Impacts of Traffic Interventions on Road Safety: An Application of Causal Models

Haojie Li

Imperial College London

Centre for Transport Studies

Department of Civil and Environmental Engineering

Submitted for the Diploma of the Imperial College (DIC),

PhD degree of Imperial College London

Copyright Declaration

Acknowledgements

This thesis would not have been possible without the assistance of a number of people and institutions.

First, I am most grateful to Professor Daniel Graham and Dr Arnab Majumdar who co-supervised me throughout this research. I express my thanks to them for their constant guidance, advice and support.

I am also grateful to the China Scholarship Council and the Lloyd's Register Foundation Transport Risk Management Centre for providing financial support for this project.

I would also like to thank all those institutions and people who provided access to data for this thesis and who kindly answered all my related questions: Dr Rosa Sanchis-Guarner-Herrero of the London School of Economics; the Ordnance Survey and the Office for National Statistics.

I would also like to thank my family, particularly my wife Yao Lin, who supports me all the time.

Finally, my special thanks go to the staff and students of the Centre for Transport Studies at Imperial College London who supported me throughout the period of this research.

The research presented here is my own, except where the work of others has been referenced.

Haojie Li

Publications

Journal

Li, H., D. J. Graham, and A. Majumdar. The effects of congestion charging on road traffic casualties: A causal analysis using difference-in-difference estimation. Accident Analysis and Prevention, Vol. 49, 2012, pp. 366-377.

Li, H., D. J. Graham, and A. Majumdar. The impacts of speed cameras on road accidents: an application of propensity score matching methods. Accident Analysis and Prevention, in press.

Conference Proceedings

Li, H., and D. J. Graham. The Effects of Congestion Charging on Road Traffic Casualties: A Causal Analysis using Difference-In-Difference Estimation. Transportation Research Board 90th Annual Meeting, Washington, DC, 2011.

Tuerk, A., D. J. Graham, and H. Li. Impact of congestion charging on traffic volumes: evidence from London. Transportation Research Board 91th Annual Meeting, Washington, DC, 2012.

Li, H., D. J. Graham, and A. Majumdar. The impacts of speed cameras on road accidents: an application of propensity score matching methods. Transportation Research Board 92th Annual Meeting, Washington, DC, 2013.

Abstract

This thesis is concerned with the causal relationship between traffic interventions and road safety. It focuses on two issues that have been overlooked in the existing empirical literature: the establishment of a causal link between traffic interventions and road traffic accidents, and the application and development of formal causal approaches, which have not yet been applied in the field of road safety.

In the past decades substantial studies have been conducted to investigate the risk factors contributing to road accidents. It has been shown that the frequency and severity of road accidents are associated with various factors, including traffic characteristics, road environment and demographic characteristics. However, the existence of a causal link between traffic interventions and road accidents remains unclear due to the complex character of traffic interventions. Meanwhile, the lack of formal causal models makes it difficult fully to address issues such as confounding effects and regression to the mean bias.

This thesis begins by reviewing and discussing different types of traffic interventions in order to demonstrate the chains through which traffic interventions are related to road safety. To address the shortcomings in empirical literature, three models for causal inferences are discussed: the difference-in-difference method, the propensity score matching method and Bayesian methods.

These formal causal approaches are then applied to three empirical studies: the London congestion charging scheme, speed limit enforcement cameras, and the road network design. The conventional models are also employed and compared with formal causal models.

List of Tables

List of Figures

List of Acronyms

AADF: Annual Average Daily Flow

AADT: Annual Average Daily Traffic

AIC: Akaike Information Criterion

ATE: Average Treatment Effect

ATET: Average Treatment Effect on the Treated

BE: Bayesian Estimation

BIC: Bayesian Information Criterion

CIA: Conditional Independence Assumption

CSC: Common Support Condition

DDD: Difference-in-Difference-in-Difference

DfT: Department for Transport

DIC: Deviance Information Criteria

DID: Difference-In-Difference

DW: Durbin-Watson

EB: Empirical Bayesian

ERP: Electronic Road Pricing

FB: Full Bayesian

FCSs: Fatal and Serious Collisions

GIS: Geographical Information System

GLMs: Generalized Linear Models

GWPR: Geographically Weighted Poisson Regression

GWR: Geographically Weighted Regression

IMD: Index of Multiple Deprivation

KSI: Killed and Seriously Injured

LCC: London Congestion Charge

MCMC: Markov Chain Monte Carlo

MLE: Maximum Likelihood Estimation

NB: Negative Binomial

OD: Origin-Destination

ODPM: Office of the Deputy Prime Minister

OLS: Ordinary Least Square

ONS: Office for National Statistics

OS: Ordnance Survey

PDF: Probability Density Function

PICs: Personal Injury Collisions

PSM: Propensity Score Matching

RCM: Rubin Causal Model

RMSPE: Root Mean Squared Prediction Error

RTM: Regression-To-Mean

SAR: Spatial Autoregressive

SEM: Spatial Error Model

SPFs: Safety Performance Functions

SUV: Sports Utility Vehicle

TfL: Transport for London

V/C: Volume-to-Capacity ratio

VKT: Vehicle Kilometres Travelled

VMT: Vehicle Miles Travelled

ZINB: Zero-Inflated Negative Binomial

ZIP: Zero-Inflated Poisson

Chapter 1: Introduction

Over the last two decades there has been a small but growing body of research seeking to determine the effects of traffic interventions on road safety. This thesis contributes to the literature on this theme by focusing on the issue of establishing a causal link between traffic interventions and road traffic accidents and by addressing some of the methodological limitations of previous work. This introductory chapter presents the background, motivation and objectives of the thesis. The first section provides a brief review of evaluation studies of traffic interventions and explains why the estimation results are inconsistent and have been questioned. Section 1.2 explains the motivation and objectives of the thesis, while the research contributions are presented in section 1.3. An outline of the thesis is provided in the last section.

1.1 Background

Road accidents place a great burden on individuals, property and society. During the last few decades, considerable research has been conducted to identify important factors related to the occurrence of road accidents, including traffic characteristics, road characteristics, socio-economic and environmental factors (Baruya, 1998; Ossiander and Cummings, 2002; Taylor et al., 2002; Martin, 2002; Golob et al., 2004; Lord et al., 2005; Kononov et al., 2008; Noland and Quddus, 2005; Abdel-Aty and Radwan, 2000; Noland and Quddus, 2004; Graham and Glaister, 2003; Wedagama et al., 2006; Dissanayake et al., 2009; Graham et al., 2008; Quddus, 2008; Wier et al., 2009). In recent years, some researchers have paid close attention to the impact of road safety measures on the incidence of road accidents (Pau and Angius, 2001; Galante et al., 2010; Allpress and Leland Jr, 2009; Elvik, 2001; Goldenbeld and van Schagen, 2005; Hess and Polak, 2003; Newstead and Cameron, 2003; Chen et al., 2002; Christie et al., 2003; Mountain et al., 2005; Shin et al., 2009; Keall et al., 2001; Gains et al., 2004, 2005; Jones et al., 2008). Most of these studies, however, make associational inferences instead of causal inferences. That is to say that, association, or correlation, is a relationship between two or more variables, while causation implies that the change in one thing directly causes a change in the other. In other words, causal relationships between one variable and another cannot be obtained only from the observed association between them.

It is worth noting that the road safety measures that have been evaluated in previous studies are different from the traffic interventions in this thesis. "Traffic interventions", as discussed in this research, are defined as policies, legislation and enforcement, the construction of road networks, and other general-purpose measures which directly or indirectly affect traffic condition, drivers' behaviour and the travel environment. Traffic interventions are different from other road safety measures in that they may influence traffic conditions, the model split of transport, and other aspects. The implementation of traffic interventions, therefore, can have a direct or an indirect impact on road accidents, regardless of whether that impact is expected or unexpected. It is more complicated to estimate the causal effects of traffic interventions than measures designed for road safety. A better understanding of this causal relationship, however, would help policy makers to evaluate the safety outcomes of traffic interventions and hence enhance the prevention of road accidents.

There are currently two approaches widely used in previous evaluation studies for road safety: observational before-after control methods (Goldenbeld and van Schagen, 2005; Christie et al., 2003; Cunningham et al., 2008; Gains et al., 2004; Jones et al., 2008) and empirical Bayes (EB) methods (Hauer, 1997; Hauer et al., 2002; Persaud et al., 2009; Persaud and Lyon, 2007; Sayed et al., 2004; Hirst et al., 2004). The problem with the conventional before-after studies is that they have failed fully to address issues such as confounding effects, regression to the mean, and time trend effects. Recently, Bayes approach has become popular as a statistically defensible method that can deal with key issues evident in observational before-after studies, particularly as a means of increasing the precision of estimation and correction for the regression to the mean bias. A review on the EB before and after safety studies is conducted by

Persaud and Lyon (2007). They discuss the basics of the EB approach and justify the validity of and need for the EB approach, and address the critical issues in the interpretation of EB evaluations. The EB approach requires the use of a reference group similar to the treatment group, however. While the validity of the EB approach relies heavily on the availability of an appropriate reference group, there are rarely studies looking at the suitability of candidate reference groups.

In previous studies, the inferences made regarding the safety effects of road safety measures have often been inconsistent and debatable. For example, numerous studies have been conducted to investigate the effect of safety cameras, with results showing that the implementation of safety cameras has significantly reduced vehicle speed and casualty numbers near camera sites (Goldenbeld and van Schagen, 2005; Hess and Polak, 2003; Newstead and Cameron, 2003; Chen et al., 2002; Christie et al., 2003; ARRB Group Project Team, 2005; Mountain et al., 2004; Cunningham et al., 2008; Mountain et al., 2005; Shin et al., 2009; Keall et al., 2001; Gains et al., 2004, 2005; Jones et al., 2008). Despite the apparent wealth of empirical evidence, however, there is still debate about the effectiveness of speed cameras. Opponents of speed cameras argue that there has been no proper independent study of speed camera effectiveness using a controlled sample of the population and proper scientific techniques (ABD, 2011). While this point of view may appear somewhat pessimistic in light of the many empirical studies that have been conducted on this theme, it is in fact true that the existing research has failed fully to address issues of confounding, selection bias and reverse causality. This is due in part to the fact that formal causal approaches to inference, used routinely in other areas of science such as medicine and epidemiology, have not yet been adopted.

Rubin (1973a, b, 1974, 1977, 1978) developed potential outcomes models for observational analysis of causal effects. Rubin proposed that the causal inferences can be made by comparing potential outcomes, which can be defined as outcomes for the same unit given various doses of exposure to the treatment. These models, labelled the Rubin Causal Model (RCM) by Holland (1986), are widely applied in both the statistics and econometrics literature. Indeed, the RCM approach has a long-standing precursor in the statistical literature in the early work on randomised experiments such as Fisher (1925) and Neyman (1923). In econometrics, meanwhile, the first attempts to evaluate labour market programmes by Ashenfelter (1978) and Ashenfelter and Card (1985) mark the first application of RCM models. RCM has become an important tool and has been applied in many areas, such as labour economics, public health, industrial investment, etc. Despite the superiority of the causal modelling, however, such techniques have not been tried in road casualty analysis, especially the evaluation of traffic interventions.

The lack of formal causal models in the evaluation of traffic interventions means that there is uncertainty regarding the size of the causal effects. The causal relationship between traffic interventions and road casualties is thus a worthwhile area of research. The next section introduces the specific gaps in the literature that form the motivation of this thesis, as well as defining the thesis' objectives.

1.2 Motivations and Objectives

The first motivation for this thesis regards the causal link between traffic interventions and road safety. A substantial number of studies have been conducted to investigate the risk factors contributing to road casualties. It has been shown that the frequency and severity of casualties are associated with various factors, including traffic characteristics, road environment and demographic characteristics (Baruya, 1998; Ossiander and Cummings, 2002; Taylor et al., 2002; Martin, 2002; Golob et al., 2004; Lord et al., 2005; Kononov et al., 2008; Noland and Quddus, 2005; Abdel-Aty and Radwan, 2000; Noland and Quddus, 2004; Graham and Glaister, 2003; Wedagama et al., 2006; Dissanayake et al., 2009; Graham et al., 2008; Quddus, 2008; Wier et al., 2009). On the other hand, many studies have focused on the relationship between traffic interventions and traffic flow, travel modes, environment and business matters (e.g. Eliasson and Mattsson, 2006; Olszewski and Xie, 2005; Tuerk and Graham, 2010; Wichiensin et al., 2007). Despite the fact that several studies have evaluated the effects of traffic interventions on the number of accidents and injuries (Hyatt et al., 2009; Quddus, 2008; Noland et al., 2008; Mountain et al., 2005; Shin et al., 2009; Gains et al., 2004, 2005; Jones et al., 2008), the assignment of these interventions are not randomized and the observational inferences are not always of high quality. The nature of the causal link between traffic interventions and road casualties remains unclear, therefore. Although the complex character of traffic interventions makes it difficult to generalize about their effects, a better understanding of the causal relationship would help policy makers to evaluate the safety outcomes of interventions and hence improve the prevention of road accidents. In this thesis, formal causal models are applied to establish a causal link between traffic interventions and road casualties, and to address the shortcomings of previous literature by studying this link.

The second motivation for this thesis regards the dataset used for road safety analysis at the aggregate-level. One issue which is very critical in all road safety analysis is the selection of appropriate traffic exposure variables. Traffic exposure is the most important factor influencing traffic crash counts, however there is not currently an appropriate variable that can be used to control for the traffic exposure in area-level analyses. In analysis at the disaggregate (unit) level, where the study object is usually road sections or intersections, the annual average daily traffic (AADT) or vehicle miles travelled (VMT) is preferred as the traffic exposure variable (Huang et al., 2010; Marshall and Garrick, 2011; Jones et al., 2008). At the aggregate (area) level, however, these variables are not always available and, although proxy variables for traffic exposure have been developed (Graham and Glaister, 2003), they entail some limitations. A failure properly to control for the traffic exposure could bias the inferences drawn from studies.

Another issue concerns the usage of data about road network characteristics. A detailed dataset of the road network, including road class, road length and node information can be obtained from Ordnance Survey (OS) Meridian^{TM}2. Although this dataset has been used in several studies in the UK (Noland and Quddus, 2004; Haynes et al., 2007; Graham and Stephans, 2008; Jones et al., 2008), the data only covers a single year, which means the variance in the road network over time cannot be accounted for.

In this thesis, data for control variables is obtained from different sources and aggregated at the same level. In particular, a new method is proposed to construct the traffic exposure variable and a dataset of the road network covering the period 2001-2010 is constructed. Detailed discussion is provided in other chapters.

The final motivation for the thesis regards modelling the mechanism of treatment assignment. Road safety measures, the objective of which is to improve the level of road safety, are different from traffic interventions which have a more general purpose. Road safety measures, such as traffic calming and road traffic legislation and enforcement, are usually implemented at locations and at times where a high number of road casualties is observed. This causes a common phenomenon called regression to the mean (RTM), also called selection bias, when repeated measurements are made on the same unit. Since time series, or longitudinal, datasets are usually employed in the analysis of road accidents, it is crucial to consider the effect of this selection bias, especially in estimates of the effectiveness of countermeasures. Even so, incorrect inferences are still being made due to the failure to recognize this bias. If no correction has been applied, the expectations about the effectiveness of safety improvements will be unrealistic.

Traditionally, the empirical Bayes method is used to increase the precision of estimation and to correct for selection bias (Lord and Park, 2008; Persaud and Lyon, 2007). The effect of selection bias can be also reduced by introducing comparison or control groups (Barnett et al., 2005), a technique widely used in before-after control studies (Goldenbeld and van Schagen, 2005; Christie et al., 2003; Cunningham et al., 2008; Gains et al., 2004; Jones et al., 2008). Usually, a reference or control group is employed to estimate the counterfactual outcomes of the treatment group. A critical issue which has been inadequately addressed in previous studies in this area is the selection of this reference or control group. To the best of my knowledge, there is currently no clear criterion by which to select the reference or control group similar to the treatment group. This is largely because of the lack of a solid understanding of the mechanism behind the treatment assignment.

The motivations described above give rise to the following main objectives of this thesis:

(1) To explore various factors affecting road traffic accidents and traffic interventions.
(2) To investigate the causal relationship between traffic interventions and road traffic accidents using empirical studies.
(3) To model the assignment mechanism of traffic interventions by applying causal evaluation techniques and identifying confounding factors.
(4) To control for traffic exposure and confounding factors, using a time-series database containing information about road casualties, geographic and road characteristics, by collecting and organising data previously unavailable.
(5) To provide policy-makers with a better understanding of the causal effects of traffic interventions on road safety, and hence to help policy-makers to evaluate the safety outcomes of traffic interventions and hence improve the prevention of road accidents.

1.3 Research Contributions

Despite the fact that great effort has been made to understand the factors affecting road casualties, we know little about how traffic interventions impact road safety. This thesis contributes to the literature on this theme by focusing on the causal link between traffic interventions and road casualties. Traffic interventions impact road safety directly or indirectly by influencing traffic conditions, travel modes, driving environment and behaviours. Hence, the causal relationship between road safety and general-purpose traffic interventions is not straightforward and is difficult to establish. Policy makers, whether as part of national government (e.g. Department for Transport (DfT)) or local authorities, can have a better understanding of the safety outcomes of traffic interventions, and hence improve the prevention of road accidents. In this thesis three empirical studies are presented to show the causal effect on road safety of three traffic interventions: the London congestion charge, speed limit enforcement cameras, and the road network design in the UK.

The second contribution we make in this research relies on the application and development of formal causal approaches which have not previously been applied in the field of road safety. These causal approaches yield a solid understanding of the mechanism behind the treatment assignment and, in so doing, allow us to find out and control for confounding factors that may bias estimation. They can also address a critical issue in previous before-after control studies regarding the selection of reference or control groups. The causal models employed here can be used to obtain clear criteria-based evidence on the selection of units into treatment or control groups. A full DID model, PSM method and Bayesian methods are used in this thesis to evaluate the causal effects of traffic interventions on road casualties, with the conventional EB approach also being further developed by combining it with a propensity score.

One of the direct outputs of this research is the creation of a national panel dataset containing detailed information about traffic characteristics, geographic and road characteristics, demographic information, and road accidents data. This comprehensive database will be useful to researchers in road safety related studies by providing them with previously unavailable data. In doing so, the thesis is also able to address issues with the traffic exposure and road characteristic variables that were discussed in 1.2 above. In relation to traffic exposure the thesis adopts an approach based on traffic assignment to estimate the traffic exposure at ward level. The idea of this method is that trips generated between origin-destinations (ODs) are assigned to transportation networks and aggregated in each ward. In relation to the data for road network characteristics, the thesis uses panel data of the road network (obtained from OS Meridian ™ 2 for the period from 2001 to 2010) to account for effects due to the variation in road characteristics over time.

1.4 Structure of the Thesis

The thesis is organized into eight chapters, including the present one. There are three review chapters (chapters 2 to 4) and three empirical chapters (chapters 5 to 7). A brief summary of each chapter is given below.

Chapter 2 provides a review of the literature on the factors affecting road casualties. Evidence is provided on the effects on road casualties of various factors, including traffic characteristics, road network and infrastructure, and demographic environmental characteristics. The chapter also discusses the conventional methods for road casualty analysis, with the application of these methods, both in linear and non-linear models, being presented.

Chapter 3 defines and discusses different types of traffic interventions. It demonstrates the chains through which traffic interventions are related to road safety. Results from previous evaluation studies are presented to show the safety effects of traffic interventions. Methodological shortcomings in the literature are also discussed.

Chapter 4 provides a historical review of literature on techniques for causal inferences. A comparison of the econometric models and the conventional Bayesian framework are presented. Three models for causal inferences are discussed in detail: the difference-in-difference (DID) method, propensity score matching (PSM) method and

Bayesian methods. To address the issue of the selection of proper control groups, we propose the EB approach using a propensity score.

In chapter 5, the DID method is applied to identify the impacts of the London congestion charge on road traffic accidents within the central London charging zone. A full DID model is developed that is integrated with generalized linear models, such as Poisson and Negative Binomial regression models. Covariates are included in the model to adjust for factors that violate the parallel trend assumption, which is critical in the DID model.

In chapter 6, the effects of speed limit enforcement cameras on road safety are examined. The PSM method is applied to select proper reference groups, and hence control for selection bias. The EB method and a simple before and after approach are also employed and compared with the PSM method. 771 sites and 4787 sites for a period of 9 years are observed for the treatment and reference groups respectively.

Chapter 7 investigates how changes in road network characteristics affect road safety at ward level in the UK. Panel data of road networks from 2001 to 2010 is used to generate traffic flow data at ward level. The full Bayes (FB) method is applied to estimate the causal relationship between various risk factors and road traffic accidents. A panel data semi-parametric model is also used to interpret the treatment effect heterogeneity of the continuous treatment.

Chapter 8 presents some concluding remarks, the limitations of the thesis and some potential directions for future research.

Chapter 2: Literature Review

This chapter provides a critical review of the literature on traffic accidents in order to identify the factors affecting the occurrence and severity of road traffic accidents. The methods used in previous road accident analyses are also discussed. A brief introduction is given in the first section, followed by a review of the literature on various factors affecting road accidents. The discussion of statistical models for accident analysis is provided in section 2.3. A conclusion is presented in the last section.

2.1 Introduction

Factors affecting road accidents have been the focus of considerable research in transport studies during the past decade. Most studies have shown that a broad range of factors can affect road accidents and many inferences have been drawn based on the exploration of traffic characteristics, primarily, traffic speed, density and flow. Taking each of these characteristics in turn, it can be assumed that increased speed would lead to more severe accidents (Ossiander and Cummings, 2002; Taylor et al., 2002). The relationship between the accident rates and density follows a U-shape relationship (Zhou and Sisiopiku, 1997) and it has been suggested that low traffic flow can induce both a higher accident rate and more severe accidents (Martin, 2002). It is essential to consider and examine these factors when conducting road safety analysis. In addition to traffic characteristics, other factors affecting traffic accidents include road infrastructure, and demographic and environmental characteristics. Failure to control for variables affecting road accidents, i.e. confounding factors, can bias the results of any analysis of traffic interventions on road traffic accidents.

When it comes to statistical analysis, two approaches have been used to analyse road crash[1] frequency data in previous studies. The first approach is to develop a continuous variable and then to apply linear regression models. The second is to employ Generalized Linear Models (GLMs) to model the number of accidents. Whilst the linear regression model, using Ordinary Least Square (OLS), can account for autocorrelation and heteroskedasticity, it violates the strict OLS requirement for normality. On the other hand, although the GLMs, such as the Poisson and Negative Binomial (NB) regression, can accommodate nonnegative discrete data, accounting for autocorrelation and heteroskedasticity with such models is complicated. Noland and Karlaftis (2005) tested the sensitivity of crash models to alternative specifications, including linear regression models and GLMs, and suggested that NB models were more robust than linear regression models.

In recent years, more attention has been given to the issue of reducing the severity of road accidents. Insights into the safety effectiveness of traffic interventions also require a more comprehensive understanding of the relationships between crash severities and vehicle, roadway and human factors. Over the last decade, various approaches have been employed to analyse crash severity data. Discrete outcome models play a dominant role in the literature. The dependent variables used in such models may have either binary or multiple response outcomes. Dependent variables with multiple response outcomes can be further treated as ordinal and unordered. Since this research mainly focuses on the safety effects of traffic interventions using aggregate-level data, such discrete outcome models are not applicable. In this chapter, a brief discussion of methodologies for casualty severity analysis is provided.

[1] "Crash" and "accident" are similar terms and will be used interchangeably in this thesis.

2.2 Factors Affecting Road Casualties

In this section, a review of the literature on various factors affecting the occurrence and severity of road accidents is provided. Three types of factors are discussed, including traffic characteristics, road network and infrastructure, and demographic and environmental characteristics.

2.2.1 Traffic Characteristics

Speed, flow and density are three main characteristics affecting the occurrence and severity of road accidents. Traffic interventions may also affect traffic characteristics. In order fully to understand how traffic interventions act on road accidents, it is worthwhile to review the literature on the relationship between traffic characteristics and road accidents.

Numerous studies have been conducted to investigate how these factors are associated with road accidents. One common assumption is that high speed, free flow and low density are associated with more severe accidents. Meanwhile other studies have also found that the frequency of road accidents is associated with traffic speed. It seems logical that driving faster is more likely to lead to an accident. Others believe that the crash involvement rate depends on the deviation from the mean speed rather than absolute speed of the traffic. Although there is no conclusive answer to this problem, it is explicit that accident severity increases with pre-crash speed.

Baruya (1998) investigated speed-accident relationships on European roads based on accident and speed data collected from four European countries. A Poisson regression model was employed to analyse the discrete data with a negative relationship being found between mean speed and accident frequency. An ecological study designed by Ossiander and Cummings (2002) examined the effect of increased speed limits on freeways in Washington State on the incidence of crashes. Twenty years of data were analysed using Poisson regression with findings suggesting that an increased speed limit was related to a higher rate of fatal crashes and more deaths on highways. A similar result obtained by Aljanahi et al. (1999) also suggested that reducing the speed limit could result in fewer accidents. Taylor et al. (2002) collected data from 174 road sections across UK, including accident data, traffic flow and speed data. The authors used GLMs to link accident frequency with traffic speed, flow and other factors. Their results showed that accident frequency was positively related with the mean traffic speed.

One limitation of these studies is the use of the Poisson regression model. In deciding the best fit for the data in this model, it is necessary to undertake comparative analysis using other models, such as the NB regression. The data used in these studies were all cross-sectional in nature and it is essential to account for the spatial correlation when modelling the accident data.

As discussed before, it seems reasonable to relate the number of accidents with the traffic flow. Research by Martin (2002) described the influence of the hourly traffic volume on accident rates and severity. Data from 2000 km of French interurban motorways over a two year period was used. The results showed that accident rates were highest in light traffic (1000 to 1500 vehicles/h) and accident severity was also greater when hourly traffic was light. Golob et al. (2003; 2004) analysed crash data for 1998 in California State using a tool that monitors the real-time safety level of traffic flow. Their results showed that there was a relationship between traffic flow and accident rates and traffic volume had more influence on accident severity than speed. Ivan et al. (2000) employed Poisson models to regress highway accident rates on traffic density[2], land use and light conditions. The paper demonstrated that traffic intensity, such as the volume/capacity ratio, played an important role in predicting accident rates and interpreting the causes of high accident rate hotspots. Lord et al.

[2]Traffic density is defined as vehicles per km.

(2005) conducted a detailed investigation into the relationships of crashes and traffic flow characteristics on freeway segments which are located in rural and urban area of Montreal, Quebec. Predictive models of three different functional forms were developed for rural and urban freeways. Single-vehicle and multiple-vehicle crashes were analysed separately to get a proper explanation. Their results showed that traffic volume is inadequate as the only explanatory variable and that functional forms including traffic density and V/C ratio performed better in predicting single- and multi-crashes. The models suggested that accident risk and the number of accidents increased as the traffic density and V/C ratio increased.

Another important traffic characteristic that has been examined by many studies is congestion. It is generally believed that the crash frequency increases as congestion levels rise; however severity levels are not expected to be affected. Traffic flow conditions change as traffic becomes congested and hence affects accidents. Shefer and Rietveld (1997) noted that congestion led to lower numbers of fatalities since lower speeds reduced the likelihood of fatal accidents. The increased travel time by car due to congestion can also result in a shift from private car use to public transport, thus indirectly improving safety levels. Kononov et al. (2008) explored the relationship between road safety and congestion with the application of safety performance functions (SPFs). The authors observed that the total number of accidents increased with congestion. Noland and Quddus (2005), meanwhile, conducted a disaggregated spatial analysis to examine the effect of congestion on road safety in the Greater London area. Although no conclusive results were found, they suspected that congestion may affect crash severity more on high speed roads than in urban conditions. Overall, then, these results would tend to suggest that congestion charging has the potential for improving safety as well as benefiting mobility. On the other hand, a spatial analysis of the M25 motorway in England by Wang et al. (2009), employing a precise congestion measurement and spatial models to explore the impact of congestion on road accidents, found little or no impact due to the mixed effects of traffic congestion.

Despite varying findings from the different studies, it is obvious that traffic characteristics can have a major impact on both the occurrence of road traffic accidents and their severity. A good understanding of how these factors work can guide traffic interventions so that they can affect both the number and severity of road accidents.

2.2.2 Road Network and Infrastructure

Road infrastructure and its impact on traffic accidents has been another focus of road safety research in recent years. Numerous studies have noted that infrastructure characteristics can affect road traffic accident rates and should be included as vital variables in road safety analysis.

Abdel-Aty and Radwan (2000) estimated the frequency of accident occurrence on a principal arterial in Central Florida by employing negative binomial models. This paper highlighted the importance of the road infrastructure characteristics, such as the degree of horizontal curvature, the number of lanes, shoulder widths and the road section’s length. The results showed that people driving on a road with narrow lane and shoulder width, a larger number of lanes and reduced median width were more likely to be involved in accidents. Noland (2003) investigated the effects of infrastructure changes on road traffic accidents while other factors that may affect the occurrence of such accidents were controlled. The variables on infrastructure characteristics included lane miles, number of lanes for different types of road and the proportion of each type of road. Noland’s results suggested that certain changes in highway infrastructure in the US between 1984 and 1997 had the effect of increasing absolute total number of traffic casualties. Another spatially disaggregate analysis of road casualties in England undertaken by Noland and Quddus (2004) examined the effects of road characteristics and land use on road casualties. Their results suggest that an increased length of “B” road[3] can increase serious injuries, although the coefficients for other types of road were not significant. Noland and Oh (2004) examined how changes in road infrastructure can affect the occurrence of road accidents. The authors used fixed effect NB regression to estimate county-level time-series data in Illinois, USA. The authors found that increased road accidents were associated with increased number of lanes, increased lane widths and decreased outside shoulder width. One suggestion made by them was the need to account for time-variant factors. Amoros et al. (2003) aggregated accident data by road type within a number of counties in France and subsequently analysed this data using the NB regression. One of the findings in this study was that the difference in accident numbers and their severity between the counties depended on the type of road.

2.2.3 Demographic and Environmental Characteristics

In addition to their impact as vehicle drivers, the populace impacts road traffic accidents in other ways. Factors such as population, employment, age and gender can reflect the social structure and economic activities of an area with an attendant impact on accidents. Furthermore, environmental characteristics, such as land use, are the principal determinant of trips and may also influence accident rates.

Previous studies have been published on the relationship between accidents and demographic and environmental characteristics. Zajac and Ivan (2003) evaluated the effect of different types of roadways and area type features, such as the land use type, on the injury severity of pedestrian accidents in rural Connecticut. The authors found that downtown fringe, village and low-density residential areas generally experienced lower injury severity. Graham and Glaister (2003) investigated the impact of land use mix, urban scale and density on pedestrian casualties. They employed a disaggregated

[3]In the UK, all roads can be classified into Motorway, A road, B road and Minor road.

model to explore the influence of the local environment on pedestrian accidents. Their results suggest that residential zones tend to have higher incidences of pedestrian casualties than economic districts.

In addition, population and employment were also found to play important roles in explaining the influence on pedestrian casualties. In research on the influence of land use on non-motorized accidents in Newcastle-upon-Tyne (UK) conducted by Wedagama et al. (2006), primary functional land use, population density and junction density were treated as explanatory variables. Both pedestrian and cyclist casualties during working hours were positively associated with retail land use in the city centre. Dissanayake et al. (2009) examined the feasibility of using land use factors to analyse child pedestrian casualties on road. The authors used a geographical information system (GIS) technique to allocate accident data to the map of Newcastle-upon-Tyne, where land use types were divided by trip attractors and generators. Results from six GLMs showed that child pedestrian accidents were related to secondary retail and high residential land use types. Clifton and Kreamer-Fults (2007) employed multivariate models to estimate crash severity and crash risk exposure near public schools in Maryland, USA. Transit (i.e. public transport) access, commercial access, recreation facilities and population density were found to be positively associated with higher aggregate crash severity.

Graham et al. (2005; 2008) analysed child pedestrian casualties in England at the ward level. They found child-related accidents were positively related to socio-economic deprivation. A dummy independent variable to measure the volume of through traffic was generated by a gravity model using population and employment in proximate wards and found to be significant. Bedard et al. (2002) analysed crashes involving single vehicle in the USA and suggested that older drivers and female drivers were at a greater risk of being fatally injured in an accident compared to younger and male drivers. Another analysis by Eluru et al. (2008) examined nonmotorist injury severity in accidents in the USA and found that the elderly were more injury-prone. In the research by Quddus (2008), results from spatial models suggested that older people were associated with fewer traffic injuries but with more serious injuries. A recent result obtained by Wier et al. (2009) using data from San Francisco indicated that variables, including percentages of neighbourhood commercial land use and residential-neighbourhood commercial land use, employment, percentages of people living in poverty and percentages of people aged 65 and over, were positively associated with vehicle-pedestrian crashes. Table 2.1 summarizes previous studies on various factors affecting road accidents.

Table2.1 A summary of studies on factors affecting road accidents

Authors	Modelling Approach	Data and Units of Analysis	Factors Affecting Road Accidents	Main findings
Baruya (1998)	Poisson Regression	203 links from 4 European Countries for the 1990s	Traffic Speed	A negative relationship was found between mean speed and accident frequency
Aljanahi et al. (1999)	Poisson Regression	1 county of the UK and 1 county in Bahrain, 1987-1990	Traffic Speed	Reducing the speed limit could result in less accidents
Ossiander and Cummings (2002)	Poisson Regression and Negative Binomial Regression	Freeways of Washington State, 1974-1994	Traffic Speed	An increased speed limit was associated with a higher fatal crash rate and more deaths on freeways
Taylor et al. (2002)	Principal components analysis, Generalised Linear Modelling	174 road segments from rural roads in England	Traffic Speed	Accident frequency was positively related with the mean traffic speed
Ivan et al. (2000)	Non-linear Poisson Regression using quasi-likelihood estimation techniques	17 US rural two-lane highway segments, 1997-1998	Traffic density and land use	Traffic intensity played an important role in predicting accident rates and interpreting the causes of high accident rates point
Martin (2002)	Poisson Regression and Negative Binomial Regression	2000 km of French interurban motorways	Traffic Flow	Accidents rates were highest in light traffic and accident severity was greater when hourly traffic was light

Table2.1 A summary of studies on factors affecting road accidents (Continued)

Authors	Modelling Approach	Data and Units of Analysis	Factors Affecting Road Accidents	Main findings
Golob et al. (2004)	Nonlinear canonical correlation analysis	All police-reported cases on the California State Highway System, 1998	Traffic Flow	There was a relationship between traffic flow and accident rates. Volume had more influence on accident severity than speed
Golob et al. (2003)	Nonlinear canonical correlation analysis	6 major freeway in California, 1998	Traffic flow, weather and lighting conditions	
Lord et al. (2005)	Generalized Estimating Equation	A rural section and an urban section in Canada, 1994-1998	Traffic flow, density and V/C ratio	Accident risk and the number of accidents increased as density and V/C ratio increased
Noland and Quddus (2005)	Negative binomial regression	15366 spatial units in the Greater London area, 1999-2001	Traffic Congestion	Congestion may affect crash severity on high speed roads rather than in urban conditions
Shefer and Rietveld (1997)	Piece-wise, linear speed-density function	Simulated dataset	Traffic Congestion	Congestion led to lower numbers of fatalities
Wang et al. (2009)	Poisson based models using a full Bayesian estimation	M25 motorway in England, 2004-2006	Traffic Congestion	Little or no impact was found due to mixed effects of traffic congestion
Kononov et al. (2008)	Safety performance functions	Multilane freeways in Colorado, California, and Texas	Traffic Congestion	Congestion charging could have the potential for safety improvement as well as mobility benefits

Table2.1 A summary of studies on factors affecting road accidents (Continued)

Authors	Modelling Approach	Data and Units of Analysis	Factors Affecting Road Accidents	Main findings
Noland and Quddus (2004)	Negative binomial regression	8414 wards of England, 1999	Road characteristics and land use	Increased length of B roads could increase serious injuries
Amoros et al. (2003)	Negative binomial regression	8 counties in France, 1987-1993	Road types	Difference in accidents and severities between counties depended on the type of road
Noland and Oh (2004)	Fixed effect NB regression	102 counties in Illinois, 1987-1994	Road network infrastructure and geometric design	Increased road accidents were associated with increased number of lanes, increased lane widths, and decreased outside shoulder width
Abdel-Aty and Radwan (2000)	Negative binomial regression	A principal arterial in Central Florida, 1992-1994	Road infrastructure characteristics	Narrow lane and shoulder width, larger number of lanes and reduced median width were more likely to induce accidents
Noland (2003)	Fixed effect NB regression	50 US states, 1984-1997	Road infrastructure characteristics	Changes in highway infrastructure had the effect of increasing total traffic casualties
Clifton and Kreamer-Fults (2007)	Ordinary least squares linear regression	The State of Maryland, 2000-2002	Environmental context	Transit access, commercial access, recreation facilities and population density were found to be positively associated with higher aggregate crash severity
Graham et al. (2008)	Generalized Linear Models	Cross-section, 7927 Census Area Statistic wards in England, 1998-2002	Deprivation	Child-related accidents were positively associated with deprivation

Table2.1 A summary of studies on factors affecting road accidents (Continued)

Authors	Modelling Approach	Data and Units of Analysis	Factors Affecting Road Accidents	Main findings
Zajac and Ivan (2003)	Ordered probit modelling	Roadways in rural Connecticut, 1989-1998	Roadway and the area type	Low-density residential areas generally experienced lower injury severity
Bedard et al. (2002)	Multivariate logistic regression	US traffic fatalities, 1975-1998	Drivers' age and gender	Older drivers and female drivers were at greater risk of being fatally injured in an accident
Wier et al. (2009)	Least squares regression	176 San Francisco, California census tracts, 2001-2005	Land use	Proportions of land area zoned for neighbourhood commercial use and residential-neighbourhood commercial use had a positive association with vehicle-pedestrian injury collisions
Quddus (2008)	Classical spatial models and Bayesian hierarchical models	633 wards in London, 2000-2002	Road infrastructure, socio-economic and traffic conditions	Road casualties were positively associated with road length and traffic flow, but negatively associated with population aged 60 or over.
Eluru et al. (2008)	Mixed generalized ordered response logit model	60 areas across the US, 2004	Drivers' age and casualty type	The elderly were found to be more injury-prone
Wedagama et al. (2006)	Generalized Linear Models	Two zones in Newcastle upon Tyne, 1998-2001	Urban land use	Pedestrian and cyclist casualties during working hours were positively associated with retail land use
Graham and Glaister (2003)	Generalized Linear Models	8414 wards of England, 1999-2000	Urban scale, density and land use	Residential zones tended to have higher incidences than economic district

2.3 Models for Road Casualty Analysis

In this section, a review of models for road casualty analysis is provided. Casualty frequency models are discussed first, followed by a description of spatial accident models. Models for casualty severity analysis are also reviewed.

2.3.1 Casualty Frequency Models

Two frequently used approaches for casualty frequency analysis, linear regression and GLMs, are discussed and compared in this section. A brief description of GLMs, including Poisson, NB and Zero-inflated models, is also provided here.

2.3.1.1 Linear Regression Models

Traditionally researchers studied accident rates in a two-stage processs. The first stage was to convert the count outcomes into continuous outcomes, followed by the use of conventional regression models. The exposure rate variable was developed instead of the number of crashes to avoid violating the normality assumption in OLS. The regression structure for panel data can be defined as:

$Y_{it} = \alpha + \boldsymbol{\beta}\mathbf{X_{it}} + u_{it}$, $i = 1, \ldots, N; t = 1, \ldots, T$ *2.1*

Where,

i refers to groups;

t refers to time periods;

Y_{it} refers to the accident rate in group i, period t;

α refers to the intercept;

$\boldsymbol{\beta}$ refers to the vector of coefficients;

$\mathbf{X_{it}}$ refers to the vector of covariates in group i, period t;

The error term u_{it} can be further described as

$u_{it} = \mu_i + \lambda_t + v_{it}$ *2.2*

Where μ_i is the group-specific effect,λ_t is the time-specific effect and v_{it} is the individual effect.

To combine equation *2.1* and *2.2*, we will obtain

$Y_{it} = \alpha + \boldsymbol{\beta}\mathbf{X_{it}} + \mu_i + \lambda_t + v_{it}$ *2.3*

This model becomes a random effects model if both the time-specific effect λ_t and the group-specific effect μ_i are random, while a fixed effects model is made when the inferences are confined to the effects in the sample. The Hausman test can be used to decide which model should be employed (Hausman, 1978). Although the linear regression model, using OLS, can account for autocorrelation and heteroskedasticity, it violates the strict OLS requirement for normality. Generalised Linear Models were developed to account for this issue.

2.3.1.2 Generalized Linear Models

In recent decades, generalized linear models, such as Poisson and NB regression models, have become a popular approach to investigate the relationship between accident frequency and various covariates. The GLMs are superior to linear regression models in accommodating nonnegative discrete data. Accounting for autocorrelation and heteroskedasticity with such models is complicated, however, as will be discussed in the following sections.

(1) Poisson Models

The Poisson regression model is described as (see Lord et al., 2005)

$$P(Y \mid \mu) = \frac{\exp(-\mu)\mu^{Y}}{Y!} \qquad 2.4$$

Where $P(Y|\mu)$ is the probability of Y accidents occurring and μ is the expected number of accidents.

A generalized linear model with a Poisson distribution is given as,

$$\ln\mu_{it} = \alpha + \boldsymbol{\beta}\mathbf{X} \qquad 2.5$$

Where α is the intercept, $\boldsymbol{\beta}$ is the vector of coefficients and $\boldsymbol{X}$ is the vector of covariates. A critical limitation of the Poisson regression model, however, is that the variance of the data is confined to be equal to the mean,

$$\mathrm{Var}(Y) = E(Y) = \mu \qquad 2.6$$

It is often the case that the variance is greater than the mean, which is known as over-dispersion. To cope with this issue, a Poisson regression model with a Gamma distributed error term has been developed.

(2) Negative Binomial Models

Various factors, including data clustering and misspecification of the model, can lead to over-dispersion. It has been shown that the dispersion is largely due to the nature of crash data, which are subject to Bernoulli trials.

Since unobserved heterogeneity due to omitted variables widely exists in the crash data set, a NB error is usually introduced in the GLMs. In this model, coefficients are analysed to investigate the average relationship between dependent variables, e.g. the number of accidents, and possible covariates.

The functional form of the NB model is as follows (Lord, 2006):

$$\mu=\exp(\alpha + \beta X)\exp(\varepsilon) \qquad 2.7$$

where *exp(ε)* is the Gamma distributed error with mean 1 and variance $1/\phi$. ϕ is a Poisson-Gamma distributed variable with mean *exp(βX)* and a variance *exp(βX)(1+ exp(βX)/ϕ)* respectively.

The probability density function (PDF) of the Poisson-Gamma model is given by the following equation

$$f(y_{it} \mid \mu_{it}, \phi) = \frac{\Gamma(y_{it}+\phi)}{\Gamma(\phi) y_{it}!} \left(\frac{\phi}{\mu_{it}+\phi}\right)^{\theta} \left(\frac{\mu_{it}}{\mu_{it}+\phi}\right)^{y_{it}} \qquad 2.8$$

While GLMs employ a simple function of the dispersion parameter α, many researchers have proposed that the variance of NB model is better illustrated by a dispersion function that depends on site-specific characteristics, such as demographic and socio-economic factors. If the dispersion parameter is incorrectly estimated, any subsequent analyses of road traffic accidents can be undermined. For example, Lord (2006) investigated how the low mean problem (LMP) affected the estimation of the dispersion parameter and the effects of an unreliable dispersion parameter in the analyses of highway safety. A series of NB models with different dispersion parameters and sample size were simulated. The results showed that a dataset characterized by a small sample and the low mean problem can seriously affect the estimation of the dispersion parameter. Based on the work of Lord (2006), Mitra and Washington (2007) used an independent dataset to explore additional dispersion functions of traffic flow. The authors evaluated their models by several methods, including significance of coefficients, standard deviance and the deviance information criteria (DIC). The results indicated that the additional dispersion functions greatly depended on how the mean models were specified. If the mean function was well defined and all vital covariates were included, the extra-variance would become insignificant.

