflect the compound effects of the number of lanes, lane width and whether or not the road is divided. None of these aspects, however, were independent risk factors for MVC or SVC in Component 1 if carriageway width was removed from the model. While it would be possible to design an experiment to systematically vary all these factors, it would require a large number of experimental conditions and a large sample size to tease out the effects which would make the experiment costly to conduct. Maximum median width was also considered to be inappropriate for investigation in the driving simulator at this stage. A road with a maximum median width greater than zero metres may have a median for all or part of the road but it also might only have one traffic island along the length of the road segment. Thus it is difficult to systematically manipulate maximum median width to measure the effect on driver behaviour without further investigation using other methods (e.g. observation) to determine which particular aspects might be driving the changes in crash risk.

Criterion 2. It must be possible to design a driving simulation scenario appropriate to investigate the effect of the particular risk factor on behaviour relevant to the crash type of interest (Essential).

Simulation scenarios can be designed to measure driver responses to changes in the environment and to potential hazards (stationary or moving obstacles). It is not possible, however, to use driving simulation to measure the effect of a risk factor on pedestrian behaviour. While it is possible, in the simulator, to measure the response of a driver to a pedestrian in an environment where PVC risk is high, this is restricted to scenarios where appropriate pedestrian movements can be realistically simulated. The results from Component 1 revealed that PVC incidence was significantly lower when lanes were wider. It was hypothesised that this may be related to the extra space (field of safe travel) provided by wide lanes for pedestrians and motorists to negotiate around each other if a pedestrian is caught on the carriageway whilst attempting to cross. This would be difficult to investigate in a driving simulator as the range of pedestrian movements that can be feasibly simulated may not be realistic enough. PVC risk also increased as the number of midblock pedestrian crossings increased. While it would be possible to measure the responses of drivers to pedestrians at pedestrian crossings in a driving simulator, this risk factor was excluded from consideration because of the strong possibility that the increase in risk could be partially explained by an increase in the number of crossing pedestrians, rather than the pedestrian crossings per se (as pedestrian exposure was not included in the model). Likewise, the effects of speed limit on PVC may have been partly explained by pedestrian volumes.

Several of the risk factors that were only significantly associated with SVC (and not other crash types) may be difficult to investigate in a driving simulator at this stage, without further knowledge about why risk might be increased. For example, it is unclear why an increase in the number of bus stops was associated with SVC which makes it difficult to design an experimental study to investigate the issue further. Compared to roads with no bicycle lanes, roads with narrow bicycle lanes had significantly more SVC while roads with wide bicycle lanes had fewer SVC. Had there been an association with MVC (in particular, crashes involving vehicles and bicycles), it could be hypothesised that bicycle lanes affect the safety margins for cyclists and car drivers, however that was not the case, and it is unclear at this stage why only SVC were influenced. Finally, SVC incidence almost doubled when cars were required to share a lane with trams. This could be due to vehicles losing grip on slippery tram tracks, the presence of road furniture for vehicles to collide with (e.g. pedestrian barriers at tram stops) or may be because drivers attempt to overtake trams on the left hand side in areas where vehicles are parked (a moving vehicle crashing into a stationary vehicle is considered to be a SVC). This hypothesis could be tested in a driving simulator, however, it would require the capability to simulate moving trams. These issues may best be investigated initially using observational studies to narrow down potential mechanisms so appropriate experimental scenarios can then be designed.

An increase in the number of roundabouts was significantly associated with an increase in MVC incidence. There were only seven road segments in this study that had roundabouts and there was great variation between the roundabouts in terms of the number of arms, the number of lanes and the road geometry. Therefore, there was no typical type of roundabout on strip shopping centre road segments in Component 1 so this was deemed not to be an appropriate risk factor to investigate in the driving simulator at this time.

The roadside environment characteristics that were significantly associated with crashes may influence risk through their effect on the ability of drivers to recover if something goes wrong. For example, if there is an obstacle on the roadway or another driver cuts in front while changing lanes, the roadside environment may become a field of safe travel (that is, the driver may avoid the threat by driving onto the nature strip or footpath). Nature strips were associated with fewer MVC. The nature strip provides extra space for drivers to recover—without a nature strip drivers may have nowhere to go to avoid colliding with a moving vehicle. An increase in roadside poles and trees makes driving onto the roadside to avoid another vehicle more risky which could explain the increase in SVC when there were more roadside poles and trees. It is unknown, however, whether driving simulation is a valid tool for measuring how drivers would react to such an incursion into their field of safe travel—is the simulated roadside environment (that is, footpaths and nature strips) realistic enough that they would consider it to be a field of safe travel? Do the simulated roadside trees and poles have the same negative connotations in terms of the potential for causing injury as those in the real world? These issues have not generally been considered in simulator validation studies and validation of the driving simulator was not within the scope of this study.

The presence of nature strips and roadside poles and trees, as well as development height, could also impact crash risk through their effect on the visual complexity of the environment. A simulation scenario could be designed to test this hypothesis, however, it would be preferable to first use another research method (e.g. survey or focus group) to determine which aspects of the environment increase visual complexity (e.g. Edquist, 2009) prior to designing a simulator study to determine the effect on driver behaviour.

Thus, the risk factors that met Criterion 2 in addition to Criterion 1 were curves, different types of accesses and intersections (apart from roundabouts) and roadside parking (and clearways).

Criterion 3. There must be a theoretical rationale for why the manipulations of the risk factor that are possible in the driving simulator are hypothesised to affect driver behaviour and crash risk (Essential).

It must be possible to derive hypotheses from theories of driving behaviour as to why the risk factor is expected to change driver behaviour. The multiple comfort zone model proposes that environmental factors and other road users can impact safety margins and the time available to the driver which increases task demands. Once the change in safety margins reaches a certain threshold, the driver will perceive risk and feel discomfort leading to action to restore safety margins to a comfortable level (Summala, 2007). Workload theory also suggests that factors that make the environment more complex can affect task difficulty, effort and task performance (de Waard, 1996).

All of the risk factors that met criterion 1 and 2 can be considered as meeting this criterion. The presence of curves, different types of accesses and intersections and roadside parking (and clearways) can affect safety margins and task demands. Some affect perceived lateral safety margins (roadside parking, intersections and accesses), available stopping distances (intersections, curves) or require drivers to monitor both speed, lateral position and lateral acceleration (curves). They can also act to affect safety margins through increasing the likelihood that the driver will have to monitor and interact with other road users and obstacles, therefore increasing the opportunity for conflict (accesses and intersections, roadside parking).

Criterion 4. The aspects of driver behaviour hypothesised to be affected by the risk factor have to be measurable in the context of a driving simulation experiment (Essential)

The multiple comfort zone model predicts that when the threshold level of a safety margin is violated and risk is perceived, the driver will act to restore safety margins to an acceptable level. Task difficulty, mental workload and effort will also be affected. The measurement of these aspects of behaviour in the driving simulator was covered in detail in Section 8.1.4.

Therefore, the aspects of behaviour hypothesised to be affected by curves, accesses and intersections and roadside parking under the multiple comfort zone model are measurable in the context of a driving simulation experiment. Hence, all of the risk factors that met the first three criteria also met criterion 4.

8.2.2 Desirable criteria

After applying the four essential criteria, the risk factors identified during Component 1 of the thesis were narrowed down to a selection that could be further investigated using driving simulation. In order to select just one risk factor for the case study to be conducted in Component 2 of the thesis, three more (desirable) criteria were devised.

Criterion 5. There must be a lack of knowledge about the effect of the risk factor on driver behaviour relevant for crash risk (Desirable).

While there are still undoubtedly gaps in the knowledge, driver behaviour at curves and intersections has frequently been studied in previous research. How drivers negotiate curves has captured the attention of traffic psychologists and is arguably one of the most investigated aspects of driver behaviour using a range of methods from driving simulation to observational studies, including evaluations of potential countermeasures (e.g. Arien et al., 2013; Gawron & Ranney, 1990; Land & Horwood, 1995; Lehtonen, Lappi, Koirikivi, & Summala, 2014; Lehtonen, Lappi, Kotkanen, & Summala, 2013; Lehtonen, Lappi, & Summala, 2012; Milleville-Pennel, Jean-Michel, & Elise, 2007; Reymond, Kemeny, Droulez, & Berthoz, 2001; Shinar, Rockwell, & Malecki, 1980). Likewise, although there are still gaps in the literature, intersections (signalised, unsignalised and accesses) are arguably the second-most commonly studied aspect of the road environment in terms of the effect on driver behaviour. Studies have focused on how driver behaviour at intersections is affected by aspects such road geometry, road user type or driver characteristics (e.g. age, impairment) and also the effectiveness of potential countermeasures (e.g. Cornelissen, Salmon, & Young, 2013; Gstalter & Fastenmeier, 2010; Louveton, Bootsma, Guerin, Berthelon, & Montagne, 2012; Marti, Morice, & Montagne, 2015; Min, Min, & Kim, 2013; Montella et al., 2011; Plavsic, Klinker, & Bubb, 2010; Romoser, Pollatsek, Fisher, & Williams, 2013; Sager et al., 2014; Salmon, Lenne, Walker, Stanton, & Filtness, 2014; Werneke & Vollrath, 2012; Werneke & Vollrath, 2014; Yan & Radwan, 2007; Young et al., 2013).

In contrast, it is difficult to find studies that have investigated the effects of many other aspects of the built environment on driver behaviour, including driveways/laneways, service roads and roadside parking. Thus, these are the three risk factors identified in Component 1 that meet the first five criteria for risk factor selection.

Criterion 6. The risk factor should have different, and perhaps unexpected, effects on different crash types (Desirable).

The presence of a service road was associated with an increase in MVC on the main road (that is, on the through lanes because the small number of crashes that occurred in the service road were excluded). There was, however, no effect of service roads on SVC or PVC. This is not unexpected. Service roads separate drivers who want to access roadside businesses from through traffic. The speed of through traffic is likely to be higher when service roads are present because there is less need to monitor vehicles that are entering or leaving parking spaces, or pedestrians accessing the premises. Vehicles that slow to access the service road, or those that are leaving the service road to re-join the main thoroughfare may therefore be at risk of MVC because of the speed of through traffic.

Laneways and driveways are often difficult to see from the perspective of passing traffic, likewise, the view of the main road is often obstructed for vehicles that are waiting to turn. Hence driveways and laneways provide an opportunity for conflict between vehicles. In addition, in shopping areas, pedestrians are likely to be crossing in front of vehicles entering and exiting laneways and driveways. Yet, unexpectedly, as the number of driveway and laneways increased, the incidence of MVC actually decreased and there was no significant effect on PVC (or SVC). It is possible that drivers and pedestrians recognise the increased risk, and adjust their behaviour to account for the decreased safety margins. The effect on drivers could be measured in the simulator, however, the effect on pedestrians could not.

Roadside parking may influence crashes by several means. Parked vehicles are stationary roadside objects that vehicles could crash into. Thus, one might expect an increase in SVC where roadside parking is present. In Component 1, however, roadside parking was not associated with SVC, despite 42% of SVC on the strip shopping centre road segments involving a vehicle colliding with an object or parked vehicle. Parked vehicles, at some time, become moving vehicles that enter and leave parking spaces (perhaps unpredictably from the point of view of other road users). Thus, one might expect that the presence of parked vehicles would be associated with an increase in MVC. Parking on both sides of the road was, as expected, associated with increased MVC incidence. Unexpectedly, however, only 3% of the MVC involved vehicles entering or leaving parking spaces and only 5% involved collisions with an opening door or a person alighting or boarding the vehicle. Therefore, although parking was associated with MVC, only a small proportion of MVC were classified as parking-related crash types. Parked vehicles can also influence risk because they can obscure other moving vehicles and road users. This might also explain an increase in other types of MVC when parking is present. In addition, while PVC were not associated with the presence of parking per se, the presence of a parking clearway was associated with reductions in PVC. There are therefore a number of interesting aspects of the effect of roadside parking on driver behaviour in relation to different types of crashes that could be investigated in a driving simulator.

Thus while both laneways/driveways and roadside parking/clearways had different and potentially unexpected effects on different crash types, the multitude of ways that roadside parking can influence safety margins and the unexpected relationship with MVC and SVC make it a particularly interesting risk factor to investigate further.

Criterion 7. Is there a simple countermeasure for the risk factor, the effects of which can be measured in the same experiment? (Desirable)

If roadside parking does influence crash risk, then banning or reducing parking may be one option for reducing crashes. Retailers and business groups, however, vehemently oppose this option because of the perceived detrimental effect on business from passing trade. Thus other countermeasures might be necessary. One simple countermeasure is to reduce the speed limit where roadside parking is present in strip shopping centre road segments so that drivers have more time to respond to hazards. The effectiveness of this approach could easily be investigated in a driving simulator. In addition, this might shed some light on why PVC incidence was significantly lower in 40 and 50 km/h zones compared to 60 km/h zones.

8.2.3 Choice of risk factor

The criteria developed were used to choose one of the risk factors identified in Component 1 of the thesis for further investigation into its effect on driver behaviour in a simulator study. Based on the four essential and three desirable criteria presented, roadside parking was chosen as the risk factor for further study in this component of the thesis, in conjunction with investigating the effect of changing the speed limit as a potential countermeasure.

8.3 Roadside parking: crash risk and driver behaviour

Since roadside parking has been chosen as the risk factor to be investigated with respect to its effect on driver behaviour, it is an opportune time to reiterate what is known about the relationship between roadside parking and crash risk before progressing to a review of previous studies of roadside parking and driver behaviour.

8.3.1 Crash risk

Roadside parking is said to increase conflicts between vehicles, make it difficult for drivers to see pedestrians about to cross the road and can lead to "dooring" crashes when drivers and passengers in parked vehicles open their door into the path of cyclists (Johnson, Newstead, Oxley, & Charlton, 2013). In contrast, others have argued that roadside parking improves safety because it serves the dual purpose of slowing down through traffic (Marshall, Garrick, & Hansen, 2008) and providing a buffer between pedestrians and traffic (Dumbaugh, 2005).

The review of the relationship between characteristics of the built urban environment and traffic crashes in Chapter 3 ascertained that there was equivocal evidence surrounding the relationship between roadside parking and crash frequency, with some studies finding an increase in risk, and others failing to find any relationship (Alavi, 2013; Bonneson & McCoy, 1997; Greibe, 2003; Jonsson, 2005; Potts et al., 2007; Sawalha & Sayed, 2001). Component 1 of this thesis, however, established that the incidence of MVC was 47% higher when roadside parking was permitted on both sides of the road compared to when parking was not permitted on strip shopping centre road segments in metropolitan Melbourne (regardless of whether the type of parking was in lane, sheltered or angle). In addition, when there was a parking clearway that prohibited parking for certain time periods (usually peak traffic periods), the incidence of PVC was reduced by 34% compared to roads without a parking clearway. There was, however, no association between parking and SVC. Therefore parking does appear to influence road safety on strip shopping centre road segments in metropolitan Melbourne.

8.3.2 Observational studies

The mechanisms by which roadside parking may increase traffic crashes were investigated in an observational study conducted in Bogota, Columbia. The researchers observed how roadside parking influenced risk on urban roads with high concentrations of pedestrians and traffic (Bocarejo, Jiménez, & Torregroza-Vargas, 2013). Although there appeared to be less regulation and enforcement in Bogota than in the Melbourne metropolitan area, the types of conflict observed were still relevant for this study. The authors identified 19 hazardous situations involving roadside parking that were further classified into two higher-level categories. The first category of hazardous situation occurred when parked vehicles obstructed the vision of road users, for example pedestrians crossing the road (or vehicles turning into the road) being unable to see past a parked vehicle to see if it is safe to cross (or enter), as well as through traffic being unable to see these road users coming from behind parked vehicles. The second category involved increased conflict that occurs when road users have to stop and/or change trajectory due to parked or parking vehicles. This category includes situations where vehicles have to change lanes to avoid parked vehicles, where vehicles are entering or leaving parking spaces, when occupants of parked vehicles open the car door into the path of through traffic, when pedestrians stand between parked vehicles or on the road to wait for public transport and finally, when pedestrians that have crossed the road are forced to walk around a parked vehicle to get onto the footpath.

Therefore, according to Gibson and Crooks (1938) and the multiple comfort zone model (Summala, 2007) roadside parking has the potential to influence the field of safe travel through changes in lateral and longitudinal safety margins. Task demands are increased through the need to monitor moving vehicles, vehicles with the potential to move (perhaps unpredictably) and the potential for hidden obstacles (e.g. road users who might emerge from behind parked vehicles). Roadside parking also makes the visual environment more complex.

8.3.3 Experimental studies

A literature search revealed that there is a lack of published research into the effect of roadside parking on driver behaviour. Within the psychological literature, parking (or the avoidance of) is most often mentioned in relation to the self-regulatory behaviour of older drivers (e.g. Baldock, Mathias, McLean, & Berndt, 2006). The focus on the elderly and parking behaviour extends to experiments conducted to measure the effect of age and field of view on parking ability (Douissembekov, Michael, et al., 2015; Douissembekov, Navarro, et al., 2015). The effect of parked vehicles on the behaviour of other drivers, however, has rarely been investigated and the literature search discovered only one such study (Edquist et al., 2012).

The study was conducted in a driving simulator to measure the effect of roadside parking on driver behaviour and mental workload. The scenario was an urban four-lane road in a built-up area with a speed limit of 60 km/h. There were three experimental conditions relevant for assessing the influence of roadside parking on driver behaviour: a road with no parking bays, a road with parking bays marked on the roadside that were empty and a road with marked parking bays, 90% of which were occupied by parked vehicles. Compared to when no vehicles were present, mean speed was significantly slower and speed variability was significantly higher when 90% of parking bays were occupied. Drivers chose a lateral position further away from the kerb but there was no effect on the variability of lane position. Reaction time to a secondary choice reaction time task was slower and ratings of mental workload (measured using a modified NASA-TLX) were also significantly higher. In addition, when 90% of parking bays were occupied, reaction time to a pedestrian crossing the road was significantly slower compared to when there were no parked vehicles. This is interesting in the context of the results of thesis Component 1 which showed that PVC incidence is significantly lower when there is a parking clearway on the road segment. There was no significant difference in driving performance measures, workload or reaction times between the other two conditions where no parked vehicles were present; that is, where there were no marked parking bays or when parking bays were marked, but empty. Therefore, it was the presence of parked vehicles rather than marked parking bays that affected behaviour.

Parked vehicles therefore appear to affect perceived safety margins and lead to significant changes in vehicle control, drivers' mental workload (measured using self-report and secondary task) and response time to a hazardous event. The results of this experiment raise some interesting questions. In the previous study, 90% of parking bays were occupied by vehicles in the only condition in which there were vehicles parked on the roadside. What is the effect on driver behaviour and mental workload when the proportion of parking bays that are occupied varies? Is there a gradual consistent change in driving performance and mental workload, or is there a threshold beyond which drivers perceive risk and change their behaviour (as hypothesised by the multiple comfort zone model)?

The multiple comfort zone model also highlights the role of the driver's perception of risk and discomfort in changing behaviour, and the effect of reduced safety margins on task difficulty and mental workload. While the previous study measured workload using self-report and secondary task techniques, it is also of interest to measure if driver's self-reported feelings of risk, discomfort, and task difficulty vary in response to manipulations of the number of vehicles parked on the roadside. In addition, the previous study measured overall mental workload using a modified version of the NASA-TLX scale, however, the dimensions of the NASA-TLX are not necessarily relevant for the driving task (Pauzié et al., 2007) and the RSME scale is a more sensitive measure of overall workload (Veltman & Gaillard, 1996). Therefore the RSME was used to measure overall workload in this study.

The multiple comfort zone model (Summala, 2007) and zero-risk theory, from which it evolved (Näätänen & Summala, 1974) propose that there is threshold level of risk that must reached in order for drivers to perceive risk and act. In contrast, models such as Risk Homeostasis Theory and Risk Allostasis Theory (Fuller, 2011; Wilde, 1976) postulate that risk is continually monitored. A series of experiments that manipulated either speed or headway to a lead vehicle supported the notion that there is a threshold level of risk that drives behaviour (Lewis-Evans, 2012; Lewis-Evans et al., 2010; Lewis-Evans & Rothengatter, 2009). For example, when driving in a simulated urban environment, ratings of risk and task difficulty did not vary at low speeds (20 to 40 km/h) but did increase as speeds increased from 40 km/h to 100 km/h (Lewis-Evans & Rothengatter, 2009). It is unknown, however, whether a threshold for the perception of risk exists when the manipulation involves aspects of the physical environment (e.g. the number of parked vehicles).

It is also useful to investigate the effect of a simple countermeasure that could potentially negate some of the risk associated with roadside parking. Speed limits are often reduced in strip shopping centre road segments (which often have roadside parking), particularly those with high pedestrian volumes and a pedestrian crash problem. Therefore, a reduction in speed zone was investigated as a potential countermeasure to reduce the risk associated with roadside parking.

Another reason that the effect of speed zone is of interest is that PVC incidence was found to be significantly higher when parking clearways were absent and lower in 40 km/h and 50 km/h zones compared to 60 km/h zones in Component 1 of this thesis. The previous study found that when 90% of parking bays were occupied with parked vehicles, reaction time to a pedestrian crossing the road was significantly slower. It is possible that reducing the speed of travel might lead to faster response times to a crossing pedestrian in complex environments, for example, when parking bays are almost full.

8.4 Aims

These questions led to the formulation of the aims of this study which were to investigate the mechanisms by which roadside parking influences crash risk in urban environments and whether changing the speed limit is a valid countermeasure for reducing risk in areas with roadside parking. As one of the first studies to investigate the effect of roadside parking on driver behaviour, this study focused on examining the effect of stationary parked vehicles, rather than risk related to when vehicles enter or leave parking spaces.

Specifically, the primary aims of this study were to investigate the interactive effect of the proportion of parking bays that were occupied and speed zone on:

- driving performance (lateral and longitudinal vehicle control)
- self-reported risk, discomfort and task difficulty
- workload as measured using the RSME unidimensional scale and a secondary choice peripheral detection task

The secondary aims of this study were to:

- investigate the relationship between measures of driving performance and subjective measures
- establish whether there is a threshold level for perceiving risk in relation to the number of cars parked on the roadside
- ascertain whether drivers act to maintain a stable level of perceived risk, discomfort task difficulty and effort
- determine the effect of travel speed on driver response to an unexpected event in a highly complex environment (a pedestrian crossing the road when 90% of parking bays were occupied)

CHAPTER 9. METHODS: DRIVING SIMULATOR STUDY

9.1 Participants

The participants in this experiment were all experienced drivers because it was desirable to have a homogeneous group whose performance or workload were not impacted by lack of experience or age (e.g. Patten et al., 2006).

9.1.1 Inclusion criteria

To be eligible for inclusion in the study, participants were required to:

- be aged between 25 and 55 years,
- drive at least 5,000 km per year,
- have a full Australian driver's licence, and
- have at least five years of driving experience.

9.1.2 Exclusion criteria

People with severe motion sickness or epilepsy were not eligible to participate in the study due to the increased chance of experiencing adverse effects in the driving simulator.

9.1.3 Recruitment

Participants were recruited from staff and students of Monash University and from the MUARC Expression of Interest Database which includes the contact details of people who have expressed interest in participating in studies in the MUARC driving simulator. People listed on the database who were eligible for the study were emailed to invite them to participate. Other participants were recruited via advertisements.

9.1.4 Participant sample

Thirty-two experienced drivers completed the study: 16 males and 16 females. The average age of male participants was 40.3 years (range=28 to 54, standard deviation (sd) =9.8) and for female participants was 36.3 years (range=27 to 54, sd=8.6). On average, male participants had held their licence for 21.6 years (range 7 to 36, sd=10.2) while female participants had held their licence for 16.8 years (range 6 to 35 years, sd=7.8). The differences between males and females in terms of age and driving experience were not statistically significant ($p=0.23$ and $p=0.15$, respectively).

9.2 Materials

9.2.1 Equipment

The experiment was conducted in an Eca-Faros EF-X driving simulator with software modified for research purposes. The simulator cockpit comprised an adjustable seat with seat belt, steering wheel and indicator lever, gear shift, accelerator and brake pedals with force feedback and a dashboard with speedometer. Participants were not required to change gears while driving the simulator. The visual display comprised three monitors positioned side by side to give a forward view of 120^0. The view from a central rear vision mirror and side mirrors was also displayed on the monitors. Data were collected at 30 Hz.

Three questionnaires were administered to the participants. Questionnaire 1 gathered data on age, gender and driving experience to ensure participants met the inclusion criteria. Questionnaire 2 collected information regarding the participant's current state of wellbeing in order to assess simulator sickness symptoms (this questionnaire was administered before and after the session). Questionnaire 3 was designed to allow participants to rate their effort using the Rating Scale Mental Effort (RSME), developed by Zijlstra (1993), task difficulty, perceived risk and perceived discomfort on a 20 point scale, and their preferred and maximum travel speeds (similar to Lewis-Evans & Rothengatter, 2009) across the experimental conditions. The questionnaires are included in Appendix C.

9.3 Design

The experiment adopted a repeated measures (within-subjects) design with two repeated factors: speed zone and the proportion of parking bays that were occupied with vehicles. There were four speed zones, based on those found on urban roads where roadside parking is common in Victoria, Australia: 40 km/h, 50 km/h, 60 km/h and 70km/h. These speed zones were also chosen so the level of perceived risk increased with increases in speed (as per Lewis-Evans & Rothengatter, 2009)

There were also four levels of the experimental parking condition: empty parking bays with kerb extensions (but no vehicles) occupying 10% of the parking bays (a narrowing of the roadway where the raised kerb/footpath is extended into the roadway, refer to panel A of Figure 9.1) and conditions where either 10%, 50% or 90% of parking bays were occupied by cars, as shown in panels B, C and D of Figure 9.1, respectively. Kerb extensions were used in the condition where there were no vehicles occupying the parking bays so that the usable road width was the same as in the other conditions with parked vehicles. The locations of occupied bays were randomly chosen from all parking bays in a block for the conditions where 50% or 90% of parking bays were occupied by vehicles. For the conditions where 10% of parking bays were occupied (either with kerb extensions or vehicles), one in every ten parking bays were occupied, with between eight and ten empty parking bays between occupied locations. These conditions were imposed because it was discovered during scenario design and pilot testing that random allocation of kerb extensions or vehicles to parking bays in the 10% conditions did not generate a uniform perception of occupied bays; rather, it generated some scenarios that had long distances without any vehicles followed by groups of occupied bays. Hence this was a 4x4 factorial, repeated measures design with 16 experimental conditions in total.

Figure 9.1 Experimental conditions: parking scenarios

It was not practical to fully counterbalance the speed and parking conditions as this would have meant participants had to change their travel speed within a drive (which would have resulted in more variable driving performance). Instead, a repeated measures block design was used. Participants drove four separate drives in the simulator on the same road but the road was sign-posted with a different speed zone in each of the four drives (that is, the presentation was blocked on speed zone). The order of presentation of the speed zones was fully counterbalanced across participants to control for sequential effects. Within each drive, there was a one km section of road for each of the four parking conditions and the order of parking conditions within each drive was also fully counterbalanced.

9.3.1 Scenario design

The scenario was an urban environment that was representative of the strip shopping centre road segments with roadside parking that were included in Component 1 of this thesis. There was a road with two lanes in each direction, footpaths and multi-storey buildings built up to the footpath on both sides of the road (Figure 9.1). The road was undivided and a double white line separated traffic from opposing directions. Both of the outer lanes contained marked parking bays and there was therefore only one lane in either direction for moving traffic. Simulated pedestrians walked along the footpath and traffic travelled at the speed limit in the opposite direction to the participant's vehicle. There was, however, no traffic travelling in the same direction as the participant (so the participant could travel in free speed conditions without being influenced by other vehicles).

Each drive was comprised of six blocks, with each block separated by a signalised intersection where the traffic light was always green. The first block was 500m long with empty parking bays and a speed limit sign. The purpose of this block was for participants to have enough time to accelerate to reach the speed limit prior to reaching the blocks with the experimental parking conditions. No data were analysed from the first 500m block. The next four blocks were comprised of a one km long block for each parking condition. Driving performance data were collected when the drivers were driving through these four blocks. The sixth block was also one km long and always had 90% of parking bays occupied by cars. Approximately halfway through the sixth block, an unexpected safety-critical event occurred: a pedestrian turned and walked out onto the road in front of the participant's vehicle. The timing of the unexpected safety-critical event depended on the speed limit of the drive: as the speed zone increased, so too did the amount of time (and therefore distance) the participant was given to react to the pedestrian to account for the increased travel speed. For example, in the 70 km/h condition, the pedestrian walked out onto the road when the vehicle was at a farther distance away than in the 60 km/h condition (likewise for slower speeds).

Although each of the four drives that the participants experienced had six blocks, the experimenter stopped participants at the end of the fifth block during their first three drives. It was only in their fourth and final drive that they drove through the sixth block and encountered the unexpected safety-critical event of the pedestrian crossing the road. The reason for this was that the primary aim of the study was to measure driving performance in response to changes in speed zone and roadside parking. Measuring the response to the safety-critical unexpected event was a secondary aim of the study. Leaving the unexpected event to the last block of the last drive ensured that participants' primary driving performance was not affected by expectation of another critical event. It was suspected that if the safety-critical unexpected event occurred earlier in the scenarios that participants would drive differently whenever they saw a pedestrian on the side of the road which was undesirable. The counterbalanced presentation order of speed zone across the four drives ensured one-quarter (eight) of the participants experienced the unexpected safety-critical event in each speed zone.

9.4 Procedure

Participants attended for one session of approximately an hour. At the beginning of the session, the study was described to the participants. They were then asked to sign a consent form and given Questionnaires 1 (demographics) and 2 (wellbeing). Participants then sat in the simulator while the controls were explained to them and were given two practice drives to attempt to drive at a constant speed through an environment similar to the test scenarios. Participants were also asked to practice braking and coming to a complete stop. The second practice drive was similar to the first, but participants also had to respond to a secondary task. The secondary task was a choice peripheral detection task. At intervals, a black stylised icon of a pedestrian would appear to the left or to the right of the screen and remain for approximately five seconds. Participants were asked to respond to the stimulus as soon as possible, by turning the indicator lever in the direction of the icon.

Once the two practice drives were completed and participants were comfortable performing the required tasks, the four data-gathering drives were conducted. Participants were asked to obey road rules, to drive straight ahead at intersections and to drive as close as possible to the speed limit, otherwise, to drive as they would in the real world. Normally when drivers feel an increased sense of risk they decrease their travel speed, however, this experiment was designed to measure risk and workload as speed increased and the number of parked vehicles increased which was the reason for asking them to drive at the speed limit. Participants were asked to verbally report the speed limit when they drove past the speed limit sign in the first block. If they forgot, they were prompted by the experimenter to ensure that they knew the speed limit of the road. This was the only time participants were reminded about the speed limit. Participants were instructed that driving was their primary task, but that they should also respond to the secondary task as quickly and accurately as they could. After each drive, the participants completed Questionnaire 3.

The study was approved by the Monash University Human Research Ethics Committee (project number CF10/3352–2010001768). Participants were reimbursed 30 AUD for their participation.

9.5 Measures

9.5.1 *Self-report*

At the conclusion of each of the four drives, Questionnaire 3 was administered in which drivers were shown an image of each of the four experimental parking conditions in the order that they experienced them, and for each image, they were asked to rate how much effort it took them to drive that section of road at the speed they had just driven using the RSME scale (Zijlstra, 1993). Participants were also asked to rate, on a 20 point scale, how difficult they found driving each section of road at that speed (from not difficult to extremely difficult), how much risk they experienced (from no risk to maximum risk) and how comfortable they felt (from extremely comfortable to extremely uncomfortable). Participants were asked to report their preferred driving speed and the maximum speed at which they would travel in each experimental parking condition.

9.5.2 *Driving performance*

9.5.2.1 *Lateral position and control*

The position of the vehicle in the lane was measured in centimetres (cm) relative to the centre of the lane, where a position of zero cm corresponded to the centre of the vehicle being equidistant from each lane edge; negative values indicated that the centre of the vehicle was closer to the kerb (the left hand side of the road in Australia); and positive values indicated the centre of the vehicle was closer to the midline of the road (the right hand side in Australia). The mean lane position (MLP) and standard deviation of lane position (SDLP) were calculated for each experimental condition. The number of lane excursions was also tallied for each experimental condition, where a lane excursion was defined as any instance where any part of the vehicle crossed the lane boundary on either side.

9.5.2.2 *Longitudinal control*

Travel speed was measured throughout the drive. The mean travel speed and standard deviation of travel speed (SDSp) were calculated for each experimental condition. The difference between the mean travel speed and the speed limit (speed differential) was calculated and compared across conditions. The mean and standard deviation (SDAcc) of the pressure on the accelerator, in terms of the percentage of the maximum possible pressure, was also calculated for each condition.

9.5.3 Response to secondary task

Participants performed the choice peripheral detection task throughout each drive. The time that elapsed from presentation of each stimulus in the drivers' periphery to their response using the indicator lever was recorded. The response was defined as correct (or a hit) if they turned the indicator lever in the direction of the stimulus, incorrect if they turned the indicator lever in the opposite direction, or a miss if they did not respond to the stimulus by turning the indicator lever. For misses, the time that elapsed between presentation of the stimulus and when the stimulus disappeared was recorded.

9.5.4 Response to unexpected safety-critical event

The time that elapsed between when the pedestrian turned to start crossing the road and when the participant first depressed the brake pedal was recorded. This was defined as the reaction time to the unexpected safety-critical event.

9.6 Data analysis

Analyses were performed using IBM SPSS Statistics, Version 20 and Stata SE, Version 11.2 (StataCorp, 2009).

9.6.1 Self-reported risk, discomfort, task difficulty, effort and objectively measured driving performance

Generalized Estimating Equations (GEE, Liang & Zeger, 1986) with the identity link and an unstructured correlation matrix were used to test the interaction between, and the main effects of, the independent variables (speed zone and parking condition) on self-reported risk, discomfort, task difficulty and effort and objective measurements of MLP, SDLP, speed differential, SDSp and SDAcc. GEE were also used to investigate the relationship between driving performance measures and self-reported effort.

9.6.2 Reaction time to secondary task

Responses to the peripheral stimuli were classified into correct responses, incorrect responses and misses. There were 1408 hits, 5 incorrect responses and 18 misses.

The choice reaction times to the secondary peripheral detection task were analysed in two different ways. The first was by conducting a GEE on mean reaction times of correct responses, which is equivalent to a repeated measures ANOVA; these are the most common methods for analysing this type of data in the psychological literature. The second method involved conducting time to multiple event analysis (survival analysis) on individual reaction times which is not commonly used but is likely to be more powerful than the standard method. These methods are explained further detail below.

9.6.2.1 Analysis approach 1: GEE using mean correct reaction times for each condition

In psychological research, the most common method for analysing reaction times to a secondary task is to calculate the mean reaction time to correct responses in each of the 16 (4 levels of parking by 4 levels of speed limit) experimental conditions and to conduct a GEE (or the equivalent repeated measures ANOVA) to determine if the mean reaction time varies according to the experimental manipulations. According to the standard method, mean reaction time for hits (correct responses) was calculated for each participant in each speed-parking condition (32 participants by 16 conditions gives 512 data points). A GEE (with the identity link, normal error distribution and an unstructured correlation matrix) was used to determine the effect of speed zone and parking condition on mean correct reaction times (Liang & Zeger, 1986).

9.6.2.2 Analysis approach 2: Multiple time to event analysis

Another, potentially more appropriate, analysis technique for analysing reaction time data is time to event analysis (or survival analysis). The second analysis approach used a Cox proportional hazards regression (survival analysis) for multiple events. Each presentation of the peripheral detection stimulus was an "event" and the survival analysis approach analysed individual reaction times to every stimulus (that is, to multiple events), rather than using the mean reaction times averaged across each speed-parking condition. Incorrect responses were excluded. Therefore there were 1,426 data points for analysis using this method. Thus more information in the data is used compared to performing a GEE on the mean correct reaction times for each condition. In survival analysis, presentations of the stimulus that are not responded to (misses) are censored at the time the stimulus disappeared. Survival analysis has been recommended for use in psychological research (Landau, 2002) and for analysing hazard perception data in driving experiments (Parmet, Meir, & Borowsky, 2014). There are some examples of simple survival analysis in the psychology literature but there appear to be no examples where time to multiple event analysis has been used for situations where participants were exposed to more than one stimulus. The current study may be the first to use such an approach.

The time to event (equivalent to the reaction time) was measured from the time from presentation of the stimulus to the response (or disappearance of the stimulus, in the case of censored observations when targets were missed). The multiple time to event method for unordered events of the same type was used, with robust variance estimates, the efron method to deal with tied observations, clustered on participant number to take into account the repeated measures nature of the data (Cleves, 2009). Incorrect responses were excluded from the analysis. Although competing risk survival analysis methods can be used to accommodate competing events (that is, events which preclude the event of interest occurring, for example making an incorrect response instead of a correct one), methods for incorporating competing events in time to multiple event models have not yet been developed. This was not a major concern considering only 0.3% of the stimuli elicited an incorrect response.

The hazard is a measure of the instantaneous risk of responding to the stimulus at a particular time point, given that a response has not yet been made up until that point. The hazard ratio (HR) is a way to compare the hazard (or reaction time) between two conditions and is calculated from the Cox proportional hazards regression analysis (Cox & Oakes, 1988). The HR is equal to the hazard for the experimental condition of interest compared to the hazard for the reference (or comparison) condition. A hazard ratio greater than one indicates that the probability of responding is greater in the experimental condition of interest than for the comparison condition (and therefore, the reaction time is faster). A hazard ratio of less than one indicates that the probability of responding is lower in the experimental condition of interest than for the comparison condition (and therefore, reaction time is slower). Linear combinations of coefficients in the regression model were used to calculate pairwise comparisons between conditions. The Cox proportional hazards regression assumes that the hazards are proportional across groups—this assumption should be tested after estimation of the model (Cox & Oakes, 1988).

Because the multiple time to event analysis (the second analysis approach) uses more information in the data (including each individual correct reaction time rather than a mean, and including the censored time for stimuli that were missed), the method is expected to be more sensitive than the standard psychological method (GEE using mean correct reaction times).

9.6.3 Reaction time to safety-critical unexpected event

The reaction times to the safety-critical unexpected event (a pedestrian crossing the road) were analysed using Cox proportional hazards regression (survival analysis) to determine if the hazard, or the time to react, to the crossing pedestrian differed across speed zones. Linear combinations of coefficients in the regression model were used to calculate pairwise comparisons between conditions. The assumption of proportional hazards was tested after estimation of the model (Cox & Oakes, 1988).

CHAPTER 10. RESULTS: DRIVING SIMULATOR STUDY

The analyses and results of the driving simulation study are presented in this chapter. First, the effects of the experimental manipulations of parking condition and speed zone on self-report measures are presented, followed by their effect on driving performance measures. Next, the effect of parking condition and speed zone on reaction time to the secondary choice peripheral detection task is described. The relationship between driving performance measures and self-report measures is also analysed. Finally, the effect of speed zone on the reaction times to the safety-critical unexpected event (a pedestrian crossing the road) are presented.

10.1 Self-reported measures

In general, mean ratings for the four self-report measures of perceived risk, discomfort, task difficulty and effort increased as the level of occupied parking bays increased and as speed limit increased, with the differences between conditions getting larger as environments become more complex. Patterns of responses were quite similar between the four different measures. There was a high degree of correlation between the four self-report measures (Table 10.1).

Table 10.1 Correlation between self-report measures: perceived risk, discomfort, task difficulty and effort

	Perceived risk	Discomfort	Task difficulty	Effort
Perceived risk	1.000			
Discomfort	0.802	1.000		
Task difficulty	0.893	0.840	1.000	
Effort	0.789	0.686	0.789	1.000

10.1.1 Self-reported risk

Self-reported risk increased as the proportion of occupied parking bays increased and the speed limit increased. When more parking bays were occupied, an increase in speed limit led to a greater increase in perceived risk than when fewer parking bays were occupied (refer to Figure 10.1 which shows the mean and 95% CI of self-reported risk for each condition calculated from the raw data). GEE analysis revealed a significant interactive effect of parking condition and speed zone on perceived risk ($p<0.0005$). Any increase in the proportion of parking bays that were occupied by vehicles led to a statistically significant increase in perceived risk in all but one situation (when the perceived risk when 10% of parking bays were occupied by kerb extensions or vehicles in 60 km/h zones were compared). There were also statistically significant increases in feelings of risk whenever speed increased in all parking conditions (p-values range from <0.0005 to 0.039), with the one exception being that when 10% of parking bays were occupied by vehicles, feelings of risk in 50 km/h zones did not significantly differ to those in 60 km/h zones ($p>0.05$).

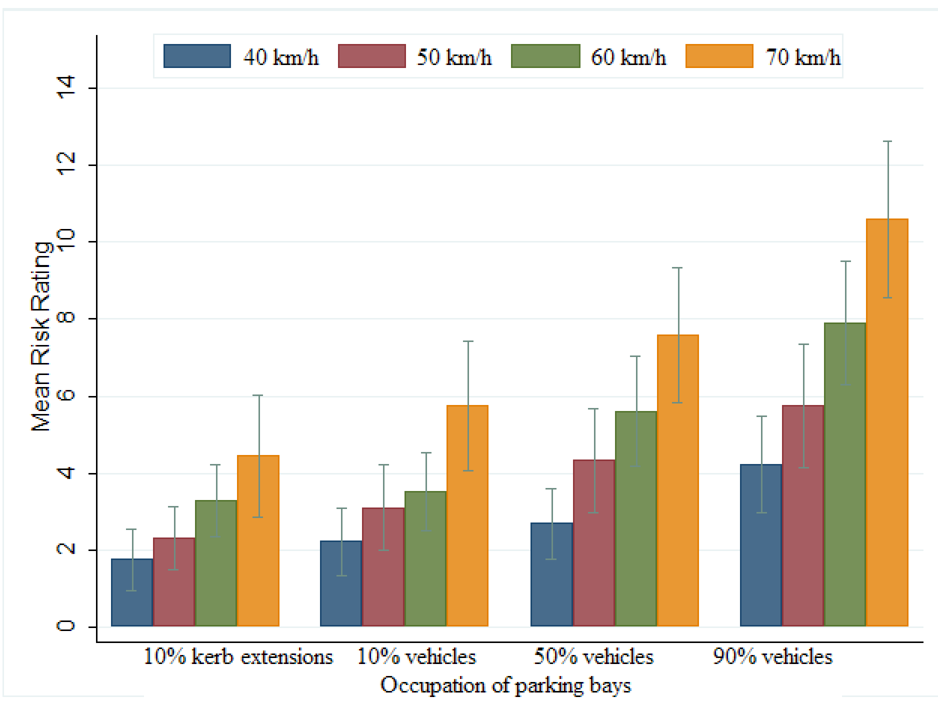


Figure 10.1 Mean (95% confidence interval) ratings of risk (out of 20) across speed and parking conditions

10.1.2 Self-reported discomfort

Figure 10.2 displays the mean ratings of discomfort (with 95% CIs) across speed zone and parking conditions (calculated from the raw data). Similar to the other self-report measures, there was a statistically significant interactive effect of speed zone and parking condition on self-reported discomfort ($p<0.0005$). Again, the effect of manipulating the speed zone on ratings of discomfort were ascertained for each parking condition.

If 90% of parking bays were occupied, any increase in speed led to a significant increase in self-reported discomfort (p-values range from $p=0.003$ to $p=0.02$).

When 50% of parking bays were occupied, discomfort was rated as significantly higher when the speed limit was 70 km/h compared to all other speed zones ($p\leq0.001$) and significantly lower when the speed limit was 40 km/h compared to than all other speed zones ($p\leq0.002$). There was no significant difference in self-reported discomfort between the 50 km/h and 60 km/h zones when 50% of parking bays were occupied by vehicles ($p>0.05$).

When 10% of parking bays were occupied, discomfort was significantly higher when travelling at 70 km/h than for lower speeds. ($p\leq0.001$). Discomfort was also significantly higher when travelling at 60 km/h compared to 40 km/h ($p<0.005$). There were no other statistically significant differences in self-reported discomfort when 10% of parking bays were occupied ($p>0.05$).

Finally, when there were no parked vehicles but kerb extensions were located in 10% of parking bays, ratings of discomfort were significantly higher when travelling at 70 km/h compared to slower speeds ($p<0.005$), however, there was no significant difference between the other speed zones ($p>0.05$).

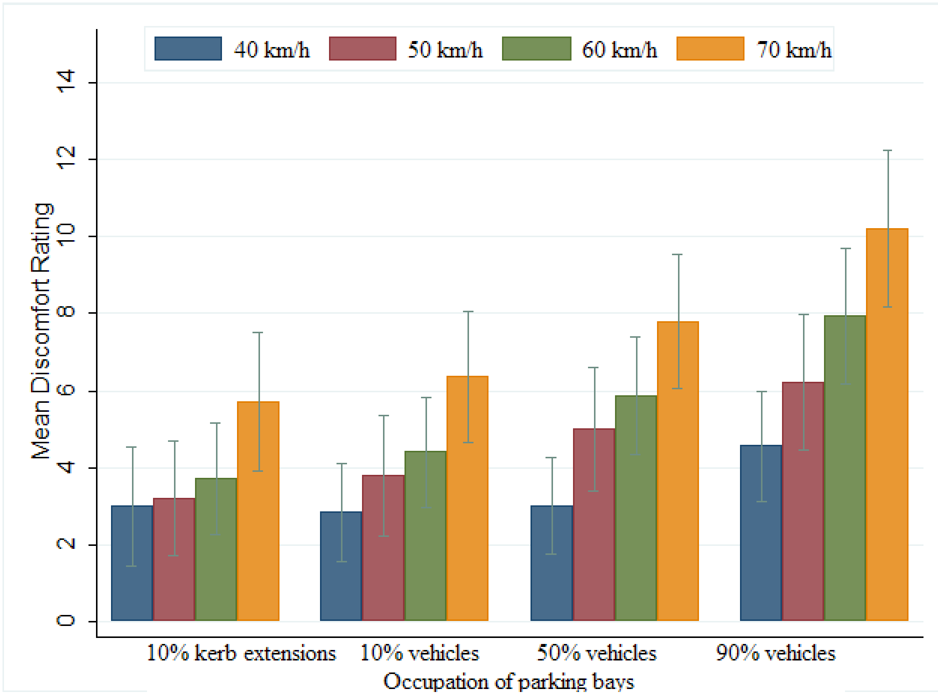


Figure 10.2 Mean (95% confidence interval) ratings of discomfort (out of 20) across speed and parking conditions

10.1.3 Self-reported task difficulty

Figure 10.3 shows the mean ratings of task difficulty (with 95% CIs), calculated from the raw data, across the experimental conditions. Ratings of task difficulty increased as the proportion of parking bays that were occupied increased and as the speed zone increased. GEE analysis revealed a significant interactive effect of parking condition and speed zone on self-reported task difficulty and discomfort (both $p<0.0005$). Further investigation therefore assessed the effect of speed zone on self-reported effort for each level of the parking condition.

When 90% of parking bays were occupied, any increase in speed resulted in a statistically significant increase in self-reported task difficulty (p-values ranged from $p<0.005$ to $p=0.038$).

When 50% of parking bays were occupied, task difficulty was rated as significantly lower when the speed limit was 40 km/h compared to than all other speed zones (all $p\leq0.007$) while ratings of task difficulty were significantly higher when the speed limit was 70 km/h compared to all other speed zones (all $p\leq0.007$). There was no significant difference in self-reported task difficulty between the 50 km/h and 60 km/h zones when 50% of parking bays were occupied by vehicles ($p>0.05$).

When cars occupied 10% of the parking bays, only the 70 km/h speed zone was associated with a significant increase in task difficulty or discomfort (p=0.013); differences between all other speed zones were not significantly different (p>0.05).

Finally, when there were no parked vehicles and kerb extensions were located in 10% of the parking bays, task difficulty was rated as significantly higher when travelling in a 70 km/h zone compared to 40 km/h (p=0.001) and 60 km/h zones (p=0.013). There were no other statistically significant differences between speed zones when kerb extensions were present but there were no parked cars (p>0.05).