(3) Zero-inflated Poisson and NB Models (ZIP and ZINB)

It is often the case that data for accidents is commonly characterized by an excess of zeros, i.e. there are more zeros than expected in the Poisson or NB models. To account for this, a zero-inflated model has been developed. Johnson and Kotz (1968) first proposed a modified Poisson distribution to deal with excess zeros:

$$P(n)=\alpha+(1-\alpha)e^{-\lambda} \quad n=0 \qquad 2.9$$

$$P(n)=(1-\alpha)\frac{\lambda^{n}e^{-\lambda}}{n!} \quad n>=1 \qquad 2.10$$

The Vuong statistic (Vuong, 1989) is used to measure whether the ZIP or ZINB is appropriate for fitting the datasets.

In zero-inflated theory, excess zeros come from two different distributions, e.g. two Poisson distributions. A dual-state process is thus applied by many researchers to handle the excess number of zeros (See Lord et al., 2004 for a detailed discussion). The validity of such progress should not rely on the spatial or temporal scale selected for analyses, however. By comparing the empirical data and simulating several models for different spatial scales, Lord et al. (2004) found that a dual-state process is not feasible for use in road traffic accident studies, which means that the ZIP and ZINB models may statistically fit well but cannot interpret the underlying crash generation process. This is because the dual-state process requires there to be some intersections or road segments with no accidents. However, there is no such situation in reality, so the dual-state process misinterprets the nature of road accidents. This leads to an important issue in statistical analyses: statistical modelling is not solely about maximizing the statistical fit, but also must account for the nature of the data generated. The authors suggested four possible reasons for excess zeroes: (1) the spatial or temporal scale chosen for analyses is too small; (2) under-reporting of crashes; (3) sites characterized by high risk and low exposure; (4) omission of important variables that interpret the crash process. One solution for (1) is to choose an appropriate spatial or temporal scale. To capture the heterogeneity when estimating Poisson and NB models, an extra term, usually known as fixed or random effects, can be introduced to account for the omitted variables and under- or misreporting of crashes. Lord (2007) also concluded that zero-inflated models should be avoided for modelling highway vehicle crash data.

2.3.2 Spatial Accident Models

In conventional GLMs, the relationship between the dependent and independent covariates is assumed to be consistent across the geography of the study area when estimating parameters. This assumption may be violated, however, because the accident rate is likely to be affected by many spatial factors, e.g. demographic and land use characteristics. To address this problem, spatial models have been developed.

Moran's I test for the residuals obtained from OLS estimation can be used to detect the presence of spatial correlation. Moran's I is described as:

$$I = \frac{N}{\Sigma_i \Sigma_j w_{ij}} \frac{\Sigma_i \Sigma_j w_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\Sigma_i (X_i - \bar{X})^2} \qquad \textit{2.11}$$

Where N is the number of spatial units indexed by i and j; X is the dependent or independent variable of interest; $\bar{X}$ is the mean of X; and w_{ij} is a matrix of spatial weights. The value of Moran's I lies in the range of -1 to +1, with 0 indicating a random spatial pattern. A Z score value can be calculated to indicate whether or not to reject the null hypothesis "there is no spatial clustering". Should spatial clustering exist, then this needs to be accounted for by the use of spatial models.

2.3.2.1 Conventional Spatial Models

Two spatial models are suitable for controlling spatial correlation, the spatial autoregressive (SAR) model and the spatial error model (SEM).

The SAR model is described by:

$$Y_i = \rho WY_i + \beta X_i + \varepsilon_i \qquad \textit{2.12}$$

Where Y is a vector of the cross-sectional dependent variable, WY is a spatially lagged variable with a weight matrix W, ρ is the coefficient for the lagged variable, β is the vector of coefficients, X is the vector of covariates and ε is a normally distributed random error term with zero mean and variance σ^2.

The SEM model is described as (Anselin, 1988):

$Y_i = \beta X_i + u_i$ *2.13*

$u_i = \lambda W u_i + \varepsilon_i$ *2.14*

Where u_i is an error term to account for spatial correlation and λ is the spatial autoregressive coefficient.

In order to model count data, the count dependent variable is converted into a continuous variable by dividing it by an appropriate exposure variable and then a SAR or SEM model is applied. The SAR and SEM models for count data can be expressed as:

$\ln(Y_i/E_i) = \rho W \ln(Y_i/E_i) + \beta X_i + \varepsilon_i$ *2.15*

and

$\ln(Y_i/E_i) = \beta X_i + u_i$ *2.16*

$u_i = \lambda W u_i + \varepsilon_i$ *2.17*

Where *E* is the exposure variable. The Maximum Likelihood method can be used to estimate these models. As discussed earlier, the conventional spatial methods are more suitable for continuous data. Hence Bayesian hierarchical methods have been developed to deal with non-negative random count data.

2.3.2.2 Spatial Models Using Bayesian Methods

In recent years Bayesian methods have become popular in the analysis of road traffic safety. There are two kinds of Bayesian method: Empirical Bayesian and full Bayesian. Persaud et al. (2007) conducted a study aiming to capitalize on the experience gained from two decades of before-after safety studies using Empirical Bayesian methods. The authors confirmed that the results from EB methodology are more valid than those produced by conventional ones. Another study by Park et al.

(2009) employed a fully Bayesian multivariate approach to evaluate the safety effects of deceasing the speed limit using accident data obtained from expressways in Korea for 13 years. Compared to the Empirical Bayesian methods, the full Bayesian methods led to more precise safety effectiveness estimates of the expected number of crashes.

A fully Bayesian framework is based on the posterior distribution of model parameters:

$$\pi(\theta \mid D) = \frac{L(D \mid \theta)\pi(\theta)}{m(D)} \qquad 2.18$$

Where D is the observed data set, θ is the vector of parameters, $\pi(\theta|D)$ is the posterior distribution, $L(D|\theta)$ is the maximum likelihood function, $m(D)$ is the marginal distribution of data D. A prior distribution is assigned to each parameter as a priori information and this information can be derived from history data or expert opinions.

The Bayesian hierarchical model with the Poisson specification is given as (Persaud et al., 2007)

$Y_i \sim \text{Poisson}(\theta_i)$

$\ln\theta_i = \ln(E_i) + \alpha + \beta X_i + \eta_i + \sigma_i$ *2.19*

where η_i is the spatial effect and σ_i is the unobserved heterogeneity. The advantage of this model is that spatial variation can be distinguished from over-dispersion.

2.3.2.3 MLE and Bayesian

There are two methods to determine the parameters in the regression models: the traditional classical Maximum Likelihood Estimation (MLE) and the Bayesian Estimation (BE). A Bayesian model makes inferences based on posterior estimates, which reflects the probabilities of interest to the analyst, i.e. the probability of the null hypothesis being true. In contrast, MLE on parameters provides the probability of observing data, given that parameters take on specific values. The classical MLE and Bayesian models are also related. MLE makes inferences based on the likelihood of data, while in Bayesian models, the likelihood of the observed data given parameter θ, is employed to calibrate the prior inference $\pi(\theta)$, with the updated knowledge summarized in the posterior density $\pi(\theta|x)$. Bayesian methods will be further discussed in chapter 4.

2.3.2.4 Geographically Weighted Regression Model

A common problem when using simple GLMs to analyse spatial data is that one model is assumed to fit all. Geographically Weighted Regression (GWR) allows for the specification of models to vary over space.

Hadayeghi (2010) employed the Geographically Weighted Poisson Regression (GWPR) model to analyse the spatial relationship between the number of zonal crashes and potential transport planning variables, using the 2001 collision data for the City of Toronto. The results showed that the GWPR is suitable for illustrating the spatial relationships and is superior to conventional GLMs in predicting the number of crashes. The basic idea of GWPR is that the observations near point *i* have more influences on the estimation of coefficients than those located further from *i*. The following model form was used by Hadayaghi (2010):

$$\ln Y_i=\alpha(u_i)+ \beta(u_i)X_i \qquad 2.20$$

Here, $u_i(=(u_{xi}, u_{yi}))$ denoting the coordinates of *i*th point.

One important step in the implementation of GWPR is the spatial kernel function and the band width, which determines the number of observations around each subject point and the distance decay in the weighting function.

The estimator from Generalized Weighted least square is

$$\beta(u_{xi}, u_{yi}) = (X^t W(u_{xi}, u_{yi}) X)^{-1} X^t W(u_{xi}, u_{yi}) y \qquad 2.21$$

where W is an n x n matrix,

$$\mathbf{W}(u_{xi}, u_{yi}) = \begin{matrix} w_{i1} & 0 & 0 & \cdots & 0 \\ 0 & w_{i2} & 0 & \cdots & 0 \\ 0 & 0 & w_{i3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & w_{in} \end{matrix}$$

Where w_{in} is the weight of the data at point n on the calibration of the model around point i. In the global OLS model every observation has a weight of unity, so w_{in} equals to one.

For GWR models, there are several choices for defining the diagonal elements of the weighting function, including: bi-square nearest neighbour function, the exponential function and the Gaussian Function. Generally, these functions are based on the distance d_{ij}. For example, the weights from the exponential kernel function are calculated as:

$$W_j(s) = \exp(-d_{ij}/\gamma) \qquad 2.22$$

Where d_{ij} is the distance from calibration location i to location j, and γ is the kernel bandwidth parameter.

2.3.3 Model Selection and Goodness of Fit

A crucial criterion for statistical modelling is model parsimony, i.e. to maximize the statistical fit of the model whilst minimizing its complexity. In order to do this, the Bayesian Information Criterion (BIC) and Deviance Information Criterion (DIC) have been developed to reward fidelity of the model fit to the data whilst penalizing increasing model complexity.

Adding parameters in MLE can increase the likelihood as well as the complexity of models. The BIC, developed by Schwarz (1978) to penalize additional parameters, is formulated as follows:

$$\mathrm{BIC} = n \cdot \ln(\hat{\sigma}_e^2) + k \cdot \ln(n) \qquad 2.23$$

Where n is the number of data points, $\hat{\sigma}_e^2$ is the error variance and k is the number of free parameters to be estimated. Given alternative models, one with a lower BIC value is preferred.

The DIC is stemmed from the Akaike Information Criterion (AIC) and is useful in Bayesian model selection. In order to define the DIC, initially consider the following:

$$D(\theta)=-2\log[p(y|\theta)]+2\log[f(y)] \qquad 2.24$$

Where $D(\theta)$ is the deviance, *p(y|θ)* is the likelihood function for the observed outcomes *y* conditional on the parameters, and *f(y)* is some standardizing function of the data alone

then DIC is defined as the following:

$$\mathrm{DIC}=D(\bar{\theta})+2p_D=\bar{D}+p_D \qquad 2.25$$

Where $D(\bar{\theta})$ is the deviance evaluated at $\bar{\theta}$, the posterior means of the parameters of interest.

p_D is the effective number of parameters for the model. $\bar{D}$ is the posterior mean of the deviance *D(θ)* and a measure of how well the model fits the data; the larger $\bar{D}$ is, the worse the model fit. Again, given alternative models, the one with a lower DIC value is preferred.

2.3.4 Casualty Severity Models

There have been numerous studies that apply crash severity models to investigate the relationship between crash severity and risk factors including traffic conditions, driver and vehicle characteristics, and geometric features (Helai et al., 2008; Xie et al., 2009; Yamamoto and Shanker, 2004; Rifaat et al., 2011; Abdel-Aty and Abdelwahab, 2004; Quddus, et al., forthcoming; Ye and Lord, forthcoming). There is also a large body of literature on the methodological techniques (see Savolainen et al., 2011 for a thorough review).

As discussed earlier, statistical models that analyse crash severity rely on the nature of the dependent variable and other issues associated with the available data, such as the size of the sample. Savolainen et al. (2011) summarises the development of research and current ideas as it relates to the statistical analysis of crash severities, highlighting the strengths and weaknesses of each method and identifying areas for future research. The authors provide a summary of various statistical methods employed to study crash severities (see Table 1 in Savolainen et al., 2011). Discrete outcome models, such as the logit and probit models are most commonly used. In general, nominal and ordinal models are two main types of discrete outcome models that have been used for the analysis of crash severity.

Few studies have directly compared different crash severity models, and there is no consensus on which model is the best. There are though a number of critical issues in applying and developing a statistically defensible method to study crash severity, as discussed below.

(1) Spatial and temporal correlation. It is very likely that correlation exists among individuals involved in the same crash, at the same intersection, on the same roadway segment, and under the same weather condition. As presented in last section, various spatial methods can be applied to control for spatial correlation for crash frequency data. However, the model structure becomes more complex when discrete data are involved. It is critical to control for

temporal and spatial correlations to improve the precision of estimates and resulting inferences. For example, Helai et al. (2008) develop a Bayesian hierarchical binomial logistic model to account for within-crash correlations.

(2) Endogeneity. The endogeneity problem refers to the correlation between explanatory variables and the outcomes or unobserved heterogeneity. For example, motorcycle riders who wear helmets may also have good driving behaviour, which cannot be captured in the model. In this case, the effects of using helmet on preventing motorcycle fatalities can be overestimated. Incorrect inferences and evaluation results of interventions can be caused if such endogeneity is ignored. Certain researchers have made attempts to account for the endogeneity problem (Winston et al., 2006; Lee and Abdel-Aty, 2008; Paleti et al., 2010).

(3) Underreporting of crashes. Statistical models generally require that the selection of sample data from a population is random and that the probability of being sampled is equal for each crash. For example, although STATS 19 data provides a detailed source of accident data, it is prone to underreporting, especially for slight injury crashes. Compared to hospital records, the reporting rate to police is around 40-60% for slight injuries from 1996-2004 (DfT, 2006). The estimated reporting rates also vary across different road user groups, with the most underreported group, vehicle occupants, at around 50% (DfT, 2006). Reasons for the underreporting could be:

- People are unaware that injury accidents should be reported;
- Some people who are affected by drugs or alcohol do not want to report to police;
- The injury is not apparent.

Underreporting may lead to misleading conclusions when setting targets for accident prevention, particularly because it is far more likely to underreport less severe crashes than more severe crashes.

There are also other important issues that need to be accounted for carefully in crash severity analysis (see Savolainen et al, 2011, for more details). However, the focus of this research is on the effects of traffic interventions on road safety at an aggregate-level, where such discrete outcome models may be not applicable. Much additional work is needed in the future to develop causal models for evaluating effects on crash severities of traffic interventions at a disaggregate-level.

2.4 Summary and Areas for Contribution

Much effort has been devoted to understanding the factors affecting road accidents, however, inferences on the impact of traffic interventions on road safety are not conclusive due to the lack of formal causal approaches. It is valuable, therefore, to explore the safety outcomes following implementation of such traffic interventions, but the relationship between them is not straightforward and difficulties lie in establishing causal relationships due to the confounding effect of other variables. It is essential, therefore, to have a good knowledge of and control for factors that affect road safety.

In this chapter, studies on various factors affecting road accidents have been reviewed. This is a topic that has become the focus of substantial research during the last decade. Most previous studies have shown that a broad range of factors could affect road accidents. Many inferences have been drawn based on the exploration of traffic characteristics, including speed, density and flow. For example, some studies found increased speed is related to fewer accidents, while other studies found the opposite. A mixed relationship was also found between density and road safety in the literature, depending on the measurements of density and types of accidents. In terms of traffic congestion, most of earlier studies suggest that there is a negative relationship between traffic congestion and road accidents due to lower speeds in congested situations. Recent studies, however, found that congestion could increase accidents. There is also debate in that some of the factors related to road characteristics have a mixed effect on road safety. For example, some studies found road horizontal curvature to be negatively associated with road safety, while recent studies found it to be protective. It can be seen from the literature that traffic characteristics have mixed effects on road safety.

Because of the nature of accident data, GLMs are widely used to model the number of accidents. Compared to the linear regression model, the Poisson and NB regression can accommodate nonnegative discrete data very well. Although those GLMs with the assumption of NB error distribution can take account of uncorrelated heterogeneity, they may not be able to explain the effect of spatial correlation. To address this problem, spatial models, such as SEM and Bayesian hierarchical models have been developed. A considerable number of studies have investigated the relationship between crash severity and various factors. Most of these studies use disaggregated data, while this research is conducted at an aggregate-level. Further research on the effects of traffic interventions on road safety at a disaggregated-level is a priority for future research.

Previous research has generally suffered from certain methodological weaknesses. The most important of these is the lack of a well-developed method to distinguish the causal relationship from general association. Statistical models used to draw causal inferences are distinctly different from those used to draw association inferences. Association between two variables does not automatically imply that one causes the other. Most of the research on the effect of traffic interventions has analysed the effect of an intervention by simply conducting a before-after study. This has an underlying limitation in that it is impossible to distinguish the time trend and regional effects from the intervention. In addition, the absence of important control variables, such as demographic and economic characteristics, can also affect the accuracy of estimation. In light of these weaknesses in previous research in this area this thesis seeks to make a contribution to the research literature by developing a formal causal model. This is a topic that will be considered further in Chapter 4.

Chapter 3: Traffic Interventions and Road Traffic Safety

In this chapter, traffic interventions are defined and classified according to their primary purposes. Three types of interventions are discussed: traffic interventions affecting exposure risk, road plans and construction, and, road traffic legislation and enforcement. A review of the literature on these traffic interventions and road traffic safety is provided, showing the chains through which traffic interventions impact on road safety. Methodological issues in previous studies are also discussed in this chapter.

3.1 Introduction

Traffic interventions can be considerably varied in their nature and therefore a pre-requisite to the analysis of their effects is to define clearly what is meant by such an intervention. The traffic interventions discussed in this research are defined as policies, legislation and enforcement, new road construction or the change of an existing road network, and other general-purpose measures which directly or indirectly affect traffic conditions, drivers' behaviour and the travel environment.

The complex character of traffic interventions makes it difficult to generalize about their effects. Traffic interventions affect road safety by affecting traffic flow, the travel modal split, and other aspects. The effect on road safety is largely determined by the way in which these interventions are designed and implemented.

A considerable body of research has evaluated the effects of traffic interventions on the number of accidents and injuries (Hyatt et al., 2009; Quddus, 2008; Noland et al., 2008; Mountain et al., 2005; Shin et al., 2009;Gains et al., 2004, 2005; Jones et al., 2008). Most of these previous studies, however, are empirical in nature and the methodologies used to assess their effects have considerable drawbacks, with the consequence that the understanding of these effects remains uncertain. As mentioned above, these traffic interventions can be complicated and influence the occurrence and severity of traffic accidents indirectly. The relationship between traffic interventions and road traffic accidents is sometimes too indirect to be estimated. For this reason, conventional evaluation methods for road safety analysis may not be applicable. Furthermore, randomized controlled trials are not easy to conduct in such a way as to evaluate the effects of traffic interventions on road safety. The following chapter will discuss specific causal models for the analysis of the effects of traffic interventions on road safety. In the following sections, specific interventions will be discussed to demonstrate how road safety is affected by traffic interventions. A review of previous studies will also be provided.

3.2 Traffic Interventions and Road Safety

Three types of traffic interventions are considered in this study. The first type involves interventions affecting exposure to risk by adjusting traffic volume and travel modes.

As discussed earlier, traffic interventions usually affect road safety indirectly by influencing other factors, among which traffic volume is the most important factor affecting the accidents number. The study by Elvik and Vaa (2004) indicates that an increase of 100% in traffic volume will lead to an increase of 80% and 25% in the number of injury accidents and fatal accidents respectively, given that everything else remains unchanged. Therefore, it can be assumed that the number of road traffic accidents can be reduced by limiting the amount of traffic flow.

The risk of road traffic accidents varies among different travel modes. High risks are usually related to individual travel modes, such as cycling, walking and riding motorcycles. Passengers using public transport have a relatively low risk (Elvik, 2004). The number of injury accidents can be reduced by a transition from individual to public travel modes. The interventions that cause such a change in the travel modes may also affect the number of accidents.

Second, road construction and plans can also influence the accidents number by affecting the traffic distribution. More traffic can be induced on the road network by increasing road capacity in areas with capacity problems. New roads are usually safer than older roads (Elvik, 2004). Thus road construction and plans face a tradeoff between increased traffic and reduced accident rate per km driven.

Another intervention concerns road traffic legislation including a number of Acts of Parliament and Statutory Regulations. It is widely recognized that enhanced traffic legislation and enforcement is a cost-effective way to prevent road accidents and achieve improvement in road safety, especially when it is targeted at drink driving, speeding and non-use of seat belts (KfV, 2007). Different from other types of traffic intervention, road traffic legislation and enforcement that aims to prevent hazardous driving behaviors are expected to have a proven, direct relationship with road safety. It is also expected that drivers who have been convicted for numerous violations will have a higher accident rate per driver than drivers who have been convicted for few or no violations (West et al., 1992; Smiley et al., 1989; Evans and Wasielewski, 1983). The estimation of such effects may be biased, however, due to the regression to the mean effect since traffic legislation is usually enforced at locations and at times where violations are most likely to happen. Chapter 4 discusses the regression to the mean bias. Figure 3.1 below indicates the mechanism for reducing the number of accidents.

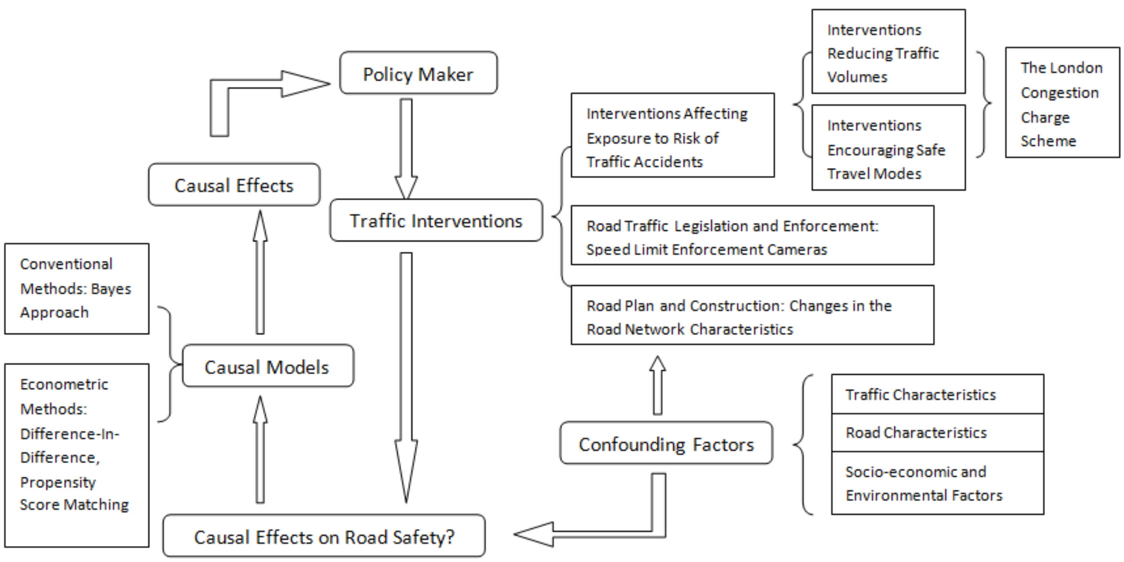

Figure 3.1 The framework of the causal analysis of effects of traffic interventions on road safety

3.3 Traffic Interventions Affecting Exposure to Risk of Road Traffic Accidents

As discussed earlier, the single most important factor affecting road safety is the traffic volume in both the short- and long-run. The greater the traffic on any given road, the greater the number of traffic accidents expected to occur, if all other conditions remain unchanged (Golob et al., 2003; Martin, 2002; Dixit et al., 2011; Lord et al., 2005). Golob et al. (2003) find that traffic flow conditions are highly related to the likelihood of traffic accidents. Their results show that road safety is affected by several key traffic flow elements, including the median speed and the mean volume, as well as the temporal variations in speed and volume. As described in chapter 2, Martin (2002) investigates the relationship between accident rates and hourly traffic volume, and examines the impact of traffic on accident severity. The result suggests that the rates of both injury accidents and property damage-only accidents are negatively related to the traffic volume.

The risk of road traffic accidents also varies considerably between different travel modes. Elvik (1996) shows that all forms of individual transport involve a higher risk of road traffic accidents than public transport. The risk of injury is particularly high for walking, cycling and riding mopeds and motorcycles. The risk of road traffic accidents of different travel modes have been investigated in a number of studies (White, 2004; Leigh and Wilkinson, 1991; Crandall and Graham, 1989). Table 3.1 (Elvik, 2004) shows the relative risk of injury for different travel modes in six different countries, estimated on the basis of injuries recorded in the official accident record in these countries and travel behavior surveys made in the same countries. The risk for a car driver is set equal to 1.00 and the risk of other travel modes is estimated as a ratio to the risk for a car driver.

Table 3.1 Relative risk of injury of different methods of transport in different countries (Elvik, 2004)

Relative risk of injury in different countries. Drivers' risk= 1.00						
Means of travel	**Norway**	**Denmark**	**Sweden**	**The Netherlands**	**Germany**	**Great Britain**
Pedestrian	**4.35**	**6.65**	**4.13**	**6.07**	**3.50**	**7.15**
Cyclist	**3.90**	**7.76**	**5.73**	**5.67**	**9.50**	**14.02**
Moped/mc	**8.30**	**29.94**	**17.87**	**197.60**	**31.25**	**20.26**
Car driver	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Car passenger	**0.75**	**1.94**	**0.87**	**1.13**	**1.50**	**1.25**
Bus	**0.25**	**0.12**	**0.13**	**0.20**	**0.13**	**0.59**
Tram	**0.60**		**0.87**	**0.02**	**0.25**	
Train	**0.05**	**0.04**	**0.13**	**0.02**	**0.05**	**0.22**

As shown in table 3.1, in all countries all types of public transport have a lower risk of injury than car drivers. However, car drivers have a lower risk of injury than pedestrians, cyclists and motor-cyclists. It can be speculated that the number of traffic injuries could be reduced if more trips were made using public transport other than private travel modes. It is worth noting, however, that not all the journeys can be done using public transport. Nevertheless, the number of traffic injuries as well as the property-damage only accidents can be reduced by encouraging people to choose travel modes, which have the lowest risk of road traffic accidents.

At the outset, it is important to note that there is no standard package of interventions suitable for all contexts and countries. Interventions proven to be effective in one setting may be not applicable elsewhere and will need careful adaption and evaluation. As discussed earlier, interventions can be developed from two aspects to manage exposure to the risk of road traffic accidents: reducing traffic volumes and encouraging the use of safer modes of transport. One key characteristic of such interventions is that the objectives are usually complicated. There may also be overlap between interventions for these two purposes. Improving road safety is not the only objective, and sometimes, it is not the most important objective of these interventions. For instance, one main objective of imposing taxation on car users is to reduce fuel consumption and air pollution, but it remains the case that road safety is directly or indirectly affected by these interventions.

3.3.1 Traffic volume reduction measures

Traffic volumes on roads can be influenced using a number of measures:

1. Road plans and road construction

Road plans and road construction, including new road construction and the improvement of existing roads, affect road capacity in a given area and the roads level. Both elements have an impact on traffic volume as well as the traffic density and speed. This is discussed in greater detail in section 3.4.

2. Regulating commercial transport

Commercial transport can represent a high risk on roads due to the time pressure often experienced by their drivers and the large vehicles used. The objectives of regulating commercial transport usually include reducing pollution and noise and improving safety on the roads (Elvik, 2004). The effect on road safety of the regulation of commercial transport has been studied (Evans, 1994; Elvik, 1997; Phillips and McCutchen, 1991), mostly from the perspective of the consequences of deregulation of commercial goods transport.

In general, the introduction of regulating commercial transport can influence the number of commercial transport vehicles and other motor vehicles involved in injury accidents. The deregulation of commercial transport is related to a slight increase in the number of injury crashes. There are two reasons for this. First, deregulation may cause an increase in traffic volume. Second, deregulation may boost the establishment of new companies. It has been shown that newly established companies experience a higher injury accident rate than older ones (Corsi and Fanara, 1989).

3. Road pricing

Road pricing refers to paying for the use of public roads, based on the extent of use and the costs this imposes on society. In effect, road users are charged for using motor vehicles and using a given road at a given time. Road pricing can impact on the traffic volume, the traffic distribution as well as the travel modes split.

It is worth noting that road pricing is not primarily intended as a road safety measure. Rather, it is regarded as a traffic control measure designed to either reduce or spread the peak hour traffic. The impact of road pricing on accidents is difficult to estimate due to contradictory factors. Usually, road pricing leads to less overall traffic and hence will reduce the number of accidents. There may be an increase in the number of accidents, however, due to an increase in the speed of the remaining traffic.

There are several forms of road pricing and studies have been conducted to investigate the effects of road pricing on not only traffic flow or speed, but also road safety.

(a) Toll roads

Toll roads charge users for the cost of road construction and maintenance without taxing on non-users. It has been introduced in many countries and currently in the UK tolls are only collected on a single major road, the M6 Toll, and a small number of tunnels and bridges.

(b) Gasoline prices and motor vehicle taxation

The gasoline price could be related to traffic accidents through a series of chains, as shown in figure 3.2.

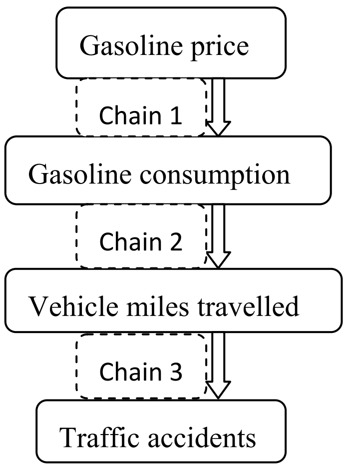

Figure 3.2 The relationship between the gasoline price and traffic accidents

The first chain is the relationship between the gasoline price and gasoline consumption. It has been shown that estimates of the elasticity of demand are in the range of -0.1 to -0.5in the short-run, and -0.6 to -1.2 in the long-run (Grabowski and Morrisey, 2004). These estimates suggest that a 10 percent increase in the price of gasoline leads to a 12 percent reduction in gasoline consumption in the long term. Shipping companies may use trains more and trucks less because of sustained higher gasoline prices. People may buy smaller and more fuel-efficient cars, or even turn to motorcycles. Employees may also want to move house to places less distant from their work.

The second chain is between gasoline consumption and vehicle miles travelled (VMT). It is easy to understand that a reduction in gasoline consumption can be explained as less VMT but this inference may not hold in the long run. For instance, people may switch from a sports utility vehicle (SUV) to a more fuel-efficient car. Hence, the relationship between gasoline consumption and VMT may be attenuated due to the conversion of vehicle types.

The last chain considers the relationship between the VMT and the number of accidents. The probability of being involved in a crash increases as people drive more. Dee (2001) conducted a state level study in the USA of the relationship between VMT and traffic casualties among younger drivers using data from 1982 to 1998. The result indicates that a 10 percent increase in VMT was associated with a 2 percent increase in teenage motor vehicle casualties.

Based on the above three chains, it may be hypothesized that an increase in the gasoline price is expected to reduce motor vehicle accidents. For example, higher gasoline price will reduce consumption, which in turn reduces VMT and ultimately reduces traffic casualties. However, in the long run, this relationship may become ambiguous due to people changing their travelling behaviours and vehicles. For instance, a more fuel efficient vehicle or other transport modes may be chosen because of the higher gasoline price and the long-term effect may be weakened. In contrast, people moving home to be near their workplace would enhance the long-term effect, because this would reduce VMT and consequently decrease the risk of a traffic accident.

To date, only a limited number of studies have investigated how road traffic safety is affected by a gasoline tax. Researchers have generally shown that there does seem to be a relationship between fuel prices and road casualties, although the direction of the relationship depends on the type of accidents.

Leigh and Wilkinson (1991) employed a random effects model to investigate how gasoline prices and taxes affected traffic casualties in the US between 1976 and 1980. They concluded that a 10 percent increase in the gasoline tax was associated with a decrease in the accident rate by 1.8 to 2.0 percent. Considering the fact that gasoline tax occupies only 16 percent of the total gasoline price, this effect is remarkable. Leigh et al. (2008) evaluated the effect of a 20 percent increase in gasoline prices on public health in the US. A simulation-based partial equilibrium model was applied and estimates on price elasticity of motor vehicle accidents were drawn from the literature. By controlling for other factors, the authors found 1994 fewer deaths from vehicle crashes.

Grabowski and Morrisey (2004) estimated the effect of gasoline prices on motor vehicle fatalities per capita and per VMT by applying a fixed effects model using 1983-2000 monthly panel data in the US. They used a time series variable to control for seasonal variation and national temporal trends. Cross-sectional differences were also accounted for. Their results suggest that motor vehicle fatalities increased by 2.3 percent when gasoline prices decreased by 10 percent. Grabowski and Morrisey (2006) further generated panel data on the total number of traffic fatalities for the 48 continental U.S. states for the period between1982 and 2000. Their results suggest that exogenous increases in state gasoline taxes were plausibly associated with fewer traffic fatalities.

Hyatt et al. (2009) investigated the relationship between motor vehicle injury and mortality rates and gasoline prices in the US between 1992 and2007. By using monthly gasoline price and fatality panel data, they found higher gasoline prices were related to increased motorcycle casualties. More individuals used motorcycles as their main commuting mode and traffic fatalities and injuries shift to motorcycles. However the rate of motorcycle-related casualties per vehicle remains stable despite the increase in gasoline prices. This means the increase in motorcycle casualties was more a factor of the increasing number of motorcyclists on the road during the period of the study.

Most of previous studies have examined only fatal accidents and ignore the total number of traffic accidents. Chi et al. (2009) analysed the effect of gasoline prices on traffic accidents, which are further discriminated by age and gender. They employed time geography theory to investigate the impact of gasoline price on total traffic accidents. Their results suggest that a negative relationship exists between gasoline prices and traffic accident rates in the short run.

In summary, gasoline prices can affect road safety through four factors: travel frequency and distance, commuting mode, driving behaviour and residential relocation. First, as gasoline prices rise, people may drive less and for shorter distance, as well as seeking to make trips more efficient. Second, people would choose more fuel-efficient vehicles, which are lighter and more vulnerable if involved in an accident (White, 2002). However, this effect has been found to be insignificant when other safety measures are implemented at the same time (Leigh and Wilkinson, 1991). Other commuting modes, such as bicycle, motorcycle, walking and public transportation may be chosen instead of private driving. Third, higher gasoline prices could induce workers who live far from workplaces to relocate their place of residence. Fourth, people tend to drive more slowly and avoid sudden acceleration and deceleration, which make trips safer. All of these factors could decrease the distance travelled and in turn reduce the likelihood of traffic accidents. Therefore, higher gasoline prices are expected to be negatively related to road accidents.

Turning to vehicle taxation, there is no conclusive evidence regarding the effect of changes in vehicle taxation on accidents. Vehicle taxation may affect the number of accidents indirectly by affecting travel demand. It is possible that the number of traffic accidents can be affected by taxing the purchase, use and ownership of motor vehicles as follows. First, the travel demand can be changed by influencing the number of vehicles purchased and affecting driving distances. Second, by imposing high taxes, there are fewer vehicles with high risk of road traffic accidents on roads. Finally, the use of safety equipment on roads can be increased by the use of tax rebates for such equipment.

One study by Fridstrom and Rand (1993) investigated the impacts of different changes in vehicle taxation in Norway on the total number of trips. The results indicate that vehicle taxation has an impact on travel demand. The number of person kilometers driven by car would increase by 25-30 per cent, if vehicle taxes were abolished. Because there would be more individual cars, the number of vehicle kilometers travelled would increase by 35-40 per cent, which leads to an increase in the number of injury accidents by 25-30 percent.

Another study by Pirdavani et al. (2012) evaluates the road safety effect of increasing the fuel price by 20% in Flanders, Belgium. They find significant effects of increasing the fuel cost on reducing driving distances. A 20% increase in fuel price is expected to reduce the annual vehicle kilometers travelled by 5.02 billion, which leads to a reduction in total number of accidents by 2.83 per cent.

(c) Congestion charging

Congestion pricing or congestion charging is a system of surcharging road users to regulate traffic demand, making it possible to manage congestion without increasing supply. The congestion charging scheme is currently applied in a small number of cities including London, Singapore, Milan and Stockholm. Congestion charge can be classified into four different types: area wide congestion pricing, which charges for being inside an area; a cordon area around a city center, which charges for passing the cordon line; corridor or single facility congestion pricing, where access to a lane or a facility is priced; and a city center toll ring, with toll collection surrounding the city.

It has often been debated whether congestion charging is a regressive or a progressive measure. The answer relies on how costs and benefits are balanced. Eliasson and Mattsson (2006) investigated the equity effects of congestion pricing, showing that when assessing the impacts of congestion charging, it is important to take into account the distribution of population, the transport modes they use and the revenues allocated back to them.

Congestion charging has been recognized as a feasible means of solving congestion and environmental problems in an urban area. By increasing the generalized cost of a trip by car, travel demand, especially for the single-occupancy private vehicles, is reduced and hence the congestion problem is relieved. The inferences on overall effects of a congestion charge are not conclusive, however. Research has been undertaken to investigate the impacts of congestion charging on congestion, traffic levels, public transport, environment and business activities. Olszewski and Xie (2005) introduced the experience of the Singapore Electronic Road Pricing (ERP) system, which charges for entering an urban central area. They obtained demand elasticity values from traffic counts before and after the revisions of the charging rate and found the value for cars is higher than other vehicles, which means car drivers are more sensitive to the rate changes and more likely to switch to other travel modes. The successful experience of the Singapore ERP system suggests that road pricing can effectively control traffic demand and reduce congestion. This study is limited, however, because only data for the charged periods was used. It is difficult to understand the traffic level and drivers' travel behaviour beyond charging hours.

Research has been done on the effects of the London congestion charging scheme on road pricing, especially on whether congestion charging has an impact on traffic levels, travel behaviour and traffic patterns (e.g. Tuerk and Graham, 2010; Eliasson and Mattsson, 2006; Wichiensin et al., 2007). Recently, the safety effects of the London congestion charge have also been studied (Quddus, 2008; Noland et al., 2008; Li et al., 2012) and will be further discussed in chapter 5.

In summary, previous studies have shown that road pricing can affect traffic demand and the travel behaviour of road users, and hence have an indirect effect on road accidents. The lack of appropriate causal methods, however, makes the inferences drawn by these studies very uncertain. In this research, causal models will be applied to study the effects of the London congestion charge on road accidents.

All the above interventions control the exposure to the risk of road injury accidents by managing traffic demand. Various policy interventions are applied in traffic demand management and affect traffic condition by changing travel behaviour. Economic strategies or policies, such as road pricing and taxation, are usually applied to reduce traffic demand, or to redistribute the demand in space or in time. In doing so, it is possible to reduce the exposure to the risk of road traffic accidents. Despite that the primary objective for implementing such demand management interventions is not to reduce road accidents, their effects on road safety should not be neglected.

3.3.2 Encouraging use of safer modes of transport

As shown in Table 3.1, individual transport, e.g. pedestrians and cyclists, involves a higher risk of injury accident than travel by public transport. Changes in travel from private vehicles to the use of public transport can lead to a reduction in the number of private vehicles on roads. In doing so, the traffic volume can be effectively limited in areas where the infrastructure enables the provision of good public transport. Therefore the use of public transport in conjunction with safe walking and cycling is encouraged to improve road safety.

There are several possible strategies that may encourage the use of public transport:

1. Improving a mass transit system. The improvement rests on shorter walking distances between stops, more routes and less ticketing procedures, and safer vehicle and waiting areas (Ibrahim, 2003; Walle and Steenberghen, 2006; Amadori and Bonino, 2012; dell'Olio et al., 2011). dell'Olio et al. (2011) analyzed the quality of service expected by public transport users. The results show that the public transport factors that users most valued are waiting time, comfort and cleanliness. The authors also state that the more important factors when analyzing the desired quality from potential public transport users are waiting time, travel time and level of occupancy.
2. Better coordination between different modes of travel. This includes allowing bicycles to be carried on board trains, buses and ferries, and improving "park and ride" facilities, so that drivers can park their cars near public transport stations. It also requires a good knowledge of transfer behavior within the public transport system. Public transport systems could benefit from the improvement to the transfer experience. Guo and Wilson (2010) proposed a new method to evaluate different transfer modes. This approach was employed to analyze the London underground, which is one of the largest and most complex public transport systems in the world. This study confirmed that transfers without good coordination between different modes of travel can impose a large amount of cost

on the public transport systems. The cost can be significantly reduced by better design and plan of the transfer service.

3. Lower public transport fares and higher costs for using individual cars can encourage the use of public transport. There might be less usage of cars and more demand for public transport with an increase in car ownership costs or car usage costs, such as vehicle taxation. The usage of cars can be reduced by increased journey time. However this does not lead to significantly more public transport. Similarly, there might be less demand for public transport and more usage of cars with increased fares on public transport and longer journey time using public transport. Increasing the frequency of public transport can also increase the use of public transport and reduce car traffic.

There are only a few studies on the safety impacts of interventions which aim to manage the traffic demand and the modal split of transport (Leigh and Wilkinson, 1991; Haughton and Sarkar, 1996; Grabowski and Morrisey, 2004; Noland et al., 2008). These studies suffer from a number of methodological issues and consequently the knowledge of these effects is ambiguous. This will be discussed in detail in Chapter 5.

3.4 Road Plans and Road Construction

Safe road infrastructure rests on sound land use and road network planning. The distances between housing and work places and other locations of daily lives need to be taken into account. It is also crucial to ensure that the fastest travel route is also the safest. In other words, the travel distance on the more dangerous lower order roads is limited in favour of the safer higher order roads. It is difficult to design or reconstruct road network which meets the above conditions, especially when the existing network has been in use for a long time.

A good knowledge of the impact of road design on road safety is essential for road network planning. There are several important issues in formal road network planning.

1. Classifying roads by their function

The speed, traffic volume and level of safety measures vary between roads with different functions. Conflicts often occur between users of motor vehicles and pedestrians and cyclists, especially in residential and urban areas.

Road hierarchy is important for providing safer roads and safer design. The road classification considers location of accidents sites, land use, vehicle and pedestrian flows, and objectives such as speed control. Current road classification needs to be reconsidered to improve the existing road network. To ensure that a road reflects its true function, the number of road categories is limited and multi-functional roads should be avoided.

Usually, the main road network needs a development and reconstruction when there is a capacity problem, especially in large cities. The question is whether the increase in the VMT induced by new roads offsets the benefit of safer higher order roads. The traffic volume induced is associated with several factors, such as the size of increased road capacity, previous road categories, the time saved by developing new roads.etc.

2. Street network design

Debate continues among urban designers and transportation engineers regarding the optimal type of street pattern. This is a particularly important issue for new communities and developing communities. In the past, urban streets were constructed to carry people and goods in a safe, fast and reliable way. Because the street with grid pattern dealt with these requirements very well, it had been adopted in many urban areas for a long history (Southworth and Parthasarathy, 1996). However, street patterns with limited access, such as lollipop and loop, had become the predominant street pattern for developing communities to improve safety level and reduce congestion in suburban area.

While the social benefits and drawbacks of different types of urban forms and street patterns have been investigated in previous studies (Talen, 2006; Camagni et al., 2002), little effort has been devoted so far to the safety evaluation of different street designs. Another issue which has not been fully addressed in most previous studies relates to the safety implications of the overall community design. For example, street widening projects, usually applied for improving safety and relieving congestion, have been shown to be negatively related to road safety (Dumbaugh, 2006; Noland, 2000; Huang et al., 2002; Sawalha and Sayed, 2001; Swift et al., 2006). This is probably because too much attention has been paid to assessing how the changes affect individual road segments rather than how these changes might affect the community as a whole. The same applies when we attempt to minimize the opportunities for through traffic to improve safety on residential streets. One fact which is overlooked in this case is that limiting street connectivity in residential neighbourhoods can impact safety elsewhere (Ewing and Dumbaugh, 2009; Ewing et al., 2002).

In chapter 7, a panel analysis of ten years of UK road network data is conducted to investigate how changes in the characteristics of the road network impact on road safety outcomes.

3. Traffic calming measures

Traffic calming measures are designed for slowing down or reducing traffic volume. The objective of such measures is to improve safety for pedestrians and cyclists as well as to improve the living conditions for residents living along the road.