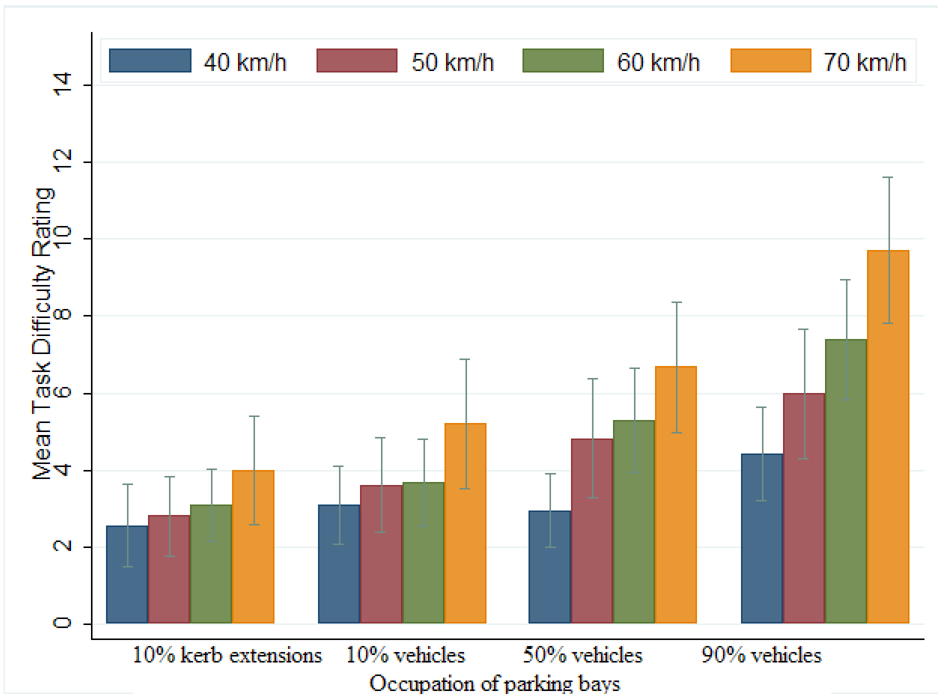


Figure 10.3 Mean (95% confidence interval) ratings of task difficulty (out of 20) across speed and parking conditions

10.1.4 Self-reported effort

Figure 10.4 displays the mean and the 95% CI participant ratings of effort on the RSME for each experimental condition, calculated from the raw data. The y-axis represents the mean effort rating, with the verbal anchors from the RSME shown at the corresponding y-value. Ratings of self-reported effort increased as the proportion of parking bays that were occupied increased and as the speed zone increased. Using the RSME descriptive anchors, ratings of effort ranged from between almost none and a little for the scenarios with lower speed limits when few parking bays were occupied to between rather much and considerable for the most complex scenario when 90% of parking bays were occupied by vehicles and the speed limit was 70 km/h.

GEE analysis revealed a significant interactive effect of parking condition and speed zone on self-reported effort ($p<0.0005$). Comparisons were performed to assess the effect of speed zone on self-reported effort for each level of the parking condition.

When kerb extensions were present or when 10% of parking bays were occupied, the effort required to drive at 70 km/h was significantly higher than the effort required to drive at the lower speed zones ($p\leq0.005$), which did not differ significantly from each other ($p>0.05$; left hand panels of Figure 10.4).

When 50% of parking bays were occupied, effort was rated as significantly higher when driving at 70 km/h than for lower speeds ($p\leq0.004$), and as significantly higher in the 60 km/h speed zones relative to lower speed zones ($p<0.02$). Effort required to drive at 40 km/h and 50 km/h did not differ significantly ($p>0.05$) when 50% of parking bays were occupied with vehicles.

When 90% of parking bays were occupied, any increase in speed led to a significant increase in self-reported effort (p-values ranged from $p<0.0005$ to $p=0.017$).

Furthermore, as the proportion of parking bays that were occupied increased, the difference in self-reported effort between speed zones generally became larger. For example, increasing the speed zone from 60 km/h to 70 km/h when 90% of parking bays were occupied led to a greater increase in effort (12.0 units, 95% CI 5.8 to 18.1) than when 50% of parking bays were occupied (8.3 units, 95% CI 2.6 to 14.0).

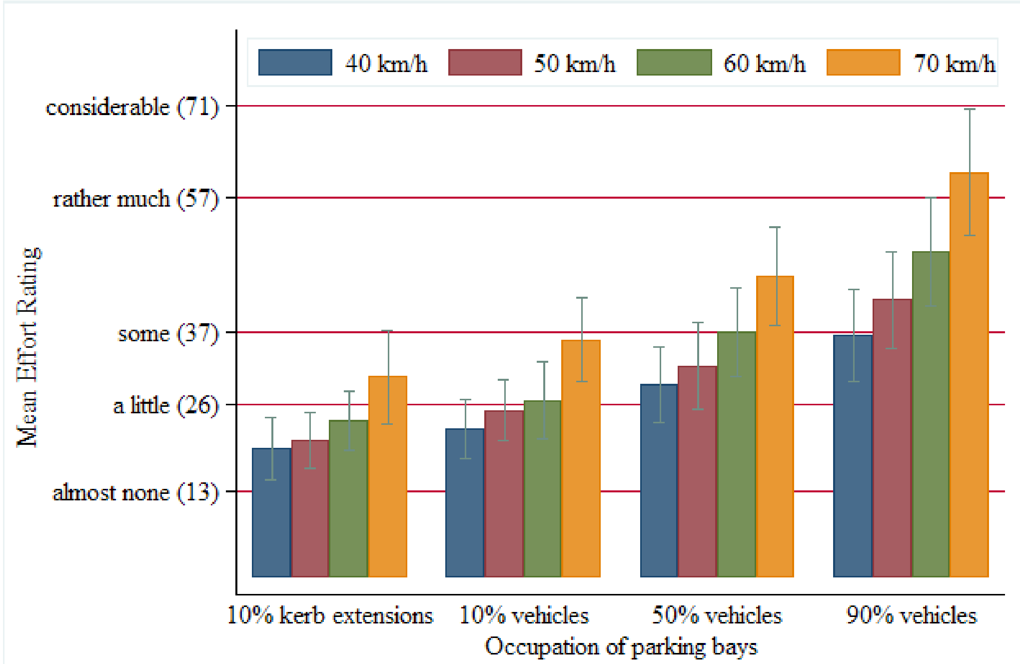


Figure 10.4 Mean (95% confidence interval) ratings of effort (maximum 150) across speed and parking conditions

10.1.5 Self-reported preferred and maximum travel speeds

Participants were asked to report their preferred travel speed for each parking condition, as well as the maximum speed at which they would travel. The mean preferred and maximum travel speeds are displayed in Figure 10.5 for each parking condition. Preferred travel speed and maximum travel speed both decreased significantly as the level of occupied parking bays increased (both $p<0.001$). The approximate level of risk, discomfort, task difficulty and effort that were perceived when driving at the self-reported preferred speed in each parking condition were estimated from Figures 10.1 to 10.4 to establish whether the preferred speed in a particular environment was chosen to keep subjective levels constant. In addition, the preferred speed was compared to the threshold for when changes in speed led to significant changes in risk, discomfort, task difficulty and effort.

There was no significant difference ($p>0.05$) in the preferred speed when kerb extensions were present (mean=59.2 km/h, 95% CI 56.7 to 61.7 km/h) compared to when 10% of parking bays were occupied (mean=57.2 km/h, 95% CI 55.0 to 59.4 km/h). According to Figure 10.4, travelling at approximately 60 km/h in these parking conditions corresponds to an effort rating of less than "a little". Risk, discomfort and task difficulty were all rated as between three and four out of 20 when speed was approximately 60 km/h and 10% of parking bays were occupied by cars or kerb extensions. In addition, the preferred speed was related to the threshold level for perceiving a change in effort, task difficulty and discomfort. These measures did not significantly increase with increases in speed limit until the speed exceeded 60 km/h (approximately equal to the preferred speed). Perceived risk, however, significantly increased with any increase in speed limit so the threshold for perceived risk was unrelated to preferred speed when 10% of parking bays were occupied by cars or kerb extensions.

When 50% of parking bays were occupied, mean preferred speed (51.1 km/h, 95% CI 48.9 to 53.3 km/h) was significantly lower than the two conditions with fewer parking bays occupied ($p<0.005$) and corresponded to a rating of between "a little" and "some" effort (refer to Figure 10.4). Risk was rated as just over 4 out of 20 while discomfort and task difficulty were rated as approximately 5 out of 20 when speed was approximately 50 km/h and 50% of parking bays were occupied. Moreover, the preferred speed was related to the threshold level for perceiving a change in effort: effort did not significantly increase with increases in speed limit until the speed limit increased beyond 50 km/h (approximately equal to the preferred speed). Perceived risk and task difficulty, however, increased with any increase in speed limit, and discomfort did not significantly increase until the speed limit was beyond 60 km/h so the preferred speed was not the threshold for significant changes in risk, discomfort or task difficulty when 50% of parking bays were occupied by vehicles.

When 90% of parking bays were occupied, mean preferred speed (46.7 km/h, 95% CI 44.0 to 49.4 km/h) was significantly lower than for all other parking conditions (range: $p<0.005$ to $p=p=0.012$), corresponding to an effort rating of more than "some" (refer to Figure 10.4). The comparable risk rating was approximately 5.5, while discomfort and task difficulty were rated as 6 out of 20 when speed was between 40 km/h and 50 km/h and 90% of parking bays were occupied. All of the self-report measures significantly increased when there was any increase in speed limit beyond 40 km/h (the lowest speed used in this study) when 90% of parking bays were occupied.

Thus while participants preferred lower travel speeds as the proportion of occupied parking bays increased, the reduction in preferred speeds did not serve to maintain a stable rating of risk, discomfort, task difficulty or effort. In addition, there is some evidence that the preferred speed is the threshold beyond which an increase in speed will consistently lead to a significant increase in perceived effort, but not perceived risk, discomfort or task difficulty.

Maximum travel speeds followed a similar pattern (except that maximum travel speeds were higher than preferred travel speeds).

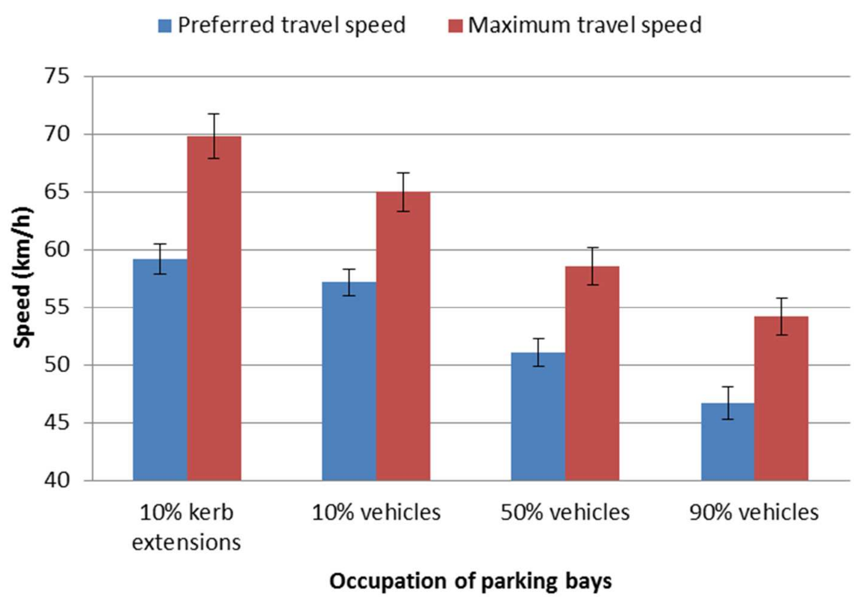


Figure 10.5 Mean (standard error) self-reported preferred speed and maximum speed for each parking condition

10.2 Driving performance

10.2.1 Lateral position and control

Measures of lateral position and control included the number of lane excursions, MLP and SDLP. GEEs were conducted to determine if lateral position and control differed according to parking condition and/or speed zone.

10.2.1.1 Lane excursions

There were only two lane excursions (instances where any part of the vehicle exceeded the lane boundaries). Both involved the same participant exceeding the lane boundaries during the same drive (40 km/h); once in the block when 10% of parking bays were occupied and once in the block when 50% of parking bays were occupied. So few lane excursions occurred that no statistical analyses could be conducted.

10.2.1.2 Mean lane position

There was a significant effect of parking condition on MLP ($p<0.001$). Figure 10.6 shows that MLP was 10.4cm left of the centre of the travel lane (closer to the kerb) when there were no vehicles present and moved 8.5cm ($p<0.001$) further away from the kerb (1.9cm left of the centre of the lane) when 10% of the parking bays were occupied with vehicles. When 50% or 90% of the parking bays were occupied by vehicles, MLP was slightly to the right of the centre of the lane (1.3cm right and 0.6cm right, respectively) and while this was significantly different from the two conditions in which 10% of parking bays were occupied ($p=<0.001$), the 50% and 90% occupied conditions did not differ in terms of mean lane position ($p>0.05$). There was no significant effect of speed zone on MLP nor was there a significant interaction between speed zone and parking condition ($p>0.05$).

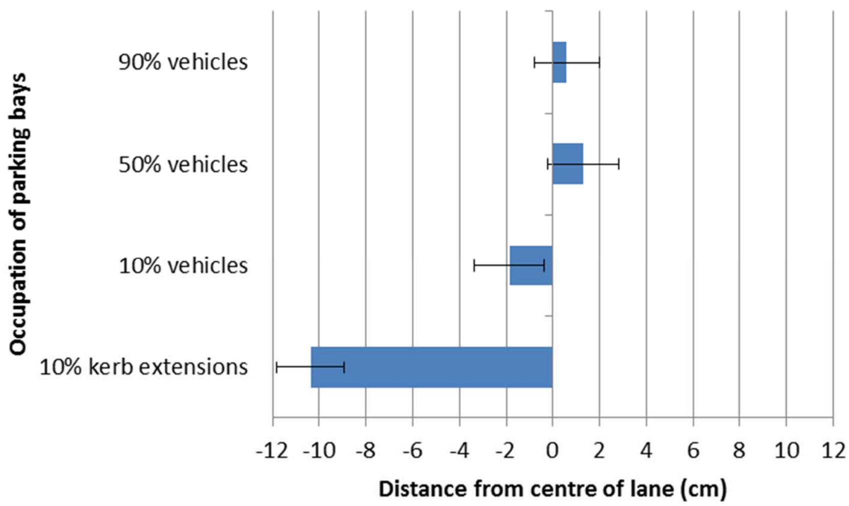


Figure 10.6 Mean (standard error) lane position across parking conditions (zero= middle of the lane of travel, negative values indicate a position closer to the kerb and positive values indicate a position closer to the centre-line)

10.2.1.3 Standard deviation of lane position

There was also a significant effect of parking condition on SDLP ($p<0.001$) (refer to Figure 10.7). SDLP was significantly higher when there were no vehicles present (18.9 cm) than all conditions when there were vehicles present (16.5 to 17.3 cm, $p<0.001$) whereas the SDLP did not differ according to whether 10%, 50% or 90% of parking bays were occupied by vehicles ($p>0.05$). There was no significant effect of speed on SDLP nor was there a significant interaction between speed and parking ($p>0.05$).

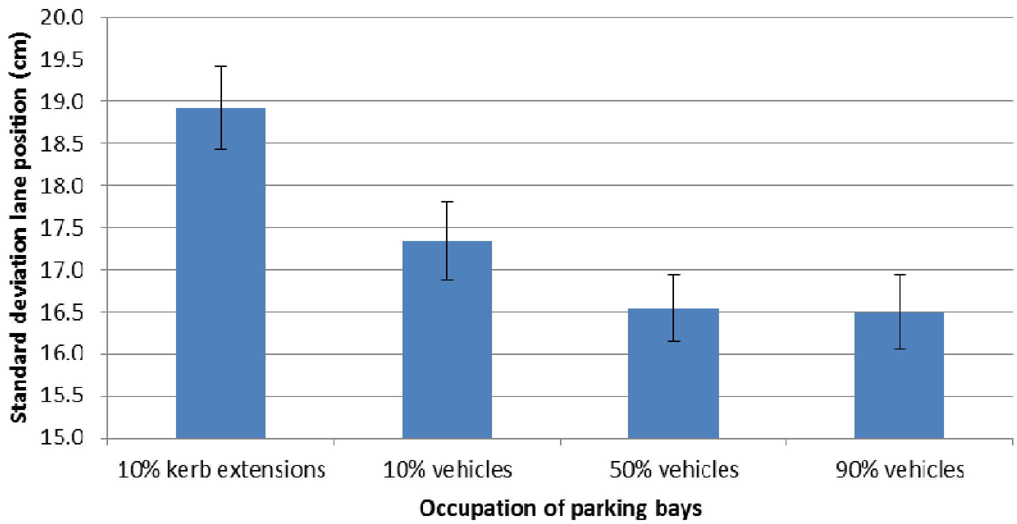


Figure 10.7 Mean (standard error) standard deviation of lane position across parking conditions

10.2.2 Longitudinal control

Participants were instructed to drive as close to the speed limit as possible. Measures of longitudinal control demonstrate how successful the participants were at performing the task. Speed-related performance measures comprised the speed differential and SDSp. The SDAcc (where pressure is measured as the percentage of maximum possible pressure) is an objective indication of how much physical effort it took the participant to try to drive at the speed limit. GEEs were conducted to determine if longitudinal control performance differed according to parking condition and/or speed zone.

10.2.2.1 Speed differential

The difference between mean travel speed and the speed limit varied according to both speed zone ($p=0.005$) and parking condition ($p<0.001$), although there was no statistically significant interaction between the two ($p>0.05$). The difference between the mean speed and the speed limit for each parking condition is shown in Figure 10.8. The speed differential was significantly greater ($p<0.001$) when 10% of parking bays were occupied with either kerb extensions or vehicles (approximately +0.5 km/h) than when 50% or 90% of parking bays were occupied by vehicles (approximately zero km/h). There was no significant difference between the two conditions with the lowest level of occupied parking bays ($p>0.05$) or the two conditions with the most occupied parking bays ($p>0.05$).

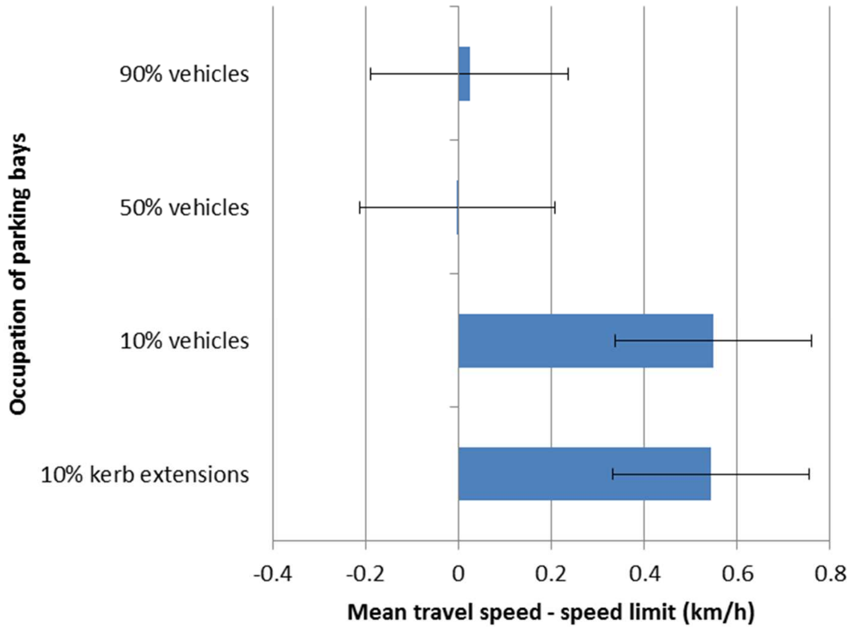


Figure 10.8 Mean difference (standard error) between the mean travel speed and the speed limit for each parking condition

The difference between the mean speed and the speed limit for each speed zone is shown in Figure 10.9. The difference between the mean travel speed and the speed limit was significantly higher when the speed zone was 70 km/h (approximately +1.0 km/h) compared to the three lower speed zones ($p<=0.02$). There was no significant difference between the speed differential for the three lower speed zones ($p>0.05$).

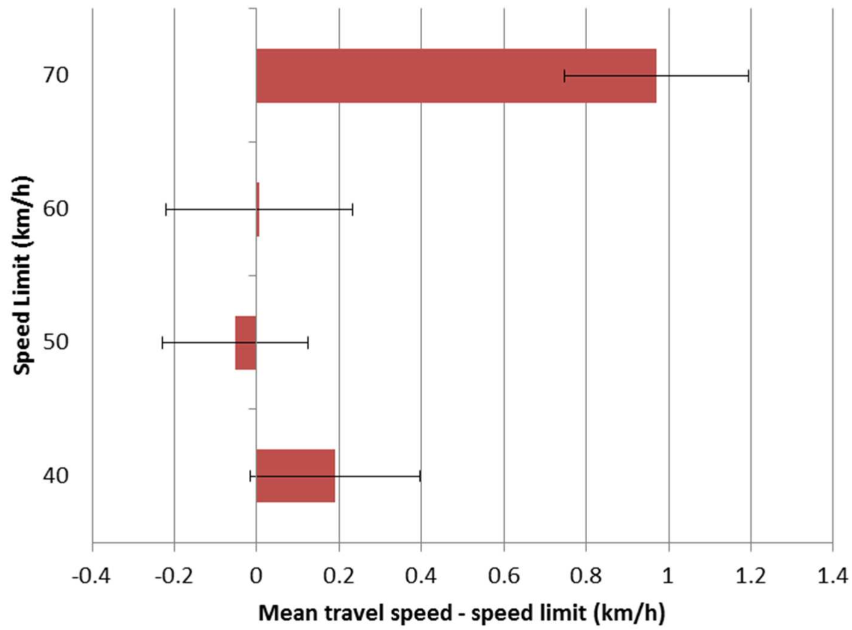


Figure 10.9 Mean difference (standard error) between the mean travel speed and the speed limit for each speed zone

10.2.2.2 Standard deviation of speed

Figure 10.10 displays the standard deviation of speed for each combination of parking conditions and speed zones. Visual inspection indicates a trend for speed variability to be highest in the 70 km/h condition, with no consistent pattern immediately apparent across the different parking conditions within speed zone. GEE analysis found that there was a significant interactive effect (p=0.027) of parking condition and speed zone on the standard deviation of speed but no significant main effects of speed zone or parking condition. Pairwise comparisons conducted to further investigate the significant interaction failed to reveal any consistent pattern of effect of parking condition and speed zone on speed variability.

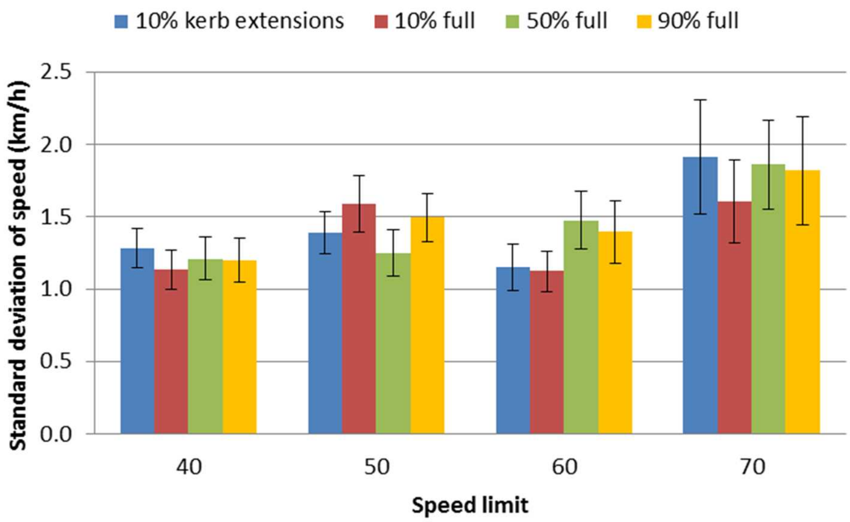


Figure 10.10 Standard deviation of speed (standard error) across parking conditions and speed zones

10.2.2.3 Standard deviation of accelerator pressure

The SDAcc varied significantly with speed zone ($p=0.005$, refer to Figure 10.11). The variability in pressure on the accelerator pedal was significantly higher when participants were attempting to drive at 70 km/h ($p<=0.003$) and significantly lower when they were trying to drive at 40 km/h ($p<=0.003$). Variability in the pressure on the accelerator pedal did not differ significantly between 50 km/h and 60 km/h zones ($p>0.05$). There was no significant main effect of parking condition on SDAcc, nor was there a significant interactive effect of speed zone and parking condition ($p>0.05$).

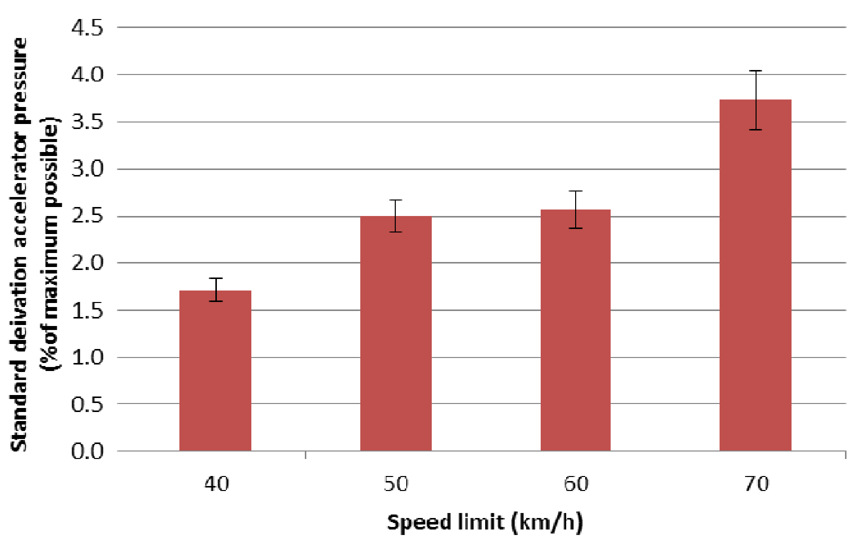


Figure 10.11 Standard deviation (standard error) of accelerator pressure by speed zone

10.3 Relationship between driving performance and self-reported effort

One of the secondary aims of this study was to explore how driving performance was related to self-report measures. The four self-report measures were highly correlated so only one was chosen for this analysis. Ratings of effort were chosen because this variable had a wider distribution of scores and the data were closer to normally distributed than the other self-report measures. This is most likely because the RSME is a validated scale with verbal anchors at multiple points on a 150 point scale, compared to the other measures which only had two anchors on a 20 point scale.

A GEE (with the identity link and exchangeable correlation matrix) was conducted to determine how effort was related to the driving performance measures in an attempt to determine if there were some aspects of the driving task that contributed more to perceived effort than others. Driving performance measures that were assessed were those related to lateral position and control (MLP and SDLP) and longitudinal control (speed differential, SDSp and SDAcc).

Manual backwards stepwise regression revealed that the only driving performance measures that were independently associated with perceived effort were SDLP and SDAcc. None of the other measures were significantly related to perceived effort.

SDAcc was positively associated with effort. As the variability in the pressure applied to the accelerator increased (as a proportion of maximum possible pressure), perceived effort increased. For every percentage point of extra variability in the pressure on the accelerator pedal, mean self-reported effort increased by 3.4 units (95% CI 1.39 to 5.41).

SDLP was negatively associated with effort, that is, as SDLP (weaving) increased, self-reported effort decreased. For every extra cm of SDLP, mean effort ratings decreased by 3.05 units (95% CI 2.14–3.97).

10.4 Secondary task performance

The median choice reaction time (in msec.) to the peripheral stimuli presented in the secondary task for each speed zone and parking condition is displayed in Figure 10.12. The IQR is also shown, with error bars representing the difference between the 25^{th} percentile (lower) and the 75^{th} percentile (upper) values. There is a trend for reaction time to increase as more of the parking bays are occupied and for reaction time to decrease as speed limit increases. Two different analysis methods were used to analyse these results and these are reported in the next sections.

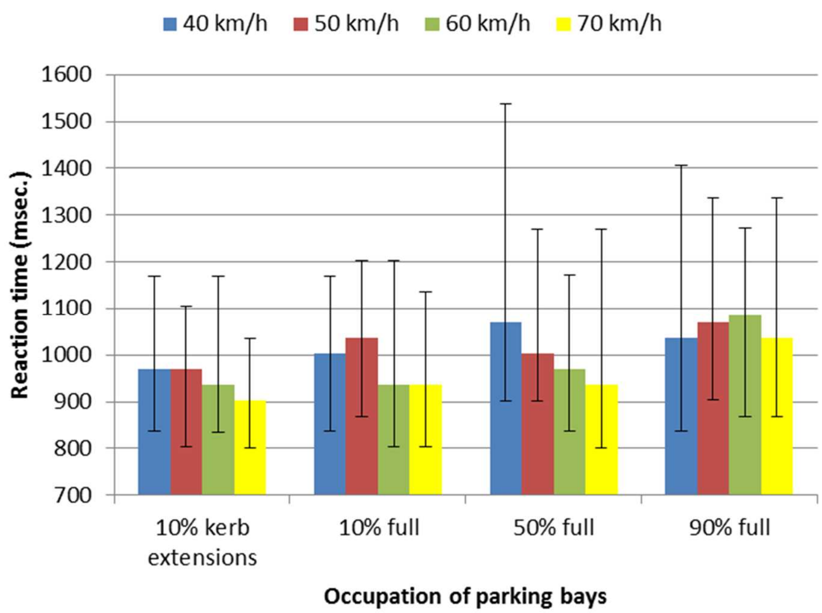


Figure 10.12 Median (interquartile range) reaction time to peripheral stimulus by speed zone and parking conditions (incorrect responses excluded)

10.4.1 Analysis approach 1: GEE using mean correct reaction times

The mean reaction times of the correct responses to the secondary task were calculated for each participant for each speed and parking condition. The mean correct choice reaction time data were normally distributed so a GEE was conducted with the identity link and an unstructured correlation matrix to determine if mean correct choice reaction time differed across speed zone and parking condition.

The analysis revealed there was no significant interactive effect of speed and parking condition on the mean choice reaction time of the correct responses to the peripheral detection task ($p>0.05$). Mean choice reaction time of the correct responses did not differ significantly across the different speed zones ($p>0.05$). There was, however, a significant effect of parking condition ($p<0.001$) on mean correct choice reaction times.

The mean choice reaction time of correct responses did not differ significantly between the conditions where 10% of the parking bays were occupied with kerb extensions or vehicles (999.6 msec vs. 1048.4 msec; $p>0.05$), nor between the conditions in which 50% and 90% of the parking bays were occupied (1152.2 msec vs.

1208.2 msec; $p>0.05$). Mean choice reaction times of correct responses were significantly faster ($p<0.05$) when 10% of the parking bays were occupied (either with kerb extensions or vehicles; 1024.1 msec) than when 50% or 90% of the parking bays were occupied (1179.7 msec).

10.4.2 Analysis approach 2: Multiple time to event analysis

Reaction times for all correct responses and censored times for missed stimuli (censored at the time at which the stimulus disappeared) were analysed in a multiple time to event analysis to determine if there was a difference in the probability (or hazard) of responding across parking conditions and speed zones.

Similar to the first analytical approach, the second analytical approach revealed that there was no significant interactive effect of speed and parking on response times ($p=0.51$). Unlike the first analysis approach, however, there were significant main effects of both speed zone ($p=0.01$) and parking condition ($p<0.001$). The proportional hazards assumption was not violated ($\chi^2(6)=4.83$, $p=0.57$).

Table 10.2 presents the hazard ratios comparing the reaction times for different speed zones where the ratio is equal to the hazard for the faster speed condition compared to the hazard for the slower speed condition. The reaction time hazard (the probability of responding at each time point given no response had occurred to that time) in the 70 km/h speed zone was 23% higher than in the 40 km/h speed zone and 20% higher than in the 50 km/h speed zone. A higher reaction time hazard means the probability of responding at any point in time was higher (given that the person had not yet responded) which corresponds to a faster reaction time. Therefore reaction times were significantly faster in the 70 km/h speed zone compared to the 40 km/h and 50 km/h speed zones. None of the other pairwise comparisons were statistically significant, although there was a trend for reaction times to decrease as speed limit increased.

Table 10.2 Hazard ratios comparing the reaction times for different speed zones

	Hazard ratio (95% CI) for speed zone		
Compared to	50 km/h	60 km/h	70 km/h
40 km/h	1.02 (0.89–1.16)	1.13 (0.98–1.30)	1.23 (1.06–1.42)
50 km/h		1.11 (0.96–1.28)	1.20 (1.06–1.37)
60 km/h			1.09 (0.94–1.26)

Table 10.3 presents the HRs comparing the reaction times for different parking conditions where the ratio is equal to the hazard for the condition where more parking bays were occupied compared to the hazard when fewer parking bays were occupied. Every pairwise comparison was statistically significant. The hazard (the probability of responding at each time point given no response had occurred to that time) became significantly lower as the proportion of bays that were occupied increased; that is, reaction times became slower as more parking bays were occupied. For example, the reaction time hazard was 32% lower when 90% of parking bays were occupied compared to when 10% of parking bays were occupied by vehicles.

Table 10.3 Hazard ratios comparing the reaction times for different parking conditions

	Hazard ratio (95% CI) for parking condition		
Compared to	10% occupied	50% occupied	90% occupied
Kerb extensions	0.84 (0.75–0.95)	0.68 (0.61–0.77)	0.57 (0.47–0.69)
10% occupied		0.81 (0.73–0.90)	0.68 (0.58–0.80)
50% occupied			0.84 (0.74–0.94)

10.4.3 Distance travelled between stimulus onset and reaction

To augment the reaction time results, the distance that the vehicle would travel between stimulus onset and participants' median reaction time if they were travelling at the speed limit (as instructed) was calculated (refer to Figure 10.13). Even though reaction times to the peripheral stimulus were significantly faster when participants were driving in 70 km/h zones than 40 km/h or 50 km/h zones, the distance they would travel during the time it took them to react becomes longer as the speed limit increased. Therefore, even if participants can react to a non-safety critical stimulus faster when travelling faster, this do not compensate for the high speed travel.

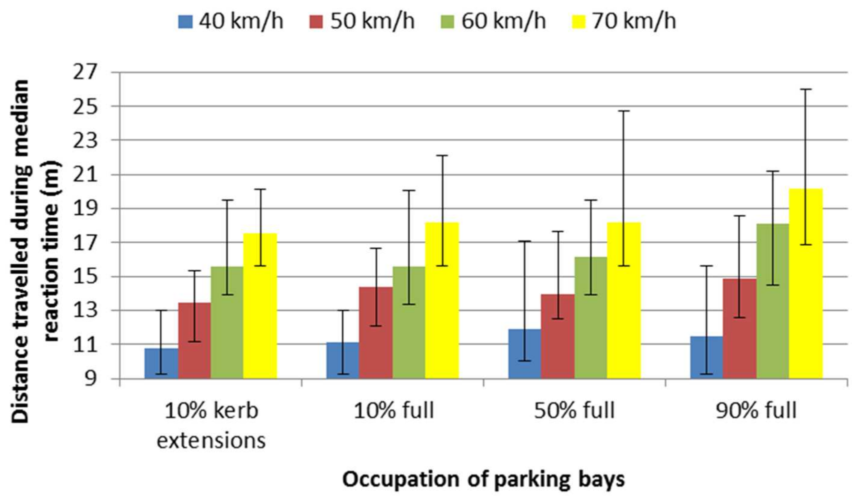


Figure 10.13 Distance travelled during median reaction time to peripheral stimulus if travelling at the speed limit

10.4.4 Comparison of results from GEE and multiple time to event analysis

It is interesting to compare the results from the two different analysis approaches for the choice reaction time to the secondary event. The multiple time to event analysis using individual correct response times and censored times for missed stimuli detected differences between all pairwise comparisons of the parking conditions, whereas the GEE using mean correct reaction times only detected a difference between the two conditions in which 10% of the parking bays were occupied (with kerb extensions or vehicles) and the conditions where 50% or 90% of parking bays were occupied. Likewise, the multiple time to event analysis detected a significant difference between reaction times in the fastest speed zone and the two slowest speed zones. In contrast, the GEE did not detect any significant differences in reaction time between speed zones.

Thus, although methods akin to the first approach (e.g. the equivalent method of repeated measures ANOVA using mean correct responses) are the most common methods used for analysing reaction time data, this comparative analysis demonstrates that survival analysis for time to multiple events is a more powerful and sensitive technique than GEE or ANOVA. This is due to two main differences between the data that were analysed in the different approaches. The standard methods exclude data from stimuli where the participant did not make a response, whereas, these data are included in the survival analysis as censored observations (where the time to censoring is equal to the time the stimulus was presented for). The multiple time to event analysis also includes each individual reaction time observation, rather than conducting analyses on the mean reaction time per condition. Thus, there is more information in the dataset.

10.5 Reaction time to the safety-critical unexpected event

The safety-critical unexpected event involved a pedestrian walking along the footpath who turned and walked onto the road in front of the participant's vehicle halfway through the last block of their final drive. The event was unexpected because the participant had completed three drives of the same route already, with no such event occurring. Furthermore, the participants regularly saw simulated pedestrians walking along the footpath throughout the scenario, so the mere presence of a pedestrian was not an indication that they were about to cross in front of the driver. For all participants, 90% of parking bays were occupied by vehicles when the pedestrian crossed the road, while the speed zone varied across participants (eight participants experienced the unexpected event in each speed zone).

One of the 32 participants did not brake when the pedestrian walked onto the road and hence collided with the pedestrian at full speed (40 km/h). At the conclusion of the experiment, the experimenter asked the participant if they saw the pedestrian before they collided with them. The participant responded that they had seen the pedestrian turn and walk onto the road but had assumed that the pedestrian would either stop, turn around and return to the side of the road they had come from, or walk faster and get out of the way. Therefore, they chose not to brake to slow down or stop. All participants had been instructed to drive as they would in the real world. The single participant who chose not to respond admitted they would not drive like that in the real world. Therefore, because the participant ignored the experimenter's instructions and behaved atypically, their data were excluded from the analysis of the reaction time to the safety-critical event (the unexpected pedestrian crossing).

The median time to press the brake pedal after the pedestrian turned, across speed zones, is shown in Figure 10.14 for the 31 participants who responded to the event. The interquartile range (the 25th percentile to the 75th percentile) is also displayed using error bars. Median reaction time to an unexpected safety-critical event was slowest for participants who were driving in the 70 km/h speed zone (2440.1 msec.) and fastest for those travelling in the 60 km/h speed zone (2106.2 msec.). There does not appear to be a great deal of difference between the median reaction times for the three slowest speed zone conditions.

Cox proportional hazards regression detected a significant difference between the slowest (40 km/h) and the fastest (70 km/h) speed zone conditions: the reaction time hazard was 68% lower when travelling at 70 km/h compared to when travelling at 40 km/h (HR=0.32, 95% CI 0.11–0.99, p=0.047). That is, participants responded faster to the crossing pedestrian when they were travelling at 40 km/h than when they were travelling at 70 km/h. There was also an indication that the hazard of responding was 63% smaller when travelling at 50 km/h compared to 40 km/h (HR=0.37, 95% CI 0.12–1.15, p=0.086). While this did not reach the standard level for statistical significance, with only eight participants per condition (or 7 in the 40 km/h condition), there was not a great deal of power to detect effects. Hence a p-value of less than 0.10 was considered worthy of attention. None of the other pairwise comparisons between speed zone conditions detected any significant difference. The assumption of proportional hazards was not violated by these data ($\chi^2(3)=5.69$, p=0.13).

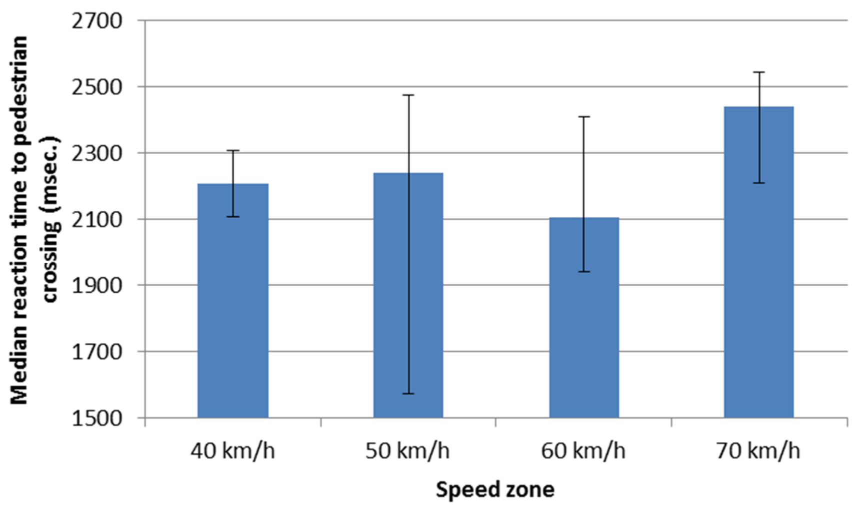


Figure 10.14 Median reaction time (interquartile range) to unexpected safety-critical event

CHAPTER 11. DISCUSSION: UNDERSTANDING RISK FACTORS

11.1 Statement of results

This study was designed to investigate the effect of roadside parking (which was associated with increased MVC in Component 1) on driver behaviour and whether reducing the speed limit may ameliorate the risk associated with roadside parking. The primary aims of this study were to investigate the effect of the proportion of parking bays that were occupied and speed limit on driving performance, perceived risk, discomfort, task difficulty and effort, and performance on a secondary choice peripheral detection task. The results will be discussed with reference to relevant previous research although it must be noted that there is only one previous study that specifically investigated the effect of roadside parking on driver behaviour (Edquist et al., 2012).

An increase in the proportion of kerbside parking bays that were occupied by parked vehicles led drivers to choose a lane position further away from the kerb (towards the centre line) and to weave less within the lane. Drivers' feelings of risk, discomfort, task difficulty and effort also increased. Previous research found a similar change in lane position and increase in workload when 90% of parking bays were occupied by vehicles compared to when no vehicles were present (Edquist et al., 2012). Post-hoc analysis of their data also discovered that SDLP was lower when vehicles were present (Rudin-Brown, Edquist, & Lenné 2014).

The most noticeable movement in lane position away from the kerb occurred between the condition when there were no parked vehicles (but 10% of parking bays featured kerb extensions) and the condition when 10% of parking bays were occupied by vehicles. Compared to when 10% of parking bays were occupied by vehicles, conditions in which more than half of the parking bays were occupied led to an additional (but smaller) change in lane position away from the parked vehicles on the kerbside to a position slightly closer to the road centre-line and therefore also the oncoming traffic. It is hypothesised that the presence of vehicles and the resulting choice of lane position closer to oncoming traffic led drivers to perceive they had less lateral space available for their field of safe travel, even though the actual width of the road and lane did not change. This perceived reduction in the field of safe travel led to drivers trying harder to maintain their lane position to avoid both parked vehicles to their left and oncoming traffic to their right. Consequently there was a reduction in weaving within the lane. The finding that effort significantly increased as SDLP decreased supports this. Prior studies have also found that SDLP decreased as mental load increased (Engström, Johannson, & Östlund, 2005) and that an effective narrowing of road width lead to increased lateral control (de Waard, 1996).

These results can be explained using the multiple comfort zone model: the presence of parked vehicles reduced lateral safety margins resulting in feelings of risk and discomfort (Summala, 2007). This increased the difficulty of the driving task and the effort required to perform the task. These feelings prompted a change in behaviour (moving away from parked vehicles) in response to the changed lateral safety margins.

Higher speed limits (and therefore travel speed) also led to increased perception of risk, discomfort, task difficulty and effort which was in agreement with previous research into increases in speed within this range on urban residential roads (Lewis-Evans & Rothengatter, 2009). There were no vehicles travelling on the same side of the road as the driver, so faster speed did not affect the probability of colliding with a vehicle in the same lane. Faster speed did, however, reduce the time available to respond to a change in safety margins. Drivers were least successful at driving at the speed limit in the 70 km/h condition—this decrease in primary task performance was likely due to increased workload. In addition, the variability in pressure applied to the accelerator (a measure of physical effort made to drive at the speed limit) was highest at 70 km/h and lowest at 40 km/h. Perceived effort significantly increased along with increases in the standard deviation of accelerator pressure. Thus, both mental and physical effort were highest when the speed limit was 70 km/h. The scenario used in this experiment was a four-lane undivided urban road and the speed limit almost never exceeds 60 km/h on these types of roads in Melbourne, Australia so the participants would have been unfamiliar with driving at 70 km/h on undivided urban roads.

In normal driving situations, drivers reduce perceptions of risk by lowering their travel speed (e.g. Gibson & Crooks, 1938; Summala, 2007). In this study, drivers were instructed to drive at the speed limit so they could not reduce their travel speed to reduce risk. Drivers were asked to drive at the speed limit so they would travel at a range of speeds in order to investigate the relationship between travel speed and behaviour and subjective measures of risk, discomfort, task difficulty and effort. Drivers reported their preferred speed for each parking condition and the maximum speed at which they would travel in that environment. As the proportion of occupied parking bays increased, preferred and maximum travel speeds decreased. There was some evidence for an association between preferred speed and the threshold beyond which changes in speed limit caused significant changes in perceived effort. In general, increases in speed limit only led to significant increases in perceived effort when the speed limit was faster than the preferred speed in that parking environment. Changes in speed limit below the preferred speed did not significantly affect ratings of effort. This relationship would not have been observed if drivers had not been instructed to travel at the speed limit, as they would not have attained speeds beyond their preferred speed. This finding supports the concept of a threshold feeling of workload and that workload is moderated by speed choice. The reduced safety margins due to the number of vehicles parked on the roadside influenced preferred speed and when drivers were forced to travel in excess of their preferred travel speed, feelings of effort significantly increased.

Despite the high correlation between the self-report measures, the threshold relationship between preferred travel speed and effort was not observed for the other self-report measures of risk, discomfort or task difficulty. In particular, almost all experimental manipulations (changes in the proportion of occupied parking bays or changes in speed limit) led to significant increases in ratings of risk. A previous simulator study where participants drove on a residential urban road at speeds between 20 km/h to 100 km/h found that perceptions of risk did not start to increase until participants were driving at speeds above 40 km/h, that is, the threshold for consciously perceiving risk was between 30 to 40 km/h (Lewis-Evans & Rothengatter, 2009). The current study has shown that if there is a threshold travel speed below which risk is not consciously perceived on complex urban roads, the threshold must also lie below 40 km/h (the lowest speed limit in this study). This study also showed that there is no threshold for the effect of the number of parked vehicles on perceptions of risk, because any increase in the proportion of parking bays that were occupied (from a minimum of none through to a maximum of 90%) led to a significant increase in perceived risk.

The results of this study have implications for theories of driver behaviour that propose that drivers act to maintain a stable level of risk or workload (e.g. Fuller, 2011; Wilde, 1976). The changes in behaviour that were observed with changes in the proportion of occupied roadside parking bays did not serve to maintain a stable level of risk or workload (effort). Drivers' choice of preferred speed also did not lead to a stable level of risk or effort across different parking environments. For example, the preferred speed when no vehicles were present corresponded with a rating on the RSME of less than "a little" and a risk rating of approximately 3.5, while preferred speed when 90% of parking bays were occupied by vehicles corresponded to "more than some" effort and a risk rating of greater than 5.5. Thus the hypothesis that drivers change their behaviour to maintain a stable level of perceived risk or workload was not supported in this study.

The secondary task paradigm was also used to measure workload in this study. It was demonstrated that the multiple time to event analysis technique (Cleves, 2009) was more sensitive to changes in reaction time across conditions than commonly used methods for analysing repeated measures data (e.g. GEE and the equivalent ANOVA).

Reaction time to the secondary task was expected to become slower as workload increased. As expected, reaction time to the choice PDT became significantly slower as the number of parked vehicles on the roadside increased which matched the subjective ratings of effort (workload). If workload increased as speed limit increased, reaction time would be expected to become slower. Unexpectedly, however, reaction time to the choice PDT was fastest in the highest speed limit condition which was the condition that was rated as requiring the most effort. The arousal hypothesis is one potential source of explanation as it is possible that drivers were more aroused when driving at higher speeds, which may have improved their ability to respond quickly to an expected stimulus. Previous research, however, found that neither arousal nor reaction time to an auditory stimulus differed across speed zones even after two hours of driving (Törnros, 1995). The finding that the reaction time to an unexpected stimulus (a pedestrian crossing the road) was significantly slower in the 70 km/h condition compared to the 40 km/h condition also renders the arousal hypothesis unlikely.