Traffic calming applies techniques which discourage traffic from entering certain areas (e.g. residential area) and install physical infrastructures for reducing speed, which include:

(a) Narrowing of streets.
(b) Speed breakers.
(c) Chicanes and raised pedestrian crossings.
(d) Block or restrict access.
(e) Giving priority to pedestrians and bicyclists.

A number of studies have evaluated the safety impacts of these traffic calming strategies (Pau and Angius, 2001; Galante et al., 2010; Allpress and Leland Jr, 2009; Elvik, 2001). For instance, Allpress and Leland Jr (2009) evaluate two novel interventions designed for controlling traffic speed within an open road where drivers were required to decrease their speed. Two different measures were implemented at the entrance to the road and required drivers to pass between a 3.5 m wide passage of either evenly or decreasingly spaced cones. The findings indicate that both arrangements of cones are effective, convenient, and cost-effective strategies. Another meta-analysis of 33 studies by Elvik (2001) evaluates the impacts of area wide traffic calming interventions on road safety in urban residential areas. Area wide traffic calming measures are particularly designed to discourage non-local traffic from using residential streets and reducing the speed of the remaining traffic. In general, evaluation studies show that area wide traffic calming leads to a reduction of the accidents number by around 15% in the whole treated area.

These traffic calming measures are often implemented with speed limits of 30 km/h, but they can be designed to achieve various levels of appropriate speed. Without securing compliance with speed limits and other road safety rules, engineering measures alone will not produce satisfactory results. Thus, various road safety laws and enforcement, including speed limit enforcement, have been used widely to prevent dangerous driving behaviour.

3.5 Road Traffic Legislation and Enforcement

Road user behavior is closely related to road safety. In order to make road user behavior as predictable and safe as possible, governments issue rules to regulate traffic behavior. It is widely recognized that traffic legislation and enforcement, particularly when it is implemented for preventing drink driving, speeding, and the non-use of seat belts, is a very important and effective method to improve road safety in a short period of time.

Traffic violations can be prevented by increasing the objective and subjective chance of being caught. The objective chance of being caught is determined by the density of actual police controls along any given road segment. The subjective chance of being caught is estimated by drivers based on the objective chance and information from media, such as television and newspaper, as well as from colleagues and friends. Drivers will avoid traffic offences, when they see the chance as being sufficiently high.

To increase the effectiveness of traffic legislation and enforcement, it is important that traffic legislation and enforcement focus on traffic offences that have a proven, direct relationship with road safety, such as drink driving, non-use of seat belts and speeding, and at locations and times where traffic offences have the most impact on road safety. To increase the credibility and acceptance of traffic legislation, it is also important to avoid the impression that traffic legislation is implemented to make profit for governments. Ideally, the income generated from fines of traffic offences should be used in activities for improving road safety. It is also recommended that regular feedback should be made to show the positive effects of traffic legislation to the general public.

In this section, traffic legislation and enforcement targeted to speeding, drink driving and the non-use of seat belts are introduced.

1. Speeding

As outlined in chapter 2, there is a clear relationship between the vehicle speed and the occurrence and severity of road accidents (Baruya, 1998; Ossiander and Cummings, 2002; Taylor et al., 2002; Aljanahi et al., 1999).High vehicle speed and large speed differences between vehicles make driving situations difficult to predict and control. Higher speed is assumed to lead to less reaction time, and more severe consequences when a crash takes place. Therefore the occurrence and severity of traffic injuries can be lessened by controlling vehicle speed.

Various methods can be applied to enforce speed limit compliance. Auto speed enforcement is considered to be the most effective measure due to the high enforcement density, and hence the high objective chance of being caught. Efficiency is further enhanced if the finepayment is also automated. Fixed and mobile speed cameras are widely applied as automatic speed enforcement in many countries.

In the UK, the safety camera programme is run by local partnerships. Strict guidelines are established to determine where to put the cameras based on historical accident numbers and the prevalence of speeding. The cameras are clearly signed so drivers can see them in advance. The income generated from fines is used to develop the programme as well as other road safety activities. A pilot scheme was started in 2000 with eight partnerships. There are 38 partnerships and more than 4000 camera sites involved in this programme by the end of 2004.

There are a number of studies on the safety effectiveness of speed enforcement cameras in the UK (Goldenbeld and van Schagen, 2005; Hess and Polak, 2003; Newstead and Cameron, 2003; Chen et al., 2002; Christie et al., 2003; ARRB Group Project Team, 2005; Mountain et al., 2004; Cunningham et al., 2008; Mountain et al., 2005; Shin et al., 2009; Keall et al., 2001; Gains et al., 2004, 2005; Jones et al., 2008). These studies will be reviewed in detail in Chapter 6. In general, the results show a significant reduction in both speed limit violations and crash numbers at camera sites.

There is still debate about the effectiveness of speed cameras, however, especially in relation to the issue of regression to the mean. Furthermore, the assignment mechanism behind the speed cameras programme remains unclear. A series of issues related to speed cameras have never been touched before, such as how effectiveness varies spatially and temporally and under what conditions speed cameras perform most effectively. To date there has been no independent study using advanced state-of-the-art causal methodologies to answer these questions. In this research, we apply both conventional methods and formal causal models to evaluate the safety effectiveness of the speed cameras in the UK. The discussion is presented in chapter 6.

2. Drink driving

The risk of a road crash, as well as the severity of the injuries that result from crashes can be influenced by alcohol impairment. The frequency of drink driving as well as the probability of being caught varies between countries. Despite that drink-driving offences are much less common than speeding violations, it is still a vital risk factor for road traffic accidents.

A summary of measures that can be used to prevent drink driving is outlined below:

(a) Set blood alcohol limits. The limits should be consistent with current epidemiology information concerning the relationship between alcohol and crash involvement. Upper limits of 0.05 g/dl for the general driving population and 0.02

g/dl for young drivers are generally considered to be the best practice at present (European Commission, 2001).

(b) Enact drink-driving laws. This includes a series of laws: those that establish a lower legal limit for blood alcohol content for inexperienced or younger drivers than for more experienced, older drivers; laws that specify minimum legal drinking age, an age below which the public consumption or purchase of alcohol drinks is illegal; laws that require installation of "alcohol ignition interlocks", an equipment which requires a driver to take a breath test before starting a car.

(c) Enforcement of drink driving laws: Random breath testing is widely used as an effective method for drink driving enforcement by the police. Should drivers exceed the blood alcohol limit, then legal prosecution, suspension of the driving license and, in extreme cases, imprisonment, are applied.

(d) Implement a graduated driver-licensing system for new drivers which sets a period during which restrictions are placed on any unsupervised driving. These restrictions should include a prohibition against driving after drinking any alcohol.

3. Seatbelts and child restraints

Seatbelts can significantly reduce the severity of injuries when traffic accidents take place. They are more effective, however, at preventing fatalities rather than severe injury since a fatal crash is closely related to head injury and internal torso injury that seatbelts can effectively prevent. The effect of seatbelts is partly dependent on the collision speed, e.g. the effectiveness is greater at lower speeds. This highlights the importance of wearing seatbelts on urban roads where traffic speed is relatively low.

Measures that can be employed to improve the use of seatbelts and child restraints include:

(a) Making the use of seatbelts and child restraints mandatory by law. These laws then need to be strictly enforced with associated public information and awareness campaigns.

(b) Encourage primary enforcement where a driver can be stopped solely for not wearing a seatbelt. This is more effective than secondary enforcement where a driver can only be stopped if another offence has been committed. For example, Houston and Richardson Jr (2002) explore how the change from an existing secondary seat belt law to primary enforcement affects road safety by examining traffic casualties in California from 1988 to 1997. The results show a significant reduction in traffic injuries due to the implementation of primary enforcement. However the impact on fatalities is not significant.

(c) Encourage the use of the appropriate type of child restraint. It is important to consider the age and weight of the child when choosing the type of restraint. Child restraints should be placed correctly. For example, it could be dangerous to place child seats in front of air bags. Ekman et al. (2001) examine long-term effects of enforcement and promotion of child restraint use in motor vehicles in Sweden. Fatality data indicates a reduction of 2.8% in average year, and 76% over the period from 1970 to 1996. It is also suggested that a much better improvement is made in areas where local authorities implemented early with safety measures, such as safety belt loan schemes and those having an organized safety-promotion program.

Road traffic enforcement and legislation aim to prevent traffic accidents by prohibiting particularly dangerous behavior and regulating driving behavior so that it becomes predictable and homogeneous. Traffic offences can be penalized by fines, traffic tickets or imprisonment, and the withdrawal of driving licenses. Traffic safety would be improved with better respect for traffic enforcement.

Despite the wealth of empirical evidence, there is still debate about the effectiveness of road safety enforcement and legislation. First, the conventional methods used in previous research have not fully addressed issues such as regression to the mean and confounding factors. Another issue is in relation to the cost-effectiveness of road safety enforcements. Only very few studies refer to the economic assessment, or attempt to assess potential improvements in cost-effectiveness. In this research, the focus is on the first issue; the second issue poses an interesting question for future research.

3.6 Concluding Remarks

In this research, traffic interventions are defined as policies and laws aiming to improve road safety levels, and other general-purpose measures which directly or indirectly affect traffic condition, drivers' behaviour and the travel environment. In this chapter, three types of traffic interventions are discussed: interventions affecting the exposure to risk of traffic accidents, road planning and road construction, and road traffic enforcement and legislation. We show the chains through which road safety is related to these interventions. While a lot of research has investigated the relationships between road safety and various factors, such as traffic flow, road characteristics, and demographic characteristics, there is not yet a good understanding of the causal link between traffic interventions and road safety.

There are two main reasons for the difficulty in evaluating this causal effect. As discussed earlier, the objectives of most interventions are very complicated. Improving road safety is not the only objective, and in many cases, it is not the most important objective of these measures. For example, the London congestion charge scheme was introduced to control the traffic flow entering the city centre and reduce traffic congestion. Fuel taxation, meanwhile, is usually used as an effective means to managing individual travel demand. Such interventions, however, may have an indirect effect on road safety by managing traffic demand and the split of travel modes. The causal links between such interventions and road safety is not straightforward, however.

Furthermore, the evaluation results obtained in previous studies are uncertain due to the lack of formal causal models. There are currently two main approaches for causal analysis in road safety literature, the before-after control methods and empirical Bayes methods. The before-after control methods, which usually simply use GLM, may suffer from the problems of regression to the mean and confounding factors. The empirical Bayes approach is advocated for its ability to control for the regression to the mean effect. However, there are also restraints with this approach. First, it requires a large sample to enable a reference group which is similar to the treatment group. Second, in previous research, not only is there insufficient justification of the selection of control groups, but how the treatment and control groups are matched is also unclear. In the light of these problems, formal causal models can be applied to quantify the effects on road safety of traffic interventions. In the next chapter, techniques for causal inferences are discussed and models which can be applied to road casualty analysis are developed. The applications of such models are presented later through three case studies: the London congestion charge, speed limit enforcement cameras, and road network design.

Chapter 4: Econometric and Statistical Modelling for Causal Inference

In the past two decades, numerous studies have been conducted to assess the econometric and statistical analysis of the evaluation of interventions or programmes. This has become an important tool and has been applied in many areas, such as labour economics, public health and industrial investment. Despite the superiority of causal modelling in dealing with issues such as confounding factors, such techniques have not been applied in road traffic accident analysis, especially the evaluation of traffic interventions. In this thesis, causal models will be applied to the traffic interventions discussed in chapter 3, in the following chapters. The objectives of this chapter are to provide both an historical review of the literature on techniques for causal inferences, and to develop appropriate techniques that can be applied to road traffic accident analysis.

4.1 Introduction

At the outset, the difference between correlation and causation needs to be clarified. "That correlation is not causation is perhaps the first thing that must be said" (Barnard 1982, p.387), i.e. statistical models used to draw causal inferences are distinctly different from those used to draw associational inferences.

Correlation, which can also be called association, is a relationship between two or more variables. Causation, on the other hand, implies that the change in one variable directly causes changes in another. In other words, causal relationships from one variable, A, to another, B, cannot be obtained only from the observed association between them. The reason for this is that the observed association between A and B could be reverse causation (B causes A), bidirectional causation (A causes B and B causes A) or the confounding effect of a third variable, C, even if A and B have no causal relationship.

The aim of this research is to evaluate the causal effects of traffic interventions on road traffic accidents and any attendant casualties. The units exposed to interventions are usually individuals, intersections, roads, cities, counties, etc. The interest in any evaluation is a comparison of the safety outcomes, usually the number of casualties, for the same unit with and without the intervention.

The literature on the evaluation of causal effects can be reviewed in both econometrics and statistics. Applications in econometrics can be traced to the evaluation of labour market programmes by Ashenfelter (1978) and Ashenfelter and Card (1985), while the statistics literature dates from the analysis of randomized experiments by Fisher (1925) and Neyman (1923). Most of the early theory focused on the use of conventional methods, such as fixed effect methods using panel data and instrumental variables methods, which deal with the problem of endogeneity. Later, Rubin (1973a, b, 1974, 1977, 1978) developed the Rubin Causal Model (RCM), which became the dominant approach to causal analysis in both the statistics and econometrics literature. In this model, Rubin proposed an idea called potential outcomes, which are defined as outcomes for the same unit given different levels of exposure to the treatment. The advantage of using potential outcomes rests on the allowance for general heterogeneity in the effects of the intervention. Another attraction of the potential outcome setup is that the parameters of interest can be defined. The RCM is discussed in the following section.

It is worth noting that there are two scenarios under which causal analyses are conducted. The most straightforward case is randomized experiments, where the treatment is assigned randomly and is independent of the potential outcomes and covariates. In randomized experiments, the average treatment effect is simply the difference in outcomes between the treatment and control groups. Many randomized experiments have been conducted in the area of economics in recent years (Banerjee, Duflo, Cole and Linden, 2007; Angrist, Bettinger and Kremer, 2005; Duflo, 2001; Miguel and Kremer, 2004). The popularity of randomized experimental evaluations, however, is restricted due to ethical issues and the high cost of implementation.

It is more common that evaluations are conducted with observational data, i.e. on an empirical basis. Because using observational data poses challenges in terms of estimating causal effects, it is necessary to introduce the assumption of unconfoundedness. This assumption is of great importance, since it can serve to adjust differences in observed pre-treatment variables between treatment and control groups. The most well-known methods for estimating causal effects with the unconfoundedness assumption are matching methods, which make comparisons in pairs of the matched treatment and control units. The propensity score matching (PSM) method will be introduced later in this chapter. What makes the PSM method attractive is that it requires the treatment assignment mechanism to be modelled (i.e. the conditional density of assignment to the treatment given pre-treatment characteristics) and in so doing gives clear criteria based evidence for the selection of units into either treatment or control groups.

Besides randomized experiments and observational data with the unconfoundedness assumption, there are various methods for special cases. These mostly rely on the availability of additional data in a specific form, such as panel data. In this thesis, an approach called difference-in-difference, which has been widely applied in econometrics, will be introduced.

Besides the RCM there are also other alternative methods which can be used for testing and estimating causal relations, such as structural equation model (SEM). SEM is a statistical approach using statistical data and quantitative causal assumptions. The SEM and RCM are closely related to each other (see Pearl, 2000 for detailed discussion). This thesis will only focus on the RCM.

Finally, it is worth introducing another widely used technique for causal analysis, the Bayesian approaches. Compared to econometric causal models, which are based on the potential outcomes framework, the Bayesian approaches usually make causal inferences using realized outcomes. The use of empirical Bayes (EB) methods in before-after evaluation studies has become very popular due to their ability to cope with key issues such as the regression to the mean bias and time trend effects. With the advances in statistics and the availability of a software package in WinBUGS, the full Bayes (FB) method has become increasingly popular. The most important advantage of the FB method is that it requires less data and can accommodate complex posterior distributions of outcomes. In this research, the results from Bayesian approaches will be compared with those from formal causal models. Discussion of the advantages and limitations of both methods will be provided.

This thesis will also focus on the practical issues raised by the implementation of these methods. Binary treatments are the main object of study, although multi-valued and continuous treatments will also be discussed. The applications of these causal models will be shown in the following chapters.

The remainder of this chapter is organized as follows. In section 4.2 the framework of the Rubin potential outcomes is introduced. The propensity score matching methods are discussed in section 4.3, followed by a detailed description of the difference-in-difference methods in section 4.4. Discussions of the Bayesian approaches are presented in section 4.5. Conclusions are provided in the final section.

4.2 The Potential Outcomes Framework: the Rubin Causal Model

It has already been noted that the potential outcomes framework, based on the Rubin causal model (RCM), is the dominant approach in modern causal analysis and programme evaluation.

4.2.1 Notations and Estimators

In presenting the RCM, it is necessary to introduce relevant notation. D_i is an indicator of treatment enrolment for individual or unit i. To facilitate understanding, consider only binary treatments.

$$D_i = \begin{cases} 1, \text{if unit i recieved the treatment} \\ 0, \text{other wise} \end{cases}$$

Let $Y_i(D_i)$ be the potential outcomes for individual i. Therefore, $Y_i(0)$ denotes the level of outcome that individual i would attain if not exposed to the treatment. Likewise, $Y_i(1)$ denotes the level of outcome that individual i would attain if exposed to the treatment. Since individual i can be either treated or not, we can only observe one of these two potential outcomes. If individual i enrols in the treatment, $Y_i(1)$ will be realized and $Y_i(0)$ will be the counterfactual outcome and vice versa. The potential outcomes can be described as:

$$Y_i(D_i)=D_i*Y_i(1)+(1-D_i)*Y_i(0) \qquad \textit{4.1}$$

The advantages of the potential outcomes approach are that it simplifies the modelling by separating the potential outcomes from the assignment mechanism, and additionally it does not require a specific regression function for estimating causal effects. For example, in terms of the realized outcomes, a regression function is usually written as:

$$Y_i=\alpha+\beta\mathbf{X_i}+\gamma D_i+\varepsilon_i \qquad \textit{4.2}$$

Where γ is interpreted as the causal effect of interventions. It is not clear, however, whether this effect is constant and this approach makes an independent assumption between D_i and ε_i. This assumption then bundles a number of conditions, such as the exogeneity assumptions and a well-structured function form. In contrast, the potential outcomes approach can define and make inferences about the causal effect without considering these assumptions and the specific form of the regression function.

Two popular estimators have been studied in previous research: the average treatment effect (ATE) and the average treatment effect on the treated (ATET). These two effects can be described as:

γ_{ATE}= E[Y(1)-Y(0)], γ_{ATET}= E[Y(1)-Y(0)|D=1].

In practice, control groups are usually selected from untreated units to construct counterfactual outcomes. Comparisons of the average outcomes between treated and control units, however, do not usually give an unbiased estimation of ATE or ATET:

$$E[Y(1)|D=1]-E[Y(0)|D=0]=E[Y(1)-Y(0)|D=1]+\{E[Y(0)|D=1]-E[Y(0)|D=0]\} \quad 4.3$$

In the above equation, the term in curly brackets is not zero for most cases due to selection bias, i.e. the treatment assignment is usually associated with the potential outcomes that individuals could attain, with or without being exposed to the treatment. The treatment assignment mechanism and selection bias are discussed in the next section.

4.2.2 Treatment Assignment Mechanisms

The treatment assignment can be discussed in terms of the probability of individuals being selected in the treatment, which is associated with observed covariates and potential outcomes. Three types of assignment mechanisms are introduced: randomized experiments, unconfounded assignment and other assignment mechanisms.

The first type of assignment mechanisms is randomized experiments, where the probability of assignment to treatment does not depend on potential outcomes.

$(Y_i(1), Y_i(0)) \perp\!\!\!\perp D_i$

Then *E[Y(0) |D=1]= E[Y(0) |D=0]* and therefore

$$E[Y(1)|D=1]-E[Y(0)|D=0]=E[Y(1)-Y(0)|D=1]+\{E[Y(0)|D=1]-E[Y(0)|D=0]\}= E[Y(1)-Y(0)|D=1] \qquad \textit{4.4}$$

Equation 4.4 is an unbiased estimator of ATET. Randomized experiments can also be conducted with a finite number of strata, known as a stratified experiment. Randomized experiments are straightforward and allow the greatest reliability and validity of statistical estimates of causal effects. Whilst they are a valuable tool for treatment evaluation, it is not always feasible to implement a randomized experiment due to high costs and ethical issues.

The second class of assignment mechanisms, unconfounded assignment, still needs the requirement that the probability of receiving the treatment is dependent on the potential outcomes, however, conditional on covariates $\boldsymbol{X_i}$ (Rosenbaum and Rubin, 1983), i.e.:

$(Y_i(1), Y_i(0)) \perp\!\!\!\perp D_i|\mathbf{X_i}$

Unconfounded assignment may also be referred to as the selection on observables (Heckman et al., 1997), exogenous (Manski et al., 1992), and conditional independence (Lechner, 2001). Unconfounded assignments, however, are not as straightforward as randomized experiments. Numerous studies have been conducted based on this assignment and the most prominent approaches are matching methods. The propensity score matching method is discussed in section 4.3.

The third class of assignment mechanisms is the one with a certain amount of dependence on potential outcomes. As discussed previously, the estimator of causal effect could be biased under this assignment mechanism. There are a number of methods which can deal with this problem, however, including instrumental variables, regression discontinuity and difference-in-difference. In section 4.4, the difference in-difference is discussed in detail.

4.2.3 Other Important Issues

Most previous research has focussed on average treatment effects, although it is also possible to assess the quantile treatment effects. The whole marginal distributions of Y(0) and Y(1) can be defined as:

$F_{Y(0)}(y) = P(Y(0) \leq y) = P(Y(0) \leq y|D=0) = P(Y \leq y|D=0)$ *4.5*

$F_{Y(1)}(y) = P(Y(1) \leq y) = P(Y(1) \leq y|D=1) = P(Y \leq y|D=1)$ *4.6*

Doksum (1974) and Lehman (1974) define:

$\gamma_q = Q_\theta(Y(1)) - Q_\theta(Y(0))$ *4.7*

as the θ-th quantile treatment effect, where θ is a quantile index between 0 and 1. Note that γ_q is the difference between quantiles of the two marginal potential outcome distributions, which does not identify the quantiles of the individual level effect:

$\gamma'_q = Q_\theta(Y(1)-Y(0))$ *4.8*

In other words, the difference of quantile is not the quantile of difference.

Before introducing the advantage of employing the idea of quantile treatment effect, another important issue needs to be highlighted: hypothesis testing in causal analysis. Most null hypotheses made in evaluation studies are that the average effect of interest is zero. This hypothesis is widely used because the average treatment effect is asymptotically normally distributed with zero asymptotic bias, and it can be tested using standard confidence intervals.

Even if the average effect is zero, however, it is interesting to investigate whether there is any individual affected by the treatment. It is particularly important for policy makers to understand the distribution of treatment effects, not simply the mean. If the policy makers know the distribution of treatment effects given information on covariates, they can redefine the unit of interest and hence increase the cost-effectiveness of the treatment. A limited number of studies have analysed the quantile treatment effect and heterogeneity in the treatment effect (Bitler et al., 2002; Abadie et al., 2002; Crump et al., 2008; Manski, 2004). This problem is out of the scope of this research, however, although it is one that would reward future investigation.

Another issue relates to the interaction between treated and untreated individuals. It is assumed that treatments applied to one unit have no effect on another unit. This assumption is referred to as the Stable-Unit-Treatment-Value-Assumption (Rubin, 1978). The focus of this thesis is on cases where this assumption is maintained.

4.3 Unconfounded Assignment: Propensity Score Matching

In this section, we first introduce the conditions under which propensity score matching methods can be used to evaluate the effect of interventions. Practical issues, such as matching algorithms and quality assessment are also discussed. A case study of speed cameras is presented in chapter 6.

4.3.1 Motivations

The fundamental problem of causal inference is that it is impossible to observe the outcomes of the same unit in both treatment conditions at the same time (Holland, 1986). One possible solution to this problem is to employ a control group of untreated units and simply estimate the difference in mean outcomes as the treatment effect. As discussed previously, however, this approach is only valid under the particular circumstances of randomized experiments where the treatment and control groups are identical. In most cases, treatment assignments are not random and the inferences could be biased due to confounding factors. That is the selection of treated units is affected by a vector of covariates ***X***, which affect the possibilities arising from treatment exposure and outcomes. For example, the sites of speed cameras are largely decided based on historical casualty numbers, and thus sites with more casualties historically are more likely to be selected as camera sites. In this case, the randomized experiment is not valid. To relax the strict requirement for randomized experiments, unconfounded assignment is proposed by Rosenbaum and Rubin (1983). In the context of unconfounded assignment, it is possible to account and adjust for differences in pre-treatment covariates and outcomes between treatment and control groups in order to properly estimate the effect of treatment.

Matching is one such approach based on unconfounded assignment. The basic idea behind matching is to match each treated unit to an untreated unit with the same values on observed characteristics, such as a vector of covariates ***X***. The matching approach becomes more difficult to implement as the number of observed covariates used increases, however. This obstacle can be overcome by matching on a single index instead of multiple dimensions. The most well-known index is the propensity score and the matching method is then termed propensity score matching (PSM). The PSM method has been widely used as a tool of evaluation in econometrics (Heckman et al., 1997; Rudner and Peyton, 2006; Hirano and Imbens, 2001; Dehejia, 2005; Dehejia and Wahba, 2002; Kurth et al., 2006; Lechner, 2001; Abadie and Imbens, 2004, 2009), but the validity of this approach rests on two assumptions which we will discuss in the next section.

4.3.2 Assumptions

The first assumption for an unconfounded assignment approach is known as the Conditional Independence Assumption (CIA), which assumes all observed differences in characteristics between the treated and untreated units are controlled for, and the outcomes that would result in the absence of treatment are the same for both groups. The CIA creates a selection process analogous to that of randomized experiments. More generally, the distribution of the counterfactual outcomes for treated and untreated groups are the same. In these circumstances it is possible to infer the counterfactual outcomes and the treatment effect can be estimated by the differences between treatment and control groups.

The CIA can be described as:

$(Y(1), Y(0)) \perp\!\!\!\perp D|\mathbf{X}, \forall\mathbf{X}$ (Unconfoundedness)

The CIA is a strong assumption and requires a very rich dataset of high quality. We return to this later in this chapter. It is clear that this assumption would not be feasible in the case of a high dimensional vector $\boldsymbol{X}$. For example, if $\boldsymbol{X}$ is a vector of n binary covariates, the number of possible matches will be 2^n. This increases the complexity in matching. To tackle this problem, Rosenbaum and Rubin (1983) proposed the idea of balancing scores, suggesting that if potential outcomes are independent of treatment conditional on covariates $\boldsymbol{X}$, they are also independent of treatment conditional on a balancing score, such as the propensity score. The propensity score, $P(D=1|\boldsymbol{X})=P(\boldsymbol{X})$, is the probability of being enrolled in a treatment given observed covariates $\boldsymbol{X}$. The CIA based on the propensity score can thus be described as:

$(Y(1), Y(0)) \perp\!\!\!\perp D|P(\mathbf{X}), \forall\mathbf{X}$ (Unconfoundedness given the propensity score)

The second problem is regarding the chance of finding a match for each individual with the same propensity score. It is likely that there is no match in the control group with a similar propensity to that of any treated individual. So it requires that individuals with the same $\boldsymbol{X}$ values have a positive probability of being in both treated and untreated groups. In other words, the proportion of treated and untreated individuals must be greater than zero for every possible value of $\boldsymbol{X}$. This condition is usually referred as the common support condition or overlap condition, which ensures sufficient overlap in the characteristics between the treated and untreated units to find adequate matches. The overlap condition can be described as：

$0<P(D=1|X)<1$ (Overlap Condition)

There are cases, however, where no overlap can be found for certain propensity scores. If so, some treated individuals would be dropped and the treatment group refined as one in which treated individuals fall within the common support. On the one hand, this reinstates the need for the overlap condition and hence increases the strength of matching. On the other hand, the loss of an element of the treated population may cause problems in evaluation, because the treatment effects are estimated based on a sub-sample instead of the whole population. Whether this problem is severe depends on the proportion of the treated sample that is lost. We will discuss approaches for quality tests of this assumption in section 4.3.5.

4.3.3 Data Requirements

It is explicit that the validity of PSM largely relies on adequate data regarding matters affecting participation and outcomes. In this section, we discuss several important issues on the data used in PSM.

The first issue is regarding the choice of covariates. It is suggested that omitting important covariates can cause serious bias in estimation (Heckman et al., 1997). Only covariates that affect both treatment participation and potential outcomes should be included. Usually, however, the researcher has no precise knowledge of these factors and, although all the covariates could be included, this could generate problems with the common support (Bryson et al., 2002). Another reason for avoiding over-parameterized models is that although the inclusion of non-significant covariates will not affect the unbiasedness and consistency of the estimates, it can increase their variance. Augurzky and Schmidt (2000) show that including the full set of covariates in small samples could cause problems in terms of higher variance. We discuss the choice of covariates in the next section.

It has also been pointed out that it is important for data for both treatment and control groups to be derived from the same sources, so that the measures used are identical or similar (Heckman et al., 1999). If data for the treatment and control groups are derived from different sources, it is particularly critical to guarantee that the covariates are constructed in the same way. It is, of course, easier to justify the CIA and the matching procedure with better and more informative data but it has been suggested that 'too good' data is not helpful either, since the overlap condition may fail with the result that the matching cannot be implemented, if $P(\boldsymbol{X})=0$ or $P(\boldsymbol{X})=1$ for some values of $\boldsymbol{X}$ (Heckman et al., 1998). In this case, matching conditional on those $\boldsymbol{X}$ values cannot be used because individuals with such $\boldsymbol{X}$ values either always or never receive treatment. Some randomness is needed, therefore, to guarantee that individuals with identical characteristics can be observed in both treatment states (Heckman et al., 1998).

The PSM method is often described as a 'data hungry' method in terms of both the number of covariates and the sample size. If the untreated group is large enough, it is still possible to find adequate matches even if the average characteristics are very different. One study by Zhao (2000) suggests that the performance of matching estimators depends largely on the data structure. We will return to this in chapter 6.

4.3.4 Implementation of PSM

In this section, we consider the practical issues when implementing matching estimators. Three steps for implementing the PSM method are discussed: modelling for estimating propensity scores, matching algorithm and estimating treatment effects.

4.3.4.1 Estimating Propensity Score

Model Choice. The first task in the PSM approach is to estimate the propensity score. Because linear probability models produce predictions outside the [0, 1] bounds of probability, the discrete choice models such as logit and probit models are usually used. For binary treatment, logit and probit models usually yield similar results, hence the choice between them is not critical (see further discussion of this point in Smith, 1997). When the treatments have multiple levels, however, a model that can represent multiple choices is preferred. Multinomial logit and multinomial probit models are two possible options (see Imbens, 2000, and Lechner, 2001, for more examples).

Covariate Choice. As discussed previously, the selection of covariates included in the model is very crucial in PSM. The problem would be less complicated if precise criteria for treatment participation were available. Where such criteria are not available, it is still possible to choose covariates based on previous empirical findings. In addition, there are strategies for the selection of the covariates to be used in estimating the propensity score (Heckman et al., 1998 & 1999).

The first approach relies on the statistical significance of the covariates. Only those covariates that are statistically significant are kept in the propensity score estimation model. To do this, one starts with a parsimonious model and iteratively adds covariates to the specification. Another method for specifying the propensity model is called 'leave-one-out cross-validation' (Black and Smith, 2003). Beginning with a minimum model specification, blocks of additional relevant covariates are then added in the model. The resulting root mean squared errors are calculated to assess goodness of fit. Augurzky and Schimdt (2000) point out that it has to be kept in mind that the main purpose of the propensity score estimation is not to predict selection into treatment as closely as possible but to balance all covariates. It is important, therefore, to ensure that the covariates included in the model are based on sound theory to characterize the participation model.

4.3.4.2 Matching Algorithm

Once the propensity score is estimated, the next steps are to match comparison units with treated units and to select the matching algorithm. In general, the treatment effect can be estimated as $Y_i(1)$-$Y_{j(i)}(0)$, where $Y_{j(i)}$ is the outcome for the comparison unit j that is matched with the treated unit i. In the early stage, treated units were paired with those in the comparison on a one-to-one basis. For each treated unit, only the unit in the comparison group with the most similar propensity score is matched with that treated unit. Usually, such pairwise matching is performed without replacement, which means each comparison group member can be used as a matched unit only once. The problem is that the matching performance could be poor when the sample size of the control group is small or when there is little common support for two groups (Dehejia and Wahba, 2002). More studies use matching with replacement and other matching algorithms other than pairwise matching.

In contrast to matching one treated unit with only one comparison unit, using all comparison units that are sufficiently close to a given treated unit is a more stable approach. To account for the sampling error, it is important to include only those comparison units that are close, to within a certain tolerance, to a given treated unit. Below, the most commonly used matching algorithms are discussed.

Nearest Neighbour Matching: one of the most straightforward matching methods is nearest neighbour matching. Units from the comparison group with the closest propensity score are chosen as matches for given treated units. As discussed previously, this matching algorithm can be performed with or without replacement. Matching with replacement can increase the matching quality and decrease the bias (Smith and Todd, 2005).

Caliper and Radius Matching: It is likely that nearest neighbour matching performs poorly if the closest neighbour is distinct. To avoid this problem, a tolerance on the maximum propensity score distance can be formulated. The idea of this matching approach is that it uses not only the nearest neighbour but all comparison members within the tolerance level. One issue in caliper and radius matching noted by Smith and Todd (2005), however, is that it is difficult to know what level of tolerance should be used.

Stratification and Interval Matching: The idea behind this matching algorithm is that the common support of the propensity score is partitioned into a set of strata and the treatment effects are evaluated with each strata by taking the mean difference in outcome between treated and comparison individuals. The question is how many strata should be used. Cochrane and Chambers (1965) suggest that five subclasses are enough to remove 95% of the bias associated with one single covariate. A more precise way to justify the choice of number of strata is to check the balance of the propensity score and the covariates. First, the propensity score is checked to confirm whether it is balanced within each stratum. Once this condition is met, the covariates are checked to confirm if they are balanced. If not, then the model for estimating the propensity score needs to be re-specified, such as by adding interactions or higher order terms (Dehejia and Wahba, 1999).

Kernel and Local Linear Matching: The matching algorithms discussed above employ only a limited number of comparison individuals to construct the counterfactual outcomes for treated units. Kernel and local linear matching are nonparametric matching estimators using a weighted average of all untreated individuals to construct the counterfactual outcome of each treated individual. Those untreated units with the closest propensity score to the treated one are assigned the highest weight. One major advantage of this matching algorithm is the lower variance, because more information is employed. The overlap condition is important for this approach so as to avoid poor matches. One interpretation is that kernel matching can be seen as a weighted regression of the counterfactual outcome on an intercept with weights (Smith and Todd, 2005). The choice of the bandwidth parameter is important and there has been some discussion on this issue (Silverman, 1986; Pagan and Ullah, 1999). High bandwidth values lead to small variance, but, probably, a biased estimate as a trade-off.

Given all the above choices, which algorithm is most appropriate? Asymptotically, all approaches should yield the same results when the dataset is large enough. The size of the data sample differs in empirical studies, however, and the choice of matching algorithms can be important (Heckman et al., 1997). In general, the performance of different algorithms depends largely on the data structure at hand (Zhao, 2000). For example, it makes more sense to match with replacements if there are only a limited number of comparison units. If there are adequate comparison units, however, it might be more appropriate to use kernel matching. Generally, it is sensible to employ and compare different algorithms. If results are similar, the choice is not important. Otherwise, more investigation is necessary to reveal the source of the disparity.

4.3.4.3 Estimating Treatment Effects Using PSM

The final step is to evaluate treatment effects using the estimated propensity score and the appropriate matching algorithm. The effects can be calculated by averaging the differences in outcomes between treated units and matched comparison units.

$$\gamma_{ATET} = E[Y(1)\text{-}Y(0)|D=1] = \frac{1}{N}\sum_{i=1}^{N}(Yi(1) - Yj(i)(0)) \qquad 4.9$$

A number of statistical software programs are available to perform matching and evaluate average effects. A frequently used program, **psmatch2**, has been developed by Leuven and Sianesi (2003) and can be installed in Stata. All matching algorithms can be implemented in this program. Functions, such as common-support graphing (**psgraph**) and covariate balance tests (**pstest**) are also included in **psmatch2**.

It is also important to estimate the standard errors to indicate the sampling error. Bootstrap methods are widely used to obtain standard errors in PSM and can be easily implemented in **psmatch2** or the Becker and Ichino (2002) PSM estimation program. This method will be discussed in the next section.

4.3.5 Quality Assessment

As explained earlier, the validity of an unconfounded model relies heavily on two key assumptions, the CIA and overlap assumptions. In this section, we discuss the methods for checking these two assumptions and assessing the matching quality.

4.3.5.1 The Conditional Independence Assumption

The CIA assumes that the treatment assignment depends only on observable covariates and that the outcomes in the absence of treatment are independent of treatment status. Although the CIA cannot be tested directly, methods are available for making it more plausible. First, it is important that the model is correctly specified and all relevant covariates that affect the assignment and potential outcomes are included. Guidelines for model specification and covariate selection have been discussed in previous sections.

Since it is assumed that, besides treatment status, there is no difference in characteristics between the treatment and comparison groups, tests need to be performed to check whether matching succeeds in balances characteristics across the treatment and comparison units.

$D \perp\!\!\!\perp \boldsymbol{X} \mid P(D=1|\boldsymbol{X})$

In other words, there should be no statistically significant differences between the covariate means of the treatment and comparison units. There are various ways to check the balance of covariates but the most widely used approach is a two-sample t-test. It is possible that balance cannot be fully achieved across the whole matching sample. If so, the sample can be divided into stratums and covariates will tend to be better balanced within each stratum (Dehajia and Wahba, 1999). Another balancing indicator suggested by Rosenbaum and Rubin (1985) is the standardized bias, which is defined as the difference in means of covariate X scaled by the square root of the average of their sample variances.

If the quality indicators are still not satisfactory, the participation model should be revised, e.g. adding higher order and interaction terms. If differences still remain after re-specification, it may suggest a failure of the CIA and alternative methods should be considered.

4.3.5.2 Common Support Condition

Verification of the common support condition is very important in assessing the performance of the propensity score matching estimation. This condition assumes that the units with the same observed characteristics have a positive probability of being in both treated and untreated groups: $0<P(D=1|X)<1$. There are several ways to check the overlap and the region of common support between treatment and control groups, where a visual inspection of the propensity score distribution for both groups is suggested as the most straightforward approach (Lechner, 2000).

Besides visual tests, there are also formal techniques to determine the region of common support. The most frequent approach is based on the minima and maxima of the propensity score for the treatment and comparison groups. Observations with propensity scores smaller than the minimum and larger than the maximum in the opposite groups are discarded from analysis. It is worth noting that the overlap condition is more crucial for kernel matching than it is for nearest neighbour matching, since all untreated observations are used to estimate the counterfactual outcome in kernel matching, while nearest neighbour matching only uses the closest observation.

One problem that may arise if there is only limited overlap between both groups within the common support region or if the density in the tails of the distributions are very thin. To overcome such problems, Smith and Todd (2005) suggest a trimming method to determine the common support. With the trimming procedure, the region of common support is defined as points whose values have positive density within both the D=1 and D=0 distributions. Additionally, it is required that the densities of these points exceed zero by a threshold amount. No matter which approach is applied, it is recommended that a visual analysis is made beforehand.

Both the maxima and minima approach and trimming requires that observations fall outside the defined overlap region to be discarded, and thus the treatment effects cannot be estimated for those units. If the proportion of lost units is too large, the estimated effect on the remaining units may not be representative (Bryson et al., 2002). In such circumstance it is necessary to inspect the characteristics of discarded units since supplementary information could be obtained for interpreting the estimated treatment effects.

The attractiveness of the propensity score method relies on the fact that it is able to construct counterfactual outcomes in the absence of randomised experiments. The estimations could be implausible, however, if the assumptions underpinning this method fail to hold. If the conditional independence assumption is affected by unobserved factors, the matching estimator may be seriously biased. To correct for this bias, the propensity matching method is combined with the difference-in-difference approach (Heckman et al., 1997). In the next section, we will discuss the difference-in-difference approach and its use in the matching method.

4.4 Difference-In-Difference

Since the work by Ashenfelter (1978) and Ashenfelter and Card (1985) the difference-in-difference (DID) method has become one of the most popular tools for evaluation studies. This section provides an overview of the standard DID approach as well as some of its derivatives. It will also discuss the limitations of the DID method.

4.4.1 Notations and Model Specification

The basic idea of the DID method is that observations are collected for two groups for two periods. One of the groups is the treatment group which is exposed to the treatment in one period. The other group is the control group which receives no treatment during both periods. In the case where the same units within a group are observed in each time period, the average gain over time in the non-exposed (control) group is extracted from the gain over time in the exposed (treatment) group. This double differencing, the so called "difference-in-difference" method, removes biases in the second period comparison between the treatment and control groups that could be the result of permanent differences between those groups, as well as biases from comparison over time in the treatment group that could be the result of time trends unrelated to the treatment (see Abadie, 2005; Finkelstein, 2002; Card and Krueger, 1994 for a more detailed discussion).

Assume that n individuals are observed in two time periods, t=0, 1 where 0 indicates a time period before the treatment group receives treatment, i.e. pre-treatment, and 1 indicates a time period after the treatment group receives treatment, i.e. post-treatment. Every group is indexed by the letter $i=T, C$ where T indicates the treatment group, and C indicates the control group; let Y_{0T} and Y_{1T} be the outcome for the treatment group before and after treatment, respectively, and let Y_{0C} and Y_{1C} be the corresponding outcome for the control group.

Under the basic DID approach, the outcome Y_{it} is modelled by the following equation (Ashenfelter and Card, 1985)

$$Y_{it} = \alpha + \beta T_{it} + \gamma G_{it} + \delta(T_{it} \cdot G_{it}) + \varepsilon_{it} \qquad 4.10$$

Where α is the constant term, β is the time trend, γ is the specific group effect, δ is the treatment effect we are interested in and ε_{it} is a random, unobserved term which contains the error caused by omitted covariates. T_{it} is the time-specific component, which takes the value 1 if Y_{it} is observed in the post-treatment period and 0 otherwise. G_{it} is a group-specific component, which is 1 if Y_{it} is an observation from the treatment group and 0 otherwise. $T_{it} \cdot G_{it}$ is an interaction term which indicates a treated individual after the intervention.

Three assumptions are necessary for DID to provide an unbiased consistent estimate of the treatment effect.

(1) The model is correctly constructed, in the sense that the function and covariates added into the equation are correct.

(2) The error term has an expectation of zero and is distributed independently of the covariates.

(3) The third assumption, which is also critical in DID estimation is that the treatment group and control group will follow the same trend over time in the absence of the treatment, also known as the parallel trend assumption,

$$E[Y'_{1T} - Y'_{0T}] = E[Y_{1C} - Y_{0C}] \qquad 4.11$$

where Y'_{it} represents outcomes of the treatment group in the absence of treatment.

4.4.2 DID Estimator

According to these assumptions, we could attain expected values of outcomes Y_{it} given by following equations.

$$E[Y_{0T}] = \alpha+\gamma$$

$$E[Y_{1T}] = \alpha+\beta+\gamma+\delta$$

$$E[Y_{0C}] = \alpha$$

$E[Y_{1C}] = \alpha+\beta$

First we consider a single difference estimator, which compares the difference only in the treatment group before and after treatment.

$\hat{\delta}_1 = \overline{Y}_{1T} - \overline{Y}_{0T}$

$E[\hat{\delta}_1] = E[\overline{Y}_{1T}] - E[\overline{Y}_{0T}]$

$= \alpha+\beta+\gamma+\delta - (\alpha+\gamma)$

$= \beta+\delta$

So we can conclude that the estimator $\hat{\delta}$ will be biased if a time trend exists, because we may treat the time trend as part of the treatment effect.

Next consider another estimator based on comparing the average difference between the treatment and control groups.

$\hat{\delta}_2 = \overline{Y}_{1T} - \overline{Y}_{1C}$

$E[\hat{\delta}_2] = E[\overline{Y}_{1T}] - E[\overline{Y}_{1C}]$

$= \alpha+\beta+\gamma+\delta-(\alpha+\beta)$

$= \gamma + \delta$

This estimator is also biased due to the specific group effect.