Another possibility is that the faster reaction time as speed increased may have been due to the choice of secondary task stimulus for this experiment. The stimulus in the choice PDT task was an icon (black silhouette) of a pedestrian that appeared in the periphery to the right or the left of the simulator screen at various intervals. The stimulus remained visible until the driver responded, or five seconds had passed, whichever came first. For the stimulus to remain in the same location of the screen, it would have appeared to travel at the same speed as the participant's vehicle. Therefore, as travel speed increased, the PDT stimulus would have been moving faster relative to the background scene which may have made the stimulus appear more salient. It is also possible that the stimulus was more difficult to detect against a background of parked vehicles than buildings, which is an alternative explanation as to why the expected increase in reaction time with increase in parked vehicles was observed. Anecdotally, participants did report finding it difficult to see the stimulus against a background of parked vehicles, although this was interpreted at the time by the experimenter as being due to increased workload. Therefore, although the choice PDT was demonstrated to be a sensitive measure of cognitive workload in past driving simulator research (Edquist et al., 2012; Stinchcombe & Gagnon, 2010) it may not be an appropriate choice for experiments where independent variables include travel speed or changes to the roadside which are perceived in the periphery (where the stimulus was presented). Previous research may not have discovered this problem because speed was not an independent variable and the findings relating to the roadside complexity were in the direction expected according to workload theory. It is recommended that future experiments to investigate the effect of changes in the road and roadside environment or travel speed on driving behaviour should use other types of secondary tasks to measure workload, for example, auditory detection tasks (e.g. Törnros, 1995).

11.2 Methodological strengths and limitations

This study was a rigorously designed investigation of the effect of roadside parking and travel speed on driver behaviour and perceptions of risk, discomfort, task difficulty and effort. The effect of travel speed on response time to an unexpected safety-critical event (a pedestrian crossing the road) in a highly complex environment (when 90% of parking bays were occupied by vehicles) was also measured. Compared to a similar previous study in which the order of presentation of conditions was pseudo-random (Edquist et al., 2012), the current study was fully counterbalanced, which is a more rigorous design that reduces the chance that the order of presentation of stimuli affected the results of the experiment. In the current study each participant only experienced the safety-critical event once, compared to the previous study in which each participant experienced the safety-critical event four times throughout the session. This ensured that the event was truly unexpected[4]. As the unexpected safety-critical event only occurred at the very end of the session in this study, it also ensured that drivers did not change their behaviour whenever they saw a pedestrian walking along the footpath during the study.

In addition to the possible problems with the choice of secondary task (discussed above), the study had some potential limitations. While the results provide some support for the multiple comfort zone theory, they do not provide direct evidence of the posited temporal relationship between feelings of risk, workload and behaviour change. The multiple comfort zone model hypothesises that changes in the environment cause changes in safety margins which lead to increased feelings of risk and discomfort, and that drivers modify their behaviour as a result. Task difficulty and workload are also impacted (Summala, 2007). This study has provided support for the notion that manipulating the environment and travel speed affect feelings of risk and workload and lead to changes in behaviour. Thus, risk, workload and behaviour are associated, however, with the current experimental design, it is impossible to determine if the change in risk preceded the change in behaviour (which is essential for the relationship to be causal). Risk, workload and behaviour would have to be measured at a finer temporal resolution for this relationship to be examined.

This study has measured changes in driver behaviour as a result of changes in the environment, which raises the issue of the link between driver behaviour and crash risk. In-depth crash investigations have established that driver behaviour is a major contributor to traffic crashes (Treat et al., 1979) and the measures of driver performance used in this study are well-accepted surrogate measures of crash risk (Rudin-Brown & Lenné 2010). It is intuitive that exceeding the lane boundaries will increase the risk of a collision, likewise, lane position and variability are important surrogate measures of crash risk because choosing a lane position closer to hazards reduces the margin for error and increases the probability of a collision (Tijerina, Kiger, Rockwell, & Wierwille, 1996). Slower reaction times to a hazard increase the

[4] Anecdotal evidence supports the assertion that the event was truly unexpected in this study. Many participants were visibly shocked and some expressed their shock verbally when the pedestrian walked onto the road at the end of the last drive.

probability of collision (because of the greater distance travelled prior to the driver's reaction) while travel speed affects crash risk due to the effect on stopping distance. Both reaction time and travel speed also influence the speed at impact, and therefore, the severity of a crash.

The relationship between these measures and objective crash risk, however, is extremely difficult to quantify because it requires detailed information about driver behaviour prior to the crash that is rarely available. For example, the probability that a lane excursion will result in a crash is related to the probability that there is an object (moving or not) in the other lane (either adjacent or opposing) or on the roadside, and the avoidance behaviour of all road users. Regan and Hallett (2011, p 282) acknowledged that: "Perhaps the greatest difficulty in interpreting driving performance deficits, however, is in knowing to what extent a given reduction in driving performance (e.g. a 20% increase in lateral lane excursions) translates into increased crash risk. Valid algorithms for linking the two remain to be developed."

Attempts have been made to establish a quantitative relationship between SDLP and crash risk. Crash risk rises exponentially with blood alcohol content (BAC) (Borkenstein, Crowther, Shumate, Ziel, & Zylman, 1964) and SDLP also increases as BAC rises (Louwerens, Gloerich, De Vries, Brookhuis, & O'Hanlon, 1987). From this information, the relationship between SDLP and crash risk has been inferred (Owens & Ramaekers, 2009). The inferred relationship has been used to estimate the crash risk due to the use of various sedating drugs for which the dose-dependent relationship with SDLP is known. The inferred relationship between SDLP and crash risk, however, is likely only transferable to risk factors that affect driver-state by similar mechanisms to BAC (e.g. sedating drugs and sleepiness) and have similar effects on basic operational control and vigilance. In crashes involving very impaired drivers, it is likely that the perceptual abilities and lack of operational control of the impaired driver is the main contributing factor to the crash. In contrast, most crashes are a result of complex interactions between road system components (road users, vehicles and the environment), hence it is unlikely that simplistic relationships between driver operational control and crashes are valid.

Despite the lack of quantitative relationships between driver behaviour and crash risk, there is still immense value in measuring driver behaviour in environments that differ according to the risk factors present. While a precise calculation of change in objective risk is impossible, determining how drivers respond to changes in the environment may illuminate the mechanisms behind the increase in crash risk (e.g. a change of lane position may reduce the risk of some crashes but increase risk of others).

11.3 Future research

This study demonstrated that the changes in lane position and weaving displayed by drivers when parked vehicles were present did not serve to cancel out the increased perception of risk. It is possible this was because there was only one lane in which they could travel and so drivers chose their lane position to balance the reduced safety margins associated with parked vehicles on the left and oncoming traffic on the right. It would be interesting to determine how manipulating the design of the road affects perceived risk and behaviour when parked vehicles are present. For example, if there was a median present that separated the driver from oncoming traffic, would that reduce the perceived risk and workload associated with the movement away from vehicles parked on the roadside? Alternatively, on multi-lane roads with kerbside parked vehicles, how far away does the driver have to move laterally before the perceived risk is reduced to the same level as when there are no parked vehicles present?

The observed change in lane position suggests that drivers perceived the risk associated with parked vehicles as higher than the risk associated with oncoming traffic when at least half of the parking bays were occupied by vehicles. Drivers give a wider margin to unpredictable road users (Gibson & Crooks, 1938) and it is possible they perceived the parked vehicles as less predictable than the oncoming traffic. The movement away from the kerb was relatively large even when only 10% of parking bays were occupied by vehicles. This may have been warranted given that 8% of MVC in Component 1 involved vehicle doors opening, and vehicles entering or leaving parking spaces while only 2% were head-on crashes. It has been established, however, that perceived risk is not a good predictor of objective risk (e.g. Charlton, Starkey, Perrone, & Isler, 2014; Tarko, 2009). An experiment designed to manipulate the objective risk of different types of crashes (e.g. by systematically manipulating the number of parked vehicles along with the number of oncoming vehicles) could shed further light on the factors affecting perceived risk.

This discovery that there is a potential association between preferred speed and the threshold beyond which changes in speed limit led to significant changes in perceived effort (but not other subjective measures) has interesting implications for a dissociation between thresholds for perceived risk and workload that should be explored in future research.

The scenarios used in this study only included stationary parked vehicles. While stationary parked vehicles can be an object that a vehicle collides with in a SVC, the presence of roadside parking can also become a hazard when vehicles enter or leave parking spaces or when vehicle doors are opened. This is another interesting avenue for further research as almost 8% of the MVC in Component 1 of this thesis were these types of crashes. To obtain a full picture of the effect of roadside parking on driver behaviour, future studies should be conducted to investigate how drivers respond to vehicles leaving and entering car parking spaces, and doors of parked vehicles being opened as they pass. Another avenue for further investigation with regard to roadside parking, driver behaviour and MVC risk is that when there are two lanes of through traffic and one is occupied by parked vehicles, vehicles slowing to turn at an intersection or to look for parking spaces may be at increased risk of being involved in a rear-end collision. Future research should also investigate the effect of roadside parking on crashes involving turning vehicles as roadside parking may also obstruct a driver's view of vehicles turning onto the road ahead of them, or the increased workload may reduce drivers' ability to detect such hazards. It is hypothesised that the impact of roadside parking on crashes involving a vehicle turning from a side road would be similar to the influence on crashes involving emerging pedestrians.

Participants in this study were all experienced drivers. Given the high crash rate amongst young, inexperienced drivers it would be interesting to determine how manipulating the number of vehicles parked on the roadside affects inexperienced drivers' behaviour and perception of risk and workload.

Finally, this research was conducted in a driving simulator, a research method that is effective for measuring relative changes in behaviour due to manipulations of the environment in a controlled manner. It would be useful to validate these results using observational research methods to determine to what extent these behaviours (particularly the lane position and weaving) are observed in real-world conditions.

11.4 Implications

There are opposing arguments as to whether on-street parking should be permitted, and if so, on what types of roads. Arguments for and against roadside parking centre around safety, mobility, land use, convenience and cost (Marshall et al., 2008). It is argued that on-street parking is more efficient than off-street parking because it is used more, it is cheaper and it requires less land than off-street parking. Retailers and business owners prefer on-street parking for the convenience of their customers and to increase passing trade but roadside parking leads to problems with safety and mobility (Box, 2004).

The results of this simulator study contribute to understanding how driver behaviour contributes to crash risk associated with roadside parking. The observation that drivers moved their lane position further away from stationary vehicles parked on the roadside may explain why roadside parking was not associated with the frequency of SVC in Component 1 of this thesis—that is, drivers moved away from parked vehicles in order to reduce their chance of colliding with them. Yet, this change in lane position meant that they moved closer to oncoming vehicles, particularly when more than 50% of the parking bays were occupied with vehicles. This may put drivers at higher risk of a head-on MVC when stationary vehicles are parked on the roadside. In addition, the increased workload drivers are under when stationary vehicles are parked on the roadside may reduce their capability to avoid other types of MVC that are not so obviously linked to lane position (e.g. rear-end crashes and crashes with vehicles emerging from side accesses).

Roadside parking significantly affected perceptions of risk, discomfort, task difficulty, effort and led to a change in lane position and the amount of weaving within the lane. The biggest change in lane position occurred when the proportion of occupied parking bays increased from none to 10% and then 50%. When the proportion of occupied parking bays increased from 50% to 90%, there was no further change in lane position, most likely because it would have been unsafe for them to move even further away from the kerb towards the road centreline. Yet there was a large and significant increase in perceived risk, discomfort, task difficulty and workload. Hence, it may be optimal to restrict the amount of on-street parking such that no more than half of the roadside has parking bays. Though many busy arterial roads have parking clearways that do not permit parking during peak travel times, a permanent reduction in roadside parking availability in busy strip shopping centre road segments is likely to be resisted strongly by retailers and customers.

Thus it is necessary to determine if other measures, apart from reducing roadside parking, can be implemented to reduce crash risk. Although changing the speed limit did not change the behavioural response to roadside parking, it did have a strong effect on perceived risk, discomfort, task difficulty and effort and reaction time to an unexpected emerging pedestrian. The speed at which drivers preferred to travel became slower as the number of cars parked on the roadside increased. In addition, it was shown that a change in speed limit only led to an increased in perceived effort when travel speed was greater than the preferred speed. When at least half of the parking bays were occupied by vehicles, preferred speed was approximately 50 km/h. Therefore, as the next step in the research and countermeasure development cycle for injury prevention, a trial could be conducted of reducing the speed limit to 50 km/h on four-lane undivided strip shopping centre road segments in metropolitan Melbourne on which roadside parking is permitted. Initially, the effect on behaviour could be measured in an observational study and long-term effects on crashes could also be measured. Given that drivers reduce their speed when roadside parking is present, this countermeasure should be acceptable to drivers. If this recommendation was enacted upon, it would involve reducing the speed limit of 48 of the 70 four lane undivided strip shopping centre road segments that were included in Component 1 of this thesis.

CHAPTER 12. SUMMARY, CONCLUSIONS AND IMPLICATIONS

The overall aim of this thesis was to develop and apply a multidisciplinary approach to identify and understand the aspects of the built urban environment that influence crash occurrence. An over-arching framework was developed to conceptualise the road system within the broader social and physical environment that included the components of the road system (road users, vehicles and roads), the interaction between them and the resulting contribution to exposure and crash risk. An understanding of each of the components and their interactions is essential for improving road safety, particularly, the influence of the built environment on driver behaviour and crash risk. Ewing and Dumbaugh (2009) stated: “If safety is to be meaningfully addressed, we must begin to develop our understanding of how the built environment influences … both the incidence [of] traffic-related crashes, injuries, and deaths, as well as the specific behaviors that cause them” (p.363).

Cross-sectional studies conducted to identify risk factors associated with crashes have been criticised because they cannot provide any information about causative mechanisms (Elvik, 2006; Hauer, 2010). It is ill-advised to attempt to develop countermeasures before fully understanding the reasons why crash risk is increased. Further scientific experiments are rarely conducted to investigate causative mechanisms, for example, to measure how risk factors affect driver behaviour. This underscores the need for a multidisciplinary approach because no single disciplinary or methodological approach can fully address these issues. Cross-sectional studies and driving simulation (or other behavioural research methods) are complementary in that the questions that are left unanswered by cross-sectional modelling or other analytical epidemiological studies (that is, why are some factors associated with crashes?) become the research questions to be addressed using behavioural research methods such as driving simulation.

Research Component 1 of this thesis sought to identify risk factors for crashes on complex urban road segments. A literature review established that, beyond the effect of traffic volume and intersections, there was a lack of strong evidence regarding the influence of the built urban environment on crash risk. Furthermore, the effect of the surrounding built environment has been neglected, presumably either because of the challenges in measuring the roadside environment, or because researchers did not recognise the potential for an effect on crashes.

This thesis was novel in the development of a comprehensive list of characteristics of the built urban environment, including the road, roadside and human activity, with the potential to influence crash occurrence. Sources were found for the majority of risk factors, although the lack of data to describe the exposure of vulnerable road users was identified as a concern for research into influence of the urban environment on public health including road trauma. This highlighted a pressing need to collect better exposure data for vulnerable road users (pedestrians and cyclists). A cross-sectional study was conducted using an innovative phased modelling approach developed to identify the characteristics of the built environment that were significantly associated with crash occurrence, specifically, MVC, SVC and PVC. The results established that, in addition to traffic exposure and road design, characteristics of the roadside environment, and facilities and amenities were associated with crash frequency on strip shopping centre road segments in metropolitan Melbourne and that the risk factors differed by crash type.

Research Component 2 of this thesis comprised a case study to demonstrate the use of behavioural research methods to investigate the behavioural mechanisms underlying crash risk. Driving simulation was chosen as the research method for the high level of control it provides to manipulate the environment and measure the subsequent effects on driver behaviour. Roadside parking was chosen for the case study for a number of reasons. In Research Component 1, the presence of roadside parking was associated with increased MVC frequency and the frequency of PVC was reduced on road segments with parking clearways but neither roadside parking nor parking clearways were associated with SVC frequency. Roadside parking was amenable to manipulation in driving simulation scenarios and the aspects of behaviour that were hypothesised to be affected by roadside parking (according to the multiple comfort zone model (Summala, 2007)) could be measured in the simulator. There was also a lack of experimental evidence about the effect of roadside parking on driver behaviour. In addition, the effect changing the speed limit, which is a potential countermeasure to ameliorate the risk associated with roadside parking, was investigated.

The simulator study revealed that drivers chose a lane position further away from the kerb and weaved less within their lane as the number of cars parked on the roadside increased. Perceived risk, discomfort, task difficulty and effort also increased when there were more parked cars present. Increasing the speed limit of the road segment led to significantly increased ratings of risk, discomfort, task difficulty and mental effort and an increase in the physical effort required to drive at the speed limit. Drivers also reacted slower to an unexpected safety-critical event when travelling at 70 km/h compared to when travelling at 40 km/h. An increase in speed beyond that preferred in a given parking environment led to significant increases in the effort required to drive at that speed which indicates that there is a threshold feeling of workload (in relation to speed) and that drivers choose their speed to moderate their workload. No such threshold relationship was discovered for ratings of risk, discomfort or task difficulty and speed. Any increase in the number of parked vehicles on the roadside was associated with increased perceived risk. It was also found that as the number of parked cars varied, drivers' change in behaviour and their choice of preferred speed did not serve to maintain a stable level of risk or workload, which has implications for theories of driver behaviour.

Methodological contributions of Research Component 2 include the demonstration of the superior power and sensitivity of the multiple time to event technique for analysing reaction time data as compared to GEE or ANOVA for repeated measures using correct reaction times, and the discovery that visual peripheral detection tasks are inappropriate to be used as secondary tasks when experimental manipulations involve changes in speed or changes to the roadside in the drivers' periphery. Recommendations were made for countermeasures to address crash risk on roads with roadside parking in complex urban areas.

As a whole, this thesis demonstrates a rigorous scientific process for applying two complementary methodological approaches (epidemiology/statistics and human factors/psychology) to address the identification of risk factors and their mechanisms as applied to the environment/road user interface depicted in the conceptual framework for this thesis. The process of identifying risk factors and investigating their mechanisms, which is portrayed as one step in the research and countermeasure development cycle for injury prevention, thus has two crucial components that require different methodological approaches. The two components are so vitally important for injury prevention that it is proposed that they be represented as distinct components of the cycle. Figure 12.1 depicts the proposed amendments to the research and countermeasure development cycle for injury prevention with the identification of risk factors and the investigation of mechanisms represented as separate components. In addition, the methodological approaches used to conduct research in each of the phases are shown, shaded according to the methodological approach used in that phase, with specific application to the investigation of the influence of the built environment on driver behaviour and crash risk. Yellow represents epidemiological/statistical methods that use real-world data to describe the size of a problem and identify risk factors. Blue represents behavioural research methods used to investigate mechanisms, and to develop and test countermeasures. Green represents research questions that can be addressed by both approaches—post-implementation evaluation of countermeasures can be performed using behavioural methods that measure the effect of the countermeasure on driver behaviour (impact evaluation) or epidemiological/statistical methods that measure the effect of the countermeasure on crashes and injuries (outcome evaluation).

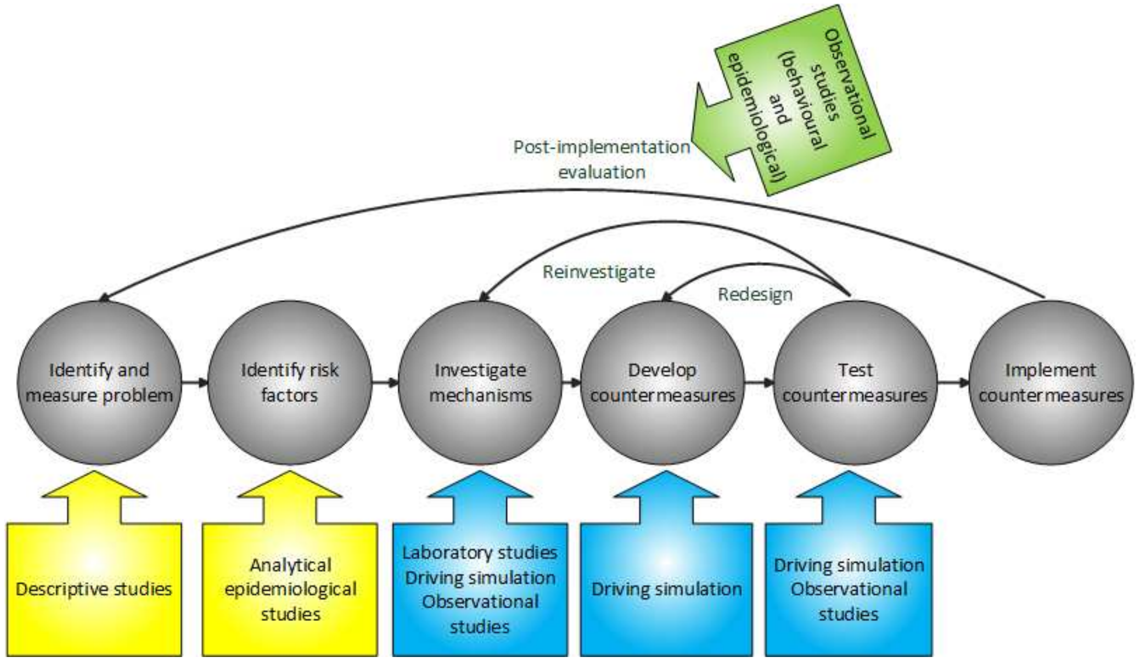


Figure 12.1 Modified research and countermeasure development cycle for injury prevention showing research methods appropriate for each stage

Both cross-sectional studies and driving simulation are commonly used methods within the field of road safety but the novel contribution of this thesis was the synergistic combination of these methods to both identify and further investigate risk factors for crashes. The identification of roadside parking as a risk factor for MVC in the cross-sectional study was important but the subsequent driving simulation investigation provided experimental evidence regarding how roadside parking influences driver behaviour and crash risk.

The success of the multidisciplinary process for investigating the effect of the built environment on road safety has implications for future research and practice in this area. The final and perhaps most important recommendation of this thesis, is that future research into the effect of the built environment (including the road, roadside and human activity) on traffic crashes should employ a variety of appropriate and complementary research methods to determine what factors increase risk and why. This recommendation may not be as simple to implement as it appears. First, it will require that road authorities and research funders recognise that moving straight from identifying risk factors to proposing solutions without first understanding the behavioural mechanisms is to risk implementing ineffective countermeasures, or worse, countermeasures that increase crash risk. Second, for such multidisciplinary research to be feasible requires either researchers who are trained in more than one discipline (which is relatively unusual), or alternatively, strong and effective collaboration between researchers from different disciplines. For such collaboration to be successful, the researchers from different disciplines must have mutual respect, a shared language for communicating (free of the jargon specific to that discipline), and a deep understanding of what each can contribute in terms of the research questions that can be answered using the methods of each discipline.

REFERENCES

AASHTO. (2010). *Highway Safety Manual* (First ed. Vol. 1). Washington DC: American Association of State Highway and Transportation Officials.

Abdel-Aty, M., Devarasetty, P. C., & Pande, A. (2009). Safety evaluation of multilane arterials in Florida. *Accident Analysis & Prevention, 41*(4), 777-788.

Abdel-Aty, M., Pande, A., Uddin, N., Dilmore, J., & Pemmanaboina, R. (2005). *Relating crash occurrence to freeway loop detectors data, weather conditions and geometric factors*. Report for Florida Department of Transportation: Department of Civil & Environmental Engineering, University of Central Florida.

Abdel-Aty, M., & Radwan, A. E. (2000). Modelling traffic accident occurrence and involvement. *Accident Analysis & Prevention, 32*(5), 633-642.

Affum, J., & Goudens, R. (2008). *Application of Netrisk road safety network assessment tool: The North coast hinterland district case study*. Paper presented at the 23rd ARRB conference - Research partnering with practitioners, 30th July - 1st August, 2008, Adelaide, Australia.

Alavi, H. (2013). *Assessing pedestrian crash risk and injury severity in concentrated urban environments.* (PhD), Monash University. Retrieved from http://arrow.monash.edu.au/hdl/1959.1/894433

Aljanahi, A. A., Rhodes, A. H., & Metcalfe, A. V. (1999). Speed, speed limits and road traffic accidents under free flow conditions. *Accident Analysis & Prevention, 31*(1-2), 161-168.

Anastasopoulos, P. C., & Mannering, F. L. (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accident Analysis & Prevention, 41*(1), 153-159.

Anastasopoulos, P. C., & Mannering, F. L. (2011). An empirical assessment of fixed and random parameter logit models using crash- and non-crash-specific injury data. *Accident Analysis & Prevention, 43*, 1140-1147.

Anastasopoulos, P. C., Mannering, F. L., Shankar, V. N., & Haddock, J. E. (2012). A study of factors affecting highway accident rates using the random-parameters tobit model. *Accident Analysis & Prevention, 45*, 628-633.

Anastasopoulos, P. C., Tarko, A. P., & Mannering, F. L. (2008). Tobit analysis of vehicle accident rates on interstate highways. *Accident Analysis & Prevention, 40*(2), 768-775.

Ando, S., Iwata, T., Ishikawa, H., Dakeishi, M., & Murata, K. (2008). Effects of acute alcohol ingestion on neuromotor functions. *Neurotoxicology, 29*(4), 735-739.

Annett, J. (2002). Subjective rating scales: Science or art? *Ergonomics, 45*(14), 966-987.

Arien, C., Jongen, E. M., Brijs, K., Brijs, T., Daniels, S., & Wets, G. (2013). A simulator study on the impact of traffic calming measures in urban areas on driving behavior and workload. *Accident Analysis & Prevention, 61*, 43-53.

Australian Automobile Association. (2006). *Star Ratings Australia's national network of roads* AusRAP: Australian Road Assessment Program.

Australian Bureau of Statistics. (2007). Australian Standard Geographical Classification (ASGC), Jul 2007, cat. no. 1216.0. Retrieved October 7th, 2011, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/Previousproducts/1216.0Main%20Features9999Jul%202007?opendocument&tabname=Summary&prodno=1216.0&issue=Jul%202007&num=&view=

Australian Bureau of Statistics. (2008a). Population by Age and Sex, Regions of Australia. 'Table 1. Estimated Resident Population by Age, Victoria, Males - June 30 2007', data cube: Excel spreadsheet, cat. no. 3235.0. Retrieved June 30th, 2012, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3235.02007?OpenDocument

Australian Bureau of Statistics. (2008b). Population by Age and Sex, Regions of Australia. 'Table 3. Estimated Resident Population by Age, Victoria, Persons - June 30 2007', data cube: Excel spreadsheet, cat. no. 3235.0. Retrieved 20th June, 2012, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3235.02007?OpenDocument

Australian Bureau of Statistics. (2008c). Socio-economic Indexes for Areas (SEIFA), Data only, 2006, 'Table 2: Statistical Local Area (SLA) Index of Relative Socio-economic Advantage and Disadvantage, 2006', data cube: Excel spreadsheet,cat. no. 2033.0.55.001. Retrieved 28th October, 2011, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/2033.0.55.0012006?OpenDocument

Australian Bureau of Statistics. (2012a). Year Book Australia, 2012. Population size and growth, cat. no. 1301.0. Retrieved July 10th, 2015, from http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1301.0~2012~Main%20Features~Population%20size%20and%20growth~47

Australian Bureau of Statistics. (2012b). Year Book Australia, 2012. Transport accidents, injuries and fatalities, cat. no. 1301.0. Retrieved June 13, 2015, from http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1301.0~2012~Main%20Features~Accidents,%20injuries%20and%20fatalities~189

Australian Bureau of Statistics. (2013). Australian Social Trends, April 2013. The 'average' Australian, cat. no. 4102.0. Retrieved June 13, 2015, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Main+Features30April+2013

Australian Bureau of Statistics. (n.d.). National Regional Profiles: Past Issues, 2004-2008. Retrieved 21st October, 2011, from http://www.abs.gov.au/ausstats/abs@.nsf/NRPPastIssues!OpenView&Start=1&Count=1000&Expand=1.1.2.1&RestrictToCategory=Main%20Areas#1.1.2.1

Australian Transport Council. (2001). *The National Road Safety Strategy 2001-2010* Australian Transport Safety Bureau. Canberra, Australia.

Australian Transport Council. (2011). *National Road Safety Strategy 2011-2020* Australian Transport Safety Bureau. Canberra, Australia.

Avelar, R., Dixon, K., Brown, L., Mecham, M., & Van Schalkwyk, I. (2013). Influence of land use and driveway placement on safety performance of arterial highways. *Transportation Research Record, 2398*, 101-109.

Babyak, M. (2004). What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. *Psychosomatic Medicine, 66*(3), 411-421.

Baldock, M. R. J., Mathias, J. L., McLean, A. J., & Berndt, A. (2006). Self-regulation of driving and its relationship to driving ability among older adults. *Accident Analysis & Prevention, 38*, 1038-1045.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: Wiley.

Bocarejo, J. P., Jiménez, A., & Torregroza-Vargas, N. (2013). *How street parking impacts on road safety.* Paper presented at the 13th World Conference of Transport, July 15-18, Rio de Janeiro, Brazil.

Boha, R., Molnar, M., Gaal, Z. A., Czigler, B., Rona, K., Kass, K., & Klausz, G. (2009). The acute effect of low-dose alcohol on working memory during mental arithmetic: I. Behavioral measures and EEG theta band spectral characteristics. *International Journal of Psychophysiology, 73*(2), 133-137.

Bonneson, J., & Ivan, J. (2013). Theory, explanation and prediction in road safety: Promising directions *Transportation Research Circular E-C179*. Washington D.C.: Transportation Research Board.

Bonneson, J. A., & McCoy, P. T. (1997). Effect of median treatment on urban arterial safety: An accident prediction model. *Transportation Research Record, 1581*, 27-36.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. Chichester, U.K.: John Wiley & Sons.

Borkenstein, R. F., Crowther, F. R., Shumate, R. P., Ziel, W. B., & Zylman, R. (1964). *The Role of the Drinking Driver in Traffic Accidents*: Department of Police Administration, Indiana University.

Box, P. C. (2004). Curb-parking problems: Overview. *Journal of Transportation Engineering, 130*(1), 1-139.

Brookhuis, K. A., & de Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis & Prevention, 42*, 898-903.

Brown, H. C., & Tarko, A. P. (1999). Effects of access control on safety on urban arterial streets. *Transportation Research Record, 1665*, 68-74.

Budd, L., Scully, J., & Newstead, S. (2011). *Evaluation of the Crash Effects of Victoria's Fixed Digital Speed and Red-Light Cameras* Report #307. Monash University Accident Research Centre.

Bureau of Infrastructure Transport and Regional Economics. (2009). *Cost of road crashes in Australia 2006* Report 118. Canberra, Australia.

Cai, H., Lin, Y., & Mourant, R. R. (2007). *Study on driver emotion in driver-vehicle-environment systems using multiple networked driving simulators* Paper presented at the DSC-NA 2007 Iowa City, September 12-14.

Cairney, P. (2008). *Road surface characteristics and crash occurrence: a literature review* Publication No. AP–T96/08. Sydney, NSW: AustRoads

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (2nd ed.). New York: Cambridge University Press.

Cameron, M. (1992). *Accident Data Analysis to Develop Target Groups for Countermeasures. Volume 1: Methods and Conclusions* Monash University Accident Research Centre, Report No. 46.

Cercarelli, L. R., Rosman, D. L., & Ryan, G. A. (1996). Comparision of accident and emergency with police road injury data. *The Journal of Trauma: Injury, Infection, and Critical Care, 40*, 805-809.

Chang, L. (2005). Analysis of freeway accident frequencies: Negative binomial regression versus artificial neural network. *Safety Science, 43*(8), 514-557.

Chang, L., & Chen, W. (2005). Data mining of tree-based models to analyze freeway accident frequency. *Journal of Safety Research, 36*, 365-375.

Charlton, S. G., Starkey, N. J., Perrone, J. A., & Isler, R. B. (2014). What's the risk? A comparison of actual and perceived driving risk. *Transportation Research Part F: Traffic Psychology and Behaviour, 25*(PtA), 50-64.

Chayanan, S., Nebergall, M., Shankar, V., Juvva, N., & Ouyang, Y. (2003). *Interaction between the roadway and roadside - An econometric analysis of design and environmental factors affecting segment accident rates* Research Report WA-RD 562.1. Seattle, WA, USA: Washington State Transportation Center

Clancy, T. A., Rucklidge, J. J., & Owen, D. (2006). Road-crossing safety in virtual reality: A comparison of adolescents with and without ADHD *Journal of Clinical Child and Adolescent Psychology, 35*, 203-215.

Cleves, M. (2009). How do I analyze multiple failure-time data using Stata. Retrieved July 9th, 2013, from http://www.stata.com/support/faqs/statistics/multiple-failure-time-data/#same

Cornelissen, M., Salmon, P. M., & Young, K. L. (2013). Same but different? Understanding road user behaviour at intersections using cognitive work analysis. *Theoretical issues in ergonomics science, 14*(6), 592-615.

Cox, D. R., & Oakes, D. (1988). *Analysis of Survival Data*. London, UK: Chapman and Hall.

D'Elia, A., & Newstead, S. (2015). Comparison of Victorian road trauma trends using traditional and alternative measures of serious injury derived from linked data. *Journal of the Australasian College of Road Safety, 26*, 19-25.
Das, A., Abdel-Aty, M., & Pande, A. (2010). Genetic programming to investigate design parameters contributing to crash occurrence on urban arterials. *Transportation Research Record, 2147*, 25-32. doi: 10.3141/2147-04
Das, A., & Abdel-Aty, M. A. (2011). A combined frequency-severity approach for the analysis of rear-end crashes on urban arterials. *Safety Science, 49*(8-9), 1156-1163.
Davis, G. A. (2001). Relating severity of pedestrian injury to impact speed in vehicle-pedestrian crashes: Simple threshold model. *Transportation Research Record, 1773*, 108-113.
de Waard, D. (1996). *The measurement of drivers' mental workload.* (PhD), University of Groningen, Haren. Retrieved from http://dissertations.ub.rug.nl/faculties/ppsw/1996/d.de.waard/
Department for Transport. (2014). *Reported road casualties Great Britain: 2013* Annual Report. London, UK: Department for Transport.
Department of Environment Land Water & Planning. (2015). Victoria Planning Provisions. Retrieved 27th July, 2015, 2015, from http://planningschemes.dpcd.vic.gov.au/schemes/vpps/52_28.pdf
Department of Justice. (n.d.). Liquor Licences and Applications Online. Retrieved February 2, 2011, from https://liquor.justice.vic.gov.au/alarm_internet/alarm_internet.asp?wci=start_alarm_internet
Department of Sustainability and Environment. (n.d.). Interactive Map. Retrieved 30th August, 2010, from http://www.dse.vic.gov.au/about-depi/interactive-maps
Department of Transport. (2009). *Victorian Integrated Survey of Travel and Activity 2007 (VISTA 2007)*. Melbourne, Australia: Department of Transport
Dissanayake, D., Aryaija, J., & Wedagama, D. M. P. (2009). Modelling the effects of land use and temporal factors on child pedestrian casualties. *Accident Analysis & Prevention, 41*(5), 1016-1024.
Douissembekov, E., Michael, G. A., Roge, J., Bonhoure, P., Gabaude, C., & Navarro, J. (2015). Effects of shrinkage of the visual field through ageing on parking performance: A parametric manipulation of salience and relevance of contextual components. *Ergonomics, 58*(5), 698-711.
Douissembekov, E., Navarro, J., Michael, G. A., Bonhoure, P., Gabaude, C., & Roge, J. (2015). Parking manoeuvres differ among drivers with narrower and wider field of view in the presence of a spatial reference. *Applied Cognitive Psychology, 29*(2), 309-313.
Dumbaugh, E. (2005). Safe streets, livable streets. *Journal of the American Planning Association, 71*(3), 283-298.
Dumbaugh, E., & Li, W. (2011). Designing for the safety of pedestrians, cyclists and motorists in urban environments. *Journal of the American Planning Association, 77*(1), 69-88.
Dumbaugh, E., Li, W., & Joh, K. (2013). The built environment and the incidence of pedestrian and cyclist crashes. *Urban Design International, 18*(3), 217-228.
Editorial. (2010). Save our cities. *Nature, 467*, 883-884.
Edquist, J. (2009). *The Effects of Visual Clutter on Driving Performance.* (PhD), Monash University, Monash University.
Edquist, J., Rudin-Brown, C. M., & Lenne, M. G. (2012). The effects of on-street parking and road environment visual complexity on travel speed and reaction time. *Accident Analysis & Prevention, 45*, 759-765.
El-Basyouny, K., & Sayed, T. (2009). Accident prediction models with random corridor parameters. *Accident Analysis & Prevention, 41*(5), 1118-1123.

Elvik, R. (2006). Laws of accident causation. *Accident Analysis & Prevention, 38*, 742-747.
Elvik, R. (2011). Assessing causality in multivariate accident models. *Accident Analysis & Prevention, 43*, 253-264.
Engström, J., Johannson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*(2), 97-120.
Ewing, R., & Dumbaugh, E. (2009). The built environment and traffic safety: A review of empirical evidence. *Journal of Planning Literature, 23*(4), 347-367.
Fildes, B., & Lee, S. (1993). *The speed review: Road environment, behaviour, speed limits, enforcement and crashes*. Report CR 127. Canberra, Australia: Federal Office of Road Safety.
Fuller, R. (2011). Driver control theory: from task diff iculty homeostasis to risk allostasis. In B. E. Porter (Ed.), *Handbook of Traffic Psychology* (pp. 13-26). London: Academic Press.
Gawron, V., & Ranney, T. (1990). Curve negotiation performance in a driving simulator as a function of curve geometry. *Applied Ergonomics, 21*(1), 33-38.
Gawron, V. J. (2008). *Human performance, workload, and situational awareness measures handbook* (2nd ed.). Boca Raton, Florida: CRC Press.
Gibson, J. J., & Crooks, L. E. (1938). A theoretical field-analysis of automobile-driving. *The American Journal of Psychology, 51*, 453-471.
Global Road Safety Facility, & The World Bank. (2014). *Transport for Health: The Global Burden of Disease from Motorized Road Transport.* Seattle, WA: IHME; Washington DC: The World Bank: Institute for Health Metrics and Evaluation,.
Google. (n.d.). Google Maps. from http://maps.google.com.au/intl/en/help/maps/streetview/
Greibe, P. (2003). Accident prediction models for urban roads. *Accident Analysis & Prevention, 35*(2), 273-285.
Gross, F., & Donnell, E. T. (2011). Case-control and cross-sectional methods for estimating crash modification factors: Comparisons from roadway lighting and lane and shoulder width safety effect studies. *Journal of Safety Research, 42*(2), 117-129.
Gross, F., & Jovanis, P. P. (2007). Estimation of the safety effectiveness of lane and shoulder width: Case-control approach. *Journal of Transportation Engineering, 133*(6), 362-369.
Gruenewald, P. J., Millar, A. B., Treno, A. J., Yang, Z., Ponicki, W. R., & Roeper, P. (1996). The geography of availability and driving after drinking. *Addiction, 91*(7), 967-983.
Gstalter, H., & Fastenmeier, W. (2010). Reliability of drivers in urban intersections. *Accident Analysis & Prevention, 42*(1), 225-234.
Hadayeghi, A., Shalaby, A. S., & Persaud, B. N. (2003). Macrolevel accident prediction models for evaluating safety of urban transportation systems. *Transportation Research Record, 1840*, 87-95.
Haddon, W. J. (1970). On the escape of tigers: An ecologic note. *American Journal of Public Health, 60*(12), 2229-2234.
Haddon, W. J. (1972). A logical framework for categorizing highway safety phenomena and activity. *The Journal of Trauma, 12*(3), 193-207.
Hamann, C., & Peek-Asa, C. (2013). On-road bicycle facilities and bicycle crashes in Iowa, 2007-2010. *Accident Analysis and Prevention, 56*, 103-109.
Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built environment affects physical activity: Views from urban planning. *American Journal of Preventive Medicine, 23*(2, Supplement 1), 64-73.
Harris, M. A., Reynolds, C. O., Winters, M., Cripton, P., Shen, H., Chipman, M., . . . Teschke, K. (2013). Comparing the effects of infrastructure on bicycling injury at

intersections and non-intersections using a case-crossover design. *Injury Prevention, 19*, 303-310.
Hart, S. G. (2006). *NASA-Task Load Index (NASA-TLX); 20 years later.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting.
Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North Holland Press
Hauer, E. (1992). Empirical bayes approach to the estimation of "unsafety": The multivariate regression method. *Accident Analysis & Prevention, 24*(5), 457-477.
Hauer, E. (1999). *Safety in geometric design standards*. University of Toronto, Canada.
Hauer, E. (2010). Cause, effect and regression in road safety. *Accident Analysis & Prevention, 42*(4), 1128-1135.
Haynes, R., Lake, I., Kingham, S., Sabel, C., Pearce, J., & Barnett, R. (2008). The influence of road curvature on fatal crashes in New Zealand. *Accident Analysis & Prevention, 40*(3), 843-850.
Heger, R. (1998). Driving behavior and driver mental workload as criteria of highway geometric design quality. *Transportation Research Circular*(E-C003), p. 43:41-10.
Henley, G., & Harrison, J. E. (2009). *Serious injury due to land transport accidents, Australia 2006-2007* Report no. 53. Cat. no. INJCAT 129. Canberra: Australian Institute of Health and Welfare.
Hennekens, C. H., & Buring, J. E. (1987). *Epidemiology in Medicine* (1st ed.). Boston, USA: Little, Brown & Co. .
Hilbe, J. M. (2012). *Negative binomial regression* (2nd ed.). New York: Cambridge University Press.
Hill, A. B. (1965). The environment and disease: Association or causation. *Proceedings of the Royal Society of Medicine, 58*(5), 295-300.
Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Hoboken, N.J.: Wiley.
Ivan, J. N., Wang, C., & Bernardo, N. R. (2000). Explaining two-lane highway crash rates using land use and hourly exposure. *Accident Analysis & Prevention, 32*, 787-795.
Jackett, M. (1993). Accident rates on urban routes - 1992 update. *IPENZ Transactions, 20*(1), 10-16.
Jahn, G., Oehme, A., Krems, J. F., & Gelau, C. (2005). Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*, 255-275.
Jama, H. H., Grzebieta, R. H., Friswell, R., & McIntosh, A. S. (2011). Characteristics of fatal motorcycle crashes into roadside safety barriers in Australia and New Zealand. *Accident Analysis & Prevention, 43*(3), 652-660.
Johnson, M., Charlton, J., & Oxley, J. (2010). The application of a naturalistic driving method to investigate on-road cyclist behaviour: A feasibility study. *Road and Transport Research, 19*, 32-41.
Johnson, M., Charlton, J., Oxley, J., & Newstead, S. (2010). Naturalistic cycling study: Identifying risk factors for on-road commuter cyclists. *Annals of Advances in Automotive Medicine, 54*, 275-283.
Johnson, M., Newstead, S., Oxley, J., & Charlton, J. (2013). Cyclists and open vehicle doors: Crash characteristics and risk factors. *Safety Science, 59*, 135-140.
Jonah, B. A. (1986). Accident risk and risk-taking behaviour among young drivers. *Accident Analysis & Prevention, 18*, 255-271.
Jonsson, T. (2005). *Predictive models for accidents on urban links: A focus on vulnerable road users.* (PhD), Lund University. (Bulletin 226)

Karanicolas, P. J., Farrokhyar, F., & Bhandari, M. (2010). Blinding: Who, what, when, why, how? *Canadian Journal of Surgery, 53*(5), 345-348.
Karlaftis, M. G., & Golias, I. (2002). Effects of road geometry and traffic volumes on rural roadway accident rates. *Accident Analysis & Prevention, 34*(3), 357-365.
Keigan, M., Cuerden, R., & Wheeler, A. (2009). *Analayis of police collision files for pedal cyclist fatalities in London, 2001-2006* Published Project Report Berkshire, United Kingdom Transport Research Laboratory
Kim, D., & Washington, S. (2006). The significance of endogeneity problems in crash models: An examination of left-turn lanes in intersection crash models. *Accident Analysis & Prevention, 38*, 1094-1100.
Knappe, G., Keinath, A., Bengler, K., & Meinecke, C. (2007). *Driving simulator as an evaluation tool - Assessment of the influence of field of view and secondary tasks on lane keeping and steering performance*. Paper presented at the 20th International technical conference on the enhanced safety of vehicles, June 18-21, Lyon, France.
Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. In D. L. Damos (Ed.), *Multiple task performance* (pp. 279-328). London: Taylor & Francis.
Kumara, S. S. P., & Chin, H. C. (2005). Application of Poisson underreporting model to examine crash frequencies at signalized three-legged intersections. *Transportation Research Record, 1908*, 46-50.
Land, M., & Horwood, J. (1995). Which parts of the road guide steering? *Nature, 377*(6547), 339-340.
Landau, S. (2002). Using survival analysis in psychology. *Understanding Statistics, 1*(4), 233-270.
Larsson, P., Dekker, S. W. A., & Tingvall, C. (2010). The need for a systems theory approach to road safety. *Safety Science, 48*, 1167-1174.
LaScala, E. A., Gerber, D., & Gruenewald, P. J. (2000). Demographic and environmental correlates of pedestrian injury collisions: A spatial analysis. *Accident Analysis & Prevention, 32*, 651-658.
LaScala, E. A., Johnson, F. W., & Gruenewald, P. J. (2001). Neighbourhood characteristics of alcohol-related pedestrian injury collisions: A geostatistical analysis. *Prevention Science, 2*(2), 123-134.
Lassarre, S., Papadimitriou, E., Yannis, G., & Golias, J. (2007). Measuring accident risk exposure for pedestrians in different micro-environments. *Accident Analysis & Prevention, 39*(6), 1226-1238.
Lee, J. (2000). *Econometric analysis of the effect of roadway geometric and roadside features on run-off roadway accident frequencies and severities.* (PhD thesis), University of Washington, Seattle.
Lee, J., & Mannering, F. (2002). Impact of roadside features on the frequency and severity of run-off roadway accidents: an empirical analysis. *Accident Analysis & Prevention, 34*(2), 149-161.
Lee, R. (2003). Linking a Nation: Australia's transport and communications 1788-1970. Chapter 10: Mobility culture in mid-twentieth century Australia. Retrieved July 10th, 2015, from http://www.environment.gov.au/resource/linking-nation-australias-transport-and-communications-1788-1970-1
Lehtonen, E., Lappi, O., Koirikivi, I., & Summala, H. (2014). Effect of driving experience on anticipatory look-ahead fixations in real curve driving. *Accident Analysis & Prevention, 70*, 195-208.
Lehtonen, E., Lappi, O., Kotkanen, H., & Summala, H. (2013). Look-ahead fixations in curve driving. *Ergonomics, 56*(1), 34-44.
Lehtonen, E., Lappi, O., & Summala, H. (2012). Anticipatory eye movements when approaching a curve on a rural road depend on working memory load.

Transportation Research Part F: Traffic Psychology and Behaviour, 15(3), 369-377.
Lenné , M. G., Corben, B. F., & Stephan, K. (2007). Traffic signal phasing at intersections to improve safety for alcohol-affected pedestrians. *Accident Analysis & Prevention, 39*, 751-756.
Lewis-Evans, B. (2012). *Testing models of driver behaviour.* University of Groningen.
Lewis-Evans, B., & Charlton, S. G. (2006). Explicit and implicit processes in behavioural adaptation to road width. *Accident Analysis & Prevention, 38*(3), 610-617.
Lewis-Evans, B., De Waard, D., & Brookhuis, K. A. (2010). That's close enough-A threshold effect of time headway on the experience of risk, task difficulty, effort, and comfort. *Accident Analysis & Prevention, 42*(6), 1926-1933.
Lewis-Evans, B., & Rothengatter, T. (2009). Task difficulty, risk, effort and comfort in a simulated driving task - Implications for risk allostasis theory. *Accident Analysis & Prevention, 41*, 1053-1063.
Liang, K., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*, 13-22.
Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice, 44*(5), 291-305.
Lord, D., Washington, S. P., & Ivan, J. N. (2005). Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: Balancing statistical fit and theory. *Accident Analysis & Prevention, 37*(1), 35-46.
Louveton, N., Bootsma, R. J., Guerin, P., Berthelon, C., & Montagne, G. (2012). Intersection crossing considered as intercepting a moving traffic gap: Effects of task and environmental constraints. *Acta Psychologica, 141*(3), 287-294.
Louwerens, J. W., Gloerich, A. B. M., De Vries, G., Brookhuis, K. A., & O'Hanlon, J. F. (1987). The relationship between drivers' blood alcohol concentration (BAC) and actual driving performance during high speed travel. In P. C. Noordzij & R. Roszbach (Eds.), *Alcohol, Drugs and Traffic Safety* (pp. 183-192). Amsterdam: Excerpta Medica.
Lucas, R. M., & McMichael, A. J. (2005). Association or causation: evaluating links between "environment" and "disease". *Bulletin of the World Health Organization, 83*(10), 792-795.
Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1-22.
Manuel, A., El-Basyouny, K., & Islam, M. T. (2014). Investigating the safety effects of road width on urban collector roadways. *Safety Science, 62*, 305-311.
Marshall, W. E., Garrick, N. W., & Hansen, G. (2008). Reassessing on-street parking. *Transportation Research Record, 2046*, 45-52. doi: 10.3141/2046-06
Martens, M. H., & van Winsum, W. (2000). *Measuring distraction: the Peripheral Detection Task* TNO Human Factors Research Institute, Soesterberg.
Marti, G., Morice, A. H., & Montagne, G. (2015). Drivers' decision-making when attempting to cross an intersection results from choice between affordances. *Frontiers in Human Neuroscience, 8*.
McInerney, R., & Smith, G. (2009). *Saving lives through investment in safer roads: The iRAP partnership.* Paper presented at the Proceedings of the 13th Road Engineering Association of Asia and Australasia, Incheon City, Korea, September 23-26.
Melway. (2005). *Melway Greater Melbourne Street Directory, Edition 32*. Glen Iris, Victoria, Australia: Melway Publishing Pty Ltd.
Melway. (2006). *Melway Greater Melbourne Street Directory, Edition 33*. Glen Iris, Victoria, Australia: Melway Publishing Pty Ltd.