As defined in the previous section, the Difference-In-Difference estimator is the difference in average outcome in the treatment group before and after the treatment, minus the difference in average outcome in the control group before and after the treatment.

$$\hat{\delta}_{DID}=\bar{Y}_{1T}-\bar{Y}_{0T}-(\bar{Y}'_{1T}-\bar{Y}'_{0T})$$

$$=\bar{Y}_{1T}-\bar{Y}_{0T}-(\bar{Y}_{1C}-\bar{Y}_{0C})$$

$$E[\hat{\delta}_{DID}]=\alpha+\beta+\gamma+\delta-(\alpha+\gamma)-(\alpha+\beta)-\alpha$$

$$=\delta$$

We can see that this is an unbiased estimator. The DID estimator is shown graphically in Figure 4.1. According to the parallel trend assumption, the dotted line would have to hold in the absence of the treatment.

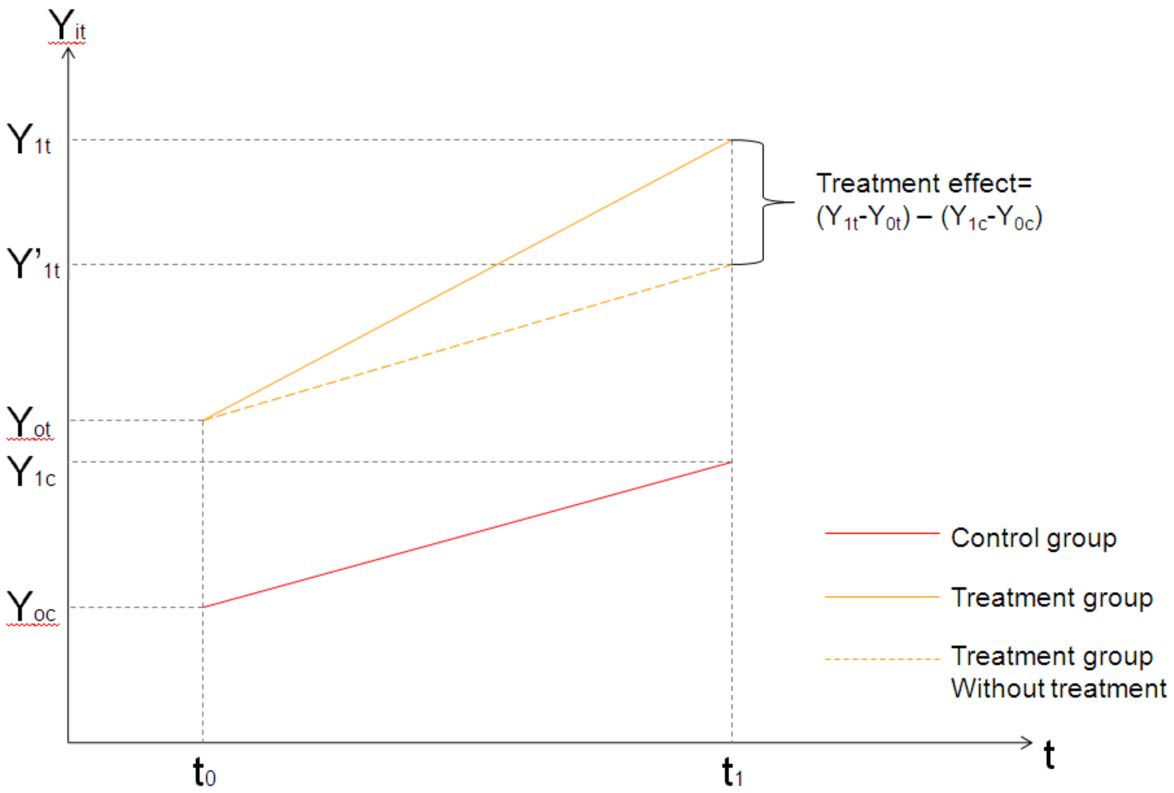

Figure 4.1 Graphic illustration for DID estimator.

The definition of treatment and control groups can be further refined to make the analysis of the treatment effect more convinced. For example, suppose one city implements the congestion charge in the central area, and the response variable is the traffic flow. One possibility is to use data from the charging area in that city, both before and after the change, with the control group being an area other than the city centre. The potential problem with this DID analysis is that other factors unrelated to the congestion charge might affect the traffic flow at the city level. It is also possible to use another city centre as the control group. The problem with this analysis is that changes in the traffic flow of the central area might be systematically different across cities due to other factors, such as income and wealth differences, rather than the congestion charge.

A more robust way is to use both a different city and a control group within the treatment city. Again, let T indicate the treatment period, let G indicate the treatment group, and let C indicate the city implementing the treatment. Then an expanded version of equation *4.11* can be described as (Gruber, 1994)

$$Y_{it} = \alpha + \beta T_{it} + \gamma G_{it} + \lambda C_{it} + \theta(T_{it} \cdot C_{it}) + \sigma(G_{it} \cdot C_{it}) + \delta(T_{it} \cdot G_{it}) + \mu(T_{it} \cdot C_{it} \cdot C_{it}) + \varepsilon_{it} \quad 4.12$$

The coefficient of interest is μ, the coefficient on the triple interaction term, $T_{it} \cdot C_{it} \cdot C_{it}$. This estimator is called the difference-in-difference-in-difference (DDD) estimate. Two potentially confounding effects can be controlled for: changes in the traffic flow across cities and changes in the traffic flow of all areas in the treatment city

4.4.3 DID Matching

As discussed in section 4.3, in the propensity score methods the conditional independence assumption is in some cases too strong, and may not hold when unobserved factors that may influence outcomes are not included in the model. The PSM estimator can be biased by selection-on-unobservables, where unobserved variables may critically determine the participation model.

The DID matching estimator can relax the strong CIA, given that pre-treatment data are available and unobserved variables are time-invariant (Heckman et al., 1997). For treated units, the dependent variable is outcome differences over pre- and post-treatment periods. The outcome difference is calculated over the same periods for comparison units. The DID matching estimator can reduce bias due to differences between treated and comparison groups, given that differences in their effects on outcomes are time invariant. Let *t* and *t'* denote the pre- and post-treatment periods respectively, and then the outcome for matching will be:

$\Delta Y_i = Y_{it'} - Y_{it}$

With this specification, the CIA can be relaxed by allowing the counterfactual outcome of the treated units to differ from the observed outcome of the untreated, with the same trend. That is:

$$E(Y_{0t'} - Y_{0t} | D=1, X) = E(Y_{0t'} - Y_{0t} | D=0, X) \qquad 4.13$$

The DID matching estimator can be calculated by estimating propensity score for both groups and applying the PSM to the differenced outcomes.

4.4.4 Issues in the DID method

Despite the benefits of the DID methods, there are several critical issues with it that are worth noting. First, the conventional DID approach strongly relies on the parallel trend assumption that the average outcomes for the treatment and control groups would exhibit a parallel time trend. Compositional differences between the treatment and control groups, however, can cause non-parallel trends in the outcomes. The DID estimator will be biased if this assumption cannot hold. There are several methods for improving the validity of the parallel trend assumption, such as including additional covariates and graphic tests. We will discuss these in the case study in chapter 5.

Another issue is regarding the serial correlation. Bertrand et al. (2004) show that the conventional DID estimation may suffer from the problem of serial correlation. The standard errors will be underestimated and the t-statistics will bias upwards in the presence of positive correlation. Over-rejection of the null hypothesis can cause false inferences regarding the effect of the treatment. Usually, the Durbin-Watson (DW) test is applied to test for the presence of serial correlation in the residuals. An ideal value of the DW test is around 2. A value which is much smaller than 2 indicates positive serial correlation. One way to correct the serial correlation is using Weighted Least Squares estimators, such as the Prais-Winsten estimator.

4.4.5 Synthetic Control Group

A key issue in causal studies is the lack of access to observe outcomes of treatment groups without the intervention. Missing counterfactual information could make the causal inferences implausible when confounders exist. A control group is then designed to provide counterfactual information.

In many evaluation programmes causal methods relying on observational data, such as propensity score approaches, panel data methods and instrumental variable methods, make a crucial assumption that the control group closely approximates the treatment group. In practice, however, it is difficult to find a control group which has the same or very similar characteristics to the treatment group.

In a DID model, outcomes for control and treatment groups from pre- and post-intervention periods are observed. The double differencing removes biases from permanent differences between control and treatment groups, as well as biases from time-trend effects. Although the DID method does not require the control group to reproduce the counterfactual outcomes for the treatment group, it relies on a parallel trend assumption that the treatment and control groups will have the same trend over time in the absence of the intervention. Another limitation in the DID method is that although the presence of unobserved covariates is allowed, the effect of these covariates is restricted to be constant over time, so that the effects can be eliminated by taking a time difference.

As an extension of the DID method, the synthetic control method constructs a weighted combination of control units to approximate the treatment unit. The evolution of the outcome for the synthetic control group is estimated as the counterfactual of what would have been observed for the treatment group in the absence of intervention. The effect of the intervention of interest can be estimated by comparing the difference in evolution between the treatment group and the synthetic control group. In recent years, a few studies apply the synthetic control method to the evaluation of an intervention. For example, Abadie and Gardeazabal (2003) evaluate the effect of the terrorist conflict on economic development in the Basque Country, using those Spanish regions not affected by the conflict as a synthetic control group. Another study estimated the impact of a tobacco control programme in California (Abadie et al., 2010). Such methods have also been applied in the evaluation of the Asian Development Bank programme (Mukherji and Mukhopadhyay, 2011). The results of these studies have shown the applicability of the synthetic control method in causal analysis.

4.4.5.1 Model Specification

The following model is developed by Abadie and Gardeazabal (2003). See also Abadie, Diamond, and Hainmueller (2007). Assume that there are T time periods and G groups, among which only the first group is exposed to the intervention and G-1 groups remain as potential control groups, which can be also defined as the donor pool. Let g= 1 denote the treated group and g= 2, 3, …, G denote the control groups; Let t= 1, ..., T_0 denote the pre-intervention periods and t= T_0 + 1, ..., T denote the post-intervention period. Let Y_{gt}^I be the outcome of group g if it received the intervention; let Y_{gt}^N be the outcome of group g if it did not receive the intervention.

One important assumption made in this method is that there is no treatment effect on any control group. This is also the fundamental assumption in all control studies.

The interest here is in the effect of the intervention on the treatment group. For $t > T_0, \delta_{1t} = Y_{1t}^I - Y_{1t}^N$, where Y_{1t}^I is the outcome observed for the treatment group during the post-intervention period, and Y_{1t}^N is the counterfactual outcome for the treatment group during the post-intervention period in the absence of the intervention. Because Y_{1t}^I is obtained from observed data, what we need to know is Y_{1t}^N. In most comparative studies the control group selected is expected to reproduce the counterfactual outcomes Y_{1t}^N, which can be expressed as $Y_{1t}^N = Y_{gt}^N (g \in \{2, 3, \dots, G\}, t=1, \dots, T)$, so that the estimator of the intervention effect $\widehat{\delta_{1t}} = Y_{1t}^I - Y_{gt}^N$ is unbiased.

As we discussed before, however, it is difficult to find a single control group to achieve this purpose, and, consequently, the synthetic control method is applied. Y_{gt}^N is modelled by the following equation

$$Y_{gt}^N = \alpha_t + \boldsymbol{\beta}_t \mathbf{X}_g + \boldsymbol{\gamma}_t \mathbf{Z}_g + \varepsilon_{gt},\ g = 1, \dots, G;\ t = 1, \dots, T \qquad 4.14$$

Where α_t is the constant, $\boldsymbol{X_g}$ is the vector of observed covariates, $\boldsymbol{\beta_t}$ is the vector of coefficients, $\boldsymbol{Z_g}$ is the vector of unobserved covariates, $\boldsymbol{\gamma_t}$ is the vector of coefficients for unobserved covariates, and the unobserved error term ε_{gt} is the heterogeneity independent across groups and in time.

Suppose that we have a set of weights $W = (w_2, w_3, \dots, w_G)$, with $w_g \geq 0$ and $\sum_{g=2}^{G} w_g = 1$. The weighted average of the control groups is described as:

$$\sum_{g=2}^{G} w_g Y_{gt}{}^N = \alpha_t + \beta_t \sum_{g=2}^{G} w_g X_g + \gamma_t \sum_{g=2}^{G} w_g Z_g + \sum_{g=2}^{G} w_g \varepsilon_{gt} \qquad 4.15$$

Provided that there is a vector $W' = (w_2', w_3', \dots, w_G')$ such that

$$\sum_{g=2}^{G} w_g' Y_{g1}^{N} = Y_{11}^{N},\ \sum_{g=2}^{G} w_g' Y_{g2}^{N} = Y_{12}^{N}\ ,\ \ldots,$$

$$\sum_{g=2}^{G} w_g' Y_{gt_0}^{N} = Y_{1t_0}^{N}\ \text{, and } \sum_{g=2}^{G} w_g' X_g = X_1 \qquad 4.16$$

When $\sum_{t=1}^{T_0} \boldsymbol{\gamma_t}'\boldsymbol{\gamma_t}$ is non-singular, it can be proved that (Abadie et al. 2009)

$$Y_{1t}^{N} - \sum_{g=2}^{G} w_g' Y_{gt}^{N} = \sum_{g=2}^{G} w_g' \sum_{s=1}^{T_0} \gamma_t (\sum_{n=1}^{T_0} \gamma_n' \gamma_n)^{-1} \gamma_s' (\varepsilon_{gs} - \varepsilon_{1s}) - \sum_{g=2}^{G} w_g' (\varepsilon_{gt} - \varepsilon_{1t}) \qquad 4.17$$

It can be also proved that the right-hand side of *equation 4.17* will approach to zero as the number of pre-intervention periods increase, and we can thereby get the estimator for δ_{1t}:

$$\widehat{\delta_{1t}} = Y_{1t} - \sum_{g=2}^{G} w_g' Y_{gt}\ (t = T_0+1,\ \ldots,\ T). \qquad 4.18$$

Let X_1 and X_g be the vectors of predictors for the treatment and control groups respectively. Let V be a positive diagonal matrix which captures the importance of every predictor. The weights W' can be calculated by minimizing:

$$\|X_1 - X_g W\| V = \sqrt{(X_1 - X_g W)' V (X_1 - X_g W)} \qquad 4.19$$

Where $w_g \geq 0$ and $\sum_{g=2}^{G} w_g = 1$. The weights $W'(V)$ depend on the diagonal matrix V. Let V' be such a diagonal matrix that the outcome for the treatment group can be best approximated by the synthetic control group determined by $W'(V')$. Let Y_{1t} and Y_{gt} be the vectors of outcomes for the treatment and control groups respectively.

$$V' = \text{argmin}(Y_{1t} - Y_{gt}W'(V))'(Y_{1t} - Y_{gt}W'(V)) \qquad 4.20$$

It needs to be noticed here that equation *4.17* requires (Y_{11}^N, ..., Y_{1T0}^N, Z_1') to fall in the convex hull of *{(* Y_{21}^N*, ...,* Y_{2T0}^N*,* Z_2*'), ..., (*Y_{G1}^N*, ...,* Y_{GT0}^N*,* Z_G*')}*. In practice, it is commonly the case that no such set of weights exists so that *equation 4.18* can hold exactly. Therefore, a weighted combination of control groups is selected such that *equation 4.18* can hold approximately. The discrepancy can now be calculated to see how well the predictors of the treatment group can be approximated by the synthetic control group. If (Y_{11}^N, ..., Y_{1T0}^N, Z_1') falls far from the convex hull of *{(* Y_{21}^N*, ...,* Y_{2T0}^N*,* Z_2*'), ..., (*Y_{G1}^N*, ...,* Y_{GT0}^N*,* Z_G*')}*, then the discrepancy would be large and the synthetic control method would not be appropriate. To make the synthetic control group close to the treatment group and reduce the interpolation biases, only groups with similar characteristics with the treatment group are chosen as potential control groups.

4.4.5.2 Inferential techniques

Abadie et al. (2009) proposed exact inferential techniques, the placebo study, to perform quantitative inferences. The idea of the placebo study is based on the classic permutation tests. Permutation tests can be used to determine whether the estimate of the intervention effects is ascribed to the randomness introduced in selecting the sample. First, a statistic is chosen that measures the effect we are interested in and the statistic is calculated from the original data. Then permutation resamples are chosen from the data to construct the permutation distribution of the statistic from resamples. Finally, the original statistic is located on the permutation distribution to get the P-value. A small P-value is evidence against the null hypothesis, indicating that the intervention has a significant effect on the population.

In a placebo study, the synthetic control method is applied to every potential control group in the donor pool and the effect of the intervention is estimated for every group. The intervention effect for the actual treatment group is compared with the distribution of the effects for placebo control groups. The null hypothesis of no intervention effect should be rejected if the effect estimated for the treatment group is not contained within a certain range of the placebo distribution (e.g. 95%). This approach can be found in the work by Abadie and Gardeazabal (2003), where the Basque Country is exchanged with a region not affected by terrorism and the synthetic control method is applied to the new 'treatment group'. Another work by Bertrand et al. (2004) applies the placebo study to time-series data to quantify the effect of serial correlation in the DID model.

4.4.5.3 Advantages of synthetic control methods

Compared to conventional comparative studies, synthetic control methods have several advantages. First, from the weights W it is possible to know exactly how much every control group contributes to the synthetic control group. Second, by comparing pre-intervention outcomes and explanatory covariates, it is explicit how close the characteristics are between the synthetic control group and the treatment group. Third, although unobserved confounders are allowed in the DID model, the effect of these confounders is restricted to be constant over time. In synthetic control methods this restriction is released. Finally, the synthetic control method can make impersonal inferences. Because this method does not need outcomes from post-intervention periods, researchers can conduct evaluation without knowing how the treatment group is affected in nature. This can help researchers to give conclusions objectively.

To summarize, in the last decade numerous studies have been conducted to evaluate the effects of traffic interventions on road safety. This existing research, however, has failed fully to address issues in causal analysis, such as confounding and selection bias. In this research formal causal approaches are employed which are used routinely in other areas of science such as medicine and epidemiology, but not yet adopted in the area of transport. In the next section, the Bayesian approaches to evaluation studies are reviewed and a conventional Bayesian model is combined with the formal casual models, such as the propensity score model.

4.5 Bayesian Approaches for Evaluation Studies

Another widely used approach for the before-after evaluation of road safety treatments is the Bayesian method. The Bayesian approach combines prior and observed data to derive an estimate for the outcomes of interest. In this section, two related Bayesian approaches are discussed, Empirical Bayes (EB) and full Bayes (FB), and their application in road safety studies is reviewed. Finally, to address the issue of selecting proper control groups, an EB method is developed by combining it with the propensity score.

4.5.1 Review of Bayesian Methods

The EB and FB are two related approaches to combining prior information and current information in order to derive an estimate for the expected safety of treated sites. In the application to road safety, the prior information is obtained from a group of similar sites and the observed information is the accident frequency for the specific site.

Although the EB and FB have a similar conceptual basis, there are differences in these two approaches. In the EB method, the prior information is obtained from a group of similar sites to estimate a sample mean and variance. A safety performance function can be also applied to establish the relationship between the accident frequency of the reference sites and various factors. An improved estimate of the long-term accident frequency can be obtained by combining the point estimates of the expected mean and the variance with the accident number of treated sites. In the FB approach, instead of a point estimate of the expected mean and variance, distributions for these statistics are estimated from a model of a reference population, and combined with the accident frequency of treated sites to estimate the long-term expected accident frequency. The estimated variance can be more accurate by using a prior distribution instead of a point estimate.

Bayesian methods have been widely used in traffic safety studies over the last two decades, especially in before-after evaluations (Hauer, 1997; Hauer et al., 2002; Li et al., 2008; Miaou and Lord, 2003; Park and Lord, 2007; Persaud et al., 1997, 2004, 2010; Persaud and Lyon, 2007; Quddus, 2008; Aquero-Valverde and Jovanis, 2009; El-Basyouny and Sayed, 2009). EB methods, in particular, have become popular as statistically defensible methods that can cope with several key issues in observational before-after studies, such as the regression to mean bias. A recent study by Persaud and Lyon (2007) reviews the use of EB in before-after safety studies, including the basics of EB evaluation and the need for and validity of the EB approach, and addresses the critical issues in the interpretation of EB evaluations.

Recently, with the advances in statistics and the availability of the software package WinBUGS (Spiegelhalter et al., 2005), FB approaches have been applied in more road safety studies. Since it requires less data and better controls for uncertainty in data, and provides more flexibility in selecting distributions for accident count (Lan et al., 2009) and more detailed causal inferences (Carriquiry and Pawlovich, 2005), the FB approach has been suggested as a useful alternative to the EB approach.

Inevitably, these two methods have been discussed and compared in recent studies. Persaud et al. (2010) compared the EB and FB approaches through two empirical applications, showing that the two approaches lead to comparable results when a large reference group is available to develop the safety performance functions (SPFs) of EB. Lan et al. (2009), meanwhile, evaluate FB methods using a simulated dataset and demonstrate how they can account for the regression to the mean bias. It was found that the FB approach could provide similar results to the EB method. The work by Huang et al. (2009) provides different results, however. They conducted an empirical analysis to evaluate different approaches, including the FB and EB methods and found that an FB approach using hierarchical models significantly outperformed the standard EB approach in correctly identifying hazardous sites.

Another important issue relates to the structure of the SPF, which may strongly affect model development and consequently the results of the before and after evaluation. Poisson and NB models have been extensively discussed in the traffic safety literature (Hauer et al., 2002; Persaud et al., 1997, 2001). Latterly random effects Poisson-Gamma model has been widely used due to its capability to deal with spatial effects (Chin and Quddus, 2003; El-Bosyouny and Sayed, 2009). The Bayesian hierarchical models offer a more flexible model structure which can cope with distributions such as the Poisson-Lognormal distribution and the hierarchical Poisson-Gamma distribution (El-Basyouny and Sayed, 2011; Miaou and Lord, 2003; Yanmaz-Tuzel and Ozbay, 2010). Different prior distributions are discussed by Yanmaz-Tuzel and Ozbay (2010) with the results suggesting that the Poisson-Lognormal model with higher levels of hierarchy and informative priors may provide more robust estimates of model parameters.

In the following sections, the validity and limitations of the EB and FB approaches will be discussed.

4.5.2 Empirical Bayes

The idea behind the Bayes approaches is that "accident counts are not the only clue to the safety of an entity. Another clue is in what is known about the safety of similar entities" (Hauer, 1997). In the EB approach, prior information is obtained from a group of sites similar to the treated group and used to calculate point estimates of sample mean and variance. Alternatively, information is acquired from a calibrated safety performance function that relates the accident frequency of the reference group to their characteristics.

4.5.2.1 Model Specification

In the EB approach (Hauer, 1997; Hauer et al., 2002), the effects of a safety treatment on the crash number can be described as

$B - A$,

where B is the counterfactual crash number in the after period without treatment and A is the observed crash number in the after period.

To estimate B, SPFs relating crashes to traffic flow and other relevant factors are used to account for the effects of the regression to the mean and changes in traffic flow. Annual SPF multipliers are estimated as the ratio of yearly observed crashes and yearly estimated crashes from the SPF. These multipliers are then used to account for the temporal effects of weather, demography and other factors. The expected annual crash number at control sites is estimated using the SPF. Then the observed crash count (x) at treated sites in the before period is combined with the sum of annual SPF estimates (p) to analyse the expected crash number (m) at treated sites before the treatment (Hauer, 1997):

$$m = w_1(x) + w_2(p) \qquad 4.21$$

where the weights w_1 and w_2 can be estimated as:

$$w_1 = p/(p + 1/k) \qquad 4.22$$

$$w_2 = 1/k\,(p + 1/k) \qquad 4.23$$

where k is the dispersion parameter of the negative binomial distribution.

The time trend effect of traffic volumes can be accounted for by applying a factor, which is the ratio of the annual SPF predictions for the after and before period. Then the expected crash number in the after period without the treatment, λ, as well as its variance can be obtained.

The estimate of λ is then summed over all treatment sites to obtain λ_{sum} and compared with the count of crashes during the after period, π_{sum}. The Index of Effectiveness (θ) is estimated as (Hauer, 1997):

$$\theta = (\pi_{sum}/\lambda_{sum})/\{1+[Var(\lambda_{sum})/\lambda^2_{sum}]\} \qquad 4.24$$

The percentage change in crashes can be described as 100(1-θ).

4.5.2.2 Empirical Bayes Using Propensity Score

In the EB approach, a reference group which is similar to the treated one is required for calibrating the SPFs. Generally, the reference group must be similar to the treatment group in terms of traffic flow, road characteristics and so on. The reference group is also used to account for changes in traffic volume. Since the validity of the EB approach relies heavily on the availability of a proper reference group, it is critical to test the suitability of candidate reference groups. One commonly used test compares time trends in accident number for the treatment and reference groups. The time trend of a good reference group should track the one of treatment group very well. For example, Hauer (1997) estimates a sequence of sample odds ratio using 1 year of before data and the following year as the after data, starting with years 1 and 2 and increasing by 1 year. The sample mean and standard error is estimated from the ratios. The candidate reference group is suitable only if this sample mean is sufficiently close to 1.0.

As discussed earlier, a good reference group must be representative of the treated entities in terms of the time trends of crash counts, traffic flow and road characteristics. However, the odds ratio approach only takes the historical crash counts into account. To address this issue, the EB method has been developed by combining with the propensity score. The propensity score can be applied to find untreated sites that are similar to treated sites in order to construct the reference group. This approach will be discussed in Chapter 6.

4.5.3 Full Bayes

Compared to the EB approach, the FB method estimates long-term expected crash frequency by directly combining the observed data at reference sites and prior information at treatment sites. Instead of a point estimate of the expected mean and its variance, a distribution of expected crash frequency is estimated.

Several studies have compared the FB approach to the EB approach (Persaud et al., 2009; Lan et al., 2009). The results show that the FB approach generally has following potential advantages. It requires smaller sample sizes and allows inference at more than one level with hierarchical models, and enables more detailed inferences, such as credible intervals and parameter distributions. The FB method can cope with distributions such as Poisson-Lognormal distribution and hierarchical Poisson-Gamma distribution, while the EB approach rests on the assumption of a negative binomial distribution. The estimation of the SPF and treatment effects is integrated in the FB approach, while these are separate tasks in the EB method. Among the above advantages, the most appealing is the data requirement issue. A reference group is not required and there is no need to collect a large data set through the integration of the estimation of the SPF and treatment effects into a single step.

4.5.3.1 Model Specification

Two random effects regression models are usually used in the FB approach, the Poisson-Gamma model and Poisson-Lognormal model. These can be described as follows:

$Y_{it} \sim \text{Poisson}(\varepsilon_i \lambda_{it})$

Where Y_{it} is the observed number of crashes at site i in year t, λ_{it} is the expected number of crashes at site i in year t, and ε_i is the random effect at site i.

For the Poisson-Lognormal model, $\varepsilon_i \sim Log\ N(0, \sigma^2)$.

For the Poisson-Gamma model, $\varepsilon_i \sim Gamma(\varphi, 1/\varphi)$ with the mean having a value of 1 and where φ is the dispersion parameter. When $E(\varepsilon)=1$ and Var $(\varepsilon)= 1/\ \varphi$, the Poisson-Gamma function becomes a NB distribution (Lord, 2006; Cameron and Trivedi, 1998).

$$\ln(\lambda_{it}) = \alpha + \boldsymbol{\beta}\mathbf{X} \qquad 4.25$$

where $(\alpha, \boldsymbol{\beta})$ are the regression coefficients and $\boldsymbol{X}$ is the vector of covariates.

To obtain the FB estimates of the unknown parameters, prior distribution is required for the hyper parameters α, $\boldsymbol{\beta}$, σ^2, φ. If prior information is available it can be used to formulate informative prior distributions, otherwise non-informative priors are typically used. A diffused normal distribution with zero mean and large variance (e.g. $N(0, 10^3)$) can be used. The posterior distributions for all parameters can be calibrated by Monte Carlo Markov Chain (MCMC) analysis available in WinBUGS.

4.5.3.2 Model Selection

The commonly used model selection criteria such as Akaike information criterion (AIC) and Bayesian information criterion (BIC) are not applicable in complex hierarchical models because parameters may outnumber the observations. Instead, the Deviance Information Criterion (DIC) proposed by Spiegelhalter et al. (2003) is applied. DIC can be described as:

$$\mathrm{DIC} = \overline{\mathrm{D}} + \mathrm{p_D}$$

Where $\overline{\mathrm{D}}$ is the posterior mean of the deviance of the model and $\mathrm{p_D}$ is the effective number of parameters in the model. As a rule of thumb, differences of more than 10 might rule out the model with the higher DIC. Differences between 5 and 10 are substantial, but differences of less than 5 indicate that the models are competitive and it could be misleading just to report the model with the lowest DIC.

4.6 Conclusions

In this chapter we present two main approaches for causal analysis: the econometric causal models based on the Rubin potential outcomes framework, and the Bayesian framework, which is based directly on realized outcomes. As discussed earlier, the potential outcomes framework offers a number of advantages over the realized outcomes framework. In potential outcome framework, the causal effects can be defined before specifying the assignment mechanism without making distribution or functional form assumptions. Further, there is no requirement on endogeneity or exogeneity of the assignment mechanism when analysing individual-specific treatment effects using potential outcomes. In terms of the realized outcomes, the causal effects are more complex to define.

Most econometric models based on the potential outcomes framework, however, heavily rely on the unconfoundness assumption, which requires that unobserved factors are independent of the treatment assignment and the potential outcomes conditional on observed covariates. To make this assumption more plausible, data with specific requirements is usually collected. For example, the PSM method is a "data hungry" method in terms of both the number of covariates and the sample size. If the untreated group is not large enough, it is not possible to find adequate matches even if the average characteristics are similar. The DID method also rests on the availability of additional data in particular forms, e.g. panel data. In contrast, the FB methods may allow the estimation of crash models with small sample sizes, although the estimation results from the FB and conventional GLMs tend to be similar with large sample sizes. This will be discussed further in the following chapters.

A key issue in control studies is the selection of control groups with similar characteristics to the treatment group. In previous research, not only is there insufficient justification of the selection criterion, but how the treatment and control groups are matched is also unclear. In this chapter an EB approach using propensity scoring is proposed to address this issue. An application of this method to the evaluation of the safety effectiveness of speed cameras will be discussed in chapter 6. In the next chapter, the DID method will be applied to study the causal effect of the London congestion charge on road traffic casualties. In chapter 7, the FB approach will be employed to investigate the effect on road safety of changes in the road network. To sum up, figure 4.2 shows how different causal methods are selected and applied in this study.

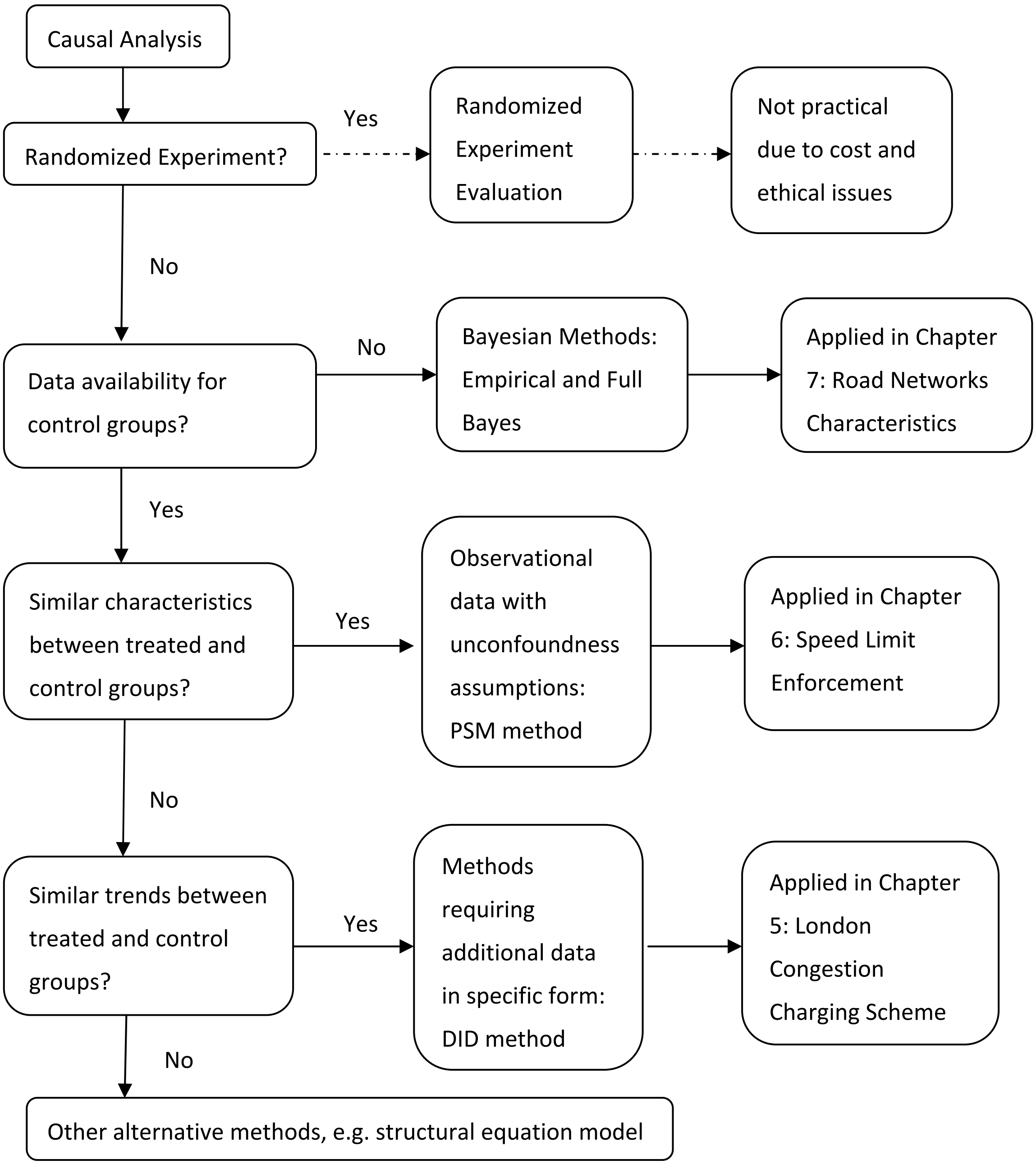

Figure 4.2 The diagram of the selection and application of causal model.

Chapter 5: A Causal Analysis using Difference-In-Difference Estimation: the Effects of Congestion Charging on Road Traffic Casualties

This chapter investigates the impacts of traffic demand management interventions on road casualties. The London congestion charge is studied using a causal difference-in-difference method. It is envisaged that by influencing travel modes and redistributing the traffic demand in space and time, traffic demand management interventions cause changes in both the number and type of casualties.

5.1 Introduction

Road pricing and taxation is a traffic intervention entailing the application of strategies and policies to reduce traffic demand, or to redistribute the demand in space or in time. These strategies and policies are usually linked to economic factors, such as fuel prices and road taxation. As fuel prices and road users' taxes rise, cars will tend to be driven less and consequently there will be less traffic congestion. It may be hypothesized that since a higher tax leads to fewer miles travelled, roads will be emptier of traffic and probably safer. There are additional complexities, however, because when road taxation is more expensive, travellers will switch to other travel modes such as bicycles and motorcycles, which may be more vulnerable to severe accidents (White, 2004; Leigh and Wilkinson, 1991; Crandall and Graham, 1989).

Many researchers have shown that there does seem to be a relationship between fuel prices and road casualties, although the direction of the relationship depends on the type of accidents analysed. Hyatt et al. (2009) investigated the relationship between motor vehicle injury and mortality rates and fuel prices. By using monthly fuel price and fatality panel data, they found higher gasoline prices were related to increased motorcycle casualties. They further explained that this increase was more a factor of the increasing number of motorcycles on the road. Grabowski et al. (Grabowski et al., 2006) generated panel data on the total number of traffic fatalities for the 48 continental U.S. states during the period 1982-2000. Their results suggested that exogenous increases in state gasoline taxes were plausibly associated with fewer traffic fatalities. Many other studies have also presented similar results (Leigh and Wilkinson, 1991; Haughton and Sarkar, 1996; Grabowski and Morrisey, 2004). All the above studies suggest that fuel tax, as one mode of road taxation, has an influence on road casualties. Similarly, we could hypothesize that congestion charging, another mode of road taxation which aims to alleviate congestion, may also affect traffic accidents. London provides a unique opportunity to study this hypothesis.

The objective of this chapter, therefore, is to test the causal effect of the London congestion charge (LCC) on road accidents. The Difference-In-Difference (DID) method is introduced as an evaluation tool to make causal inferences. The DID estimation approach is frequently applied in order to evaluate the impact of policies (Ashenfelter and Card, 1985; Card and Krueger, 1994; Finkelstein, 2002; Donald and Lang, 2007; Athey and Imbens, 2006; Abadie et al., 2010) but has, to the best of the author's knowledge, remained unused for road accident related transport research. The DID approach is applied using Generalized Linear Models (GLMs), such as Poisson and Negative Binomial models, to estimate the effect of congestion charging on the counts of accidents, which are categorized by casualty type and severity. Covariates are introduced to the DID model to adjust for factors that might lead to the violation of the parallel trend assumption in DID estimation.

This chapter is organized as follows. The literature review is presented in section 5.2. Section 5.3 describes the method and the data sources used in this analysis. Results are outlined and discussed in section 5.4. A comparative study using a synthetic control method is presented in section 5.5. The conclusions are given in the final section.

5.2 Previous Research

Causal relationships can be distinguished from pure statistical relationships if there is a plausible mechanism underpinning the relationship between target variables and the treatment (Elvik, 2011). To estimate the causal relationship between the LCC and road accidents, it is necessary to understand and reveal the mechanisms by which road pricing may affect road accidents.

Since the first congestion charging scheme was introduced by Singapore in 1975, several studies have been conducted to understand better the effects induced by congestion charging. It has been shown that congestion charging can decrease congestion effectively (Olszewski and Xie, 2005) and consequently that it can affect traffic flow conditions such as traffic volume, Volume-to-Capacity ratio (V/C) and traffic flow speed, all of which have direct impacts on the likelihood and severity of traffic casualties (Lord et al., 2005).

Tuerk and Graham (2010) conducted research on the impacts on traffic volumes of the LCC scheme. Traffic volumes crossing central London were measured by automatic traffic counters and aggregated at the hourly level. DID estimation was used to analyse the traffic data. The results showed a reduction in traffic due to the increase in the congestion fee for the LCC (7.8% for inbound vehicles).

Other papers have focused on the relationship between the congestion charge and travel mode, environment and business activity matters (e.g. Eliasson and Mattsson, 2006; Wichiensin et al., 2007). All these studies show that the congestion charge does have an effect on travel costs, travel time and the transit market. There has been little research, however, directly investigated the relationship between the LCC and road accidents in London.

Quddus (2008) conducted a time series analysis of traffic accidents in Great Britain and his results for the London congestion charge suggested an average 33% reduction of casualties after the charge was introduced. Noland et al. (2008) examined the effects of the LCC on traffic casualties by employing an intervention analysis. Data on traffic casualties used in this study was from 1991 to 2004, covering the 33 London boroughs. To account for serial correlation and seasonality effects, an intervention model (Box and Tiao, 1975) was used to analyse the effect of the congestion charge on traffic casualties. Although no significant effect was found for total casualties in the Greater London area, Noland et al.'s results suggested a significant drop in vehicle casualties and an increase in cycle casualties, which could be due to a switch in commuting modes. It is worth noting that the intervention model applied in their study cannot correct for effects due to nationwide trends which could have some broadly universal influence on accidents counts. Alternative evaluation methods need to be applied to verify the existing findings. An improved DID model which is appropriate for road accident data analysis is introduced in the next section.

5.3 Data and Model Specification

In this section, the data source is introduced and the DID model is specified in detail. Issues with implementing the DID method are also discussed here.

5.3.1 Dependent and Independent Variables

The data used for this analysis includes road accidents in the UK from 2001 to 2004. The casualty data are based on police records and collected by the UK Department for Transport (DfT) and are known as "Road accident data - GB", or the STATS 19 data base. The location of an accident is recorded using coordinates which are in accordance with the British National Grid coordinate system. Using Geographical Information System (GIS) software, such as MapInfo, every individual accident is located on the map and aggregated at the ward level. A ward is the primary unit of British administrative and electoral geography with an average area of 14 km^2.

To strengthen the parallel assumption, we add the following covariates to the DID model.

(1)The data for the road network is available from EDINA Digimap. This is used to calculate the length of road for each road type and the counts of junctions and roundabouts in every ward, since road characteristics have been proved to be related to road casualties by previous research (Noland and Quddus, 2004).

(2)The data for traffic exposure is not available at the ward level. Proxy variables are therefore employed to reflect the potential ward-level exposure. It is assumed that the internal traffic generation of ward i is proportionate to the population (P_i) and employees (E_i) in ward i. We also presume that the external traffic generation of ward I is affected by the population (PP_i) and employee (PE_i) of proximate ward j ($i \neq j$). The variable, proximate population and employee (PPE_i), which reflects the external generation of ward i can be constructed as follows (Graham and Stephens, 2008):

$$\mathrm{PPE_i} = \mathrm{PP_i} + \mathrm{PE_i} = \sum_{\mathrm{j}}^{\mathrm{i}\neq\mathrm{j}} \frac{\mathrm{P_j}}{\mathrm{d_{ij}}} + \sum_{\mathrm{j}}^{\mathrm{i}\neq\mathrm{j}} \frac{\mathrm{E_j}}{\mathrm{d_{ij}}} = \sum_{\mathrm{j}}^{\mathrm{i}\neq\mathrm{j}} \frac{\mathrm{P_j}+\mathrm{E_j}}{\mathrm{d_{ij}}} \qquad 5.1$$

Where d_{ij} is the centroid distance from ward i to ward j. The data for population and employment at the ward level is obtained from the Office for National Statistics (ONS). The data for population was further disaggregated by age cohorts and age percentage.

(3)Recent research suggests injuries of children are influenced by factors related to area deprivation (Graham and Stephens, 2008). The Index of Multiple Deprivation, published by the office for the Deputy Prime Minister (ODPM, 2004), is therefore used as a control variable. The Index of Multiple Deprivation integrates data on the following seven deprivation domain indices into one overall deprivation score: income, employment, housing and services, health, education, crime and environment.

5.3.2 Statistical Model

Due to the nonnegative integer nature of road traffic casualty count data, generalized linear models, including the Poisson and the Negative Binomial models, have been widely used to establish the relationship between traffic casualties and various risk factors.

A generalized linear model with the Poisson distribution is

$$\ln\mu_{it} = \boldsymbol{\pi}\mathbf{X}_{it} \qquad 5.2$$

The assumption that the variance is equal to the mean will be violated when the variance is significantly greater than the mean, which is also known as over-dispersion. To deal with this problem, the Negative Binomial model has been developed. A Gamma distributed error term is introduced to the Poisson regression model. The structure of the NB regression model is

$$\ln\mu_{it} = \boldsymbol{\pi}\mathbf{X}_{it} + \varepsilon_{it} \qquad 5.3$$

McDonald et al. (2000) considered the DID estimator (referred as Before-After Control-Impact in their study) when dependent variables are counts. Untransformed, log-transformed data and a generalized linear model (GLM) were applied in this study. McDonald et al. recommended the GLM for the analysis of count data because assumptions are more likely to be satisfied and interpretation of the estimated parameters is straightforward.