Melway. (2007). *Melway Greater Melbourne Street Directory, Edition 34*. Glen Iris, Victoria, Australia: Melway Publishing Pty Ltd.
Melway. (2008). *Melway Greater Melbourne Street Directory, Edition 35*. Glen Iris, Victoria, Australia: Melway Publishing Pty Ltd.
Melway. (2009). *Melway Greater Melbourne Street Directory, Edition 36*. Glen Iris, Victoria, Australia: Melway Publishing Pty Ltd.
Miaou, S. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regression. *Accident Analysis & Prevention, 26*(4), 471-482.
Miaou, S., & Lum, H. (1993). Modeling vehicle accidents and highway geometric design relationships. *Accident Analysis & Prevention, 25*(6), 689-709.
Michon, J. A. (1985). A critical review of driver behavior models: What do we know, what should we do? In L. Evans & R. C. Schwing (Eds.), *Human behavior and traffic safety* (pp. 485-520). New York: Plenum Press.
Midson, K. (2007). *Pedestrian safety in strip shopping centres*. Paper presented at the 3rd International Road Safety Conference, 29-30 November Perth, Australia.
Milleville-Pennel, I., Jean-Michel, H., & Elise, J. (2007). The use of hazard road signs to improve the perception of severe bends. *Accident Analysis & Prevention, 39*(4), 721-730.
Milton, J., & Mannering, F. (1998). The relationship among highway geometries, traffic-related elements and motor-vehicle accident frequencies. *Transportation, 25*(4), 395-413.
Min, Y.-K., Min, B.-C., & Kim, B. (2013). Changes in eye movements and driving performance at different intersections as a function of age and turn type. *International Journal of Industrial Ergonomics, 43*(4), 342-349.
Mitra, S., & Washington, S. (2007). On the nature of over-dispersion in motor vehicle crash prediction models. *Accident Analysis & Prevention, 39*(3), 459-468.
Montella, A., Aria, M., D'Ambrosio, A., Galante, F., Mauriello, F., & Pernetti, M. (2011). Simulator evaluation of drivers' speed, deceleration and lateral position at rural intersections in relation to different perceptual cues. *Accident Analysis & Prevention, 43*(6), 2072-2084.
Näätänen, R., & Summala, H. (1974). A model for the role of motivational factors in drivers' decision-making. *Accident Analysis & Prevention, 6*, 243-261.
Neale, V. L., Dingus, T. A., Klauer, S. G., Sudweeks, J. D., & Goodman, M. (2005). *An overview of the 100-car naturalistic study and findings.* Paper presented at the 19th International Technical Conference on the Enhanced Safety of Vehicles, June 6-9, Washington DC.
Newstead, S. V., & Corben, B. F. (2001). *Evaluation of the 1992-1996 Transport Accident Commission Funded Accident Black Spot Treatment Program in Victoria* Monash University Accident Research Centre, Report No. 182
Noland, R. B. (2013). From theory to practice in road safety policy: Understanding risk versus mobility. *Research in Transportation Economics, 43*, 71-84.
Noland, R. B., & Oh, L. (2004). The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. *Accident Analysis & Prevention, 36*(4), 525-532.
O'Hara, R. B., & Kotze, J. (2010). Do not log-transform count data. *Methods in Ecology and Evolution, 1*(2), 118-122.
Oppenheim, I., & Shinar, D. (2011). Human factors and ergonomics. In B. E. Porter (Ed.), *Handbook of Traffic Psychology* (pp. 193-211). London: Academic Press.
Othman, S., Thomson, R., & Lannér, G. (2009). Identifying critical road geometry parameters affecting crash rate and crash type. *Annals of Advances in Automotive Medicine, 53*, 155-165.

Owens, K., & Ramaekers, J. G. (2009). Drugs, driving, and models to measure driving impairment. In J. C. Verster, S. R. Pandi-Perumal, J. F. Ramaekers & J. J. de Gier (Eds.), *Drugs, Driving and Traffic Safety* (pp. 43-58). Basel: Birkhäuser.

Oxley, J. A., Congiu, M., Whelan, M., D'Elia, A., & Charlton, J. (2007). The impacts of functional performance, behaviour and traffic exposure on road-crossing judgements of young children. *Annual Proceedings of the Association for the Advancement of Automotive Medicine, 51*, 81-96.

Oxley, J. A., Ihsen, E., Fildes, B. N., Charlton, J. L., & Day, R. H. (2005). Crossing roads safely: An experimental study of age differences in gap selection by pedestrians. *Accident Analysis & Prevention, 37*(962-971).

Pande, A., & Abdel-Aty, M. (2009). A novel approach for analyzing severe crash patterns on multilane highways. *Accident Analysis & Prevention, 41*(5), 985-994.

Papadimitriou, E., Yannis, G., & Golias, J. (2014). Analysis of pedestrian crossing behaviour in urban areas. *International Journal of Operations Research and Information Systems, 5*, 40-55.

Parmet, Y., Meir, A., & Borowsky, A. (2014). Survival analysis: A fresh approach for analyzing response times in driving-related hazard perception paradigms. *Transportation Research Part F: Traffic Psychology and Behaviour, 25*, 98-107.

Patten, C. J. D., Kircher, A., Ostlund, J., Nilsson, L., & Svenson, O. (2006). Driver experience and cognitive workload in different traffic environments. *Accident Analysis & Prevention, 38*, 887-894.

Pauzié, A., Manzano, J., & Dapzol, N. (2007). *Driver's behavior and workload assessment for new in-vehicle technologies and design*. Paper presented at the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, June 9-12, Stevenson, WA, USA.

Persaud, B. N., Retting, R. A., & Lyon, C. A. (2004). Crash reduction following installation of centerline rumble strips on rural two-lane roads. *Accident Analysis & Prevention, 36*(6), 1073-1079.

Plavsic, M., Klinker, G., & Bubb, H. (2010). Situation awareness assessment in critical driving situations at intersections by task and human error analysis. *Human Factors and Ergonomics in Manufacturing & Service Industries, 20*(3), 177-191.

Plumert, J. M., Kearney, J. K., & Cremer, J. F. (2004). Children's perception of gap affordances: Bicycling across traffic-filled intersections in an immersive virtual environment. *Child Development, 75*, 1243-1253.

Potts, I. B., Harwood, D. W., & Richard, K. R. (2007). Relationship of lane width to safety on urban and suburban arterials. *Transportation Research Record, 2023*, 63-82.

Qin, X. (2006). Bayesian estimation of hourly exposure functions by crash type and time of day. *Accident Analysis & Prevention, 38*(6), 1071-1080.

Qin, X., Ivan, J. N., & Ravishanker, N. (2004). Selecting exposure measures in crash rate prediction for two-lane highway segments. *Accident Analysis & Prevention, 36*(2), 183-191.

Qin, X., Ivan, J. N., Ravishanker, N., & Liu, J. (2005). Hierarchical Bayesian estimation of safety performance functions for two-lane highways using Markov chain Monte Carlo modeling. *Journal of Transportation Engineering, 131*(5), 345-351.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics, SMC-13*(13), 257-266.

Read, G. J. M., Salmon, P. M., Lenné , M., & Grey, E. (2014). *Evaluating design hypotheses for rail level crossings: An observational study of pedestrian and cyclist behavior.* Paper presented at the Proceedings of the 5th Advances in Human Factors and Ergonomics Conference, 19-23 July, Krakow, Poland.

Regan, M. A., & Hallett, C. (2011). Driver distraction: definition, mechanisms, effects, and mitigation. In B. E. Porter (Ed.), *Handbook of Traffic Psychology* (pp. 275-286). London: Academic Press.

Reymond, G., Kemeny, A., Droulez, J., & Berthoz, A. (2001). Role of lateral acceleration in curve driving: Driver model and experiments on a real vehicle and a driving simulator. *Human Factors, 43*(3), 483-495.

Roberts, I., Norton, R., Jackson, R., Dunn, R., & Hassall, I. (1995). Effect of environmental factors on risk of injury of child pedestrians by motor vehicles: a case-control study. *BMJ, 310*(6972), 91-94.

Roberts, P., Cammack, M., & Rodwell, L. (2010). *A field trial of automated measurement of roadside hazard offset*. Paper presented at the Proceedings of the 24th ARRB Conference, 12-15 October, Melbourne, Victoria, Australia.

Romoser, M. R., Pollatsek, A., Fisher, D. L., & Williams, C. C. (2013). Comparing the glance patterns of older versus younger experienced drivers: Scanning for hazards while approaching and entering the intersection. *Transportation Research Part F: Traffic Psychology and Behaviour, 16*, 104-116.

Roper, P. (2003). *Data collection technology at ARRB Transport Research*. Paper presented at the Proceedings of the Australasian Road Safety Research Education and Policing Conference, 24-26 September Sydney, Australia.

Rothengatter, T. (1998). *An overview of traffic psychology: Do research and measures match?* Paper presented at the Behavioural Research in Road Safety VIII, Crowthorne: Transport Research Laboratory.

Royston, P., & Ambler, G. (1999). Nonlinear regression models involving power or exponential functions of covariates. *Stata Technical Bulletin, STB-49*, 25.

Rudin-Brown, C. M., Edquist, J., & Lenné , M. G. (2014). Effects of driving experience and sensation seeking on drivers' adaptation to road environment complexity. *Safety Science, 62*, 121-129.

Rudin-Brown, C. M., & Lenné , M. G. (2010). *Driving simulation to support road safety policy: Understanding crash risks to better inform speed setting guidelinesq.* Paper presented at the Australasian Road Safety Research, Policing and Education Conference, Canberra, ACT, Australia.

Sagberg, F. (2003). *Influencing driver behaviour and safety by road system improvements*: Norwegian Institute of Transport Economics.

Sager, B., Yanko, M. R., Spalek, T. M., Froc, D. J., Bernstein, D. M., & Dastur, F. N. (2014). Motorcyclist's lane position as a factor in right-of-way violation collisions: A driving simulator study. *Accident Analysis & Prevention, 72*, 325-329.

Salmon, P. M., Lenne, M. G., Walker, G. H., Stanton, N. A., & Filtness, A. (2014). Using the event analysis of systemic teamwork (EAST) to explore conflicts between different road user groups when making right hand turns at urban intersections. *Ergonomics, 57*(11), 1628-1642.

Salmon, P. M., Young, K. L., & Cornelissen, M. (2011). *Connect 4? The compatibility of driver, motorcyclist, cyclist and pedestrian situation awareness*. Paper presented at the Human Factors of Systems and Technology: 55th Human Factors and Ergonomics Society European Chapter Meeting, 19-21 October, Leeds, UK.

Sawalha, Z., & Sayed, T. (2001). Evaluating safety of urban arterial roadways. *Journal of Transportation Engineering, 127*(2), 151-158.

Schepers, P., Hagenzieker, M., Methorst, R., van Wee, B., & Wegman, F. (2014). A conceptual framework for road safety and mobility applied to cycling safety. *Accident Analysis & Prevention, 62*, 331-340.

Schneider, W. H., Savolainen, P. T., & Moore, D. N. (2010). Effects of horizontal curvature on single-vehicle motorcycle crashes along rural two-lane highways. *Transportation Research Record, 2194*, 91-98.

Schuurman, N., Cinnamon, J., Crooks, V. A., & Hameed, S. M. (2009). Pedestrian injury and the built environment: an environmental scan of hotspots. *BMC Public Health, 9*, 233.

Shankar, V., Mannering, F., & Barfield, W. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis & Prevention, 27*(3), 371-389.
Shankar, V., Milton, J., & Mannering, F. (1997). Modeling accident frequencies as zero-altered probability processes: An empirical inquiry. *Accident Analysis & Prevention, 29*(6), 829-837.
Shinar, D., Rockwell, T. H., & Malecki, J. A. (1980). The effects of changes in driver perception on rural curve negotiation. *Ergonomics, 23*(3), 263-275.
Shively, T. S., Kockelman, K., & Damien, P. (2010). A Bayesian semi-parametric model to estimate relationships between crash counts and roadway characteristics. *Transportation Research Part B: Methodological, 44*(5), 699-715.
Sleet, D. A., Hopkins, K. N., & Olson, S. J. (2003). From discovery to delivery: Injury prevention at CDC. *Health Promotion Practice, 4*(2), 98-102.
Srinivasan, S., O'Fallon, L. R., & Dearry, A. (2003). Creating healthy communities, healthy homes, healthy people: Initiating a research agenda on the built environment and public health. *American Journal of Public Health, 93*(9), 1446-1450.
StataCorp. (2009). Stata Statistical Software: Release 11. College Station, TX: StataCorp LP.
StataCorp. (2013). *Stata 13 Base Reference Manual*. College Station, TX.
Stefan, C., Dietze, M., Marchesini, P., Louise, W., & Candappa, N. (2010). Road Infrastructure Safety Management Evaluation Tools (RISMET): Data Systems and Requirements. Deliverable Nr 2 Retrieved 30th May, 2011, from www.eranetroad.org/index.php?option=com_docman&task=doc_download&gid=268&Itemid=53
Stevenson, M. R., Jamrozik, K. D., & Spittle, J. (1995). A case-control study of traffic risk factors and child pedestrian injury. *International Journal of Epidemiology, 24*(5), 957-964.
Stinchcombe, A., & Gagnon, S. (2010). Driving in dangerous territory: Complexity and road-characteristics influence attentional demand. *Transportation Research Part F: Traffic Psychology and Behaviour, 13*, 388-396.
Summala, H. (2007). Towards Understanding Motivational and Emotional Factors in Driver Behaviour: Comfort Through Satisficing. In P. C. Cacciabue (Ed.), *Modelling Driver Behaviour in Automotive Environments* (pp. 189-207). London: Springer.
Tarko, A. P. (2009). Modeling drivers' speed selection as a trade-off behavior. *Accident Analysis & Prevention, 41*(3), 608-616.
Theeuwes, J. (2001). The effects of road design on driving. In P. E. Barjonet (Ed.), *Traffic Psychology Today* (pp. 241-263). New York: Springer.
Tijerina, L., Kiger, S., Rockwell, T., & Wierwille, W. (1996). *Heavy Vehicle Driver Workload Assessment Task 5: Workload Assessment Protocol. Final Report Supplement.* Report no. DOT HS 808 467 (5). Washington DC: US Department of Transportation, National Highway Traffic Safety Administration.
Tingvall, C., & Haworth, N. (1999). *Vision Zero - An ethical approach to safety and mobility.* Paper presented at the 6th ITE International Conference Road Safety & Traffic Enforcement: Beyond 2000, Melbourne, Australia, 6-7 September, 1999.
Toops, H. A., & Haven, S. E. (1938). *Psychology and the motorist*. Columbus, OH: R.G. Adams.
Törnros, J. (1995). Effect of driving speed on reaction time during motorway driving. *Accident Analysis & Prevention, 27*, 435-442.
Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R. E., . . . Castellan, N. J. (1979). *Tri-level study of the causes of traffic accidents* Report No. DOT HS 805 099. U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington D.C.

Turner, B., & Tziotis, M. (2006). *Road safety engineering risk assessment - Recent and future research*. Paper presented at the Road Safety Research, Policing and Education Conference, 25-27 October, Surfers Paradise, Qld, Australia.
Turner, C., & McClure, R. (2003). Age and gender differences in risk-taking behaviour as an explanation for high incidence of motor vehicle crashes as a driver in young males. *Injury Control and Safety Promotion, 10*, 123-130.
Veltman, J. A., & Gaillard, A. W. K. (1996). *Pilot workload evaluated with subjective and physiological measures*. Paper presented at the Proceedings of the Europe Chapter of the Human Factors and Ergonomics Society Annual Meeting November 1993, Haren, The Netherlands.
VicRoads. (2004). *B-doubles and Higher Mass Limits Trucks. Information Bulletin.* . (VicRoads Publication Number 00170/2). Kew, Victoria, Australia: VicRoads.
VicRoads. (2007). *Oversized Load Carrying Vehicles*. (VicRoads Publication Number 00966/1). Kew, Victoria, Australia: VicRoads.
VicRoads. (2008a). *CrashStats User Guide Road Crash Statistics*. Kew, Victoria, Australia: VicRoads.
VicRoads. (2008b). *Register of public roads (Part A)*. Kew, Victoria, Australia: VicRoads.
VicRoads. (2010). *Traffic Engineering Manual Volume 1: Traffic Management* (4th ed.). Kew, Victoria, Australia: VicRoads.
Wegman, F., Dijkstra, A., Schermers, G., & van Vliet, P. (2006). Sustainable Safety in the Netherlands: Evaluation of National Road Safety Program. *Transportation Research Record, 1969*, 72–78.
Weller, G. (2010). *The psychology of driving on rural roads: Development and testing of a model* Wiesbaden: VS Verlag.
Werneke, J., & Vollrath, M. (2012). What does the driver look at? The influence of intersection characteristics on attention allocation and driving behavior. *Accident Analysis & Prevention, 45*, 610-619.
Werneke, J., & Vollrath, M. (2014). How do environmental characteristics at intersections change in their relevance for drivers before entering an intersection: Analysis of drivers' gaze and driving behavior in a driving simulator study. *Cognition, Technology & Work, 16*(2), 157-169.
Wilde, G. J. S. (1976). Social interaction patterns in driver behavior: An introductory review. *Human Factors, 18*, 477-492.
Williams, A. F. (1999). *The Haddon matrix: its contribution to injury prevention and control*. Paper presented at the Proceedings of the 3rd National Conference on Injury Prevention and Control, 9-12 May, Brisbane, Qld, Australia.
World Health Organization. (2010). Data Systems: A Road Safety Manual for Decision-makers and Practitioners. Retrieved May 30th, 2011, from http://rossow-beitzel.com/showmedia.php?mediaID=180&medialinkID=211
World Health Organization. (2013). Global status report on road safety 2013: Supporting a decade of action. Retrieved June 13th, 2015, from http://www.who.int/violence_injury_prevention/road_safety_status/2013/en/
Xu, X., Kouhpanejade, A., & Saric, Z. (2013). Analysis of influencing factors identification of crash rates using tobit model with endogenous variable. *Promet - Traffic - Traffico, 25*(3), 217-224.
Xu, X., Kwigizile, V., & Teng, H. (2013). Identifying access management factors associated with safety of urban arterials mid-blocks: a panel data simultaneous equation models approach. *Traffic Injury Prevention, 14*(7), 734-742.
Yan, X., & Radwan, E. (2007). Effect of restricted sight distances on driver behaviors during unprotected left-turn phase at signalized intersections. *Transportation Research Part F: Traffic Psychology and Behaviour, 10*(4), 330-344.
Young, K. L., Lenné , M. G., Beanland, V., Salmon, P. M., & Stanton, N. A. (2015). Where do novice and experienced drivers direct their attention on approach to urban rail level crossings? *Accident Analysis & Prevention, 77*, 1-11.

Young, K. L., Salmon, P. M., & Lenné , M. G. (2013). At the cross-roads: An on-road examination of driving errors at intersections. *Accident Analysis & Prevention, 58*, 226-234.

Zijlstra, F. R. H. (1993). *Efficiency in work behaviour: A design approach for modern tools.* Delft University of Technology.

APPENDICES

Appendix A

Table A.1 Glossary of abbreviations

Abbreviation	Description
AADT	Annual Average Daily Traffic Volume
ABS	Australian Bureau of Statistics
AIC	Akaike Information Criterion
AusRAP	Australian Road Assessment Program
BAC	Blood Alcohol Content
BIC	Bayesian Information Criterion
BVC	Bicycle-vehicle collision
BYO	Bring Your Own liquor licence
CBD	Central business district
CI	Confidence Interval
CN	Condition number
GEE	Generalised Estimating Equation
GLM	Generalised Linear Model
HR	Hazard ratio
IQR	Interquartile Range
iRAP	International Road Assessment Program
IRR	Incidence rate ratio
IRSAD	Index of Relative Socio-economic advantage and disadvantage
KSI	Killed or seriously injured (severe)
MLP	Mean lane position
MUARC	Monash University Accident Research Centre
MVC	Multi-vehicle collision
NASA-TLX	NASA Task Load Index
NDS	Naturalistic driving study
NZ	New Zealand
PDO	Property damage only (not severe)
PDT	Peripheral detection task
PVC	Pedestrian-vehicle collision
RSME	Rating Scale Mental Effort
sd	Standard Deviation
SDAcc	Standard deviation of accelerator pressure
SDLP	Standard deviation of lane position
SDSp	Standard deviation of speed
SLA	Statistical Local Area
SRIP	Safer Roads Infrastructure Program
SVC	Single-vehicle collision
TAC	Transport Accident Commission
TWLTL	Two-way left turn lane median
VicRoads	The Victorian State Road Authority
VIF	Variance inflation factor
VISTA	Victorian Integrated Survey of Travel & Activity
VKT	Vehicle km travelled
VMT	Vehicle miles travelled
ZINB	Zero-inflated negative binomial
ZIP	Zero-inflated Poisson

Appendix B

Table B.1 Methodological details of studies included in review of risk factors for crashes on urban roads

Reference	Crash type and severity	Study design	Location & time period	Data sources	Units of analysis (N)	Regression model	Fit, diagnostics & notes
(Abdel-Aty et al., 2009)	Total, severe (incapacitating and fatal) and rear-end. No details on minimum level of injury/damage	Cross-sectional	Florida, USA 2003-2006	Existing crash and roadway characteristic databases	Road segments on multilane arterials with partially limited access (1758)	NB regression (form 1). 27 models (3 crash types X 3 segment length X 3 land use (urban, suburban and rural)	Dispersion
(Alavi, 2013)	Pedestrian casualty crashes (weekday and weeknight)	Cross-sectional	Melbourne CBD, Australia 2000-2009	Existing databases (crashes, traffic, public transport, pedestrian monitoring, land use, travel surveys) and data collected on-site for project	CBD midblock road segments (?N)	Poisson regression (weeknight), Zero-inflated Poisson regression (weekday). Stepwise forward selection: 3 different forms (best fit chosen). Tested Poisson, , ZIP	Correlation between variables, overdispersion, spatial correlation, zero-inflation (Vuong test), McFadden's R^2, AIC

Reference	Crash type and severity	Study design	Location & time period	Data sources	Units of analysis (N)	Regression model	Fit, diagnostics & notes
(Avelar et al., 2013)	All. No details on level of injury/damage included	Cross-sectional	Oregon, USA 2004-2008	Existing databases (crashes), videos of road segments and Google Earth (road, roadside and land use)	Representative sample of homogeneous principal arterial road segments (40 urban, 82 rural)	NB regression (form 1). Stepwise selection: exposure variables first then other factors ($p<0.10$)	Dispersion and AIC
(Bonneson & McCoy, 1997)	All (includes PDO, threshold for damage not reported)	Cross-sectional	Phoenix, Arizona & Omaha, Nebraska, USA 1991-1993	Existing databases (crashes, traffic) and videos of road segments (land use, road geometry)	Urban arterial road segments between signalised intersections (189)	NB regression (form 1). Predictors screened using ANOVA (inappropriate for count data). Separate models for different median types, then combined	Pearson statistic, residual plots, dispersion, coefficient of determination R^2. NOTE: Spuriously large coefficient for undivided roads in residential/industrial areas (could be a result of combining the separate models for median types)
(Brown & Tarko, 1999)	All, KSI, PDO (threshold for damage not reported)	Cross-sectional	Indiana, USA 1991-1995	Existing databases (crashes, road inventory) and videos of road segments (detailed access control)	Representative sample of arterial streets homogeneous with respect to cross-section and traffic volume (155)	NB regression (form 2). Stepwise procedure, $p<0.10$	Tested assumption of proportionality between crashes and traffic volume, length and time. Overdispersion, plots of residuals vs predicted values

Reference	Crash type and severity	Study design	Location & time period	Data sources	Units of analysis (N)	Regression model	Fit, diagnostics & notes
(Dumbaugh & Li, 2011)	Motorist crashes, MVC, fixed-object, parked-car, BVC, PVC. No details on level of injury/damage included	Cross-sectional	San Antonio-Bexar county metropolitan region, USA 2003-2007	Existing databases (crashes, land use, street network, traffic, census)	Census block group plus 200 foot buffer (938)	NB regression (form 3)	None reported. NOTE: Problem with area-wide units of analysis and location specific risk factors – possible ecological fallacy (although location specific risk factors can be seen as indicators of land use)
(Dumbaugh et al., 2013)	PVC (all, KSI), BVC (all, KSI). No details on level of injury/damage included	Cross-sectional	San Antonio-Bexar county metropolitan region, USA 2003-2007	Existing databases (crashes, land use, street network, traffic, population census)	Census block group plus 200 foot buffer (938)	NB regression (form 3)	None reported. NOTE: Problem with area-wide units of analysis and location specific risk factors – possible ecological fallacy (although location specific risk factors can be seen as indicators of land use)

Reference	Crash type and severity	Study design	Location & time period	Data sources	Units of analysis (N)	Regression model	Fit, diagnostics & notes
(El-Basyouny & Sayed, 2009)	All. No details on level of injury/damage included	Cross-sectional	Vancouver, Canada 1994-1996	Existing data from City (little detail provided)	Arterial road corridors (58, made up of 392 road segments)	3 models: Poisson lognormal (PLN), PLN with random intercept, PLN with random parameters (form 1).	Overdispersion. Observed χ^2 and deviance information criteria used to choose best fitting model
(Greibe, 2003)	All (includes PDO, threshold for damage not reported)	Cross-sectional	Denmark 1990-1994	Existing databases (traffic volume, road geometry), videos and data collection on-site	Homogeneous urban road links (314)	Poisson regression (form 4). Manual backwards stepwise selection	Goodness of fit measured by % of systematic variation explained by model. NOTE: Can't interpret results relating to road width or number of minor exits per km (errors in table not clarified in text)
(Gruenewald et al., 1996)	SVC 8pm to 4am. No details on level of injury/damage included	Cross-sectional	California, USA April 1991-March 1996	Existing databases (crashes, retail alcohol outlets, road network, traffic, population census), telephone surveys	Areas based on population gradients (102)	SVC rate as outcome in spatial analysis models (based on linear regression taking into account spatial correlation between units)	Outliers, leverage, spatial correlation. NOTE: Linear regression may not have been appropriate, although SVC rate was normally distributed across areas.

Reference	Crash type and severity	Study design	Location & time period	Data sources	Units of analysis (N)	Regression model	Fit, diagnostics & notes
(Hadayeghi et al., 2003)	All crashes, KSI, PDO	Cross-sectional	Toronto, Canada 1996	Existing databases (crashes, travel survey, traffic network)	Traffic zones (463)	NB regression (form 5). Forward selection procedure	Pearson χ^2, R^2 alpha. Also tried spatial model—fit was better but assumes normally distributed errors so may be inappropriate. Coefficients not substantially different with spatial model
(Haynes et al., 2008)	Fatal crashes	Cross-sectional	New Zealand 1996-2005	Existing databases (crashes, digital road network, population census)	Territorial local authorities (73)	NB regression (form 5). Backwards stepwise procedure. Model built without road curvature variables, and then these were assessed in turn.	Overdispersion. NOTE: Possible ecological fallacy—curvature in area associated with crashes—doesn't mean the crashes occurred at curves

Reference	Crash type and severity	Study design	Location & time period	Data sources	Units of analysis (N)	Regression model	Fit, diagnostics & notes
(Jackett, 1993)	All. No details on level of injury/damage included	Cross-sectional	New Zealand 1987-1991	Existing databases (Land Transport Safety Authority)	Arterial or collector road segments homogeneous with respect to speed limits, traffic volumes and physical characteristics (782)	Poisson regression (form 2). Also tested linear regression and multiplicative gamma model.	Inspected residual plots to choose between model types
(Jonsson, 2005)	MVC, SVC, BVC, PVC injury crashes	Cross-sectional	Sweden 1997-2001 or 1998-2002	Existing databases (crashes, vehicle flow, street function). Vulnerable road user volumes, speed, street design and environment collected on-site	Homogeneous road links, usually the distance between two main intersections, sometimes smaller (389)	Quasi-Poisson regression with scaling factor for overdispersion (form 4 – all continuous variables entered as ln(x)). Separate models for MVC, SVC, BVC, PVC	% of total variation explained, residual analysis, leverage. SVC may not have been true SVC (other vehicles may have been involved but not recorded)

Reference	Crash type and severity	Study design	Location & time period	Data sources	Units of analysis (N)	Regression model	Fit, diagnostics & notes
(LaScala et al., 2000)	Pedestrian injuries (all) and subset where pedestrian had been drinking	Cross-sectional	San Francisco, USA 1990	Existing databases (crashes, alcohol outlets, traffic, population census), cross street density from digital maps	Census unit tracts (149)	Spatial regression model – ordinary least squares corrected for spatial autocorrelation. 2 models: all pedestrian injuries, pedestrian injuries where pedestrian had been drinking	Leverage, Cook's distance, spatial autocorrelation. NOTE: Model may not have been appropriate for count data
(LaScala et al., 2001)	Pedestrian SVC: pedestrian had been drinking or had not been drinking. No details on level of injury involved.	Cross-sectional	California, USA April 1991-March 1996	Existing databases (crashes, alcohol outlets, population census, traffic, road network), telephone surveys	Areas based on population gradients (102)	Spatial analysis models (based on linear regression taking into account spatial correlation between units). Separate models for pedestrians who had been drinking and those who had not.	None reported. NOTE: Model may not have been appropriate for count data

Reference	Crash type and severity	Study design	Location & time period	Data sources	Units of analysis (N)	Regression model	Fit, diagnostics & notes
(Lee, 2000)	Run-off road SVC. No details on level of injury/damage included	Cross-sectional	Washington State, USA 1994-1996	Existing databases (crashes, road geometry, traffic, roadside features)	Equal length road segments on principal arterial Washington State Route 3 (120)	NB regression (form 3). Poisson and ZIP also considered.	Overdispersion, Vuong statistic (to test for zero-inflation)
(Manuel et al., 2014)	All. No details on level of injury/damage included	Cross-sectional matched pairs (matched pairs were oversized and standard sized roads matched for traffic volume, speed limit, roadway type, no. lanes, spatial location)	Edmonton, Canada 2006-2010	Existing databases (crashes, land inventory), aerial photos	Two lane residential collector road segments between two intersections (212)	NB regression for matched pairs (form 1). Stepwise selection procedure, p<0.10. Tested all interactions between size of road and other variables	Scaled deviance and Pearson χ^2

Reference	Crash type and severity	Study design	Location & time period	Data sources	Units of analysis (N)	Regression model	Fit, diagnostics & notes
(Potts et al., 2007)	All, MVC, SVC. Includes all, KSI and PDO	Cross-sectional	Minnesota, USA 1999-2003 & Michigan, USA 1998-2002	Sources not identified	Arterial road segments (1153 in Minnesota, 1878 in Michigan)	NB regression (form 4). Attempted 90 models: 5 road types x 3 crash types (All, MVC, SVC) x 3 crash severities (all, KSI, PDO) x 2 locations. Models not mutually exclusive. Not all successful.	Dispersion, R^2. NOTE: Precision of estimates not reported—no standard errors or 95% CIs so can't tell if effects differ across different road or crash types. Don't know how many crashes each models is based on.
(Sawalha & Sayed, 2001)	All. No details on level of injury/damage included	Cross-sectional	Vancouver & Richmond, Canada 1994-1996	Existing databases (crashes, traffic), roadway geometry collected on-site	Arterial road segments between signalised intersections (58)	NB regression (form 1). Forward stepwise procedure.	Scaled deviance and Pearson χ^2, residual plots
(Xu, Kouhpanejade, et al., 2013)	All. No details on level of injury/damage included	Cross-sectional	Las Vegas, USA 2003-2005	Existing databases (crashes, traffic, travel speeds), access management from Google Earth	Divided major and 19 minor arterial road segments between signalised intersections (356)	Tobit regression with endogenous variable (travel speed)	None reported. NOTE: Were the distributional assumptions for Tobit regression met?

Reference	Crash type and severity	Study design	Location & time period	Data sources	Units of analysis (N)	Regression model	Fit, diagnostics & notes
(Xu, Kwigizile, et al., 2013)	All. No details on level of injury/damage included	Cross-sectional	Las Vegas, USA 2003-2005	Existing databases (crashes, traffic, travel speeds), access management from Google Earth	Divided major and 27 minor arterial road segments between signalised intersections (400)	Random coefficient simultaneous equations (ordinary least squares) – one to predict travel speed, one to predict ln(crashes)	None reported. NOTE: Linear regression inappropriate for count data (crashes).

Form 1: No. Crashes = $AADT^{\beta 1}$ x $Length^{\beta 2}$ x $\exp(\beta_3X_3+\beta_4X_4+\ldots+\beta_nX_n)$

Form 2: No. Crashes = AADT x Length x $\exp(\beta_1X_1+\beta_2X_2+\ldots+\beta_nX_n)$

Form 3: No. Crashes = $\exp(\beta_1X_1+\beta_2X_2+\ldots+\beta_nX_n)$

Form 4 regression model: $AADT^{\beta 1}$ x Length x $\exp(\beta_2X_2+\beta_3X_3+\ldots+\beta_nX_n)$

Form 5 regression model: $Vehiclekm^{\beta 1}$ x $\exp(\beta_2X_2+\beta_3X_3+\ldots+\beta_nX_n)$

Appendix C

Pre-drive questionnaire

Participant code: ______________________

Date: ______________________

Thank you for coming along today. Your involvement is greatly appreciated. For research purposes, it is important that we obtain some information concerning your background. Please answer each question as fully and as accurately as possible, and remember, all of the information that you provide will be kept confidential.

Part A - Personal Details

Are you: Male ☐ Female ☐

How old are you? ☐

Do you suffer from any form of colour blindness?

☐ Yes ☐ No

Do you suffer from any eye diseases that affect your visual acuity and/or visual field?

☐ Yes ☐ No

Do you have any neck problems that severely restrict your head movements?

☐ Yes ☐ No

Part B – Driving experience

Do you hold a current Victorian driver's licence?

☐ Yes, Probationary ☐ Yes, full ☐ No

How old were you when you when you were first licensed to drive a car?

☐

Are there any conditions on your licence?

☐ A Automatic only

☐ S Glasses or corrective lenses

☐

V Driver aids or vehicle modifications

On average, how many hours do you spend driving a car each week?

On average, how many kilometres do you drive each week?

In which environment do you drive the most?

- ☐ Metropolitan
- ☐ Residential
- ☐ Rural

Thank you.

Current Well-Being Questionnaire

Participant Code: ____________

Date: _____________ Pre-drive / Post drive (Circle)

Please indicate the extent to which each of the symptoms listed below is affecting you now.

1. General discomfort:	None	Slight	Moderate	Severe
2. Fatigue:	None	Slight	Moderate	Severe
3. Headache:	None	Slight	Moderate	Severe
4. Eye strain:	None	Slight	Moderate	Severe
5. Difficulty focusing:	None	Slight	Moderate	Severe
6. Increased salivation:	None	Slight	Moderate	Severe
7. Sweating:	None	Slight	Moderate	Severe
8. Nausea:	None	Slight	Moderate	Severe
9. Difficulty concentrating:	None	Slight	Moderate	Severe
10. Fullness of head [1]:	None	Slight	Moderate	Severe
11. Blurred vision:	None	Slight	Moderate	Severe
12. Dizzy (eyes open):	None	Slight	Moderate	Severe
13. Dizzy (eyes closed):	None	Slight	Moderate	Severe
14. Vertigo [2]:	None	Slight	Moderate	Severe
15. Stomach discomfort:	None	Slight	Moderate	Severe
16. Burping:	None	Slight	Moderate	Severe

----

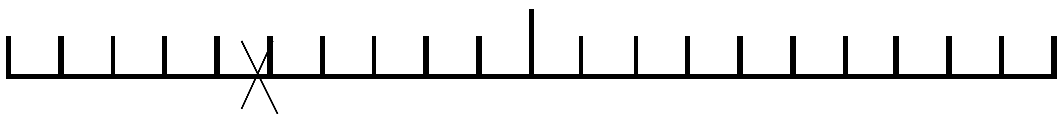
[1] Fullness of head = awareness of pressure within the head

[2] Vertigo = feeling of a loss of orientation with respond to vertical upright

Driver Behaviour in Simulated Urban Environments – Post Drive Questionnaire (A)

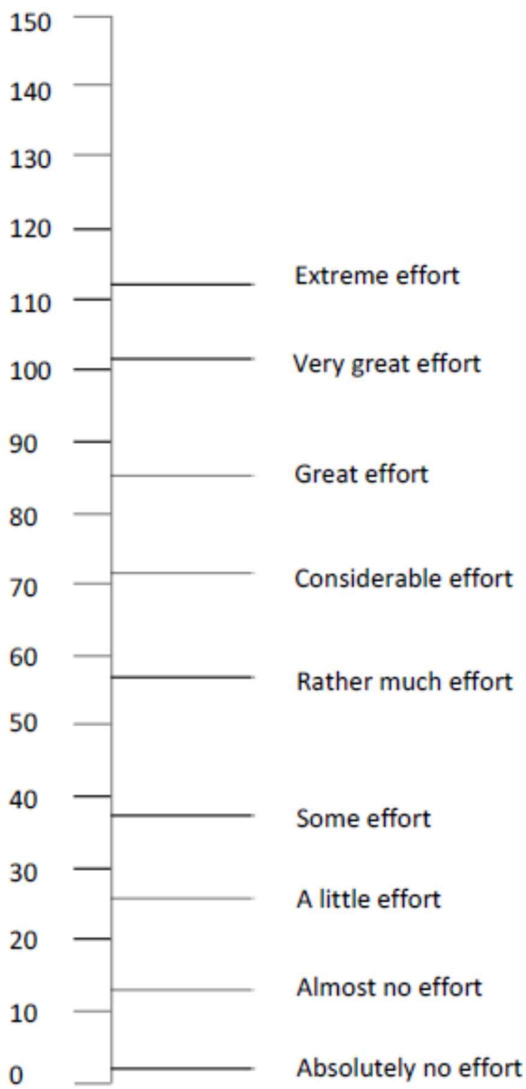
Date	
Participant Number	
Drive number	

Please answer the following questions by referring to the picture provided and marking the scales with a cross placed on the line, like this:



or this:

1. Please indicate, by marking on the vertical axis below, how much effort it took for you to drive this section of road at this speed.



2. How difficult did you find it to drive this section of road at this speed?

Not difficult **Extremely difficult**

3. How much risk did you experience driving this section of road at this speed?

No risk **Maximum risk**

4. How comfortable did you feel driving this section of road at this speed?

Extremely comfortable **Extremely uncomfortable**

5. In the real world, how often would you typically drive at this speed in this type of road environment?

a. Never
b. Seldom
c. Sometimes
d. Nearly always
e. Always

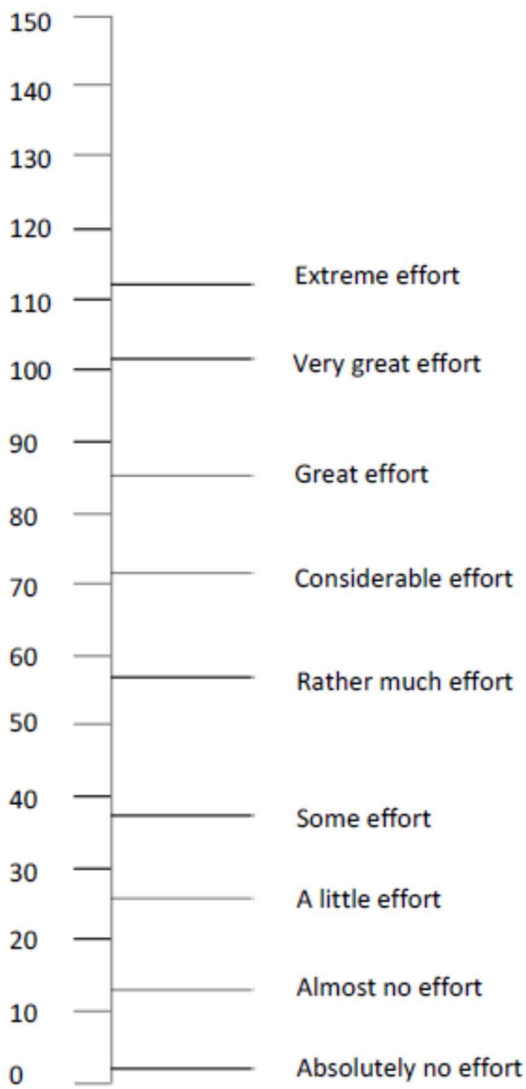
6. In the real world, how many times do you think you would have a collision, or lose control of the vehicle, if you drove in a road environment like this at this speed every day for 2 months (i.e. 60 times)?

___ (enter a number from 0 to 60)

7. In the real world, imagine if 60 drivers like you, of the same age and experience, were to each drive in a road environment like this at this speed and in these conditions (not all at once). How many do you think would have an accident or lose control of the vehicle?

____ (enter a number from 0 to 60)

1. Please indicate, by marking on the vertical axis below, how much effort it took for you to drive this section of road at this speed.



2. How difficult did you find it to drive this section of road at this speed?

Not difficult **Extremely difficult**

3. How much risk did you experience driving this section of road at this speed?

No risk **Maximum risk**

4. How comfortable did you feel driving this section of road at this speed?

Extremely comfortable **Extremely uncomfortable**

5. In the real world, how often would you typically drive at this speed in this type of road environment?

  a. Never
  b. Seldom
  c. Sometimes
  d. Nearly always
  e. Always

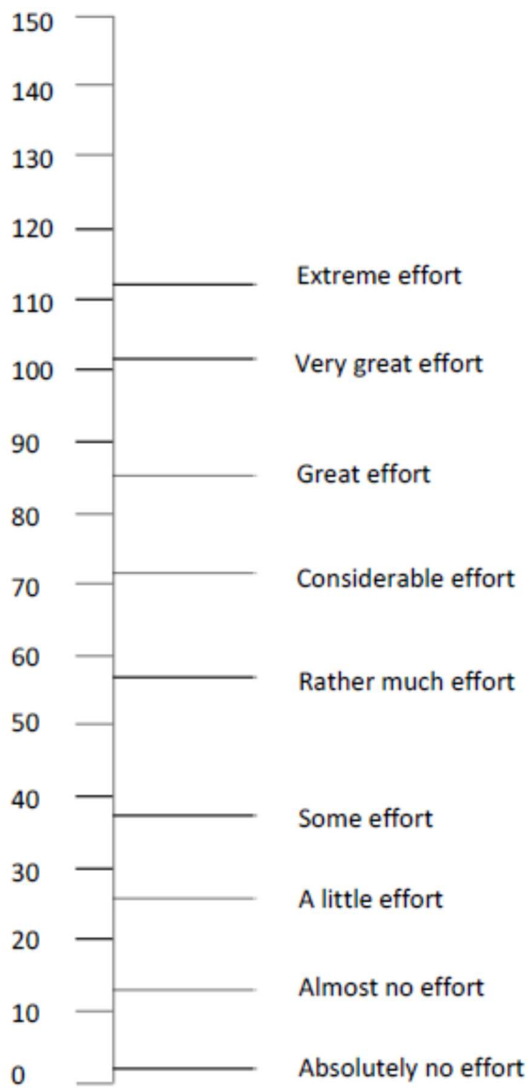
6. In the real world, how many times do you think you would have a collision, or lose control of the vehicle, if you drove in a road environment like this at this speed every day for 2 months (i.e. 60 times)?

___ (enter a number from 0 to 60)

7. In the real world, imagine if 60 drivers like you, of the same age and experience, were to each drive in a road environment like this at this speed and in these conditions (not all at once). How many do you think would have an accident or lose control of the vehicle?

____ (enter a number from 0 to 60)

1. Please indicate, by marking on the vertical axis below, how much effort it took for you to drive this section of road at this speed.



2. How difficult did you find it to drive this section of road at this speed?

Not difficult **Extremely difficult**

3. How much risk did you experience driving this section of road at this speed?

No risk **Maximum risk**

4. How comfortable did you feel driving this section of road at this speed?

Extremely comfortable **Extremely uncomfortable**

5. In the real world, how often would you typically drive at this speed in this type of road environment?

a. Never
b. Seldom
c. Sometimes
d. Nearly always
e. Always

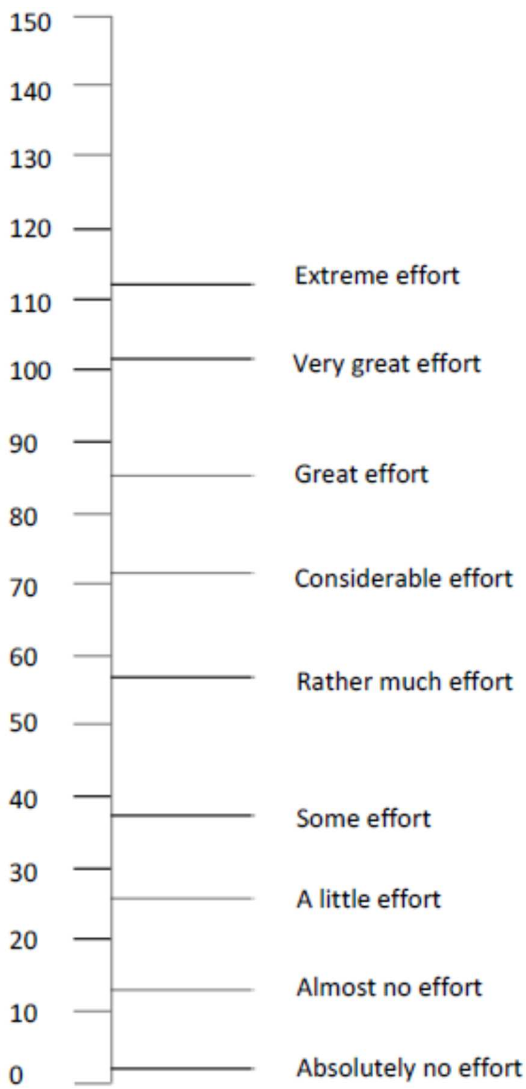
6. In the real world, how many times do you think you would have a collision, or lose control of the vehicle, if you drove in a road environment like this at this speed every day for 2 months (i.e. 60 times)?

___ (enter a number from 0 to 60)

7. In the real world, imagine if 60 drivers like you, of the same age and experience, were to each drive in a road environment like this at this speed and in these conditions (not all at once). How many do you think would have an accident or lose control of the vehicle?

____ (enter a number from 0 to 60)

1. Please indicate, by marking on the vertical axis below, how much effort it took for you to drive this section of road at this speed.



2. How difficult did you find it to drive this section of road at this s peed?

Not difficult **Extremely difficult**

3. How much risk did you experience driving this section of road at this speed?

No risk **Maximum risk**

4. How comfortable did you feel driving this section of road at this speed?

Extremely comfortable **Extremely uncomfortable**

5. In the real world, how often would you typically drive at this speed in this type of road environment?

a. Never
b. Seldom
c. Sometimes
d. Nearly always
e. Always

6. In the real world, how many times do you think you would have a collision, or lose control of the vehicle, if you drove in a road environment like this at this speed every day for 2 months (i.e. 60 times)?

___ (enter a number from 0 to 60)

7. In the real world, imagine if 60 drivers like you, of the same age and experience, were to each drive in a road environment like this at this speed and in these conditions (not all at once). How many do you think would have an accident or lose control of the vehicle?

____ (enter a number from 0 to 60)