Hence, the same specification for the DID model as used by McDonald et al. (2000) is adopted here. Based on equation *5.2* and *5.3* the basic DID model is obtained as:

$$\ln\mu_{it} = \alpha + \beta T_{it} + \gamma G_{it} + \delta(T_{it} \cdot G_{it}) + \varepsilon_{it} \qquad 5.4$$

Further covariates are then introduced to equation *5.4*. This model is called here the full DID model, compared with the basic one, with the full model being described as:

$$\ln\mu_{it} = \alpha + \beta T_{it} + \gamma G_{it} + \delta(T_{it}\cdot G_{it}) + \boldsymbol{\pi}\mathbf{X}_{it} + \varepsilon_{it} \qquad 5.5$$

The percentage change in the number of accidents due to the effect of the LCC can be obtained as

$$\frac{\mu_{1T} - \mu_{0T}}{\mu_{0T}} = e^{\delta} - 1 \qquad 5.6$$

5.3.3 Groups and Periods Selection for DID

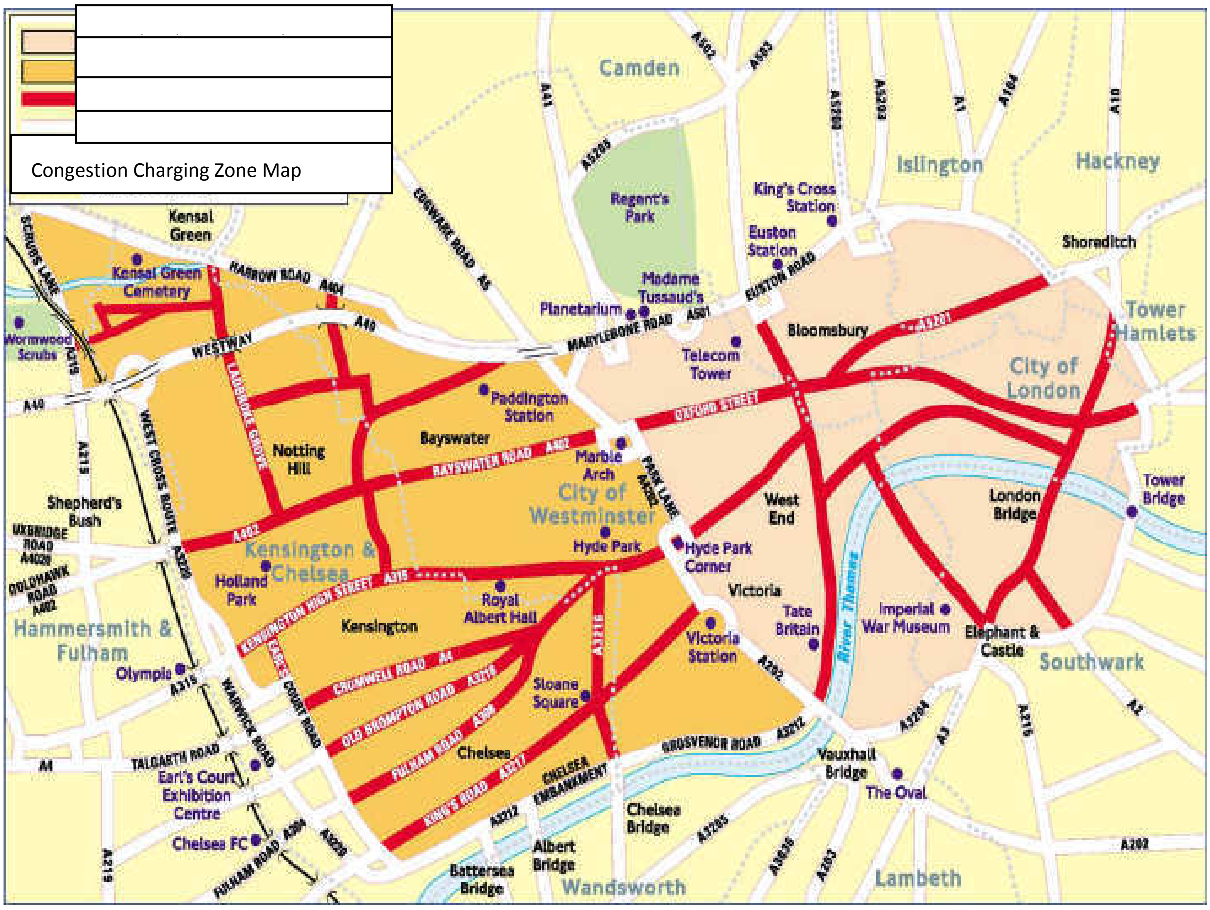

Figure 5.1 Map of the London Congestion Charging Zone.

The LCC scheme aims to reduce congestion and travel delay and thereby improve journey quality. The scheme was introduced in central London on 17 February 2003 at a flat rate of £5 per day between the hours of 7:00 am and 6:30 pm, Monday to

Friday. The charge was then raised from £5 to £8 on 4 July 2005. A western extension to the charging zone was implemented on 19 February 2007 and the charging hours were reduced to 7:00 am to 6:00 pm. The western charging zone was removed and the charge was increased to £10 on 4 January 2011. Figure 5.1 shows the area of the initial central London charging zone and the western extension. Congestion in central London reduced by up to 30 percent and average traffic speeds increased from 13km/h to 17km/h during the initial charging period (TfL, 2004). It has also been showed that the number of traffic accidents reduced significantly in both the original and extended charging zone (TfL, 2007). Periods and treatments are shown below:

(1) 2003-2004: Initial congestion charge in central London £5

(2) 2005-2006: Congestion fee increase from £5 to £8

(3) 2007-2011: Western extension of charging zone

(4) 2011-2013: Removal of the western extension. Charge increased to £10

In this study only stage 1 was investigated (i.e. the initial congestion charge in central London) because of the problem of analysing correlated multiple treatments. For instance, since the initial LCC may have impacts on the congestion charging zone which persist over time, it would be difficult to eliminate this source of confoundedness if the study were to attempt to estimate the effect of increasing the congestion fee. The focus, therefore, was on the initial LCC, with data for 2002 as the pre-treatment and for 2003 as the post-treatment.

Since the aim was to evaluate the effect of the London congestion charge the natural geographic extent of the treatment group was the central London area within the boundaries of the original LCC zone as portrayed in figure 5.1.

In respect to the control group it was important that it should exhibit independence of the treatment, i.e. the control group should not have received treatment either directly or indirectly through proximity to or interaction with treated groups. A typical example where this condition is violated is the area outside of the charging zone since it would be reasonable to speculate that the travelling and living habits of residents from the area surrounding London maybe influenced by the LCC. Because of data limitations, Wales and Scotland were excluded in this research, however. Although the DID method does not require the treatment and control groups to have the same demographic or traffic characteristics, cities with a large population and urban area are preferred as the control groups. Thus, data for conurbations which are geographically distinct from London was extracted as candidates for the control group, including the central areas of Manchester, Birmingham and Leeds.

As discussed before, the parallel trend assumption is valid conditional on the covariates *X*. To ensure the validity of this assumption, a pre-test was conducted, following Hastings (2004), using additional observations from pre-intervention years to assess whether the control group was able to mimic the temporal path of the treatment group. To give an example, Figure 5.2 shows the time trend of car casualties of London (the treatment group) and of Leeds, Manchester and Birmingham (potential control groups).

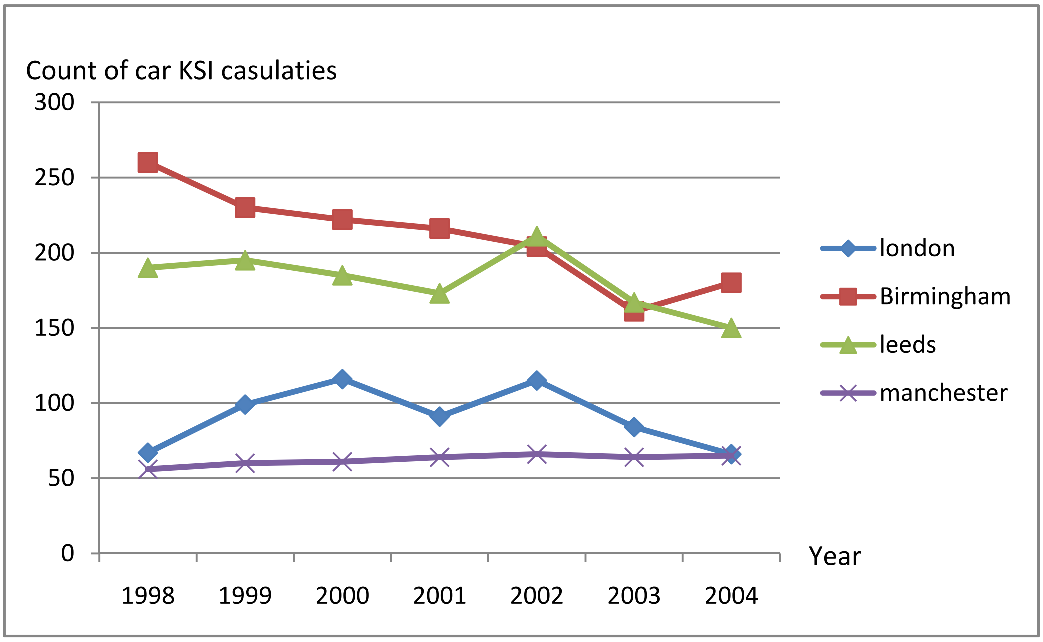

Figure 5.2 Time trend for count of car KSI casualties.

It can be seen that Leeds can best reflect the time trend of car casualties in London before the introduction of the London congestion charge in 2003. It is reasonable, therefore, to choose Leeds as the control group in the analysis of car killed and seriously injured (KSI) casualties. When the pre-test was conducted for bicycle and motorcycle casualties, however, Manchester and Birmingham were found to be the most suitable control groups. The result for control groups is presented below:

(1) Control group for car casualties analysis: Leeds

(2) Control group for bicycle casualties analysis: Manchester

(3) Control group for motorcycle casualties analysis: Birmingham

5.3.4 Issues in the DID method

There are two issues that need to be considered when using repeated data sets. The first is RTM (please refer to chapter 4 for detailed discussion). The traditional way to deal with RTM is to apply Empirical Bayes (EB). As a treated-control based approach, the full DID model introduced in this study shares the same idea that "accident counts are not the only clue to the safety of an entity; another clue is in what is known about the safety of similar entities" (Hauer, 2002). Conditional on the parallel assumption, the mean change in the control group provides an estimate of the change caused by RTM and any temporal effect. The difference between the mean change in the treatment group and the mean change in the control group is then the estimate of the treatment effect after adjusting for RTM.

Although EB has been used in many control studies, this method is not employed in this study for two reasons. First, although EB procedures can account for RTM and other effects over time, they rely on a large sample of reference groups, which ideally should have the same or similar characteristics as the treatment group but are unaffected by the treatment. This kind of reference group is not always available in practice and, in the context of this study, London is so unique in the UK that no comparable city can be found with the same or similar characteristics. In this case, the EB method is not feasible. Moreover, the word "similar" in the EB method is very ambiguous. Clearer justification for selecting reference groups is needed and this clearer definition is given in the DID method. In previous sections, two properties of the control group were discussed: (1) it should be independent of the treatment; (2) the treatment and control groups should have a parallel time trend of accidents count. Therefore, the DID method is more flexible and tractable, therefore. In section 5.5, an alternative approach for selecting the control group is discussed.

The second issue, as noted by Bertrand et al. (2004), suggests that the conventional DID estimation relying on repeated data sets may suffer from the problem of serial correlation. The standard errors will be underestimated and the t-statistics will bias upwards in the presence of positive correlation. Over-rejection of the null hypothesis can cause false inferences regarding the effect of treatment. Here, therefore, the Durbin-Watson (DW) test was applied to test for the presence of serial correlation in the residuals. The value of DW for all models was more than 1.5, which suggests no significant serial correlation is found in this study.

5.4 Main Findings

This section presents the results of our DID models. The STATS 19 data classifies the casualty by severity: KSI and slightly injured, which allows different types of casualties to be estimated separately. In this study, one casualty is defined as one accident with one or more persons injured or killed.

5.4.1 Model Selection

Results were obtained from two different specifications of the DID model. First, the basic DID model without any covariate adjustment was applied. The basic model includes only time fixed effects (CCYear), group fixed effects (CCZone), and the variable of interest (CCYear X CCZone). Next, *equation 5.5*, or the full DID model, was regressed to compare with the basic DID model by using Bayesian Information Criterion (BIC).

BIC values were obtained from all basic and full DID models for each class of injuries. The lower BIC values indicate that the full DID model is superior in interpreting the causal relationship between casualties and the LCC scheme. One possible reason is that only using the dummy variable CCZone cannot adequately explain the internal heterogeneity within the charging zone.

Both Poisson and Negative Binomial models were estimated for car, bicycle and motorcycle accidents. If the dispersion parameter is significantly greater than zero, then the Negative Binomial model provides a better fit than the Poisson model. A likelihood ratio test examined if the dispersion parameter equals zero. The associated chi-squared value and p value are included in the result and these together with the

BIC value strongly indicate that the Negative Binomial fits better than Poisson for most models, except for Models (2), (7), (8), (9) and (14).

Pearson correlation coefficients are calculated for variables and shown in Table 5.1. The highest correlation we find is under 0.8 suggesting that there is sufficient independent variance in the data.

Table 5.1 Pearson Correlation Coefficients of Covariates

	CCZoon	CCYear	CCYear X CCZone	Resident population	Resident population aged 0-15	Resident population aged 16-59
CCZoon	1					
CCYear	0	1				
CCYear X CCZone	0.5925	0.5458	1			
Resident population	-0.5605	0.0251	-0.3218	1		
Resident population aged 0-15	-0.714	-0.021	-0.4229	0.6464	1	
Resident population aged 16-59	-0.4173	0.0464	-0.2317	0.772	0.7126	1
Percentage of resident population aged 0-15	-0.5906	-0.0573	-0.3657	0.4931	0.7871	0.3275
Percentage of resident population aged 16-59	0.6589	0.0717	0.4133	-0.4502	-0.7276	-0.2543
Employee population	0.4923	0.0325	0.29	-0.1987	-0.4163	-0.098
Land area	-0.6979	0	-0.4135	0.4848	0.5468	0.3724
Employee population density	0.7153	0.001	0.4166	-0.5035	-0.6325	-0.7054
Resident population density	0.4651	0.0152	0.2874	0.2117	0.0084	0.3116
Length of minor road	-0.6636	0	-0.3932	0.6074	0.5868	0.5341
Length of motor road	-0.2944	0	-0.1744	0.1321	0.1867	0.0716
Length of A-road	0.0269	0	0.0159	0.3007	0.0811	0.3671
Length of B-road	-0.1473	0	-0.0873	0.3847	0.1798	0.4149
Density of minor road	0.5185	0	0.3072	-0.2322	-0.3477	-0.1356
Density of motor road	-0.2629	-0.0214	-0.1558	0.1137	0.1703	0.0582
Density of A-road	0.0763	-0.0064	0.0436	-0.3457	-0.3252	-0.3224
Density of B-road	0.2123	0	0.1258	0.1794	-0.0196	0.2435
Count of junctions	-0.292	0	-0.173	0.4471	0.3057	0.4536
Count of roundabout	-0.0425	0	-0.0252	0.0914	0.0573	0.0984
IMDscore	-0.6636	0	-0.3932	0.5163	0.7415	0.3825
PPE	0.5426	0	0.3585	-0.7066	-0.8116	-0.5687

Table 5.1 (Continued)

	Percentage of resident population aged 0-15	Percentage of resident population aged 16-59	Employee population	Land area	Employee population density	Resident population density
Percentage of resident population aged 0-15	1					
Percentage of resident population aged 16-59	-0.6971	1				
Employee population	-0.4788	0.4491	1			
Land area	0.4409	-0.5487	-0.1248	1		
Employee population density	-0.5235	0.5109	0.5831	-0.5466	1	
Resident population density	-0.0793	0.1528	-0.0457	-0.472	-0.0609	1
Length of minor road	0.3963	-0.4471	0.0764	0.7649	-0.5016	-0.3622
Length of motorway	0.1932	-0.2588	-0.0705	0.6128	-0.2022	-0.2484
Length of A-road	-0.0747	0.1108	0.5069	0.2826	-0.0319	-0.0779
Length of B-road	-0.0073	-0.0338	0.2299	0.2754	-0.1873	0.0419
Density of minor road	-0.3664	0.4595	0.3874	-0.5853	0.4709	0.3626
Density of motorway	0.1791	-0.2304	-0.1256	0.4587	-0.1918	-0.2047
Density of A-road	-0.3078	0.3306	0.4232	0.1797	0.4555	-0.5273
Density of B-road	-0.1284	0.1137	0.1452	-0.1163	-0.0325	0.4515
Count of junctions	0.126	-0.132	0.3606	0.4593	-0.219	-0.1875
Count of roundabout	0.0466	-0.0428	0.2447	0.4668	-0.0575	-0.0849
IMDscore	0.7658	-0.7081	-0.328	0.5237	-0.6569	-0.1709
PPE	-0.6516	0.7116	0.4434	-0.7268	0.8064	0.3397

Table 5.1 (Continued)

	Length of minor road	Length of motorway	Length of A-road	Length of B-road	Density of minor road	Density of motorway
Length of minor road	1					
Length of motor road	0.3785	1				
Length of A-road	0.5326	0.0342	1			
Length of B-road	0.3411	0.3744	0.3133	1		
Density of minor road	-0.1497	-0.3027	0.0086	-0.0624	1	
Density of motorway	0.2649	0.5106	-0.0464	0.4387	-0.2626	1
Density of A-road	0.2125	0.1129	0.4881	0.0134	-0.0117	0.0274
Density of B-road	-0.1433	0.1257	0.0261	0.7601	0.0811	0.2094
Count of junctions	0.864	0.1925	0.7753	0.4282	0.1169	0.1198
Count of roundabout	0.3035	0.4855	0.3085	0.2215	-0.2039	0.3391
IMDscore	0.6074	0.1909	0.1691	0.0473	-0.257	0.1596
PPE	-0.704	-0.2999	-0.0392	-0.1932	0.4937	-0.2633

Table 5.1 (Continued)

	Density of A-road	Density of B-road	Count of junctions	Count of roundabout	IMDscore	PPE
Density of A-road	1					
Density of B-road	-0.2167	1				
Count of junctions	0.3579	0.0016	1			
Count of roundabout	0.2714	0.0843	0.2614	1		
IMDscore	-0.1001	-0.1525	0.4085	0.1038	1	
PPE	0.1437	0.1489	-0.3421	-0.0683	-0.7267	1

5.4.2 Effects of London Congestion Charge on Road Accidents

The results show a distinct reduction in car casualties, despite an increase in bicycle and motorcycle accidents.

In terms of total car casualties, significant effects (at the 95% level) exist for the LCC with a coefficient (standard error) of - 0.054 (0.001), suggesting that the introduction of the LCC scheme reduced car casualties within the charging zone by 5.2% (Table 5.2). This is most likely due to the fact that the traffic volume has decreased since the introduction of the LCC (Tuerk and Graham, 2010), coupled with the fact that the total number of crashes decreases as the traffic volume decreases (Lord et al., 2005). The result for car KSI also shows a remarkable drop of 14.2% due to the LCC, while the reduction in slightly injured accidents is 4.6%.

According to the report from TfL (2004), there has been a decrease in the number of two-wheeled vehicles involved in accidents. The original data also suggests a reduction in the absolute number of cycle-related accidents from 1353 accidents (year 2002) to 1254 accidents (year 2003). After controlling for the time trend and regional effects, however, we find that an increase of 5.7% in total motorcycle casualties is related to the LCC, while there is an increase in bicycle casualties of 13.3% (Tables

5.4 and Table 5.6). The numbers of KSI for bicycle and motorcycle increase by 2.7% and 17.3% respectively during 2003. It is suggested that the year-on-year decrease in slightly injured two-wheeled accidents continues and is even greater in 2003 (TfL, 2004). We again fit separate models to bicycle and motorcycle slightly injured data sets. The model shows that the LCC scheme has resulted in an increase in both bicycle and motorcycle slightly injured casualties by 13.5% and 1.8% respectively.

These results are, to a large extent, consistent with the conclusions of the previous research by Noland et al. (2008). They hypothesized that this effect was down to the increasing number of two-wheeled commuters and the increased average network traffic speeds. Indeed, according to the annual report from TfL (TfL, 2004), inbound two-wheeled vehicles have increased by 15% while the average traffic speed has increased by 31% after the introduction of the LCC.

The results, therefore, suggest that the LCC scheme plays an important role in influencing traffic casualties in the London congestion charging zone.

Table 5.2 Full DID Models for Car Casualties

Full DID Models	Model 1 Car All Coef. (Std. Err.)	Model 2 Car KSI Coef. (Std. Err.)	Model 3 Car Slight Coef. (Std. Err.)
CCZoon	2.73E-01 (4.17E-02)	1.05E+00 (1.06E-03)	2.24E-01 (5.83E-02)
CCYear	-2.54E-01 (8.22E-03) *	-1.48E-01 (3.12E-02) *	-2.52E-01 (7.91E-03) **
CCYear X CCZone	-5.36E-02 (1.25E-04) *	-1.53E-01 (4.95E-02) *	-4.67E-02 (1.73E-04) *
Resident population	4.05E-04 (8.45E-05) *	8.45E-04 (3.73E-04) **	3.37E-04 (6.02E-05)
Resident population aged 0-15	-5.50E-04 (1.43E-04)	-9.65E-04 (2.03E-04) *	-4.78E-04 (1.50E-04)
Resident population aged 16-59	-1.61E-04 (2.42E-05)	-7.80E-04 (2.57E-05)	-1.16E-04 (1.48E-05)
Percentage of resident population aged 0-15	4.06E+00 (8.93E-01)	7.78E+00 (2.08E+00)	3.30E+00 (9.70E-01) **
Percentage of resident population aged 16-59	2.19E+00 (9.84E-01) **	7.32E+00 (2.52E+00)	1.87E+00 (1.06E+00)
Employee population	5.30E-06 (1.21E-06) *	7.71E-06 (3.61E-07) *	5.52E-06 (1.31E-06) *
Land area	1.31E-01 (4.05E-02)	2.60E-01 (1.49E-02) **	1.23E-01 (4.82E-02)
Employee population density	-2.31E-06 (1.33E-06) ***	-2.25E-06 (2.13E-06)	-3.69E-06 (1.61E-06) *
Resident population density	-8.01E-05 (1.50E-05) *	6.70E-05 (2.06E-05) *	-8.52E-05 (1.59E-05) *
Length of minor road	2.39E-05 (1.97E-05)	9.88E-05 (1.42E-05) *	2.32E-05 (2.25E-05)
Length of motorway	2.17E-04 (7.60E-05) *	-7.08E-04 (1.97E-04) *	-2.00E-04 (7.20E-05)
Length of A-road	1.08E-04 (3.74E-06) *	1.99E-04 (5.10E-05) *	1.02E-04 (8.51E-06) *
Length of B-road	-1.99E-04 (4.58E-05)	-1.94E-04 (1.22E-04)	-1.98E-04 (4.37E-05)
Density of minor road	1.17E-04 (1.26E-05) *	4.87E-05 (1.66E-06) *	1.19E-04 (1.47E-05) *
Density of motorway	2.40E-01 (9.79E-02) **	6.05E-01 (1.35E-01) *	2.24E-01 (9.32E-02) **
Density of A-road	8.48E-02 (7.80E-03)	-1.54E-02 (3.77E-02)	9.74E-02 (7.53E-03)
Density of B-road	6.37E+00 (4.40E-01) *	3.82E+00 (1.79E-01) *	6.45E+00 (4.16E-01) *
Count of junctions	-4.26E-03 (9.28E-04)	-9.25E-03 (1.48E-03) ***	-4.14E-03 (9.80E-04) *
Count of roundabout	9.57E-03 (1.96E-02)	-2.76E-02 (4.51E-02)	8.07E-03 (1.82E-02)
IMDscore	1.43E-02 (4.44E-03) *	1.16E-02 (3.09E-03) *	1.40E-02 (4.90E-03) *
PPE	-4.41E-04 (4.28E-04)	1.37E-04 (3.12E-04) **	-4.44E-04 (4.44E-04)
Constant	-1.80E+00 (3.49E-01) **	-9.08E+00 (2.44E+00) *	-1.48E+00 (3.65E-01) *
Obs	244	244	244
BIC	446.55	349.22	444.55
Likelihood-ratio test of alpha=0:	chibar2(01) = 76.73 Prob>=chibar2 = 0.000	chibar2(01) = 0.0e+00 Prob>=chibar2 = 0.500	chibar2(01) = 67.96 Prob>=chibar2 = 0.000

Notes: Figures are significant at: *99%, **95% and ***90%.

Table 5.3 Basic DID Models for Car Casualties

Basic DID Models	Model 4 Car All Casualties Coef. (Std. Err.)			Model 5 Car KSI Coef. (Std. Err.)			Model 6 Car Slight injured Coef. (Std. Err.)		
CCZoon	-3.81E-01	(4.92E-10)	*	7.20E-01	(2.49E-09)	*	-4.65E-01	(4.87E-10)	*
CCYear	-2.22E-01	(2.18E-11)	*	-2.01E-01	(2.18E-14)	*	-2.23E-01	(2.92E-11)	*
CCYear X CCZone	-3.17E-02	(5.43E-11)	*	-1.10E-01	(4.34E-10)	*	-1.19E-02	(1.12E-10)	*
Constant	3.94E+00	(4.92E-10)	*	6.93E-01	(3.47E-14)	*	3.90E+00	(4.87E-10)	*
Obs	244			244			244		
BIC	489.39			463.03			487.79		
Likelihood-ratio test of alpha=0:	chibar2(01) = 2245.44 Prob>=chibar2 = 0.000			chibar2(01) = 95.46 Prob>=chibar2 = 0.000			chibar2(01) = 2052.78 Prob>=chibar2 = 0.000		

Notes: Figures are significant at: *99%, **95% and ***90%.

Table 5.4 Full DID Models for Cycle Casualties

Full DID Models	Model 7 Bicycle All Casualties (Poisson) Coef. (Std. Err.)			Model 8 Bicycle KSI (Poisson) Coef. (Std. Err.)			Model 9 Bicycle Slight injured (Poisson) Coef. (Std. Err.)		
CCZoon	1.45E+00	(3.71E-01)	*	5.06E-01	(6.38E-02)	*	1.63E+00	(4.33E-01)	*
CCYear	-1.14E-01	(4.77E-04)	*	-1.73E-01	(1.71E-02)	*	-1.01E-01	(7.11E-03)	*
CCYear X CCZone	1.25E-01	(4.76E-03)	*	2.67E-02	(2.41E-02)		1.27E-01	(1.20E-03)	*
Resident population	1.14E-03	(7.83E-04)		1.29E-04	(5.42E-04)		1.34E-03	(8.51E-04)	
Resident population aged 0-15	-1.93E-04	(1.71E-03)		-1.53E-03	(8.09E-04)	***	2.91E-06	(1.92E-03)	
Resident population aged 16-59	2.39E-04	(4.26E-04)		-5.02E-04	(1.39E-03)		3.77E-04	(3.46E-04)	
Percentage of resident population aged 0-15	6.77E+00	(2.27E+00)	*	4.27E+00	(1.36E+00)	*	7.27E+00	(2.31E+00)	*
Percentage of resident population aged 16-59	5.52E+00	(2.53E+00)	**	4.95E+00	(1.59E+00)	*	5.71E+00	(2.64E+00)	**
Employee population	5.70E-06	(1.72E-06)	*	2.55E-06	(1.78E-06)		6.24E-06	(1.63E-06)	*
Land area	3.02E-02	(8.29E-03)	*	9.50E-04	(2.18E-02)		3.85E-02	(6.16E-03)	*
Employee population density	-3.41E-07	(8.63E-07)		-3.84E-06	(3.60E-07)	*	8.11E-08	(1.08E-06)	
Resident population density	-5.79E-05	(1.30E-05)	*	-7.08E-05	(3.28E-05)		-5.42E-05	(1.91E-05)	*
Length of minor road	-3.34E-05	(1.63E-05)	**	-3.21E-06	(2.01E-05)		-4.08E-05	(1.40E-05)	*
Length of motorway	-5.16E-05	(3.92E-05)		-4.63E-06	(5.31E-06)		-6.21E-05	(4.20E-05)	
Length of A-road	1.15E-06	(3.78E-05)		5.36E-05	(3.76E-05)		-9.09E-06	(3.56E-05)	
Length of B-road	-1.15E-04	(2.26E-05)	*	-2.37E-04	(8.69E-05)	*	-8.94E-05	(7.91E-06)	*
Density of minor road	3.39E-05	(3.55E-06)	*	-1.86E-05	(6.72E-05)		4.30E-05	(1.26E-05)	*
Density of motorway	4.28E-04	(1.68E-03)		1.36E-03	(1.29E-04)		-7.81E-05	(1.85E-03)	
Density of A-road	-1.21E-03	(1.50E-02)		-2.72E-02	(9.21E-03)	*	3.04E-03	(1.41E-02)	
Density of B-road	2.52E+00	(1.03E+00)	**	7.65E+00	(1.43E+00)	*	1.53E+00	(1.42E+00)	
Count of junctions	3.65E-03	(1.04E-03)	*	2.42E-03	(1.36E-04)	*	4.06E-03	(9.77E-04)	*
Count of roundabout	2.11E-02	(9.76E-02)		-1.51E-01	(1.02E-01)		5.65E-02	(9.55E-02)	
IMDscore	1.78E-02	(1.58E-03)	**	2.17E-02	(1.69E-02)		1.69E-02	(1.58E-03)	*
PPE	1.12E-03	(6.03E-05)	*	1.62E-03	(5.84E-04)	*	1.02E-03	(2.50E-05)	*
Constant	-6.22E+00	(2.88E+00)	*	-6.26E+00	(2.43E+00)	*	-6.78E+00	(2.98E+00)	*
Obs	244			244			244		
BIC	307.52			339.29			305.37		
Likelihood-ratio test of alpha=0:	chibar2(01) = 0.0e+00 Prob>=chibar2 = 0.500			chibar2(01) = 0.0e+00 Prob>=chibar2 = 0.500			chibar2(01) = 0.0e+00 Prob>=chibar2 = 0.500		

Notes: Figures are significant at: *99%, **95% and ***90%.

Table 5.5 Basic DID Models for Cycle Casualties

Basic DID Models	Model 10 Bicycle All Casualties Coef. (Std. Err.)			Model 11 Bicycle KSI Coef. (Std. Err.)			Model 12 Bicycle Slight injured Coef. (Std. Err.)		
CCZoon	1.24E+00	(2.02E-10)	*	7.42E-01	(8.99E-14)	*	1.32E+00	(1.18E-10)	*
CCYear	-7.67E-02	(5.11E-11)	*	-1.34E-01	(1.19E-17)	*	-6.45E-02	(3.63E-11)	*
CCYear X	2.95E-02	(2.08E-10)	*	2.18E-02	(5.07E-11)	*	6.81E-02	(1.20E-10)	*
Constant	1.67E+00	(1.06E-12)	*	-3.08E-02	(9.18E-17)	*	1.47E+00	(2.68E-15)	*
Obs	244			244			244		
BIC	499.39			473.27			493.89		
Likelihood-ratio test of alpha=0:	chibar2(01) = 460.26 Prob>=chibar2 = 0.000			chibar2(01) = 14.89 Prob>=chibar2 = 0.000			chibar2(01) = 409.68 Prob>=chibar2 = 0.000		

Notes: Figures are significant at: *99%, **95% and ***90%.

Table 5.6 Full DID Models for Motorcycle Casualties

Full DID Models	Model 13 Motorcycle All Casualties Coef. (Std. Err.)	Model 14 Motorcycle KSI (Poisson) Coef. (Std. Err.)	Model 15 Motorcycle Slight injured Coef. (Std. Err.)
CCZoon	2.55E+00 (1.63E-02) *	2.08E+00 (2.71E-01) *	2.62E+00 (7.59E-02) *
CCYear	-2.61E-01 (1.68E-02) *	-3.40E-01 (6.52E-03) *	-2.43E-01 (1.93E-02) *
CCYear X CCZone	5.57E-02 (1.15E-02) *	1.60E-01 (7.09E-03) *	1.83E-02 (4.59E-04) *
Resident population	2.48E-03 (2.80E-05) *	1.15E-03 (8.22E-04)	2.50E-03 (7.56E-04) *
Resident population aged 0-15	-3.60E-03 (2.39E-03)	-4.88E-04 (3.33E-03)	-3.34E-03 (1.51E-03) **
Resident population aged 16-59	-2.90E-03 (4.80E-04) *	-9.49E-04 (1.18E-03)	-2.92E-03 (3.71E-04) *
Percentage of resident population aged 0-15	8.35E+00 (6.01E-01) *	1.58E+00 (1.53E+00)	1.02E+01 (1.47E+00) *
Percentage of resident population aged 16-59	1.23E+01 (1.71E-01) *	4.00E-03 (5.24E-01)	1.46E+01 (1.03E+00) *
Employee population	4.07E-06 (5.46E-07) *	-3.18E-06 (1.18E-06) *	4.62E-06 (7.24E-07) *
Land area	2.27E-01 (8.87E-02) *	1.96E-01 (8.15E-02) **	2.43E-01 (1.09E-01) **
Employee population density	-7.51E-06 (1.86E-06) *	-4.82E-06 (8.48E-07) *	-4.57E-06 (1.12E-06) *
Resident population density	-4.92E-05 (1.21E-05) *	-9.38E-05 (2.47E-05) *	-4.50E-05 (2.81E-06) *
Length of minor road	-3.20E-05 (2.14E-05)	-5.20E-06 (3.79E-05)	-4.50E-05 (2.77E-05)
Length of motorway	-3.87E-04 (2.37E-05) *	-3.31E-04 (9.67E-05) *	-3.49E-04 (2.53E-06) *
Length of A-road	-8.66E-06 (1.11E-05)	7.69E-05 (8.73E-06) *	-2.87E-05 (2.85E-06) *
Length of B-road	3.22E-06 (1.10E-04)	7.47E-05 (1.00E-04)	-7.53E-06 (5.10E-05)
Density of minor road	4.34E-05 (7.40E-09) *	1.47E-04 (2.90E-05) *	1.95E-05 (5.34E-06) *
Density of motorway	2.89E-01 (6.58E-02) *	1.87E-01 (5.74E-02) *	2.79E-01 (6.51E-02) *
Density of A-road	7.39E-02 (1.15E-02) *	6.76E-02 (2.36E-02) *	5.30E-02 (1.30E-02) *
Density of B-road	1.38E+00 (1.36E+00)	2.26E+00 (1.06E+00) **	5.56E-01 (8.54E-01)
Count of junctions	2.45E-03 (1.27E-03) ***	-3.00E-03 (1.11E-03) *	3.96E-03 (1.72E-03) **
Count of roundabout	7.45E-02 (4.16E-03) *	1.60E-01 (7.05E-02) **	4.56E-02 (2.25E-02) **
IMDscore	2.08E-02 (4.76E-03) *	1.44E-02 (3.96E-03) *	2.34E-02 (5.74E-03) *
PPE	2.55E-04 (9.28E-05) *	-6.27E-05 (4.64E-04)	4.07E-04 (2.41E-04) ***
Constant	-1.19E+01 (2.93E-01) *	-3.68E+00 (2.00E-01) *	-1.42E+01 (7.61E-01) *
Obs	268	268	268
BIC	418.59	334.04	414.78
Likelihood-ratio test of alpha=0:	chibar2(01) = 18.07 Prob>=chibar2 = 0.000	chibar2(01) = 0.0e+00 Prob>=chibar2 =1	chibar2(01) = 5.35 Prob>=chibar2 = 0.01

Notes: Figures are significant at: *99%, **95% and ***90%.

Table 5.7 Basic DID Models for Motorcycle Casualties

Basic DID Models	Model 16 Motorcycle All Casualties Coef. (Std. Err.)	Model 17 Motorcycle KSI Coef. (Std. Err.)	Model 18 Motorcycle Slight injured Coef. (Std. Err.)
CCZoon	1.77E+00 (3.96E-08) *	1.18E+00 (1.89E-10) *	1.90E+00 (1.79E-08) *
CCYear	-2.49E-01 (4.28E-09) *	-4.05E-01 (1.03E-16) *	-2.06E-01 (3.56E-09) *
CCYear X	3.37E-02 (4.26E-09) *	4.24E-01 (2.22E-11) *	-4.90E-02 (6.65E-08) *
Constant	1.63E+00 (2.43E-08) *	1.67E-01 (1.97E-17) *	1.36E+00 (4.48E-11) *
Obs	268	268	268
BIC	453.37	471.63	475.89
Likelihood-ratio test of alpha=0:	chibar2(01) = 780.68 Prob>=chibar2 = 0.000	chibar2(01) = 38.86 Prob>=chibar2 = 0.000	chibar2(01) = 654.08 Prob>=chibar2 = 0.000

Notes: Figures are significant at: *99%, **95% and ***90%.

5.4.3 Control Variables in Full DID Models

The DID models provide some other interesting results on the effect of covariates on accident counts. The time effect variable CCYear has a significantly negative relationship with the number of traffic casualties (at the 99% level) for all models, indicating a nationwide downtrend of traffic accidents. This is in part possibly related to other traffic laws and policies, such as speed limits, seatbelt law and improvements in road infrastructures. The other dummy variable, CCZone, controlling for regional differences, also proved to be significant in most models.

The absolute numbers as well as the density of ward population and employment are used to control for traffic exposure within each ward. Most models show there are positive effects from the level of population and employment. This implies that more accidents may occur in wards with a higher number of residents and job opportunities. In contrast, coefficients for ward resident density in most models are negative, implying that regions with high resident density may experience fewer accidents. Considering many previous studies this result is unsurprising. The variable PPE, which reflects external traffic generation, shows positive effects in most models. This result suggests that a higher level of population and employment in proximate wards is related to more accidents. The ward population is further categorized by age cohorts and percentage, although these are not significant in all models.

Another set of independent variables are the characteristics of the road infrastructure, including the length and density of motorway, A-road, B-road and minor road. We find that increased car casualties are associated with motorway and A-road length. Since motorway and A-roads are designed mainly for high speed vehicles, it is reasonable to suppose that the risk for cars is relatively higher. The B-roads and minor roads are expected to have lower speeds and more cyclists and the results suggest that both B-road and minor road density have significant positive effects on cyclist casualties. The coefficients for counts of junctions and roundabouts are found to be less significant for car casualties, except cycle-related accidents. It is significant (at the 95% level) that more roundabouts and junctions in wards result in more cycle-related injuries. One interpretation may be the complexity of junctions and roundabouts and fewer safeguards, making cyclists more vulnerable.

Socio-economic deprivation has previously been shown to be positively related to road traffic casualties, and this has been confirmed by the results of this study which indicate that IMD scores have positive effects on all types of casualties.

5.5 Conclusions

This chapter has presented new evidence to show how traffic interventions affect road safety by influencing travel behaviours and mode choices. The results of the analysis also support the conclusions of chapter 4 that econometric causal models are superior to conventional statistical models in causal relationship analysis.

The Difference-In-Difference estimation method was employed, which can eliminate biases due to regional differences and nationwide trends, but this model has been further developed by combining it with generalized linear regression models. Covariates are included in the full DID to adjust for factors that violate the parallel trend assumption. By comparing BIC value it can be seen that full DID models performs better than basic ones.

The regression results suggest that there has been a significant reduction in car casualties within the congestion charging area. There has been a significant reduction of 5.2% in car casualties due to a reduction in traffic within the congestion charging area (Tuerk and Graham, 2010). Meanwhile, motorcycle- and cycle-related casualties have increased by 5.7% and 13.3% respectively after the LCC, probably because more motorcycles and bicycles have been used instead of cars (TfL, 2004). Our results are largely consistent with the conclusions of the previous research. For example, Noland et al. (2008) found a reduction of about 68 casualties per year or a drop of 3.4% and an increase of 16% in motorcycle-related casualties. Other variables have also been found to have affected the number of casualties significantly, including population, employment, deprivation and road infrastructure, thereby confirming the conclusions from previous studies.

This chapter, therefore, highlights that more attention needs to be paid to road safety strategies when introducing traffic demand management interventions such as road pricing and taxation. Policy makers need to be aware of the effect of such interventions in shifting travel modes among road users, which can create a large class of vulnerable road users who may experience a greater number of casualties.

Chapter 6: An Application of Propensity Score Matching Methods: the Impacts of Speed Cameras on Road Casualties

In contrast to some other traffic interventions, road safety law and policy sets clear standards and direct goals for reducing road casualties and improving road safety. It is, however, more likely to experience the problem of regression to the mean (selection bias) when evaluating the safety effects of the interventions. In this chapter, the propensity score matching method is applied to assess the safety effects of speed limit enforcement cameras. The empirical Bayes method is employed to validate the propensity score matching method. The results suggest that propensity score can be used as a criterion by which to select the control or reference groups in the before-after control study.

6.1 Introduction

The impacts of speed on accident severity are well known and speed limits in the UK define maximum desirable traffic speeds for the purposes of road safety. There are a number of policy measures that can be taken by governments in order to improve road safety by reducing traffic speed. An example of such a measure is that of speed limit enforcement cameras. These were first introduced in the UK in 1991 to persuade drivers to comply with these limits, and their use has grown in recent years. Numerous studies have shown that the introduction of speed cameras can help to reduce vehicle speeds as well as the number of road accidents. The main challenge in evaluating them is the construction of the counterfactual outcomes, i.e. what would have happened to the "treated" units in the absence of any treatment. Since the counterfactual outcome cannot be observed, statistical methods are used for its estimation, in particular naïve before-after and Empirical Bayes (EB) methods. Typically, a reference or control group is employed to estimate the counterfactual outcomes of the treatment group. Due to the confounding factors, however, the characteristics of treated and untreated units may differ in the absence of any treatment. In other words, the characteristics of units that are treated differ in some systematic way from those that are not treated, and those characteristics also have a bearing on the incidence of selection bias (i.e. regression-to-mean) and the severity of its impact. This means that only untreated units with similar characteristics to those treated can be used to approximate the counterfactual outcomes of the treatment group. A critical issue that has been inadequately addressed in previous studies is the selection of this reference or control group.

This study tackles this critical issue of the reference group by using the propensity score matching method (PSM), and subsequently uses this method to evaluate the effect of speed cameras on the reduction of road traffic accidents. The PSM method has become a popular approach to estimate causal treatment effects, but, to the best of the author's knowledge, remains untried in road traffic safety research. What makes the PSM method attractive is that it gives a clear criterion by which to select the reference or control group.

Unlike the traditional matching methods, which implement matching on multiple dimensions, the PSM enables matching to be reduced to a single dimension, the propensity score. This means that treated and untreated units with similar propensity scores can be compared to obtain the treatment effect. This method was used in this study to analyse a large dataset from speed camera sites in England in order to assess their efficacy.

This chapter is organized as follows. The literature review is presented in Section 6.2. The methods and data used in the analysis are described in Section 6.3 and Section 6.4. The results are presented and discussed in Section 6.5. The conclusions are given in the final section.

6.2 Literature Review

In the past decade, numerous studies have been conducted to investigate the impact of speed enforcement cameras on safety (Goldenbeld and van Schagen, 2005; Hess and Polak, 2003; Newstead and Cameron, 2003; Chen et al., 2002; Christie et al., 2003; ARRB Group Project Team, 2005; Mountain et al., 2004; Cunningham et al., 2008; Mountain et al., 2005; Shin et al., 2009; Keall et al., 2001; Gains et al., 2004, 2005; Jones et al., 2008).Table 6.1 summarizes these studies, which in general show that the implementation of speed cameras has significantly reduced vehicle speeds and the number of accidents near camera sites. In these studies, it is also suggested that the impact of speed cameras relates to various factors, such as types of road, the speed limit and site length, all of which will be adjusted for in this study. In addition, several outstanding issues which have yet to be fully addressed in the previous evaluations of the effects of speed cameras on road accidents will be addressed in this research.

Table 6.1 Summary of Studies on Safety Effects of Speed Cameras

Authors	Data and Units of Analysis	Methods	Main Findings
Goldenbeld and van Schagen, 2005	28 road sections; Speed data from 5 years after and 1 year before the enforcement; accident data from 5 years after and 8 years before the Enforcement.	Before-after study with comparison groups; long period data is used to control for RTM.	A reduction of 21% in injury accidents and serious casualties; significant decrease in mean speed and speed limit violators.
Hess and Polak, 2004	43 camera sites with 10 years of accident data in Cambridgeshire.	ARIMA/SARIMA	A reduction of 31.26% in accidents causing injury.
Newstead and Cameron, 2003	10 years of accident data for Queensland; segments of 2km, 4km and 6km are treated as treatment groups separately.	Quasi-experimental analysis with all areas outside camera sites as control groups.	A reduction of 45% in fatal crashes and a significant reduction in other kinds of crashes
Chen et al., 2002	22 km corridor with 12 photo radar locations; data from 2 years before and 2 years after is used.	Empirical Bayes with comparison groups.	A 2.8km/h reduction in mean speed; a decrease of 14% in collisions at photo radar sites, 19% reduction in collisions at non-photo radar sites and 16% total reduction along the corridor.
Christie et al., 2003	101 mobile speed camera sites in South Wales; 1996-2000 accident data extracted from STATS 19.	A circle zone based and a route based before-after analysis.	The route based method is superior than the circle based one and suggested a 51% reduction in injury crashes.
ARRB Group Project Team, 2005	28 speed camera sites in New South Wales; speed and accident data from 3 years before and 2 years after the operation of speed cameras.	Quasi-experimental analysis.	About 6 km/h fall in mean speed; Reduction of 23% for casualty crashes, 20% for injury crashes and 90% for fatal crashes.
Mountain et al., 2004	62 fixed speed camera sites with 1km upstream and downstream; accident data from3 years before and 3 years after the installation of cameras.	Empirical Bayes with UK national accident totals as comparison groups.	4.4 mph fall in mean speed; Reduction of 26% and 34% for overall injury accidents and fatal and serious accidents separately.
Christopher et al., 2008	14 corridors in Charlotte, Ohio; data from 4 years before and 1 year after the enforcement.	Before-after study with comparison groups.	Reduction of about 14% in collisions; around 7% decrease in mean speed.
Mountain et al., 2005	79 speed enforcement cameras with accidentfrom3 years before and 3 years after the enforcement.	Empirical Bayes is used to control for RTM.	A reduction of 22% in personal injury accidents; engineering schemes incorporating vertical deflections offer the largest benefits.

Table 6.1 Summary of Studies on Safety Effects of Speed Cameras (Continued)

Authors	Data and Units of Analysis	Methods	Main Findings
Shin et al., 2009	A 6.5 mile urban freeway; data covering1.5 years.	A before-after study with a comparison group; a before-after study with traffic flow correction; an Empirical Bayes before-after study	A 9mph fall in mean speed; a general reduction of 44-54% in total number of crashes.
Keall et al., 2001	Open roads with visible and hidden speed cameras; aggregated crash data for nearly 5 years.	An interrupted time-series design with a comparison group.	A reduction of 2.3 km/h in mean speed in speed camera areas; a net fall of 11% in the crash rate in the trial area.
Gains et al., 2003, 2004, 2005	24 areas with over 2300 speed camera sites in the UK; data from 3 years baseline and 3 years after programme.	A before-after study with a comparison group	A fall of around 7% in mean speed at camera sites; a33% reduction in personal injury collisions at speed camera sites.
Jones et al., 2008	29 camera locations in Norfolk, UK; monthly road traffic casualties data from 1999-2003.	A before-after study with comparison groups; effects of RTM are estimated.	A 1% decline in overall crashes and a 19% decline in crashes at speed camera sites.

Most studies to date have used naïve before-after methods with control groups (Goldenbeld and van Schagen, 2005; Christie et al., 2003; Cunningham et al., 2008; Gains et al., 2004; Jones et al., 2008). In these studies, a group of similar sites is usually selected as the control group in order to account for the general trend in accidents. This method is unable to control for the effects of the RTM, however.

The empirical Bayes (EB) method has been suggested by Hauer et al. (2002) as effective in controlling for the RTM effect. The EB approach relies on a large sample of reference groups, however, which ideally should have the same or similar traffic flow and road characteristics, i.e. the reference group must be representative of the treated sites. In previous research, not only is there insufficient justification of the selection of control groups, but how the treatment and control groups are matched is also unclear. In fact, this issue of similarity is also critical when selecting the control group in the before-after method.

Rosenbaum and Rubin (1983) propose the propensity score matching method as a solution to this problem of similarity. “Similar” groups can then be defined clearly as those with similar propensity scores, thereby avoiding selection bias and thus ensuring that the difference between the treatment and control groups can be attributed to the treatment.

A further issue that needs to be considered is that of accident migration. One manifestation of the phenomenon of accident migration is seen on roads near to speed camera sites if drivers should choose alternative routes to avoid these sites. In turn this may lead to an increase in accident numbers on alternative routes that avoid speed cameras whilst decreasing numbers at such sites. Mountain et al. (2005) show that speed enforcement cameras can affect route choice and that this has a significant effect on accidents at camera sites. The authors also suggest that the traffic flows before and after the installation of speed cameras should be monitored and accounted for in the model. Both Mountain et al.’s (2005) study and that of Christie et al. (2003) estimate the variation of speed camera effects at different distances from the cameras. There are two reasons for doing this. First, it is possible that drivers may decelerate and accelerate abruptly before and after the camera sites. This is known as the “kangaroo” effect and is another manifestation of accident migration. When this happens, it is necessary to know if there is unexpected increase in accidents upstream or downstream from the camera sites. Second, although the Department for Transport (2004) provides site selection guidelines on the length of camera sites, there is little knowledge regarding the most effective area for speed cameras. This study will use a similar definition of section length to that of Mountain et al. (2005) and Christie et al. (2003).

6.3 Method

This section discusses the details and issues when implementing the propensity score matching method. Then, an empirical Bayes model is specified to estimate the effect of speed camera. Results from both methods are compared and discussed later in the chapter.

6.3.1 Propensity Score Matching

In the case of a randomized experiment, the treatment status T_i is unconditionally independent of the potential outcomes Y_i. For non-randomized observational data, such independence cannot be achieved due to the confounding factors $\boldsymbol{X}$, that is, covariates that affect both the probability of treatment exposure and potential outcomes. Consequently, simple comparison of mean outcomes between treated and untreated groups will not in general reveal the causal effect. The conditional independence of potential outcomes and treatment status can be ensured, however, by adjusting for the vector of covariates $\boldsymbol{X}$, so that consistent causal estimates of treatment effects can be obtained.

A widely used method called the propensity score matching (PSM) is applied to evaluate the effect of fixed speed cameras on road accidents. The idea behind this method is to construct a control group that is similar to the treatment group in all relevant pre-treatment covariates $\boldsymbol{X}$. Instead of matching directly on all the covariates $\boldsymbol{X}$, PSM has the advantage of reducing the multiple dimension of matching to a single dimension, the propensity score, which is the probability of receiving a treatment. Conditional on the propensity score, differences in observed outcomes between the two groups can be solely attributed to the intervention impacts. In other words, adjusting for the propensity score is enough to eliminate the bias created by all confounding factors.

Notation

The treatment indicator is defined as T_i, where T_i=1 if individual i receives the treatment and T_i =0 otherwise. Let $Y_i(T)$ denote the potential outcome for individual i, where i =1,…, N and N denote the total population. The treatment effect for individual i can be described as:

$$\delta_i = Y_i(1) - Y_i(0). \qquad 6.1$$

In practice, the parameter of interest is usually the average treatment effect on the treated (ATT), which can be defined as:

$$\delta_{ATT} = E(\delta|T = 1) = E(Y(1)|T = 1) - E(Y(0)|T = 1) \qquad 6.2$$

For a treated unit, *Y(0)* is the counterfactual. In a random assignment experiment, the δ_{ATT} can be estimated as the simple difference in mean outcomes between those who receive the treatment and those excluded from the treatment. Such random assignment is usually not feasible, however, due to high costs and ethical issues. It is commonly the case that assignment is not random, where the possibility of receiving the treatment may be related with factors affecting *Y(0)* and *Y(1)*, so the estimation of the treatment effect using difference in outcomes between participants and non-participants could be biased. The PSM method is designed to ensure that estimates are made between comparable individuals and that the bias created by all treatment confounders is eliminated.

6.3.2 Implementing PSM

The procedure for estimation of treatment effects using PSM can be illustrated in three steps:

(1) Specify a discrete outcome model for estimating the propensity score

(2) Choose an algorithm to match treated and untreated individuals in terms of the propensity score

(3) Estimate the treatment effect

Estimating the propensity score

The first step when implementing PSM is to estimate the propensity score. For a binary treatment variable, logit and probit models are usually preferred to a linear probability model, which may generate predictions outside the [0, 1] bounds of probabilities. Because logit and probit models usually produce similar results, the choice is not critical. In this study, a logit model is used, which is described as follows:

$$P\,(T=1 \mid \mathbf{X}) = \frac{\mathrm{EXP}(\alpha+\boldsymbol{\beta}'\mathbf{X})}{1+\mathrm{EXP}(\alpha+\boldsymbol{\beta}'\mathbf{X})} \qquad 6.3$$

Where α is the intercept and $\boldsymbol{\beta}'$ is the vector of regression coefficients.

One key issue when specifying the propensity score model is the inclusion of covariates. To satisfy the unconfoundedness condition, variables influencing both the selection of treatment groups and potential outcomes should be included in the model. If there are explicit criteria used in the treatment group selection, such criteria must be included in the propensity score model. Other factors can be decided based on a sound knowledge of previous studies. It is also suggested that over-parameterized models should be avoided (Bryson et al., 2002; Augurzky and Schmidt, 2000). There are two reasons for this. First, including extraneous covariates may violate the overlap condition. Second, the inclusion of non-significant covariates can increase the variance.

In conclusion, the following covariates should be included in the model based on these criteria:

(1) Covariates that strongly influence the selection into the treatment and outcomes;

(2) Covariates which are statistically significant in the regression model;

(3) Covariates that have been suggested as important factors affecting outcomes in previous research.

Matching Algorithm

After estimating the propensity score, a matching algorithm is selected to construct the control group from non-treated individuals. The matching algorithms have been described in detail in chapter 4.

There is no theoretical guidance on how to select the most appropriate matching algorithm. Given a large sample, the result from all algorithms should be similar and therefore the choice is not critical. The matching approach can have a considerable impact on the results when only a small sample is available, however. In this case, the decision can be made based on the distribution of the estimated propensity score for treated and untreated groups. For example, if some treated units have many close neighbours while others do not, the kernel and local linear matching methods would be more appropriate. In general, the reasonable way is to try every matching approach. If there is a large difference in the results, further judgment and consideration is required.

Estimating treatment effects

Once treated units have found matches from the untreated group, the treatment effect can be evaluated by taking differences in outcomes between treated units and their matches. A number of programs are available for STATA and other statistical software programs. Bootstrapping methods are usually applied to calculate standard errors. The program used in this study is **psmatch2** in STATA, which was developed by Becker and Ichino (2002).

Difference-In-Difference (DID) Matching

The conditional independence assumption is too strong and may not hold when unobserved factors that may influence outcomes are not included in the model. However, the CIA can be relaxed by using the DID matching estimator (Heckman et al., 1997). Given data from the pre-treatment period, any time-invariant confounder can be controlled for. In the DID matching approach, the dependent variable is the difference between pre-intervention and post-intervention periods. The treatment effect can be simply calculated by applying the procedures described above.

6.3.3 Issues in PSM estimation

Two important issues in PSM estimation are discussed in this section. First, the validity of the overlap condition needs to be checked. It is assumed that there is a positive probability of receiving the treatment for both the treated and untreated units conditional on covariates ***X***. The most straightforward way to evaluate this condition is by the visual inspection of the density distribution of the propensity score for both the treatment and control groups. The histograms of the distribution of propensity scores for both groups, together with a comparison of the minima and maxima values in each group, can help to provide clear knowledge of the extent to which there is overlap in the propensity score between the two groups.

The second issue regards the assessment of the matching quality. It should be noted that the real purpose of matching is to balance the characteristics between the treatment and control groups. Rosenbaum and Rubin (1983) propose a theorem stating that after conditioning on $P\ (T=1 \mid \boldsymbol{X})$, covariates ***X*** should be independent of the treatment decision.

$$\mathbf{X} \perp \mathrm{T} \mid \mathrm{P}(\mathrm{T}=1 \mid \mathbf{X})$$

A balancing test was proposed by Rosenbaum and Rubin (1983) and applied by Dehejia and Wahba (2002). In this test, the units are first divided into blocks with similar propensity scores. Within each block, the t-test is used to determine whether the distribution of covariates ***X*** is the same in both groups. If a block has unbalanced covariates then that block is divided into smaller blocks and the evaluation is repeated. If differences still remain then the specification for the propensity score is revisited and higher-order or interaction terms are included. The process is then repeated from the beginning. Such a stratification test can be done using the program **psmatch2** in STATA.

It is worth noting that the violation of the assumptions underpinning the PSM can bias the estimates. In some cases, when there is little theoretical or empirical evidence on the nature of selection into a treatment, it is difficult to know which factors influence participation and outcomes and thus what set of covariates should be included in the propensity score model. Since estimates of treatment effects can be sensitive to the covariates used this may bias estimates. This may happen in particular with innovative treatments about which there is little prior knowledge. In such circumstances instance, pre-intervention research would be helpful to identify the covariates involved. It is also possible that the selection into treatments is driven by factors that are not observable, and thus the matching estimator may be seriously biased. As discussed earlier, with additional pre-intervention data, a modified version, the DID matching method can be applied to correct for some of this bias, as long as the effects of unobserved factors are fixed over time.

In summary, the procedures for using the PSM to evaluate the safety effects of speed cameras can be illustrated as following steps.

(1) The data for both camera sites and comparison sites, such as accident records and site information, is collected and constructed in a single data set.

(2) Covariates are selected to be included in the logit model in order to estimate the propensity score for both groups, which is the probability of being treated as camera sites.

(3) The distributions of propensity scores are compared between camera sites and comparison sites to check the overlap condition. If the condition is not satisfied, then covariates included in the logit model need to be re-selected.

(4) It is recommended that multiple matching algorithms are applied to increase the credibility of the PSM.

(5) A balancing test is conducted to test whether the camera sites and comparison sites are statistically similar after matching. If significant differences are found, the logit model is re-specified and the process is repeated from the beginning.

(6) The safety effects of speed cameras can be evaluated by taking differences in outcomes between matched camera sites and comparison sites. In the original PSM, outcomes are the observed number of accidents in the post-intervention period, while the outcome variables are the difference between pre-intervention and post-intervention in the DID matching.

Figure 6.1 shows a flow-diagram illustrating how the PSM can be applied to the estimation of the safety effects of speed cameras.

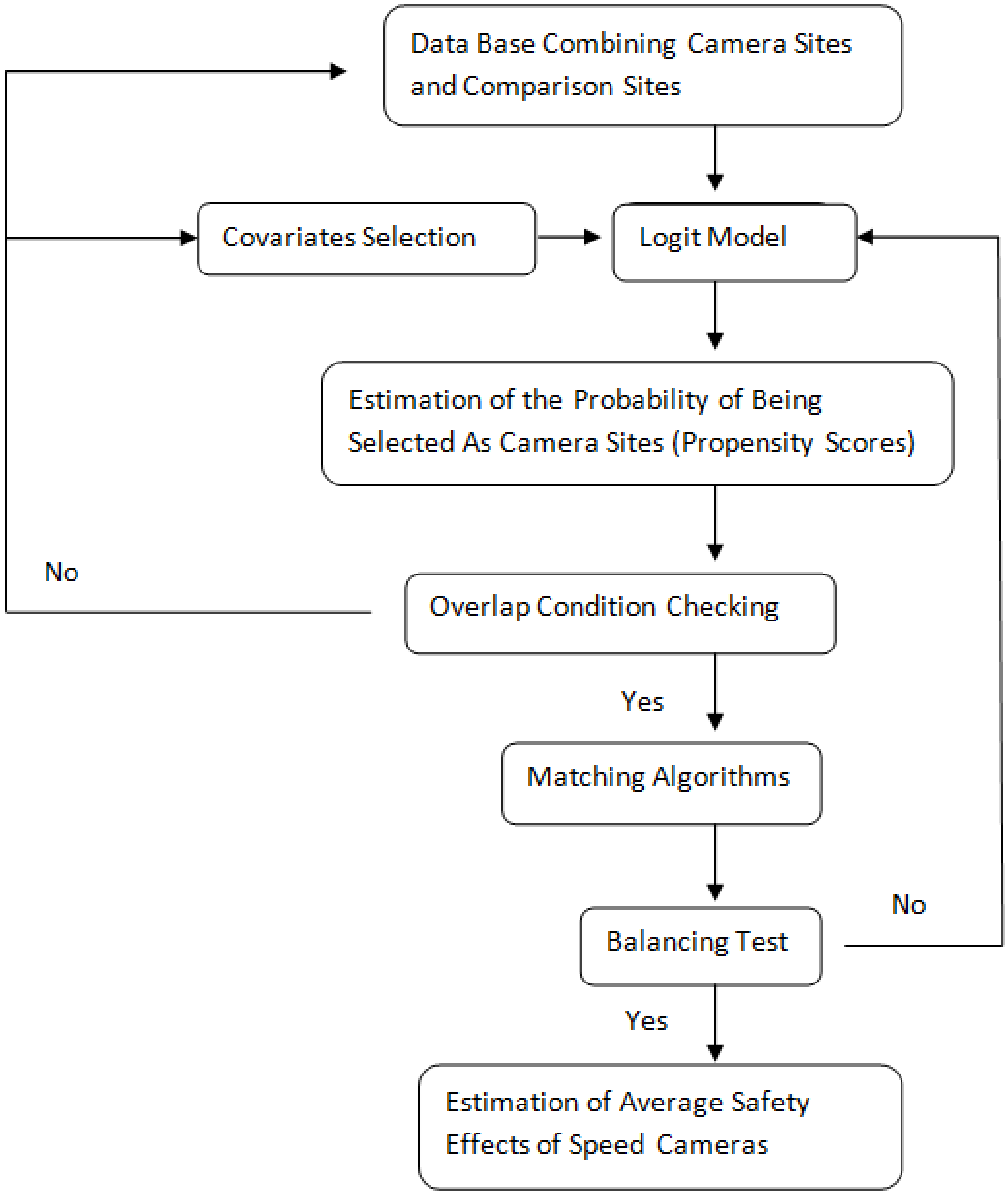

Figure 6.1 Diagram of the application of the PSM to the evaluation of safety effects of speed cameras

6.3.5 The Bayesian Approaches using Propensity Scores

Empirical Bayes (EB) methods have been introduced and widely used in before-and-after traffic safety countermeasures evaluation (Hauer, 1997; Hauer et al., 2002; Persaud et al., 2009; Persaud and Lyon, 2007;Sayed et al., 2004;Hirst et al., 2004).

Recent applications in the analysis of speed camera effects include studies by Mountain et al. (2005), Shin et al. (2009), and Gains et al. (2005). In the EB approach, the predicted number of crashes without treatment is derived by combining the observed crash counts in the before period and the expected number of crashes from Safety Performance Functions (SPFs).

The validity of the EB approach relies heavily on the availability of a proper reference group since an inappropriate reference group can bias the estimation of SPFs. The propensity score, however, can be used to find untreated sites that are similar to treated sites, thereby enabling an appropriate reference group to be constructed. In this study, two reference groups were applied in the EB approach. One reference group was selected based on the propensity score, while the other one contained all the potential reference sites. Comparisons were conducted between results from the EB and PSM models.

The SPF used in this study was based on the model proposed by Mountain et al. (1997), which has also been used in other studies (Gains et al., 2005; Mountain et al., 2004, 2005).

The number of observed crashes y can be modelled as:

$$y \sim \mathrm{Poisson}(\mu\varepsilon)$$

$$\log \mu = \alpha + \log L + \beta_1 \log V + \beta_2 \left(\frac{J}{L}\right) + \varepsilon \qquad 6.4$$

Where L is the site length, V is the AADF at each section and J is the number of minor junctions within the site length. E is a Gamma distributed random error term.

The EB estimate of total number of crashes $\widehat{\mu_B}$ in a before period of t_B years is

$$\widehat{\mu_B} = t_B \hat{\mu} \qquad 6.5$$

Then, the predicted number of crashes in the before period, $\widehat{M_B}$ can be obtained by

$$\widehat{M_B} = \rho\widehat{\mu_B} + (1 - \rho)X_B \qquad 6.6$$

where X_B is the observed number of crashes in the before period and

$$\rho = (1 + \frac{\widehat{\mu_B}}{\varphi})^{-1} \qquad 6.7$$

φ is the shape parameter for the NB distribution.

The estimate of the number of crashes in the after period had the treatment not occurred, $\widehat{M_A}$, can be calculated after adjusting the time trend effect:

$$\widehat{M_A} = (\frac{N_{A\_POP}}{N_{B\_POP}})\widehat{M_B} \qquad 6.8$$

where N_{A_POP} and N_{B_POP} are the numbers of crashes for the total population in the before and after periods, respectively.

To control for the effect of any flow changes due to the treatment, the expected flow in the after period had the treatment not occurred, V'_A can be estimated as:

$$V'_A = \left(\frac{V_{A\_POP}}{V_{B\_POP}}\right) V_B \qquad 6.9$$

where V_{A_POP} and V_{B_POP} are the traffic flow for the whole population in the before and after periods, respectively, and V_B is the observed traffic flow in the before period.

The estimate of the number of crashes in the after period can be refined as

$$\widehat{M'_A} = (\frac{V_A}{V'_A})^{\beta_1}\widehat{M_A} \qquad 6.10$$

where V_A is the observed traffic flow.

The treatment effect can then be obtained as:

$$\delta_{ATT} = \frac{\frac{X_A}{t_A} - \frac{\widehat{M'_A}}{t_A}}{X_{B/T_B}} \qquad 6.11$$

6.4 Data

6.4.1 Covariates

Clear knowledge of the criteria for treatment assignment can help determine which covariates should be included in the propensity score model and hence improve the validity of the PSM method. Currently, in the UK, formal site selection guidelines for fixed speed camera sites exist (Gains et al., 2004), as shown below.

(1) Site length: Between 400-1500 metres.

(2) Number of fatal and serious collisions (FSCs): at least 4 FSCs per km in the last three calendar years.

(3) Number of personal injury collisions (PICs): at least 8 PICs per km in the last three calendar years.

(4) 85th percentile speed at collision hot spots: 85th percentile speed at least 10% above speed limit.

(5) Percentage over the speed limit: at least 20% of drivers are exceeding the speed limit.

The first three guidelines can be thought of as primary criteria and the latter two as secondary criteria. Notwithstanding these guidelines there are sites not meeting the above criteria which may still be selected as enforcement sites for one or more of the following reasons (DfT, 2004):

• Community concern - where the local community requests the authorities enforce at a particular site because traffic speeds there are causing concern for road safety

• Collision frequency - where a site has a high incidence of PICs, but an insufficient number of FSCs collisions to meet the criteria, but where there is well-founded concern that a failure to reduce speeds or red-light running at this site will result in future increases in FSCs collisions, including fatalities

• Engineering factors - where roads (or parts of roads) do not meet minimum engineering requirements. Enforcement at such sites should be a short-term measure only until the local authority rectifies the problem.

Selection of speed camera sites, therefore, is primarily based on accident history. Accident data can be obtained from the STATS 19 database and located on the map using MapInfo. Secondary criteria such as the 85th percentile speed and percentages of vehicles over the speed limit are not normally publically available for all sites on UK roads, however. If speed distributions differ between the treated and untreated groups, then the failure to include the speed data could bias the estimation, an issue discussed in previous research (Mountain et al., 2005; Gains et al., 2004*2*). For untreated sites with a speed limit of 30 mph or 40 mph, the national average mean speed and percentages of speeding are similar to the data for the camera sites. The focus groups for this study are sites with a speed limit of 30 mph and 40 mph throughout the UK. It is reasonable to assume that there is no significant difference in the speed distribution between the treated and untreated groups and hence exclusion of the speed data will not affect the accuracy of the propensity score model.

In order to account for the "kangaroo" effect (Elvik, 1997; Thomas et al., 2008), in this study, the effective length of camera sites was determined by investigating a different range of distance to camera sites, including 200m, 500m and 1km.If a route meets a major junction or a route for another camera site, however, that route is terminated (Christie et al., 2003). Some sites, therefore, cannot be evaluated for all distance bands. The percentages of such sites are 0% for 200m sites, 3% for 500m sites and 9% for 1km sites respectively.

It is also possible that drivers may choose alternative routes to avoid speed camera sites. Accident reduction at camera sites may include the effect induced by a reduced traffic flow. The benefits of speed cameras will therefore be overestimated if traffic flow is not controlled for. The annual average daily flow (AADF) is available for both treated and untreated roads and the effect due to traffic flow was controlled for in this study by including the AADF in the propensity score model.

In addition to the criteria that strongly influence the treatment assignment, factors that affect the outcomes should also be taken into account when the propensity score model is specified. This study, therefore, further included road characteristics such as: road types, speed limit, and the number of minor junctions within the site length, which have been suggested as being important factors when estimating the safety impact of speed cameras (Gains et al., 2005; Christie et al., 2003; Mountain et al., 1997).

The final propensity score model can be described as below:

$$P\,(T=1 \mid \mathbf{X}) = \frac{\mathrm{EXP}(\alpha+\beta_1 C_{FSC}+\beta_2 C_{PIC}+\beta_3 L+\beta_4 V+\beta_5 A+\beta_6 B+\beta_7 M+\beta_8 S_{30}+\beta_9 S_{40}+\beta_{10} J)}{1+\mathrm{EXP}(\alpha+\beta_1 C_{FSC}+\beta_2 C_{PIC}+\beta_3 L+\beta_4 V+\beta_5 A+\beta_6 B+\beta_7 M+\beta_8 S_{30}+\beta_9 S_{40}+\beta_{10} J)} \qquad 6.12$$

where C_{FSC} and C_{PIC} are the FSCs and PICs in the last three years before the camera installation. The road type is defined by binary indicators, A (A road), B (B road) and M (Minor road). S_{30} (S_{40}) equals 1 if the speed limit is 30mph (40mph) and 0 otherwise. Other covariates are as defined above.

6.4.2 Sample Size

The PSM method is known as a "data-hungry method" in terms of the number of treated and untreated units. Matching can only be implemented when there is sufficient overlap between both treatment and control groups for every propensity score block. If no match can be found for treated units at some propensity scores, these treated units will be discarded and the estimation of the ATT will be biased.

Thus, a large untreated pool is required to ensure adequate matches. The literature is not explicit, however, on how large the untreated group should be. According to previous research, the ratio of the number of control group candidates to the number of treatment group members ranges from 1.5:1 to over 30:1(Hirano and Imbens, 2001; Dehejia and Wahba, 2002; Kurth et al., 2006;Smith and Todd, 2005; Peikes et al., 2008). The ratio chosen in this study was around 7:1, which was assumed to be sufficient to ensure the matching quality. Due to data restrictions, 771 camera sites from the following eight English administrative districts were included in the treatment group: Cheshire, Dorset, Greater Manchester, Lancashire, Leicester, Merseyside, Sussex and West Midlands. A total of 4787 potential control sites were selected randomly within these districts. The accident data for the three years before and after the camera installation were acquired for every site and the research period covered nine years from 1999 to 2007. Whilst concerns have been raised about the completeness and reliability of accident data in STATS19, in the case of accidents at speed camera sites, given the nature of such sites, it is likely that all accidents were captured and that the data is therefore reliable and complete.

It is worth noting that there is a difference in the data requirement for the EB and PSM methods. Although both methods rely on a large sample of untreated units, the EB method further requires that the untreated group must be representative of the treated sites in order to estimate the SPFs, which constrains the application of the EB method.

6.5 Results and Discussion

6.5.1 The estimation of propensity scores

The first step in the propensity score matching method is to estimate the probability of being selected in the treatment group. The logit model is regressed on the covariates and the covariates that influence the participation and the outcome should be included in the model. Table 6.2 shows that all covariates except minor roads are significant in the estimation of the propensity score. This is probably because there are very few speed cameras installed on minor roads in the study sample. The result confirms that the covariates included in the propensity score model are important in predicting the possibility of being selected as camera sites.

Table 6.2 The propensity score model

	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]	
Number of minor junctions	0.023	0.007	3.33	0.001	0.009	0.036
AADF in baseline years	1.30E-05	2.36E-06	5.52	0.000	8.40E-06	1.76E-05
PICs in baseline years	-0.013	0.003	-4.13	0.000	-0.019	-0.007
FSCs in baseline years	0.159	0.018	8.67	0.000	0.123	0.194
Site length	-0.141	0.064	-2.20	0.028	-0.267	-0.015
A Road	-0.377	0.128	-2.95	0.003	-0.627	-0.126
B Road	-0.307	0.135	-2.27	0.023	-0.572	-0.042
Minor Road	-0.078	0.193	-0.40	0.686	-0.457	0.301
Speed Limit 30mph	1.017	0.101	10.11	0.000	0.820	1.214
Speed Limit 40mph	0.594	0.106	5.61	0.000	0.387	0.802
Constant	-1.876	0.168	-11.14	0.000	-2.206	-1.546
Observations	5558					

6.5.2 Tests of matching quality

Before estimating the effects of speed cameras, the validity of the PSM method must be checked. One approach is through a visual inspection of the propensity score distribution for both the treatment and comparison groups. From the histograms of propensity scores for both groups, the extent to which there is overlap in the scores between the treatment and comparison groups is apparent. Observations that fall outside the region of common support must be discarded and cannot be estimated.

The estimation will be unaffected if the proportion of discarded observations is small (Bryson et al., 2002). If the proportion is too large, however, the estimated treatment effect could be inaccurate. Figure 6.2 shows the distribution of propensity scores for both groups. For the treatment and the potential comparison groups 771 sites and 4787 sites were observed, respectively. Only seven treated sites were found to be outside the region of common support and discarded. It can be concluded, therefore, that there is sufficient overlap of the distributions.

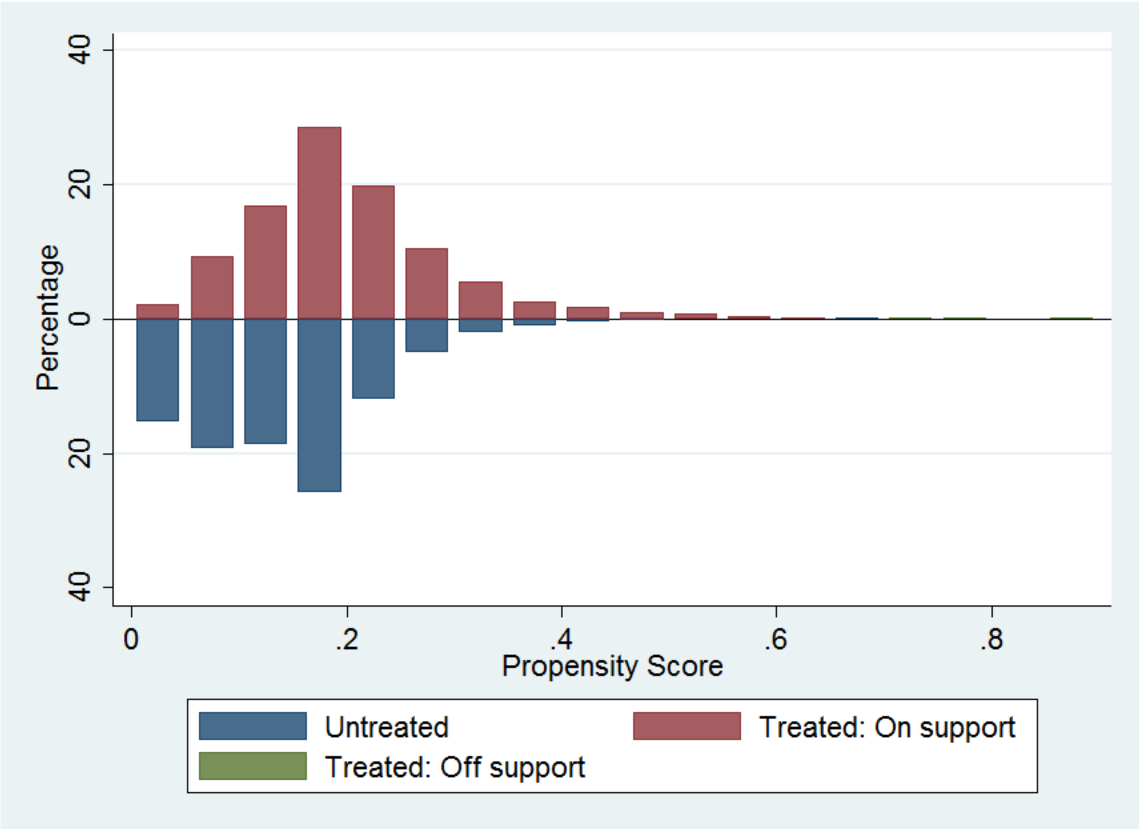

Figure 6.2 Propensity score distribution.

The next step was to perform balancing tests to assess the matching quality since these tests can verify that treatment is independent of the covariates after matching. The PSM method aims to balance characteristics between the treatment and comparison groups, i.e., there should be no significant differences between the covariate means of the treatment and comparison groups after matching. Table 6.3 shows the t-test of differences in covariate means before and after the matching. It can be seen that there are significant differences in all covariates, except site length, when using all sites as the comparison group. It is clear that the characteristics between groups are imbalanced and the estimation of the treatment effect could therefore be biased. The PSM method was subsequently used to construct matched comparison groups. Table 6.3 shows that all covariates were now balanced between the treatment and matched comparison groups and that, consequently, the bias due to the differences in observable characteristics was reduced.

Table 6.3 Checking the covariate balance between groups before and after using nearest neighbours (k=5) matching

Variable	Sample	Mean Treated	Mean Control	%bias	%reduced |bias|	t-test t	p>|t|
Number of minorjunctions	Unmatched	5.4578	3.5233	35.6	97.6	10.84	0.000
	Matched	5.3307	5.2852	0.8		0.16	0.873
AADF in baseline years	Unmatched	19039	18020	10.1	88.3	2.52	0.012
	Matched	19049	19168	-1.2		-0.22	0.823
PICs in baseline years	Unmatched	12.722	8.3347	34.7	99.2	9.60	0.000
	Matched	12.510	12.474	0.3		0.05	0.959
FSCs in baseline years	Unmatched	1.8431	1.0391	41.7	97.6	12.63	0.000
	Matched	1.7969	1.7773	1.0		0.18	0.861
Site length	Unmatched	0.7118	0.7009	2.3	-137.2	0.59	0.554
	Matched	0.7094	0.7353	-5.4		-1.05	0.294
A Road	Unmatched	0.7276	0.7984	-16.7	74.2	-4.47	0.000
	Matched	0.7279	0.7096	4.3		0.79	0.427
B Road	Unmatched	0.2101	0.1613	12.6	89.3	3.37	0.001
	Matched	0.2096	0.2148	-1.3		-0.25	0.803
Minor Road	Unmatched	0.0376	0.0230	8.5	82.2	2.42	0.016
	Matched	0.0378	0.0404	-1.5		-0.26	0.792
Speed Limit 30mph	Unmatched	0.7575	0.5118	52.7	97.9	12.90	0.000
	Matched	0.7565	0.7513	1.1		0.24	0.813
Speed Limit 40mph	Unmatched	0.1219	0.1828	-17.0	97.9	-4.14	0.000
	Matched	0.1224	0.1237	-0.4		-0.08	0.938

6.5.3 Effects of speed cameras on road accidents

The effects of speed cameras on road accidents were estimated using three different methods: a naïve before-after approach, the PSM method and the EB method. Since different algorithms can be chosen when employing the PSM method, the robustness of the results must be checked to ensure that the estimation does not depend upon the chosen algorithm. In this study, results from five algorithms, two of which were of one type (K-nearest neighbours), were compared to increase confidence in the PSM method. The matching algorithms used were: K-nearest neighbours matching (K=1), K-nearest neighbours matching (K=5), radius matching (caliper=0.05), stratification matching and kernel matching (caliper=0.05). The EB method was used and compared with the PSM method, with two reference groups used in this method. One reference group was selected based on the propensity score, while the other one contained all the potential reference sites. For FSCs, only the effect on absolute accident numbers was estimated, since the sample for FSCs was zero-inflated and a large amount of data was discarded when analysing effects on annual FSCs per km in percentages, thereby making the results unreliable.

Table 6.4 presents the estimations of the effects of speed cameras on annual PICs and FSCs per km. The observed reduction in annual PICs per km was 1.441 in absolute numbers and 30.7% in percentage terms. When applying the PSM method, the results were very similar for all five algorithms, with an average reduction in PICs of around 1.068 (25.9% in percentage terms). Such similar results indicate that the estimations are independent of the algorithms used, thereby increasing confidence in the PSM method. The results from the EB method, using all sites as the reference group, showed a reduction of 0.854 in absolute numbers and 23.3% in percentage terms, slightly lower than the results from the PSM method. Matched sites were then used as the reference group with the estimated reduction in PICs being1.026 (25.7% in percentage terms). In Table 6.4 the effect on FSCs is also analysed using the three methods. Unsurprisingly, the result from the naïve before-after approach shows the largest fall of 0.342 in absolute number. The PSM methods give a consistent estimation for all five algorithms an average reduction of 0.132. A similar result was obtained from the EB method using matched sites as the reference group, where the reduction was 0.135.

Table 6.4 Effects of Speed Cameras on Annual PICs/FSCs per km

	Changes in annual PICs per km in absolute numbers					Changes in annual PICs per km in percentage					Changes in annual FSCs per km in absolute numbers				
	Treatment Group	Control Group	Changes	S. E.	T-Stat	Treatment Group	Control Group	Percentage Changes	S. E.	T-Stat	Treatment Group	Control Group	Changes	S. E.	T-Stat
Unmatched	771	4787	-1.441	0.131	-11.02	726	4077	-30.70%	0.041	-7.5	771	4787	-0.342	0.037	-9.25
DID Propensity Score Matching															
K-nearest Neighbours Matching (K=1)	764	663	-1.035	0.21	-4.92	726	600	-29.70%	0.051	-5.83	771	663	-0.141	0.06	-2.34
K-nearest Neighbours Matching (K=5)	764	2923	-1.068	0.168	-6.33	726	2625	-24.60%	0.034	-7.21	771	1676	-0.124	0.049	-2.5
Radius Matching (Caliper=0.05)	769	4626	-1.081	0.155	-6.97	724	3921	-25.20%	0.031	-7.99	769	4626	-0.131	0.046	-2.82
Stratification Matching	769	4628	-1.042	0.15	-6.96	725	4078	-24.70%	0.029	-8.48	769	4628	-0.135	0.044	-3.05
Kernel Matching (Bandwidth=0.05)	771	4626	-1.117	0.147	-7.61	726	4077	-25.10%	0.032	-7.89	771	4626	-0.131	0.046	-3.01
Average Effect			-1.068					-25.90%					-0.132		
Empirical Bayes using all sites as reference group	771	4787	-0.854	0.102	-8.34	726	4077	-23.30%	0.037	-7.01	771	4787	-0.197	0.044	-4.42
Empirical Bayes using matched sites as reference group	764	2923	-1.026	0.127	-8.04	726	2625	-25.70%	0.036	-6.36	771	1676	-0.135	0.069	-1.98

Table 6.5 summarizes the effects of speed cameras on PICs and FSCs given different distances from camera sites. These results suggest that speed cameras are most effective up to 200 metres, where the reduction in annual PICs per km is approximately 1.350 (27.5% in percentage terms).The effectiveness decreases, however, as the distance from the camera site increases, with the estimations showing approximately 1.135 (26.4% in percentage terms) for up to 500 metres and 0.570 (18.5% in percentage terms) for up to 1km. In terms of the effects on FSCs at different distances from cameras, certain figures in Table 6.5 are insignificant. This is probably due to too few FSCs being observed to give conclusive estimates. Nevertheless, it is obvious that the reduction of FSCs within 200 metres from the camera, 0.188, is the largest. Up to 500 metres, this reduction is 0.164, whilst for up to 1 km the average reduction is 0.049, although the estimations using all algorithms are insignificant.

Table 6.5 Effects of Speed Cameras on Annual PICs/FSCs per km at Different Distance from Cameras

	Changes in annual PICs per km			Percentage Changes in annual PICs per km			Changes in annual FSCs per km		
	0 to 200m	0 to 500m	0 to 1km	0 to 200m	0 to 500m	0 to 1km	0 to 200m	0 to 500m	0 to 1km
K-nearest Neighbours Matching (K=5)	-1.372**	-1.103*	-0.598**	-27.3%**	-25.2%*	-16.6% *	-0.115	-0.149***	-0.047
Radius Matching (Caliper=0.05)	-1.387*	-1.148*	-0.720*	-29.2%*	-26.3%*	-18.7% *	-0.209***	-0.169**	-0.049
Stratification Matching	-1.324*	-1.149*	-0.467**	-25.7%*	-26.2%*	-19.1% *	-0.167***	-0.180*	-0.053
Kernel Matching (Bandwidth=0.05)	-1.318*	-1.141*	-0.496**	-27.7%*	-27.9% *	-19.4% *	-0.180	-0.161**	-0.045
Average Effect	-1.35	-1.135	-0.421	-27.50%	-26.40%	-18.50%	-0.188	-0.164	-0.048

Notes: Figures are significant at: *99%, **95%, ***90

6.5.4 Discussion

The issues of selecting control groups to account for confounding factors and how the treatment and control groups are matched are critical in assessing the impacts of road safety measures. This can be seen particularly when assessing the effect on road traffic accidents due to the introduction of speed cameras in the UK. This study introduced the PSM method to account for these two issues and then applied it to data from the UK to assess the impact of speed cameras. Similar estimation results indicate that the PSM method and the EB method are comparable. This study also shows that the characteristics of the treatment and comparison groups are well balanced after matching. The results confirm that the EB method using matched sites as the reference group is superior to the one using all sites. Therefore, the author suggests that propensity scores can be applied as the criterion when constructing the reference group. Indeed, the construction of such a reference group can be used in any road traffic safety analysis where a safety measure has been implemented, not simply for assessing the impacts of speed cameras.

This study also has two major findings on the impacts of speed cameras on accidents. The first relates to the distance at which speed cameras have their greatest impact. For both PICs and FSCs, there is a reduction in accidents, but the extent of this reduction decreases with distance from the cameras. Speed cameras, therefore are found to be most effective up to 200 metres from camera sites, although the reduction in accidents up to 500 metres is also significant. Figure 6.3 presents the cumulative reductions in annual PICs and FSCs. The cumulative reduction increases dramatically from 0 to 500 meters but this tendency reduces from 500 metres to 1km. It is obvious that the reduction in accidents due to the effect of speed cameras is negatively correlated to the distance from the camera sites. It is unclear, however, whether this relationship holds over larger distances (i.e. over 1000 metres) because data restrictions prevented reliable estimation. The suggested site length by DfT (2004) is between 400meters to 1.5 km, which tallies with the effective length estimated in this study.

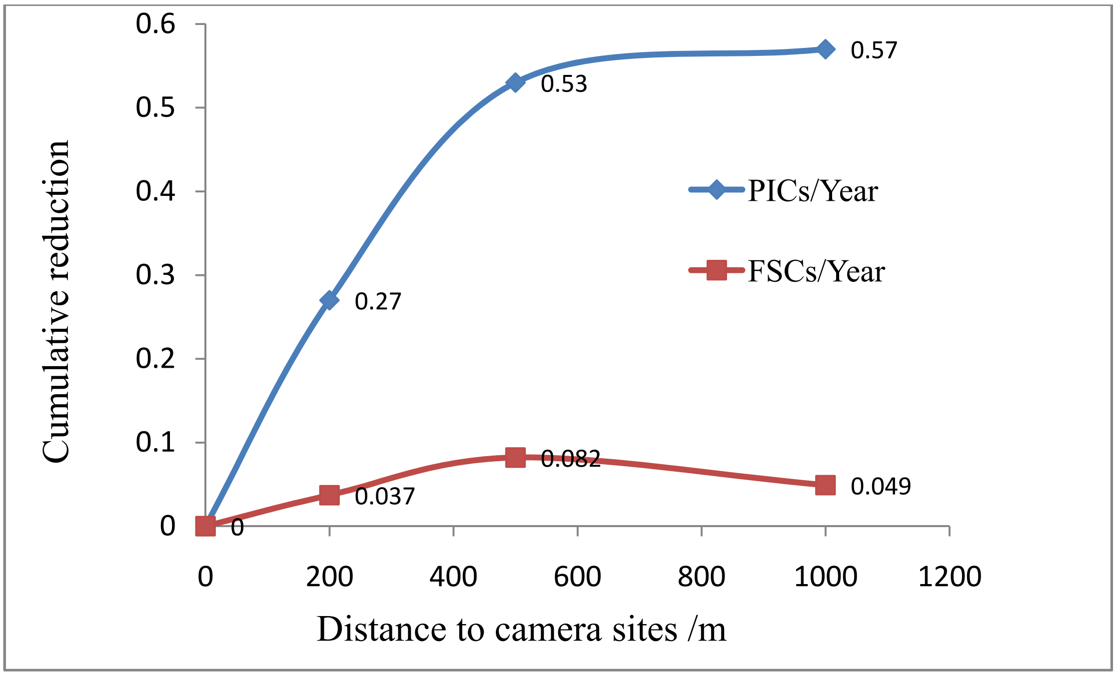

Figure 6.3 Cumulative reductions in annual PICs/FSCs

The second finding relates to accident migration. Having controlled for accident migration due to the choice of alternative routes to avoid speed cameras by including the covariate AADF, this study finds there is no evidence of the "kangaroo" effect, i.e. no increase in accidents upstream and downstream of camera sites. This is an important finding in that it shows the drivers do not alter their behaviour to deliberately decelerate and accelerate abruptly before and after the camera sites. Rather speed cameras have a constant effect on driver behaviour in reducing their speed.

6.6 Conclusion

In contrast to other types of traffic interventions, most traffic law enforcement aims to reduce casualties and thus is usually implemented in areas with high casualty records. Selection bias may, therefore, occur in the assessment of intervention effects. In this chapter, the propensity score matching method is introduced to evaluate the safety effects of traffic law enforcement.

The main problem in using the EB or other before-after control methods to estimate intervention effects is the selection of a proper reference or control group similar to the treatment group. The propensity score matching method can account for this issue and find best matches for each treated individual. The attractiveness of this method is that it overcomes the problem of selection bias and gives a clear criterion by which to select the reference or control group.

The PSM method was applied in this chapter to evaluate the safety effects of speed limit enforcement cameras. Different matching algorithms were applied to validate the PSM method. Furthermore, comparing the EB method with the PSM method gave similar results.

The key findings from this analysis are:

- There is a reduction of 25.9% in annual PICs per km at camera sites, which equals 1.068 in absolute accident numbers. The number of FSCs per km every year falls by 0.132after installing camera sites. The results are in line with the findings of previous research. For example, Mountain et al. (2004) found a reduction of 26% for overall injury accidents at camera sites. Another research by Goldenbeld and van Schagen (2005) suggested a reduction of 21% in injury accidents after installing the speed cameras.

- Speed cameras are found to be most effective within 200 metres from the camera site. This effect is substantial up to 500 metres but falls rapidly from 500 metres to 1000 metres. This is consistent with the suggested site length, which is between 400 metres to 1.5km.

- The EB method using matched sites as the reference group produces comparable results with the ones from the PSM method. This suggests that the propensity score can be used as the criteria for selecting the reference group when the EB method is employed.

The impact of speed cameras over 1000 metres cannot be estimated because of data restrictions. There might be an inflection point where the effect of speed cameras on casualties becomes zero and further research in this area is recommended.

Chapter 7: An Application of Full Bayes Models Using Panel Data: Effects of Changes in Road Network Characteristics on Road Casualties

Road network planning needs to be based on the best knowledge available of effects of road design on road safety in order to ensure a high level of road safety. This chapter analyses how changes in road network characteristics affect road casualties. In order to do this, another widely used approach for before-after evaluation studies, the Bayesian method, is applied. In addition, a panel semi-parametric model is used to estimate the dose-response function for continuous treatment variables. This chapter is organized as follows. After the introduction in section 7.1, a review of previous research is presented in section 7.2. Section 7.3 describes the data sources followed by a discussion of methods in section 7.4. Results are outlined and discussed in section 7.5. The conclusions are given in the final section.

7.1 Introduction

The statistical relationship between road casualties and the characteristics of a road network has been investigated in the literature (e.g. Noland and Quddus, 2004; Wier et al., 2009; Huang et al., 2010; Dumbaugh and Li, 2010; Marshall and Garrick, 2011; Rifaat et al., 2011; Jones et al, 2008; Quddus, 2008). Specifically, road casualties are found to be significantly associated with road network characteristics, such as road length, density and nodes. No consistent conclusion has been drawn in the literature, however, regarding strength of this association. One probable constraint in previous research is that, to the best of the author's knowledge, no longitudinal or panel data of road network has been employed. This implies that any variance in road network characteristics over time cannot be controlled for and inferences made on the impacts of road network on road casualties could therefore be biased or wrong. Another key issue, which is critical in all analyses relevant to road casualties, concerns the exposure variable. In analyses at a disaggregate level, the ideal variable used to control for risk exposure is the annual average daily traffic (AADT) for the unit of interest. In terms of aggregate analysis, however, the AADT is unavailable for an aggregate area and proxy variables are therefore usually used. There are various such proxy exposure variables, such as the usage of cars (Quddus, 2008), aggregated

AADT (Jones et al., 2008; Marshall and Garrick, 2011) and a proxy variable derived from a gravity model using data of population and employment (Graham and Glaister, 2003). When applying these variables as exposure variables, however, limitations arise.

The objective of this chapter is to investigate the relationship between road casualties and road network characteristics. In order to do this, a panel data set for the road network in England and Wales between the years 2001 to 2010 was used to account for the variance in road network over time. The exposure variable used in this analysis was the daily traffic trips generated within the study area, which was estimated based on the origin-destination (OD) data obtained from the Office for National Statistics. A two-stage regression was used, whereby the traffic trips were estimated with ordinary least square regression in the first stage and the road casualties were analysed using a panel generalized linear model in the second stage. To account for the additional error in the two-stage regression, the bootstrap technique was employed. Full Bayes hierarchical models were also employed to control for spatial correlation and compared with the traditional models.

7.2 Previous Research

As discussed in Chapter 2, over the past decade, considerable research has shown that road casualties are associated with various factors among which road network characteristics are considered as particularly important (Lord et al., 2005; Aguero-Valverde and Jovanis, 2006; Graham and Glaister, 2003; Wang and Abdel-Aty, 2006).

A spatially disaggregated analysis of road casualties in England undertaken by Noland and Quddus (2004) examined the effects of road characteristics and land use on road casualties. Their results suggest that increased length of B roads could increase serious injuries, although the coefficients for other types of road were not significant. Marshall and Garrick (2011) investigated how street network characteristics affected road safety in 24 Californian cities from 1997 to 2007. Street network characteristics, such as street network density, street connectivity and street network patterns were controlled for in this study and will be discussed in detail in the following sections. Marshall and Garrick's results suggest that road casualties for all levels of crash severity are correlated with street network characteristics. A higher density of intersection counts is associated with fewer crashes, while street connectivity (link to node ratio) is positively related to crashes. Dumbaugh and Li (2011) also find that miles of arterial roadways and numbers of four-leg intersections were major crash risk factors in Texas, using data from 2003 to 2007. Another study conducted by Rifaat et al. (2011) examined the effect on crash severity of different street patterns, including grid-iron, loops and lollipops, and mixed patterns. Although only pedestrian and bicycle crash data were analysed, the authors found significant effects of street pattern, road features and environmental conditions on crashes. One limitation of all the above research is that the spatial correlation of casualties was not examined. This could violate the traditional Gauss-Markov assumptions, and hence lead to biased inferences.

To control for spatial variation, spatial models can be applied to avoid inference errors. A recent study by Jones et al. (2008) uses district-level data to investigate the effects of various factors on traffic casualties in England and Wales. The authors found that traffic casualties were significantly related to road length, curvature, junction density and other geographical variables. Testing for spatial autocorrelation however showed no positive autocorrelation at the district level. In the study by Quddus (2008), however, a significant positive spatial correlation among ward-level traffic crashes was found in the Greater London from 2002 to 2004.Quddus applied both traditional and spatial models with the results from the traditional NB models and the Bayesian hierarchical models being very similar in suggesting that traffic crashes are associated with the road infrastructure, socioeconomic and traffic conditions. Substantial positive spatial correlation was also found in the analysis of crash data for Florida's 67 counties from 2003 to 2007 (Huang et al., 2010). One reason for the diverse results of spatial correlation tests could be due to the different spatial aggregation levels in the papers mentioned above.

If spatial correlation is present, appropriate spatial models need to be employed to account for the spatial dependence. Generally, there are two methods for spatial analysis of road casualty data, traditional econometric models and Bayesian hierarchical models. Although the results from both models are very similar in many cases, Bayesian hierarchical models have been suggested to be more appropriate and have been employed in many studies (Persaud et al., 2010; Aguero-Valverde and Jovanis, 2006; Haque et al., 2010). For example, Aguero-Valverde and Jovanis (2006) compared full Bayes (FB) hierarchical models with traditional negative binomial (NB) models using county-level crash data for Pennsylvania. The existence of spatial correlation in county-level crash data was revealed, although their results from FB hierarchical models were generally consistent with the NB estimates. This similarity in results from the NB and FB models has also been found in other studies (Quddus, 2008; Mitra and Washington, 2007). This is probably because uncorrelated heterogeneity accounts for most of the variation and the traditional NB models can sufficiently control for this effect. Another probable reason is that when there is sufficient data the NB method works well and the results from the two methods are similar. In contrast, when the sample is rather small, the prior information in Bayesian methods will dominate the analysis and Bayesian methods will be superior. The choice of method used should be made given the resources available (e.g. data and prior information).

The idea that the values of parameters could arise from distributions is a fundamental feature of Bayesian methods. Bayesian hierarchical models can accommodate distributions such as hierarchical Poisson-Gamma distribution and Poisson-Lognormal distribution (Miaou and Lord, 2003; El-Basyouny and Sayed, 2011; Siddiqui et al., 2012; Yanmaz-Tuzel and Ozbay, 2010). Different prior distributions have been discussed by Yanmaz-Tuzel and Ozbay (2010) with their results suggesting that a Poisson-Lognormal model structure with more informative priors and higher levels of hierarchy may reduce the biases in modeling parameters, hence leading to more robust estimates. In this chapter, hierarchical Poisson-Lognormal models are adopted to compare with the traditional NB models.

As discussed earlier, two issues evident in previous studies examining the relationship between the road network and road casualties remain to be adequately addressed. One critical issue is the selection of appropriate traffic exposure variables. In this chapter, a new method to construct the traffic exposure variable is proposed. The other issue concerns the usage of data for road network characteristics. Detailed data regarding the road network, including road class, road length and node information can be obtained from OS Meridian TM 2. Although this data set has been used in several studies in the UK (Noland and Quddus, 2004; Haynes et al., 2007; Graham and Stephans, 2008; Jones et al., 2008), the data availability is only for a single year and, consequently, variance in the road network over time cannot be accounted for. To overcome this problem, OS Meridian TM 2 for 2001 to 2010 (except for 2005) is employed in this study. The data set is discussed in the next section.

7.3 Data

7.3.1 Dependent Variable

The data used in this analysis includes road casualties recorded in England and Wales from 2001 to 2010. The accident data was collected from the STATS 19 data base and was further classified by severity type. The location of an accident was recorded using the British National Grid coordinate system. Each individual accident was located on a map and these casualties were further aggregated at the ward level, i.e. the primary unit of British administrative and electoral geography. Geographical Information System (GIS) software, such as MapInfo and Arcmap were used to process the data.

7.3.2 Exposure Variable

Traffic exposure is the most important factor influencing traffic crash counts, however there is no appropriate variable that can be used to control for the traffic exposure in an area-level analysis. In previous disaggregated analyses of traffic crashes, the AADT has frequently been used to indicate the traffic exposure level (Abdel-Aty and Radwan, 2000). The AADT was also employed in recently conducted area level analyses (Huang et al., 2010; Marshall and Garrick, 2011; Jones et al., 2008). In studies conducted by Marshall and Garrick (2011) and Huang et al. (2010), AADT count points were first located to each road and the average AADT was then calculated for each road. The VMT was obtained by multiplying the average AADT of each road by its centreline mile length, and all VMT values were summed to form the exposure variable for each study area. In the UK, a similar application of AADT data was conducted by Jones et al. (2008). Traffic count data supplied by the UK Department for Transport classifies the estimate of the average daily count of vehicles into six categories ranging from pedal cycles to heavy goods vehicles at 5982 survey census points. With a grid reference, each point was assigned to districts and the AADT was estimated by taking mean count values across the points located on each road class.

While the usage of the AADT data provides substantial information on the traffic flow at an area-level, there are two limitations in this approach. First, the AADT data is usually only available at a limited number of data collection points. The under-representation of roads, especially minor roads, could lead to an underestimation of traffic flow in each area. Secondly, the fact that the AADT data only accounts for motorized vehicle travel means that it cannot fully depict the level of overall activity, which consists of travels by foot, bicycle and other transit modes.

Graham and Glaister (2003) developed a proxy variable for traffic exposure using a gravity model. The idea of this approach is that external traffic generation of each ward is affected by the population and employee of its proximate wards. One question with this method, however, is how “proximate wards” are defined. This will be discussed later in this chapter.

In this thesis, an approach based on traffic assignment is proposed in order to estimate the traffic exposure at ward level. The idea underlying this method is that trips generated between origin and destinations (OD) are assigned to transportation networks and aggregated in each ward. The traffic assignment focuses on the selection of routes between OD and the traffic volume on each route in transportation networks. The centroids of wards are treated as origins and destinations, and transportation networks are constructed by links among the centroids. Figure 7.1 shows the synthetic road network constructed by ArcGIS among wards. In the conventional transport forecasting model, traffic assignment is the fourth step following trip generation, trip distribution and mode choice. In this research, a wealth of information about OD statistics was obtained from the Office for National Statistics. In this data set, the daily trips of residents or workers from residence to workplace are provided as matrices with breakdowns of the characteristics of the people and transit modes. In this research, to reflect the overall activities within each ward, the daily trips of all people and transit modes were included in the OD matrices.

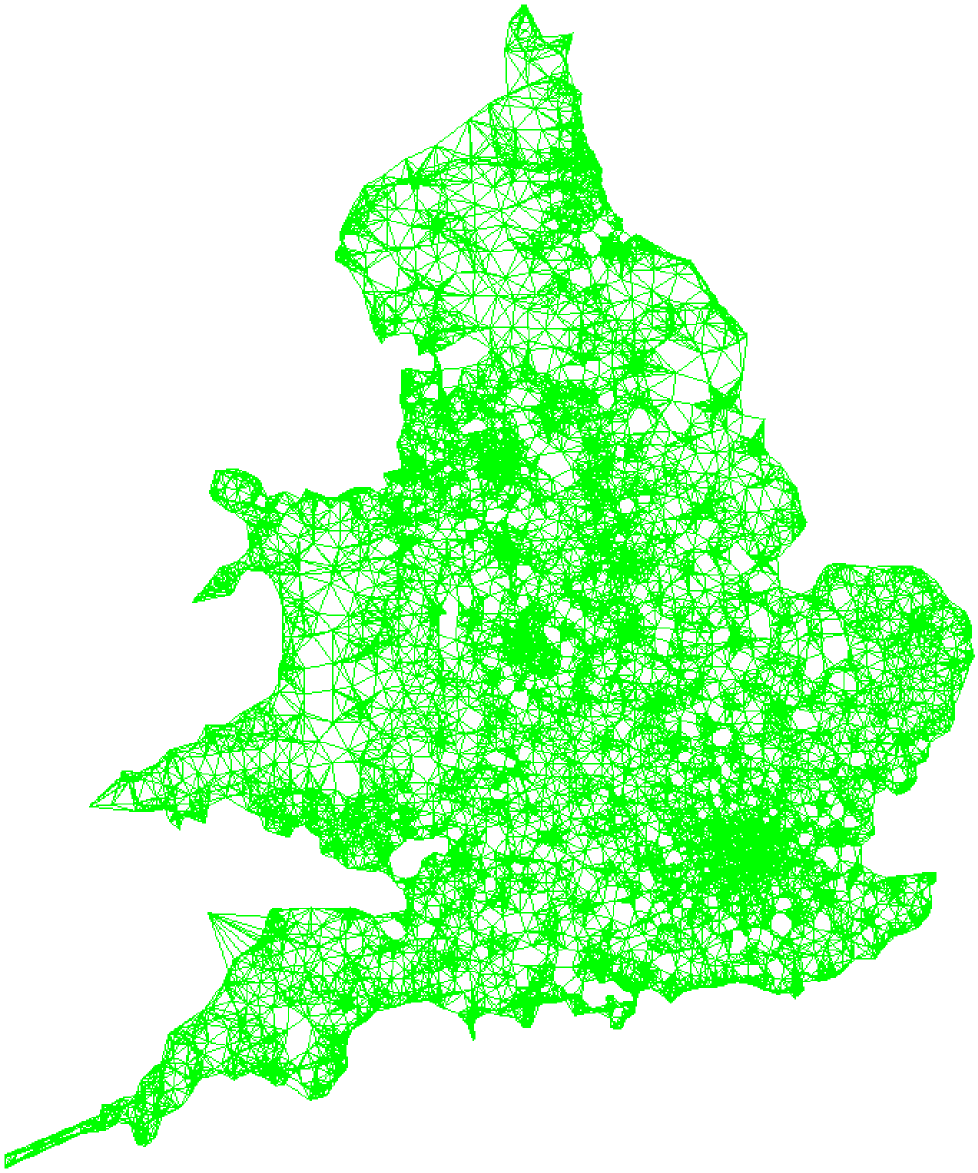

Figure 7.1 Synthetic Road Network Constructed at the Ward-Level

The traffic assignment was implemented using TransCAD. The assignment method employed was the well-known user equilibrium method proposed by Wardrop (1952). Equilibrium model is widely used for the prediction of traffic patterns in transportation networks. It is assumed that travellers will choose the shortest path from origin to destination and network equilibrium occurs when no traveller can decrease his/ her travel cost by shifting to a new path. The trips are assigned to links among the centroids and aggregated in each of the wards. The distribution of the number of trips in wards across England and Wales is presented in Figure 7.2. Subsequently the number of total trips is used as the proxy variable for traffic exposure. One issue with this approach is that the OD data was only available for 2004. The traffic exposure for other years was estimated by ordinary least square regression, discussed as follows.

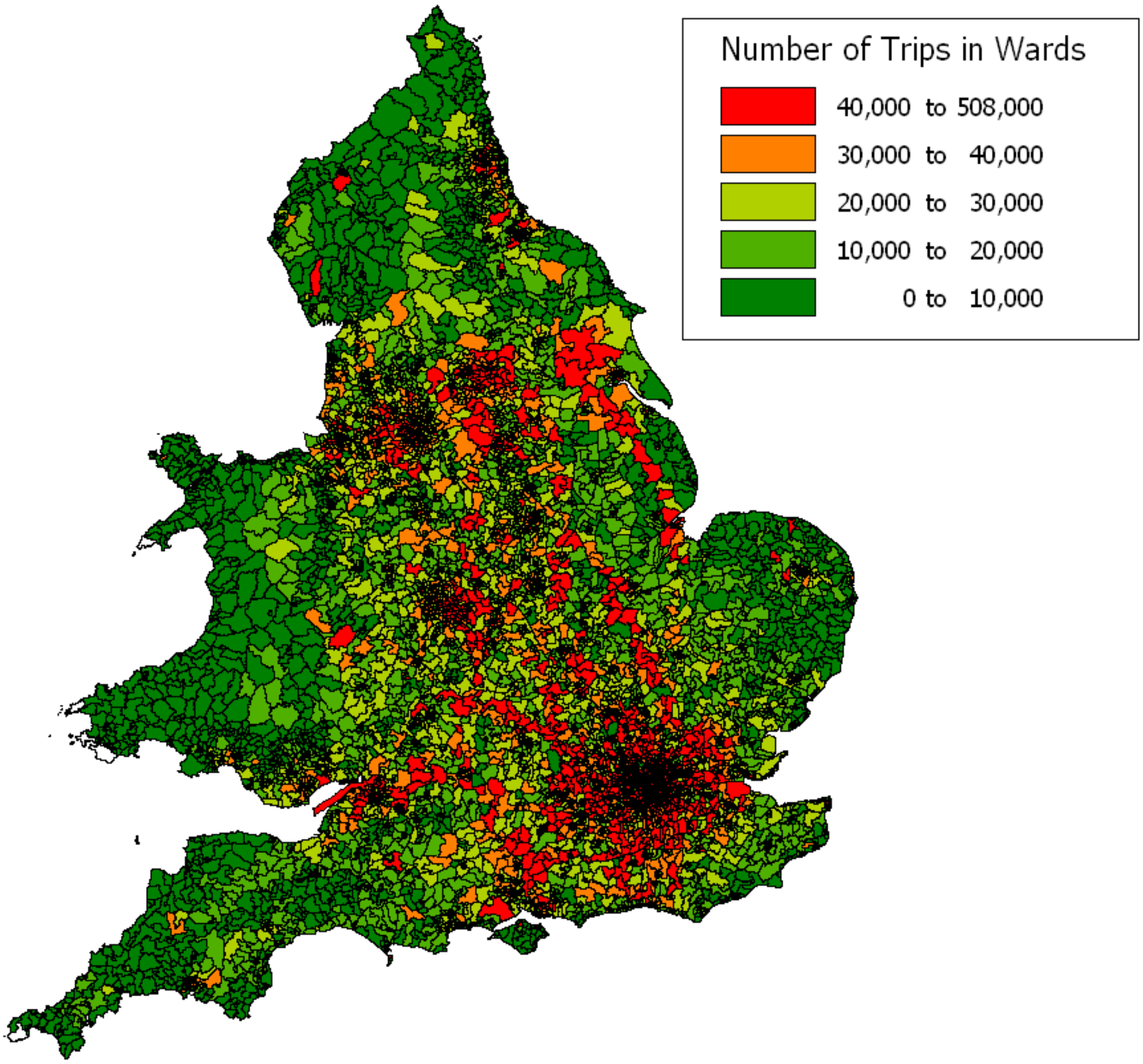

Figure 7.2 The Distribution of Trips in Wards in England and Wales

7.3.3 Road Characteristics

A major contribution of this research is that panel data of the road network was used to account for effects due to the variation in road characteristics over time. Compared to time-series and cross-section data, panel data provides several benefits. For example, panel data is able to control for spatial- and time-invariant individual heterogeneity, and it also provides more informative data thus better enabling study of the dynamics of adjustment.

In this study, detailed information regarding the road network was obtained from Ordnance Survey (OS) Meridian ^{TM} 2 for the period from 2001 to 2010 except for 2005. A set of variables was employed to describe the characteristics of the road network at ward level.

(1) Traditional road network characteristics. The length, as well as the density, of the road network was calculated according to road class, e.g. Motorway, A road, B road, Minor road. Road network nodes were defined as meeting points of two or more roads. The total number and density of nodes and roundabouts was also calculated.
(2) Connectivity and accessibility of the road network. It has been suggested that the degree of connectivity and accessibility of a road network can influence the number of crashes (Marshall and Garrick, 2011). The measure used in this study was the link-to-node ratio, which was calculated by dividing the number of links by the number of nodes. A high link-to-node value indicates a more connected road network than one with a low link-to-node value. A node with only one link, also known as a dead end, is usually associated with a residential area. The density of dead ends was used in this study as a measure of the accessibility of a network.
(3) Curvature of the road network. Road curvature has been suggested as an important factor influencing road casualties (Haynes et al., 2007; Jones et al., 2008; Quddus, 2008). The literature indicates that straighter roads have more crashes than roads with more bends. The variable used in this research to measure curvature was the number of vertices per km. This was obtained using ArcGIS and divided by the road length in each ward.

7.3.4 Socio-demographic Characteristics

Previous research has suggested an association between road traffic crashes and socio-demographic characteristics, such as population, employment and deprivation (Wier et al., 2009; Dissanayake et al., 2009).In particular, a positive relationship has been found in relation to the sizeof the population and the level of employment, which implies that more casualties may occur in areas with more residents and job opportunities. To consider this effect, the data for population and employment at the ward level was obtained from the Office for National Statistics (ONS).

Recent research also suggests that child injuries are influenced by factors related to area deprivation (Graham and Stephens, 2008). Therefore, the Index of Multiple Deprivation (IMD), published by the Office of the Deputy Prime Minister (ODPM, 2004), was used as a control variable in this study. The Index of Multiple Deprivation integrates data on the following seven deprivation domain indices into one overall deprivation score: income, employment, housing and services, health, education, crime and the environment.

7.4 Method

A two-stage regression method was employed to explore the relationships between crash counts and road network characteristics. In the first stage, the traffic exposure for all years of the analysis was estimated using OLS regression. In the second stage, the fitted traffic exposure values were included, together with other variables, in a Generalized Linear Model (GLM). One important issue that needed to be addressed was that the standard errors from the second stage regression were biased because the traffic exposure was itself estimated. To correct the standard errors, therefore, the bootstrap approach was used in both stages.

7.4.1 Model for estimating travel activities

Within-ward travel activities

The first step involved identification of the relevant variables to be used in the model for estimating travel activities. As discussed previously, trips in each ward consisted of traffic generated within the ward and traffic passing through. There were also two main types of trips categorized by destinations: home-end and work-end trips as set out below.

For trips generated within the ward, the total number of trips leaving or returning to homes or work places is likely to be related to the population, employment, deprivation and total length of roads. Home-end trips can be described as a function of:

Home-end Trips= f (Population, Deprivation, Length of roads) 7.1

While work-end trips can be described as a function of:

Work-end Trips = f (Employment, Deprivation, Length of roads) 7.2

The total trips generated in the ward can thus be described as:

Within-Ward Trips = Home-End Trips + Work-End Trips= f (Population, Employments, Deprivation, Length of roads) 7.3

Pass-through travel activities

In terms of pass-through trips, Graham and Glaister (2003) employed the idea of using the resident and employment population of proximate wards. A gravity model for trip distribution was applied in their study. It was assumed that pass-through trips are related to trips generated in proximate wards, but what these are is not clearly defined in the literature.

A study conducted by Dent and Bond (2008) investigated the commuting patterns of part-time and full-time workers in the UK. One important finding reveals that the average commuting distances in the UK were7.5 km for part-time workers and 13 km for full-time workers respectively. The average commuting distance given in the report DfT (2011), meanwhile, was 8.6 miles for 2009.

In this study, wards within a certain distance, called the "bandwidth", were taken into account when selecting proximate wards from which to estimate the pass-through trips. The bandwidth selected was13 km, which is consistent with the average commuting distance for full time workers.

The function of pass-through trips can be described as follows:

Pass-Through Trips = f (Length of roads, Sum of Employments in Neighbour Wards, Sum of Population in Neighbour Wards) 7.4

To account for both within-ward and pass-through trips, the model used to estimate total trips was developed as follows:

$$T_i = \alpha + \beta_1 IMD_i + \beta_2 RL_i + \beta_3 E_i + \beta_4 P_i + \beta_5 \Sigma\, ((E_j + P_j)/D_{ij}) \qquad 7.5$$

where T_i is the total trips in ward i, IMD_i is the IMD score for ward i, RL_i is the length of roads for ward i, E_i is the employments of ward i, P_i is the population of ward i. Ward j is a neighbour ward of ward I within the bandwidth. Both the resident and employment population of neighbour wards are indicated by E_j and P_j. D_{ij} is the distance between centroids of wards i and j.

7.4.2 Bootstrapping Generalized Linear Models

7.4.2.1 Statistical Modelling

As outlined in chapter 2, data on road traffic accident casualties is characterised by being both non-negative integer and not normally distributed. Generalized linear models, such as the Poisson and the Negative Binomial models, therefore, are usually used to establish the association between road traffic accident casualties and various factors that might affect road traffic accident casualties.

To select the model, the dispersion parameter was estimated for all types of crashes. If the dispersion parameter was significantly greater than zero, then the Negative Binomial model provided a better fit than the Poisson model. Otherwise, the Poisson model was selected.

7.4.2.2 Bootstrapping Regression Models

In the two-stage regression, the traffic exposure was estimated using OLS and included as the exposure variable in the GLM for accident analysis. It is commonly the case that the standard errors from the second-stage regression are incorrect since the traffic exposure variable is itself estimated. To correct the standard errors for the two-stage model, the bootstrap method was used. This method can be applied to produce accurate confidence intervals, standard errors and hypothesis tests.

The key analogy of bootstrap is "The Population is to the sample as the sample is to the bootstrap sample" (see Fox, 2008, pp. 590). Bootstrap estimates relevant characteristics of the population using the sample data. The sampling distribution of a statistic is then constructed empirically by resampling from the sample. The resampling procedure parallels the process by which sample observations were drawn from the population.

Let S_1 and S_2 denote the sample used in the first and second regression respectively. The algorithm for bootstrapping the two-stage regression in this study is introduced below.

(1) Use S_1 to estimate the traffic exposure regression, *Equation 7.5*,

$T_i = \alpha + \beta_1 IMD_i + \beta_2 RL_i + \beta_3 E_i + \beta_4 P_i + \beta_5 \Sigma ((E_j + P_j)/D_{ij})$, and store the estimates of coefficients $\alpha', \beta_1', \ldots, \beta_5'$.

(2) Take a bootstrap sample of S_2 and call it S_{b2}.

(3) Use S_{b2} to calculate and store the fitted value of traffic exposure $\widehat{T_i}$:

$$\widehat{T_i} = \alpha' + \beta_1' IMD_i + \beta_2' RL_i + \beta_3' E_i + \beta_4' P_i + \beta_5' \Sigma ((E_j + P_j)/D_{ij}) \qquad 7.6$$

(4) Regress crash counts on $\widehat{T_i}$ and other observed covariates in S_{b2} to estimate the vector of coefficient $\boldsymbol{\pi}$ and standard errors.

(5) Repeat steps (2)-(4) 1000times.

There are two sources of random variation in terms of bootstrap inferences. Firstly, almost all the variation among bootstrap distributions comes from the selection of the original sample from the population. This variation can be reduced, however, by using a large original sample. Secondly, bootstrap resampling randomly from the original sample may introduce additional variation. Usually, 1000 or more bootstrap resamples are required to reduce additional variation.

7.4.3 Bayesian Spatial Model

To account for possible spatial autocorrelation of crash counts among adjacent wards, Bayesian hierarchical models were employed as a comparison with the conventional NB models. In this model, the area-specific random effects are decomposed into two components. The first models the effects that vary in a structured manner in space, such as correlated heterogeneity, while the second models the effects that vary in an unstructured way between areas, such as uncorrelated heterogeneity.

The Bayesian hierarchical model can be described as:

$Y_i \sim$ Poisson (μ_i),

$\ln\mu_i = \alpha + \boldsymbol{\beta}\mathbf{X_i} + \varepsilon_i + u_i$ *7.7*

where u_i is the spatial correlated heterogeneity. The conditional autoregressive (CAR) model proposed by Besag et al. (1991) is presented below:

$[u_i|u_j, i\neq j, \tau^2_u] \sim N(\bar{u}_i, \tau^2_i)$

where:

$$\bar{u}_i = \frac{1}{\sum_j \omega_{ij}} \sum_j u_j \omega_{ij} \quad 7.8$$

$$\tau_i^2 = \frac{\tau_u^2}{\sum_j \omega_{ij}} \quad 7.9$$

ω_{ij} =1 if i ,j are adjacent, otherwise 0.

Both τ^2_ε, τ^2_u are assigned as gamma distributions with priors *Ga* ~ (0.5, 0.0005), as suggested by Wakefield et al. (2000). The vector of coefficients $\boldsymbol{\beta}$ is assumed as a highly non-informative normal distribution with N (0, 0.01) and the intercept α is assigned as a uniform prior distribution to reflect the lack of precise knowledge of the value of the coefficients.

Two separate Markov Chain Monte Carlo (MCMC) analyses with different initial values were used to assure convergence. The first 5000 samples were discarded as a burn-in and a further 20000 iterations were run for each chain. Visual examination of time series plots of the samples for each chain and the Gelman and Rubin diagnostic (Gelman and Rubin, 1992) were used to check the convergence.

The overall goodness of fit was measured by the Deviance Information Criterion (DIC), which is a generalization of the Akaike Information Criterion (AIC). Please refer to chapter 4 for detailed discussion on the DIC.

7.4.4 Dose-Response Function for Continuous Variables

In the binary treatment case, the treatment variable is denoted as $D = \{0, 1\}$. The model can be described as:

$$Y_i = \alpha + \boldsymbol{\beta} \mathbf{X_i} + \gamma d_i + \varepsilon_i,\ d_i \in D \qquad 7.10$$

The effect of the binary treatment variable *D* can be interpreted as the estimate of the coefficient.

$$E[Y(1)] - E[Y(0)] = \gamma$$

Such an estimate, however, may be inappropriate in the continuous treatment case, where *D* is allowed to be an interval $[d_1, d_2]$. It is possible that the treatment effect may vary over *D*, i.e. the relationship between the outcomes and the treatment cannot be simply expressed as a linear relationship.

In this study, a panel data semiparametric model is applied to estimate dose-response functions for continuous variables.

$$Y_{it} = \alpha_i + \boldsymbol{\beta} \mathbf{X_{it}} + f(d_{it}) + \varepsilon_{it},\ d_{it} \in D \qquad 7.11$$

where $f(d_{it})$ is assumed as a polynomial function, which can be expressed as

$$f(d_{it}) = \mu_0 + \mu_1 d_{it} + \mu_2 d_{it}^2 + \mu_3 d_{it}^3 \ldots + \mu_m d_{it}^m \qquad 7.12$$

The best fitting power, *m*, is selected by maximizing the likelihood of *equation 7.12*.

The effect on road casualties of changes in the connectivity and accessibility of road networks is of interest in this thesis. Specifically, *density of dead end* and *links per node* were treated as continuous treatment variables. The dose-response functions of these two treatments were estimated separately by holding other control variables constant and were shown by spline curves.

7.5 Results

7.5.1 Estimation of Traffic Exposure at Ward-Level

In section 7.3, the traffic exposure in each ward was estimated using OD data. Since data was only available for 2004, however, the traffic exposure for other years was estimated using OLS regression. Predictor variables included in the regression were population, employment, road length, IMD score and the employments and populations of proximate wards.

In order to assess how well the traffic exposure for other years are likely to be predicted by the model, two measures were used: the adjusted R^2 value of the model, and the signs and values of the regressors. Table 7.1 shows the results from the model with the adjusted R^2 value of 0.81 indicating that the predictors sufficiently explained the variability in the data set. The employment and population in neighbour wards are positively associated with the traffic exposure. This is consistent with our assumption that pass-through trips need to be accounted for when predicting the traffic exposure. In terms of within-ward trips, the employment population was positively related to the traffic exposure, while the effect of the resident population was less significant. This is probably because trips recorded in the OD data set are mostly commuting trips. It is not surprising, therefore, that employment population is more significant than the resident population in this model. Another finding is that less traffic activities occur in deprived areas with higher IMD scores.

Table 7.1 Model for the traffic exposure regression

	Coef.	Std.Err.	t	P>|t|	[95%	Conf. Interval]
Resident Population	-0.099	0.045	-2.20	0.028	-0.188	-0.011
Road length	0.104	0.006	18.88	0.000	0.093	0.115
IMD score	-61.685	14.163	-4.36	0.000	-89.448	-33.921
Employment Population	0.874	0.029	29.78	0.000	0.816	0.931
Sum of Pop and Emp in Proximate Wards	0.098	0.001	124.25	0.000	0.096	0.099
Constant	-3003.120	368.619	-8.15	0.000	-3725.725	-2280.515
Observations=6952						
Adj R-squared=0.81						

7.5.2 Estimation Results

As discussed in section 7.4, the traffic exposure data estimated at the first stage was then employed as the traffic exposure in the analysis of road casualties. Considering the fact that the traffic exposure data was itself estimated, the bootstrapping approach was applied to correct the standard errors. Tables 7.2, 7.3 and 7.4 show the regression results from bootstrapped models for total casualties, slightly injured casualties and killed and seriously injured casualties respectively. Both standard errors and bootstrapped standard errors are shown, with the latter, for most of the variables, being slightly larger than the former. This is because the variation of the traffic exposure predicted in the first regression was taken into account at the second stage. As expected, traffic exposure is significantly correlated with casualty numbers in all models.

The road network characteristics are divided into three categories: traditional road characteristics of a road network, the degree of connectivity and accessibility, and the curvature. As suggested in many other studies (e.g. Huang et al., 2010), road length and density are positively associated with road casualties at all severity levels. In terms of nodes, wards with higher node density were found to have fewer casualties for all categories of casualties. This is consistent with previous research (Marshall and Garrick, 2011; Ladron de Guevara et al., 2004). An unexpected result is the positive relationship between the number of roundabouts and road casualties. An inverse relationship has been found in previous research (Lord et al., 2007). One possible reason for this is that roundabouts are designed to reduce traffic speed, so there will be more minor injuries but fewer fatalities. So, if we look at total casualties, the coefficient of roundabouts may be the positive, as it includes all injuries. It is a surprise, however, that the sign for fatalities is also positive. This may be due to the fact that roundabouts are relatively more scarce compared to the number of nodes, with approximately 13000 of the former and 850000 of the latter.

It can be hypothesized that areas with a better-connected road network will have more casualties, because since pedestrians, cyclists and motor vehicles have better accessibility total traffic activities tend to be more frequent. Two variables were used as indicators of road network connectivity: the links per node and the number of nodes with one link (Chin et al., 2008).The results indicate that an increase in links per node is associated with an increase in the casualty numbers for all severities. Lower densities of nodes with one link, also known as dead ends, usually indicate limited access to streets. The results show that higher densities of dead ends are associated with fewer casualties and this will be discussed in detail in the next section.

There have been very few studies on the effect of curvature of the road network and, in those that have been conducted, different measures of curvature have been used. Our results suggest road networks with more horizontal curving are associated with fewer casualties for all severity levels. This result is consistent with previous findings (Jones et al., 2008; Quddus, 2008).The mechanisms for this could be complex, however, one possible reason is that vehicles have lower speeds when passing curving road sections.

Table 7.2 Bootstrapped Models for Total Casualties

Total	Coef.	Std. Err.	Bootstrap Std. Err		[95%	Conf. Interval]
Motorway	7.404E-02	2.348E-03	4.661E-03	*	6.944E-02	7.864E-02
Aroad	4.016E-02	8.148E-04	8.358E-04	*	3.856E-02	4.176E-02
Broad	1.953E-02	8.445E-04	8.523E-04	*	1.788E-02	2.119E-02
Minor road	8.454E-04	4.010E-04	4.040E-04	**	5.950E-05	1.631E-03
Number of Nodes	1.215E+00	3.321E-02	3.354E-02	*	1.150E+00	1.281E+00
Vertices Density	-2.549E-01	1.152E-02	1.496E-02	*	-2.774E-01	-2.323E-01
Motorway density	1.450E-01	2.184E-02	2.368E-02	*	1.022E-01	1.878E-01
Aroad density	2.276E-01	6.107E-03	6.225E-03	*	2.156E-01	2.396E-01
Broad density	9.361E-02	7.792E-03	7.992E-03	*	7.834E-02	1.089E-01
Minor road density	3.861E-02	2.570E-03	2.660E-03	*	3.358E-02	4.365E-02
Nodes density	-6.927E-03	3.563E-04	4.647E-04	*	-7.625E-03	-6.229E-03
Number of Roundabouts	5.013E-02	1.467E-03	1.625E-03	*	4.725E-02	5.300E-02
Density of Dead Ends	-1.816E-02	7.962E-04	7.713E-04	*	-1.972E-02	-1.660E-02
Links per Node	1.151E+00	2.362E-02	2.394E-02	*	1.104E+00	1.197E+00
Traffic Exposure	9.055E-03	1.324E-04	2.754E-04	*	8.796E-03	9.315E-03
Constant	-1.357E+00	6.264E-02	6.265E-02	*	-1.479E+00	-1.234E+00

*Significant at 99% level

**Significant at 95% level

Table 7.3 Bootstrapped Models for Slightly Injured Casualties

Slightly Injured	Coef.	Std. Err.	Bootstrap Std. Err.		[95%	Conf. Interval]
Motorway	7.716E-02	2.415E-03	2.580E-03	*	7.243E-02	8.190E-02
Aroad	3.725E-02	8.415E-04	8.756E-04	*	3.560E-02	3.890E-02
Broad	1.719E-02	8.732E-04	9.049E-04	*	1.548E-02	1.891E-02
Minor road	---	---	---		---	---
Number of Nodes	1.183E+00	3.421E-02	3.531E-02	*	1.116E+00	1.250E+00
Vertices Density	-2.340E-01	1.188E-02	1.265E-02	*	-2.573E-01	-2.107E-01
Motorway density	1.533E-01	2.246E-02	2.455E-02	*	1.093E-01	1.973E-01
Aroad density	2.417E-01	6.278E-03	6.334E-03	*	2.294E-01	2.541E-01
Broad density	1.048E-01	8.019E-03	8.274E-03	*	8.913E-02	1.206E-01
Minor road density	4.456E-02	2.644E-03	2.853E-03	*	3.938E-02	4.975E-02
Nodes density	-7.599E-03	3.657E-04	3.787E-04	*	-8.316E-03	-6.882E-03
Number of Roundabouts	5.345E-02	1.508E-03	1.268E-03	*	5.050E-02	5.641E-02
Density of Dead Ends	-1.892E-02	8.254E-04	8.267E-04	*	-2.053E-02	-1.730E-02
Links per Node	1.170E+00	2.475E-02	2.483E-02	*	1.122E+00	1.219E+00
Traffic Exposure	8.635E-03	1.354E-04	1.657E-04	*	8.370E-03	8.900E-03
Constant	-1.579E+00	6.569E-02	6.568E-02	*	-1.707E+00	-1.450E+00

*Significant at 99% level

**Significant at 95% level

Table 7.4 Bootstrapped Models for Killed and Seriously Injured Casualties

Killed and Seriously Injured	Coef.	Std. Err.	Bootstrap Std. Err.		[95%	Conf. Interval]
Motorway	5.059E-02	4.986E-03	6.063E-03	*	4.082E-02	6.036E-02
Aroad	5.039E-02	1.888E-03	1.975E-03	*	4.669E-02	5.409E-02
Broad	2.890E-02	2.004E-03	2.021E-03	*	2.498E-02	3.283E-02
Minor road	5.863E-03	8.496E-04	8.808E-04	*	4.197E-03	7.528E-03
Number of Nodes	1.272E+00	6.836E-02	7.099E-02	*	1.138E+00	1.406E+00
Vertices Density	-3.123E-01	2.373E-02	2.427E-02	*	-3.588E-01	-2.658E-01
Motorway density	---	---	---	---	---	---
Aroad density	1.624E-01	1.386E-02	1.626E-02	*	1.352E-01	1.896E-01
Broad density	4.901E-02	1.861E-02	2.031E-02	*	1.252E-02	8.549E-02
Minor road density	1.803E-02	6.007E-03	6.436E-03	*	6.256E-03	2.980E-02
Nodes density	-5.139E-03	8.283E-04	8.302E-04	*	-6.762E-03	-3.515E-03
Number of Roundabouts	7.506E-03	3.000E-03	3.297E-03	**	1.626E-03	1.338E-02
Density of Dead Ends	-8.905E-03	1.601E-03	1.647E-03	*	-1.204E-02	-5.767E-03
Links per Node	9.487E-01	5.000E-02	5.224E-02	*	8.507E-01	1.047E+00
Traffic Exposure	9.076E-03	2.945E-04	3.689E-04	*	8.499E-03	9.653E-03
Constant	-6.693E-01	1.431E-01	1.431E-01	*	-9.497E-01	-3.888E-01

*Significant at 99% level
**Significant at 95% level

To account for the extra variation due to spatial dependence among the observations, full Bayesian models were applied and compared with the traditional NB models, as discussed in section 7.2. As expected, the Poisson-Lognormal model accounting for spatial correlation exhibited the lowest DIC value, indicating that it performs the best among all the Bayesian models.

The results from Bayesian spatial models and traditional NB models are very similar. This is probably because the extra variation was largely due to area-specific heterogeneity, which was controlled for in both models. Results show that increased density of vertices, density of nodes, and density of dead ends are associated with reduced road casualties, while there were positive relationships between road casualties and other factors, such as road length, road density, traffic exposure and the ratio of link-to-node.

Table 7.5 Full Bayesian Models for Total Casualties

Total	PL		PL with Random Effects		PL with Random Effects and Spatial Effects	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Motorway	4.595E-02	6.096E-04	5.279E-02	8.035E-04	5.277E-02	7.982E-04
Aroad	3.240E-02	2.737E-04	3.383E-02	3.510E-04	3.377E-02	3.509E-04
Broad	2.074E-02	3.190E-04	1.650E-02	3.938E-04	1.646E-02	3.898E-04
Minor road	5.047E-03	1.115E-04	4.184E-03	1.931E-04	4.164E-03	1.642E-04
Number of Nodes	1.295E+00	8.783E-03	9.689E-01	1.608E-02	9.669E-01	1.190E-02
Vertices Density	-3.545E-01	2.994E-03	-2.475E-01	5.472E-03	-2.464E-01	3.996E-03
Motorway density	1.318E-01	6.879E-03	7.283E-02	8.226E-03	7.389E-02	8.164E-03
Aroad density	2.183E-01	1.896E-03	2.023E-01	2.344E-03	2.030E-01	2.305E-03
Broad density	8.988E-02	2.731E-03	9.462E-02	3.280E-03	9.564E-02	3.294E-03
Minor road density	5.941E-02	7.408E-04	2.934E-02	1.262E-03	3.016E-02	1.154E-03
Nodes density	-7.804E-03	1.084E-04	-2.164E-03	1.586E-04	-2.201E-03	1.488E-04
Number of Roundabout s	3.607E-02	4.188E-04	3.654E-02	5.427E-04	3.676E-02	5.242E-04
Density of Dead Ends	-1.723E-02	1.960E-04	-1.588E-02	3.234E-04	-1.576E-02	2.923E-04
Links per Node	1.429E+00	4.880E-03	1.083E+00	1.311E-02	1.073E+00	1.043E-02
Traffic Exposure	4.444E-03	2.503E-05	2.389E-03	5.268E-05	2.341E-03	5.197E-05
Constant	-1.799E+00	1.307E-02	-9.066E-01	3.572E-02	-8.799E-01	2.825E-02
DIC Value	20916		17562		17283	

Table 7.6 Full Bayesian Models for Slightly Injured Casualties

Slight	PL		PL with Random Effects		PL with Random Effects and Spatial Effects	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Motorway	4.711E-02	6.640E-04	5.630E-02	8.657E-04	5.611E-02	8.619E-04
Aroad	3.028E-02	3.135E-04	3.227E-02	3.752E-04	3.216E-02	3.823E-04
Broad	1.907E-02	3.449E-04	1.544E-02	4.220E-04	1.533E-02	4.186E-04
Minor road	4.770E-03	1.504E-04	3.919E-03	1.829E-04	3.825E-03	1.824E-04
Number of Nodes	1.326E+00	1.303E-02	1.005E+00	1.402E-02	9.917E-01	1.227E-02
Vertices Density	-3.613E-01	4.548E-03	-2.557E-01	4.695E-03	-2.510E-01	4.402E-03
Motorway density	1.395E-01	7.143E-03	6.643E-02	8.711E-03	6.804E-02	8.998E-03
Aroad density	2.253E-01	1.991E-03	2.109E-01	2.532E-03	2.118E-01	2.489E-03
Broad density	9.785E-02	2.931E-03	9.999E-02	3.548E-03	1.012E-01	3.513E-03
Minor road density	6.141E-02	9.172E-04	3.321E-02	1.250E-03	3.408E-02	1.269E-03
Nodes density	-7.982E-03	1.199E-04	-2.529E-03	1.570E-04	-2.555E-03	1.579E-04
Number of Roundabouts	3.857E-02	4.487E-04	3.831E-02	5.758E-04	3.857E-02	5.723E-04
Density of Dead Ends	-1.822E-02	2.362E-04	-1.682E-02	3.689E-04	-1.679E-02	3.199E-04
Links per Node	1.477E+00	1.115E-02	1.127E+00	1.573E-02	1.118E+00	1.182E-02
Traffic Exposure	4.243E-03	2.575E-05	2.360E-03	5.610E-05	2.308E-03	5.653E-05
Constant	-2.088E+00	2.975E-02	-1.198E+00	4.271E-02	-1.177E+00	3.315E-02
DIC Value	17584		15495		15275	

Table 7.7 Full Bayesian Models for Killed and Seriously Injured Casualties

Killed and Seriously Injured	PL		PL with Random Effects		PL with Random Effects and Spatial Effects	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Motorway	4.161E-02	1.568E-03	3.752E-02	1.994E-03	3.706E-02	1.964E-03
Aroad	4.367E-02	6.949E-04	4.227E-02	8.268E-04	4.178E-02	7.978E-04
Broad	2.924E-02	7.721E-04	2.331E-02	9.225E-04	2.285E-02	8.736E-04
Minor road	7.223E-03	3.555E-04	6.231E-03	5.009E-04	5.912E-03	4.071E-04
Number of Nodes	1.186E+00	3.104E-02	9.247E-01	4.300E-02	8.902E-01	3.329E-02
Vertices Density	-3.440E-01	1.064E-02	-2.563E-01	1.495E-02	-2.433E-01	1.113E-02
Motorway density	4.784E-02	1.918E-02	8.670E-02	2.272E-02	9.002E-02	2.269E-02
Aroad density	1.769E-01	5.337E-03	1.581E-01	6.016E-03	1.587E-01	6.062E-03
Broad density	4.258E-02	7.592E-03	5.486E-02	8.866E-03	5.720E-02	8.770E-03
Minor road density	4.484E-02	2.147E-03	1.503E-02	2.967E-03	1.580E-02	2.874E-03
Nodes density	-6.799E-03	3.023E-04	-2.029E-03	3.926E-04	-1.894E-03	3.855E-04
Number of Roundabouts	1.898E-02	1.210E-03	2.163E-02	1.440E-03	2.300E-02	1.459E-03
Density of Dead Ends	-1.183E-02	5.883E-04	-1.231E-02	7.147E-04	-1.215E-02	7.810E-04
Links per Node	1.167E+00	1.581E-02	9.695E-01	2.216E-02	9.372E-01	2.241E-02
Traffic Exposure	5.762E-03	6.467E-05	4.462E-03	1.216E-04	4.128E-03	1.253E-04
Constant	-2.984E+00	4.267E-02	-2.494E+00	6.031E-02	-2.405E+00	6.162E-02
DIC Value	10997		8228		8070	

7.5.3 Dose-Response Functions for Continuous Variables

As discussed earlier, it may not be appropriate to assume a linear relationship between the outcomes and the treatment in the continuous treatment case, where the treatment effect heterogeneity has become a primary interest of researchers. The panel data semi-parametric model was applied in this study to investigate how the treatment effect varied over the treatment doses.

Figure 7.4 displays the average non-parametric fit of links per node, with the shadow indicating the 95% confidence interval. It can be seen that there is an increase trend in the total casualties over the value of links per node, which is consistent with the results obtained from both the bootstrapped model and the FB model. It is obvious, however, that the relationship does not follow a linear trend. The treatment effect remains largely unchanged under the value of 2.5 of links per node, whilst above this value there is a significant increase. Figure 7.5 shows the marginal effect of the dead end density on total casualties. It is evident that there are more casualties in an area with lower dead end density. Specifically, the number of total casualties reaches a peak when the dead end density is near zero, which indicates a highly connected and accessible road network.

Both treatment variables indicate the degree of the connectivity and complexity of the road network. The graphs show that a better connected and more complex road network (i.e. with a higher value of links per node and lower dead end density) may experience more crashes.

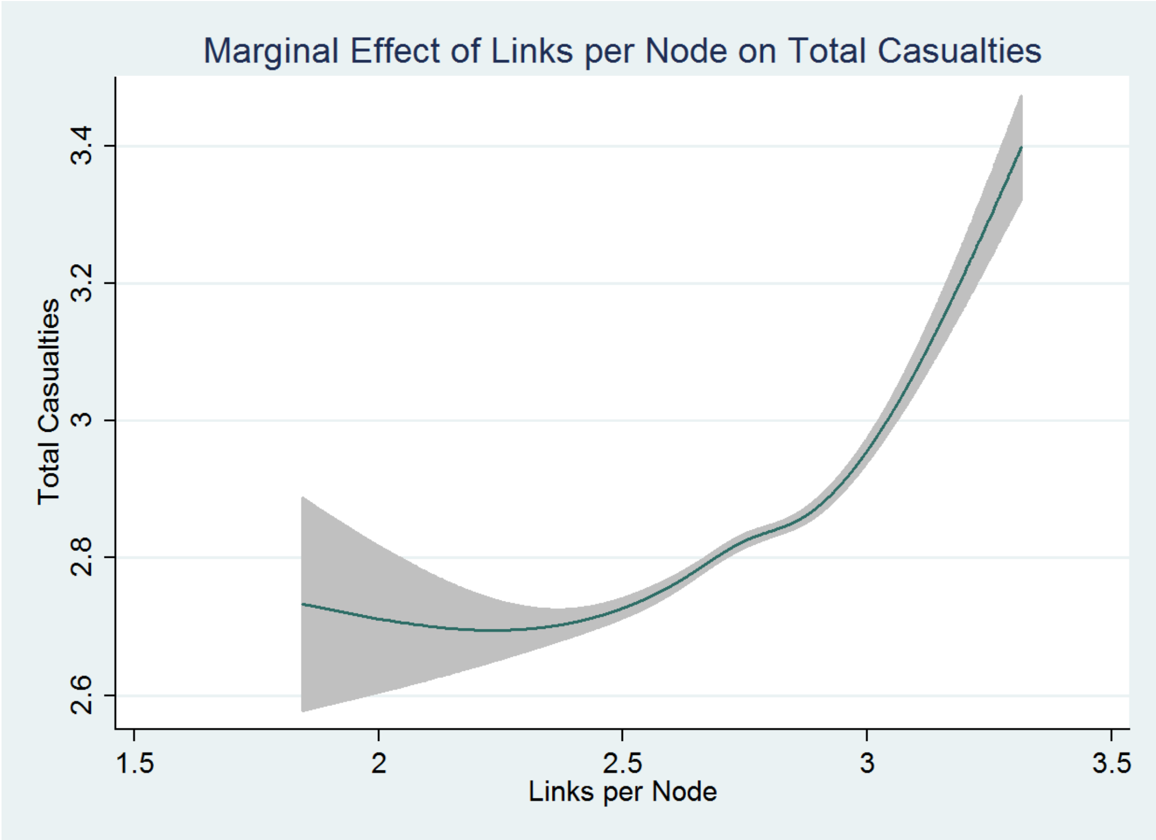

Figure 7.4 Marginal Effects of Links per Node on Total Casualties

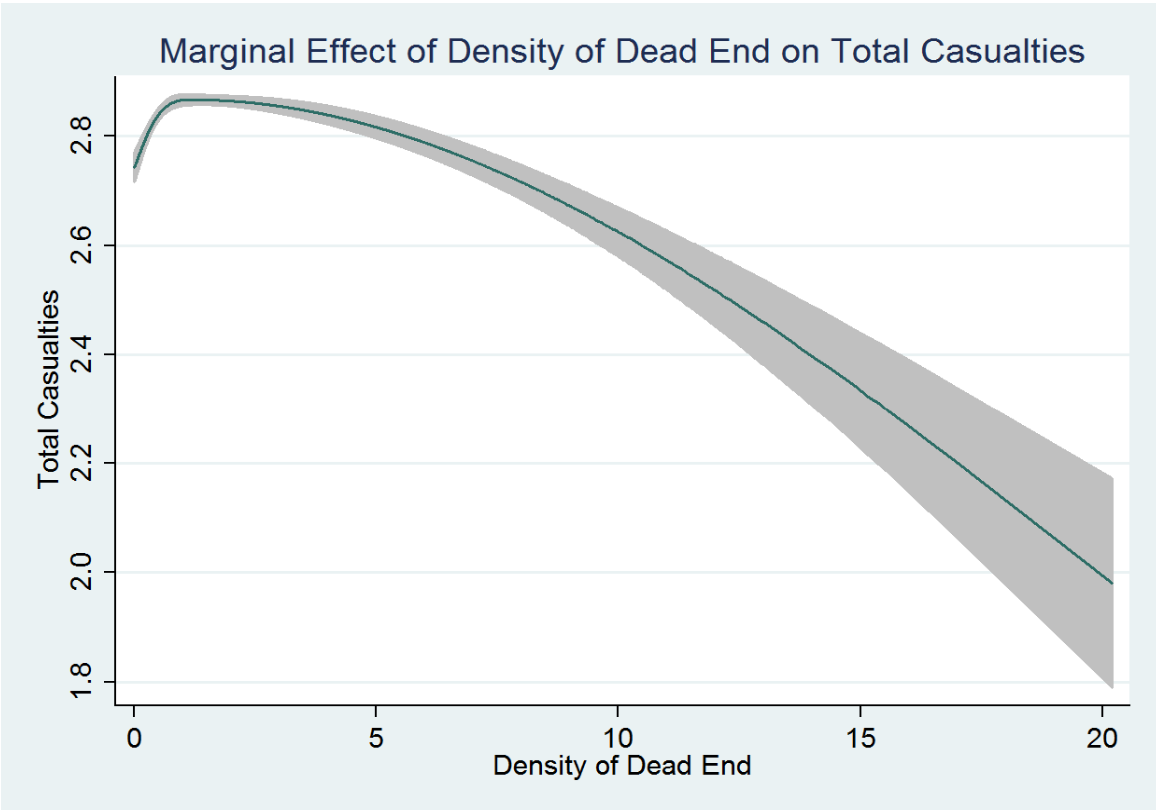

Figure 7.5Marginal Effects of Density of Dead Ends on Total Casualties

7.6 Conclusions

In this study NB models and full Bayesian models were employed to analyse the relationships between road casualties and various road network characteristics in the UK. The results are consistent with the general conclusions of previous research in this field. Furthermore, several outstanding issues in previous research have been addressed.

Several studies at an aggregation level, concerned with analysing the effects of road network factors on road casualties, have been conducted in UK. One common problem in these studies is that the data for the road network is not longitudinal due to the availability of data. In this research, OS Meridian^{TM} 2 from 2001-2010 was obtained, which enables the study to control for the variation in the road network across time. Various factors related to road networks have been proposed previously. To the best of the author's knowledge, however, these have not been examined simultaneously at the ward-level. By using GIS software, such as ArcGIS, this study has been able to generate the potential explanatory variables at the ward-level. In particular, these variables are divided into three categories: (1) traditional road characteristics, such as road length and road class; (2) accessibility and connectivity of road network; (3) curvature of road. Applying the panel data for the road network, most of the findings are in line with previous research. Marshall and Garrick (2011) include dead ends as explanatory variable in their model, however, no significant relationship is found. In this research, it has been shown that the accessibility and connectivity of road network is an important factor affecting road casualties using ratio of node to link and density of dead ends as measures. Furthermore, a panel semiparametric model was used to estimate the dose-response functions of these two continuous treatments to show the heterogeneity of treatment effects.

The traffic exposure is a critical issue in road accident analysis; however, there is a lack of appropriate variables that can be used for traffic exposure at an aggregation level. In this research, a method is proposed for constructing traffic exposure at the ward-level based on OD data. A synthetic road network was first built among wards and OD trips were assigned to this network using a traffic assignment method. Trips were aggregated at the ward-level and fitted by a regression model. The high adjusted

R^2 value indicates that the traffic exposure data was well fitted by the model, which validates the predictions of the traffic exposure. To control for the additional variation due to the traffic exposure data, a two-stage bootstrapped model was developed. Most variables in this model were still significant, despite a slight increase in standard errors.

Previous studies give diverse answers regarding the test of spatial correlation. Jones et al. (2008), for example, find no positive spatial correlation using the data at the district-level, while there is an opposite result when the data aggregation comes to the ward-level (Quddus, 2008). The test results in this study also show that there is significant spatial clustering among the observations. This confirms the contention that spatial dependence would increase at lower level of aggregation.

Chapter 8: Conclusions

This thesis contributed to the literature on the causal link between traffic interventions and road casualties by employing formal causal models. Chapters5 to 7, the three empirical chapters of the thesis, examined the effects of traffic interventions on road safety. Three interventions were focused on: the London congestion charge, speed limit enforcement cameras, and the design of the road network.

The first section of this chapter attempts to pull together the findings from the empirical chapters of the thesis. Section 8.2 explains the limitations of the study. Finally, section 8.3 sets out possible directions for future research.

8.1 Main Findings and Contributions

As discussed in chapter 1, this study contributes to the literature from three aspects. In this section the main findings of the thesis in relation to each of these main findings are summarised.

(1) The establishment of the causal link between traffic interventions and road safety.

Three empirical studies were conducted to establish the causal link between road safety and traffic interventions with different purposes: the London congestion charge, speed limit enforcement cameras, and the design of the road network.

In chapter 5, the impacts of the London congestion charge on road casualties in the central London area was investigated. The primary aim of the LCC is not to improve road safety, however the safety impact of the charge should not be neglected. As discussed in chapter3, traffic volume is the single most important factor influencing road safety (Golob et al., 2003; Martin, 2002; Dixit et al., 2011; Lord et al., 2005), although the risk of injury can also vary considerably between different travelling modes (Elvik, 2004; White, 2004; Leigh and Wilkinson, 1991; Crandall and Graham, 1989). It was envisaged that by influencing travel modes and redistributing the traffic demand in space and time, the London congestion charge would cause changes in both the number and type of casualties.

The results of this study suggest that there has been a significant reduction of 5.2% in car casualties due to a reduction in traffic within the congestion charging area (Tuerk and Graham, 2010). Meanwhile, motorcycle- and cycle-related casualties have increased by 5.7% and 13.3% respectively after the LCC, probably because more motorcycles and bicycles have been used instead of cars (TfL, 2004). Our results are largely consistent with the conclusions of the previous research by Noland et al. (2008).

The study on the LCC highlights that more attention needs to be paid to road safety strategies, especially for cycle users, when introducing traffic demand management interventions such as road pricing and taxation. Policy makers need to be aware of the potential shifts in travel modes among road users, which can increase the proportion of vulnerable road users and lead to an increase in the number of casualties.

In chapter 6, the safety effects of speed limit enforcement cameras were evaluated. There is still debate about the effectiveness of speed cameras and it has been pointed out that the existing research has failed fully to address issues of confounding and RTM effects. In this study, formal causal approaches, which have been used routinely in other areas of science such as medicine and epidemiology, were employed to estimate the effects of speed cameras. The results show an average reduction in PICs of around 1.068 (25.9% in percentage terms), and an average reduction of 0.132 in FSCs. This is in line with the findings of previous studies. For example, Mountain et al. (2005) find a reduction of 22% in personal injury accidents using the EB method. Hess and Polak (2004) employ ARIMA/SARIMA to estimate the safety effects of camera sites in Cambridgeshire and find a reduction of 21% injury accidents.

One possible phenomenon that affects the evaluation results is called accident migration. One type of this phenomenon, the kangaroo effect, may arise if drivers decelerate and accelerate abruptly before and after the camera sites. This may increase the number of casualties near camera sites. The results, however, show no evidence that the kangaroo effect causes an increase in casualties upstream or downstream of camera sites. Another type of accident migration is the choice of alternative routes to avoid speed cameras. Then accident reduction on roads with cameras could be both the effects of the installation of speed cameras and reduced traffic exposure. In this study, we include the AADF to control for the effect due to changes in traffic exposure.

One concern raised in this study is that police enforcement focusing on traffic violations should have a proven and direct relationship with road safety. The enforcement should be implemented at locations and at times where offences are expected to have the most effect on road safety. In other words, the criteria for assigning the enforcement need to be clearly justified and validated. For both PICs and FSCs, we found a greater reduction in casualties as the distance to the cameras decreased, so that speed cameras were found to be most effective at a distance up to 200 meters from camera sites. The cumulative reduction increases dramatically from 0 to 500 meters and this tendency reduces from 500 meters to 1km. The suggested effective length of camera sites by DfT (2004) is between 400 meters to 1.5 km, which is generally consistent with the effective length estimated in this study.

In chapter 7, the relationships between road casualties and various road network characteristics in UK were analysed. Road network characteristics are divided into three categories: traditional road characteristics, accessibility and connectivity of the road network and curvature of road. The estimation results for traditional road characteristics were consistent with the findings in previous studies (Quddus, 2008; Huang et al., 2010; Marshall and Garrick, 2011). A main finding in this study is that areas with a better connected road network may experience more casualties, because both pedestrian and motor vehicles have better accessibility and traffic activities tend to be more frequent. Our results also suggest more curvature in the road network is associated with fewer casualties for all severity levels.

In general, it is difficult to come up with an optimal road network, especially when the existing network has been put into use for a long time. However, a better knowledge of impacts of road network design on road safety can provide useful policy implications. For instance, in this study, the results show that a road network with high connectivity and accessibility has more casualties. Although the road network cannot be changed in the short term, it is possible for policy makers to enhance road safety countermeasures in areas with such road networks.

To summarize, despite the substantial literature that exists on the relationship between various factors and road casualties, few studies have investigated the effects of traffic interventions on road safety. The causal link between traffic interventions and road traffic accidents remains unclear. Most traffic interventions are implemented for a general purpose and impact road safety directly or indirectly by influencing traffic conditions, travel modes, driving environment and behaviours. Hence the causal relationship between traffic interventions and road safety is not straightforward and is difficult to establish. For policy makers, however, whether as part of national government or local authorities, a better understanding of the safety outcome of traffic interventions can help to improve road accident prevention when implementing interventions. The methods for evaluating such interventions are presented in detail in this thesis and can be easily followed by any policy maker.

(2) Application and development of formal causal approaches.

The second contribution made in this research relies on the application and development of formal causal approaches which have not previously been applied in the field of road safety. In this thesis the DID model, PSM method and Bayesian methods are each employed to evaluate the causal effects of traffic interventions on road casualties.

One obstacle in causal analysis is to control for confounding effects, such as the RMT effect and time trend effects. In chapter 5, a full DID model is developed to estimate the impacts of the LCC on road casualties. The DID model introduced in this study shares the same conceptual basis of the EB approaches that “accident counts are not the only clue to the safety of an entity; another clue is in what is known about the safety of similar entities” (Hauer, 2002). The DID model can control for the RMT and any temporal effect conditional on parallel assumption. Unlike the EB method, the DID does not require a large sample of reference groups which are similar to the treatment group. Moreover, the word “similar” in the EB method is very ambiguous. In contrast, a clear definition of the control group is given in the DID method: it should be independent of the treatment; and the treatment and control groups must have a parallel time trend of casualties count. The DID method is therefore more flexible and tractable. Furthermore, the DID method used in this thesis is extended by employing the synthetic control method which can account for the issue of parallel assumption.

As discussed earlier, a reference or control group is usually employed to estimate the counterfactual outcomes of the treatment group. Due to the selection bias, however, the treated and untreated individuals may differ in the absence of the treatment. Only untreated individuals with similar characteristics to those treated can be used to approximate the counterfactual outcomes of the treatment group. What has remained unclear in previous studies is how the reference or control group is selected. To address this issue, the propensity score is introduced in chapter 6 as an alternative approach for selecting the control group. The attractiveness of the PMS method is that it gives a clear criterion by which to select the control group and it enables matching to be reduced to a single dimension. Five matching algorithms and balancing tests were used to assess the matching quality. It has been shown that the characteristics of the treatment and control groups are well balanced after matching. Furthermore, the propensity score method was combined with the DID and the EB approaches separately. The results confirm that the EB using matched sites as the reference group is superior to the conventional one. It is therefore suggested that propensity score can be applied as the criterion when constructing the control or reference group.

(3) Addressing issues regarding the data employed for road casualty analysis.

Besides econometric causal models, conventional causal models are also employed in this study. In chapter 7, the FB method is used to investigate the relationship between road casualties and road network characteristics. The main contribution made in this chapter is to address two critical issues regarding the data used in road safety analysis. First, panel data for the road network is used to account for effects due to the variation in road characteristics over time. Detailed information about the road network was obtained by using OS Meridian ^{TM} 2 for a period from 2001 to 2010. Another issue which has not been studied in-depth concerns the exposure variable at aggregated area level. While the average AADT has become a dominant exposure variable in disaggregated analysis, there is a lack of an appropriate exposure variable for analysis conducted at area level. In this study, an approach based on traffic assignment was proposed to estimate the traffic exposure at ward level. Trips generated between ODs were assigned to transportation networks and aggregated in each ward. The aggregated trips were further fitted by a regression model to make predictions for other years. The high adjusted R^2 value indicates that the traffic exposure data is well fitted by the model used here, which validates the predictions of the traffic exposure. The results give rise to certain recommendation regarding the use of new traffic exposure variable in the aggregate analysis.

8.2 Limitations

There are several limitations with the empirical studies in this thesis. In the analysis of causal effects of the LCC on road safety, only the effects for a short period from 2003 to 2004 were estimated. As discussed in chapter 5, there were four periods in the implementation of the LCC. The DID model developed here was not suitable for estimating outcomes where the treatment status takes multiple values. There are two reasons for this: (a) a comparison or control group is difficult to define clearly, (b) correct statistical inferences require the joint estimation of all treatment effects. Hence this study cannot investigate the impacts of the western extension of the charging area or the increase in congestion fee.

The study of the safety effects of speed cameras also has limitations. First, due to data availability, speed cameras in only eight districts were evaluated. The estimation results obtained in this study may not, therefore, represent the effects of speed cameras in other districts. Furthermore, although the safety effects of speed cameras for different section lengths, were estimated, however the impact of speed cameras over 1000 metres is still not clear, again due to data availability. There might be an inflection point where the effect of speed cameras on casualties becomes zero. A critical issue in the application of the PSM method regards the selection guidelines for speed camera sites. Despite the availability of such guidelines in the UK, there is diversity in the implementation among local authorities. For example, speed cameras can be installed at a site where the local authority believes there is a community concern or engineering factor. The practical criteria for selecting camera sites can be different across areas. Specified camera site selection guidelines from local authorities would certainly improve the propensity score model.

There are also some limitations of the employed methods. For example, the PSM method applied in this thesis cannot adjust for spatial correlation among the camera sites from the same district. This may have influences on the estimation of safety effects of speed cameras. The methods employed in this thesis are univariate rather than multivariate count data model, which means that the correlation between the variables representing different severity levels cannot be accounted for. Furthermore, this study does not apply models for fatal accidents to avoid the zero-inflated problem due to the data availability.

8.3 Future Research

This thesis addressed some of the gaps in the literature on the casual link between traffic interventions and road safety. In this section, some further directions for future research are suggested.

A large number of the recent studies on safety evaluations have focused only on estimation of the average treatment effects on treated individuals. Although researchers have typically allowed for general treatment effect heterogeneity, there has been little formal investigation of the presence of such heterogeneity. In treatment evaluation studies, it is sometimes interesting to learn about distributional effects besides the average effects of the treatment, and to examine whether there is any subpopulation for which a programme or treatment has a non-zero average effect, or whether there is treatment effect heterogeneity. For example, a policy maker might be interested in the effect of a treatment on the lower or higher tail of the outcome distribution. In recent years, there have been some studies on this issue. Crump et al. (2006) developed two nonparametric tests for the presence of treatment effect heterogeneity. Later, Firpo (2007) proposed approaches for estimating quantile treatment effects with the restriction that the treatment assignment is based on observable characteristics. This has not been studied in the field of road safety, however. A study on treatment effect heterogeneity would make an interesting question.

Most of the recent studies have mainly focused on a binary treatment. Little attention is devoted to investigating settings with multi-valued, discrete or continuous treatments, which are common in practice. Hirano and Imbens (2004) proposed a generalization of the binary treatment propensity score, known as the generalized propensity score (GPS). They demonstrate that the GPS has many of the attractive properties of the binary treatment propensity score. The GPS has been applied by several studies in the context of evaluating active labour market policy (Flores et al., 2007; Kluve, 2009). These studies also provide scope for future research.

In this thesis the causal approaches applied are univariate rather than multivariate. Multivariate techniques allow researchers to look at relationships between variables and quantify the relationship between variables. This gives a much richer and realistic picture and provides a powerful test of significance compared to univariate models. To the best of the author’s knowledge, the multivariate techniques have not been applied in causal analysis. This would be an interesting topic for the future research.

References

1. Abadie, A., A. Diamond, and J. Hainmueller. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. Journal of the American Statistical Association. Vol. 105, 2009, pp. 493-505.
2. Abadie, A., and J. Gardeazabal. The Economic Costs of Conflict: A Case Study of the Basque Country. American Economic Review, Vol. 93, No. 1, 2003, pp. 113-132.
3. Abadie, A.. Semiparametric Difference-in-Differences Estimators. The Review of Economic Studies, Vol. 72, No. 1, 2005, pp. 1-19.
4. Abdel-Aty, M. A., and A. E. Radwan. Modeling traffic accident occurrence and involvement. Accident Analysis and Prevention, Vol. 32, 2000, pp. 633-642.
5. Aguero-Valverde, J., and P. P. Jovanis. Spatial analysis of fatal and injury crashes in Pennsylvania. Accident Analysis and Prevention, Vol. 38, 2006, pp. 618-625.
6. Aljanahi, A. A. M., A. H. Rhodes, and A. V. Metcalfe. Speed, speed limits and road traffic accidents under free flow conditions. Accident Analysis and Prevention, Vol. 31, 1999, pp. 161-168.
7. Allpress, J. A., and L. S. Leland Jr.. Reducing traffic speed within roadwork sites using obtrusive perceptual countermeasures. Accident Analysis and Prevention, Vol. 42, 2010, 377-383.
8. Amadori, M., and T. Boino. A methodology to define the level of safety of public transport bus stops, based on the concept of risk. Social and Behavioral Sciences, Vol. 48, 2012, pp. 653-662.
9. Amoros, E., J. L. Martin, and B. Laumon. Comparison of road crashes incidence and severity between some French counties. Accident Analysis and Prevention, Vol. 35, 2003, pp. 537-547.
10. Angrist, J., E. Bettinger and M. Kremer. Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia. American Economic Review, Vol. 96, 2006, pp. 847-862.
11. Ashenfelter, O., and D. Card,. Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs. Review of Economics and Statistics, Vol. 67, 1985, pp. 648-660.

12. Ashenfelter, O.. Estimating the Effect of Training Programs on Earnings. Review of Economics and Statistics, Vol. 60, 1978, pp. 47-57.
13. Augurzky, B., and C. Schmidt. The Propensity Score: A Means to An End. Working Paper, University of Heidelberg, 2000.
14. Banerjee, A., E. Duflo, E., S. Cole, and L. Linden. Remedying Education: Evidence from Two Randomized Experiments in India. Quarterly Journal of Economics, Vol. 122, 2007, pp. 1235-1264.
15. Barnard, G. A.. "Causation," Encyclopedia of Statistical Sciences, Volume 1, 1982, pp. 387-389.
16. Barnett, A. G., J. C. van der Pols, and A. J. Dobson. Regression to the mean: what it is and how to deal with it. International Journal of Epidemiology, Vol. 34, 2005, pp. 215-220.
17. Baruya, A.. Speed-accident relationships on European roads. 9th International Conference, Road Safety in Europe, 1998.
18. Bedard, M, G. H. Guyatt, M. J. Stones, and J. P. Hirdes. The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. Accident Analysis and Prevention, Vol. 34, 2002, pp. 717-727.
19. Besag, J., J. York, and A. Mollie. Bayesian image restoration with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics, Vol. 43, 1991, pp. 1-59.
20. Bitler, M., J. Gelbach, and H. Hoynes. What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments. Working paper, Department of Economics, University of Maryland, 2002.
21. Black, D., and J. Smith. How Robust is the Evidence on the Effects of the College Quality? Evidence from Matching. Journal of Econometrics, Vol. 121(1), 2004, pp. 99-124.
22. Bonsall, P., and C. Kelly. Road user charging and social exclusion: The impact of congestion charges on at-risk groups. Transport Policy, Vol. 12, 2005, pp. 406-418.
23. Box, G. E. P., and G. C. Tiao. Intervention Analysis with Applications to Economic and Environmental Problems. Journal of the American Statistical Association, Vol. 70, No. 349, 1975, pp. 70-79.

24. Camagni, R., M. Gibelli, and P. Rigamonti. Urban mobility and urban form: the social, environmental costs of different patterns of urban expansion. Ecological Economics, Vol. 40 (2), 2002, pp. 199-216.
25. Card, D., and A. B. Krueger. Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. The American Economic Review, Vol. 84, No. 4, 1994, pp. 772-793.
26. Carriquiry, A., and M.D. Pawlovich. From Empirical Bayes to Full Bayes: Methods for Analyzing Traffic Safety Data. Department Report, Iowa Department of Transportation, 2005.
27. Central London Congestion Charging Impacts monitoring Fifth Annual Report. Transport for London, 2007.
28. Central London Congestion Charging Impacts monitoring Second Annual Report. Transport for London, 2004.
29. Charlton, M., S. Fotheringham, and C. Brunsdon. GEOGRAPHICALLY WEIGHTED REGRESSION. NCRM Methods Review Papers, ESRC National Centre for Research Methods, 2005.
30. Chi, G, A. Cosby, P. Gilbert, and D. Levinson. Gasoline Prices and Traffic Safety: Age and Gender Variations. Working paper, University of Minnesota, 2009.
31. Chin, G. K. W., K. P. Van Niel, B. Giles-Corti, and M. Knuiman. Accessibility and connectivity in physical activity studies: The impact of missing pedestrian data. Preventive Medicine, Vol. 46, 2008, pp. 41-45.
32. Chin, H. C., and M. A. Quddus. Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. Accident Analysis and Prevention, Vol. 35, 2003, pp. 253-259.
33. Clifton, K. J., and K. Kreamer-Fults. An examination of the environmental attributes associated with pedestrian–vehicular crashes near public schools. Accident Analysis and Prevention, Vol. 39, 2007, pp. 708-715.
34. Corsi, T. M. and P. Fanara. Effects of New Entrants on Motor Carrier Safety. Transportation Safety in an Age of Deregulation, 1989, pp. 241-257.
35. Crandall, R. W., and J. D. Graham. The Effect of Fuel Economy Standards on Automobile Safety. Journal of Law and Economics, Vol.32, 1989, pp. 97-118.

36. Crump, R., V. J. Hotz, V. J., G. Imbens, and O.Mitnik. Nonparametric Tests for Treatment Effect Heterogeneity. Review of Economics and Statistics, Vol. 90 (3), 2008, pp. 389-405.
37. Dee, T.S.. Does setting limits save lives? The case of 0.08 BAC laws. Journal of Policy Analysis and Management, Volume 20, 2001, pp. 111-118.
38. dell'Olio, L., A. Ibeas,, and P. Cecin. The quality of service desired by public transport users. Transport Policy, Vol. 18, 2011, pp. 217-227.
39. Dent, A., and S. Bond. An investigation into the location and commuting patterns of part-time and full-time workers in the United Kingdom, using information from the 2001 Census. Office for National Statistics, 2008.
40. Dissanayake, D., J. Aryaija, and D. M. P. Wedagama. Modelling the effects of land use and temporal factors on child pedestrian casualties. Accident Analysis and Prevention, Vol. 41, 2009, pp. 1016-1024.
41. Dixit, V. V., A. Pande,M. Abdel-Aty, A. Das, and E. Radwan. Quality of traffic flow on urban arterial streets and its relationship with safety. Accident Analysis and Prevention, Vol. 43, 2011, pp. 1610-1616.
42. Dumbaugh, E., and W. Li. Designing for the Safety of Pedestrians, Cyclists, and Motorists in Urban Environments. Journal of the American Planning Association, Vol. 77, 2010, pp. 69-88.
43. Dumbaugh, E.. The design of safe urban roadsides: an empirical analysis. Transportation Research Board 85th Annual Meeting, Washington, D.C., 2006.
44. Ekman, R., G. Welander, L. Svanstro¨m, and L. Schelp. Long-term effects of legislation and local promotion of child restraint use in motor vehicles in Sweden. Accident Analysis and Prevention, Vol. 33, 2001, 793-797.
45. El-Basyouny, K., and T. Sayed. A full Bayes multivariate intervention model with random parameters among matched pairs for before–after safety evaluation. Accident Analysis and Prevention, Volume 43, 2011, pp. 87-94.
46. Eliasson, J., and L. Mattsson. Equity effects of congestion pricing Quantitative methodology and a case study for Stockholm. Transportation Research Part A, Vol. 40, 2006, pp. 602-620.
47. Eluru, N., C. R. Bhat, and D. A. Hensher. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accident Analysis and Prevention, Vol. 40, 2008, pp. 1033-1054.

48. Elvik, R., and T. Vaa. The Handbook of Road Safety Measures. Elsevier, London, 2004.
49. Elvik, R.. Area-wide urban traffic calming schemes: a meta-analysis of safety effects. Accident Analysis and Prevention, Vol. 33, 2001, 327-336.
50. Elvik, R..*Endringer i antall ulykker der godsbiler har vmt innblandet i perioder med skjerping og liberalisering* ***av*** *samferdselsloven.* Arbeidsdokument TST/0861/97. Transportskonomisk institutt, Oslo, 1997.
51. Elvik, R..*Trajkanters eksponering og risiko i vegtrajkk.* Arbeidsdokument TST/0775/96. Transportskonomisk institutt, Oslo, 1996.
52. European Commission. Commission recommendation of 17 January 2001 on the maximum permitted blood alcohol content (BAC) for drivers of motorised vehicles, 2001. Available from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:043:0031:0036:EN:PDF.
53. Evans, A. Bus accidents, bus deregulation and London. Transportation, Vol. 27, 1994, pp. 327-354.
54. Evans, L., and P. Wasielewski. Risky driving related to driver and vehicle characteristics. Accident Analysis and Prevention, Vol. 15, 1983, pp. 121-136.
55. Ewing, R., and E. Dumbaugh. The Built environment and traffic safety: a review of empirical evidence. Journal of Planning Literature, Vol. 23 (4), 2009, pp. 347-367.
56. Ewing, R., and R. Pendall. Measuring Sprawl and its Impact. Environmental Protection Agency, S.G. America. Washington, D.C., 2002.
57. Finkelstein, A.. The effect of tax subsidies to employer-provided supplementary health insurance: evidence from Canada. Journal of Public Economics, Vol. 84, 2002, pp. 305-339.
58. Firpo, S.. Efficient semiparametric estimation of quantile treatment effects. Econometrica, Vol. 75, 2007, pp. 259-276.
59. Fisher, R. A.. The Design of Experiments, 1st ed. Oliver and Boyd, London, 1925.
60. Flores-Lagunes, A., A. Gonzalez, and T. C. Neumann. Estimating the Effects of Length of Exposure to a Training Program: The Case of Job Corps. Working Paper, Princeton University, Department of Economics, Industrial Relations Section, 2007.

61. Fox, J.. Applied Regression Analysis and Generalized Linear Models. Sage, Thousand Oaks, CA, second edition, 2008.
62. Fridstrsm, L..*Lommeboka, prisen eller reisetiden: Hva far oss til b velge transportmiddel pa lengre reiser?* Foredrag ved Transportskonomisk institutts 35-ks-jubileum, 17.11, 1993.
63. Galantea, F., F. Maurielloa, A. Montellaa, M. Pernettib, M. Ariac, and A. D'Ambrosio. Traffic calming along rural highways crossing small urban communities: Driving simulator experiment. Accident Analysis and Prevention, Vol. 42, 2010, pp. 1585-1594.
64. Gelman, A., and D. B. Rubin. Inference from Iterative Simulation using Multiple Sequences. Statistical Science, Vol. 7, 1992, pp. 457-511.
65. Golob, T. F., and W. W. Recker. Relationships among Urban Freeway Accidents, Traffic Flow, Weather, and Lighting Conditions. Journal of Transportation Engineering, Vol. 129, No. 4, 2003, 342-353.
66. Golob, T. F., W. W. Recker, and V. M. Alvarez. Freeway safety as a function of traffic flow. Accident Analysis and Prevention, Vol. 36, 2004, 933-946.
67. Grabowski, D. C., and M. A. Morrisey. Gasoline Prices and Motor Vehicle Fatalities. Journal of Policy Analysis and Management, Vol. 23, No. 3, 2004, pp. 575-593.
68. Grabowski, D. C., T, Michael, and A. Morrisey. Do higher gasoline taxes save lives? Economics Letters, Vol. 90, 2006, pp. 51-55.
69. Graham, D. J., and D. A. Stephens. Decomposing the impact of deprivation on child pedestrian casualties in England. Accident Analysis and Prevention, Vol. 40, 2008, pp. 1351-1364.
70. Graham, D. J., and S. Glaister. Spatial Variation in Road Pedestrian Casualties: The Role of Urban Scale, Density and Land-use Mix. Urban Studies, Vol. 40, No. 8, 2003, pp. 1591-1607.
71. Gruber, J. The Incidence of Mandated Maternity Benefits. American Economic Review, Vol. 84, 1994, pp. 622-641.
72. Guo, Z., and N. H.M. Wilson. Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground. Transportation Research Part A, Vol. 45, 2011, pp. 91-104.

73. Hadayeghi, A., A. S. Shalaby, and B. N. Persaud. Development of planning level transportation safety tools using Geographically Weighted Poisson Regression. Accident Analysis and Prevention, Vol. 42, 2010, 676-688.
74. Haque, Md. M., H. C. Chin, and H. Huang. Applying Bayesian hierarchical models to examine motorcycle crashes at signalized intersections. Accident Analysis and Prevention, Vol. 42, 2010, pp. 203-212.
75. Hauer, E., D.W. Harwood, F.M. Council, and M.S. Griffith. The empirical Bayes method for estimating safety: a tutorial. Transportation Research Record, Vol.1784, 2002, pp. 126-131.
76. Hauer, E.. Observational Before–After Studies in Road Safety. Pergamon Publication, England, 1997.
77. Haughton, J., and S. Sarkar. Gasoline tax as a corrective tax: estimates for the United States, 1970–1991. The Energy Journal, Vol. 17, No. 2, 1996, pp. 103-126.
78. Hausman, J. A. Specification Tests in Econometrics. Econometrica, Vol. 46, 1978, pp. 1251-1271.
79. Haynes, R., A. P. Jones, V. Kennedy, I. Harvey, and A. Jewell. District variations in road curvature in England and Wales and their association with road traffic crashes. Environment and Planning A, Vol. 39, 2007, pp. 1222-1237.
80. Heckman, J., H. Ichimura, and P. Todd. Matching As An Econometric Evaluation Estimator. Review of Economic Studies, Vol. 65, 1998, pp. 261-294.
81. Heckman, J., H. Ichimura, J. Smith, and P. Todd. Characterizing Selection Bias Using Experimental Data. Econometrica, Vol. 66, No. 5, 1998, pp. 1017-1098.
82. Helai, H., C. Chor, and M. Haque. Severity of driver injury and vehicle damage in traffic crashes at intersections: a Bayesian hierarchical analysis. Accident Analysis and Prevention, Vol. 40 (1), 2008, pp. 45-54.
83. Hirano, K. and G. W. Imbens. The Propensity Score with Continuous Treatments, in Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family (eds A. Gelman and X.-L. Meng), John Wiley & Sons, Ltd, Chichester, UK, 2005.
84. Holland, P.. Statistics and Causal Inference (with discussion). Journal of the American Statistical Association, Vol. 81, 1986, pp. 945-970.

85. Houston, D. J., and L. E. Richardson Jr. Traffic safety and the switch to a primary seat belt law: the California experience. Accident Analysis and Prevention, Vol. 34, 2002, 743-751.
86. Houston, D. J., and L. E. Richardson Jr. Traffic safety and the switch to a primary seat belt law: the California experience. Accident Analysis & Prevention, 2002, pp. 743-751.
87. Hsiao, C.. Analysis of Panel Data. Cambridge University Press, 2003.
88. Hu, S., and W. Saleh. Impacts of congestion charging on shopping trips in Edinburgh. Transport Policy, Vol. 12, 2005, pp. 443-450.
89. Huang, H. F., and R. Stewart. Evaluation of Lane Reduction "Road Diet" Measures on Crashes and Injuries. Transportation Research Record, Vol. 1784, 2002, pp. 80-90.
90. Huang, H., M. A. Abdel-Aty, and A. L. Darwiche. County-Level Crash Risk Analysis in Florida Bayesian Spatial Modelling. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2148, 2010, pp. 27-37.
91. Hyatt, E., R. Griffin, L. W. Rue III, and G. McGwin Jr.. The association between price of regular-grade gasoline and injury and mortality rates among occupants involved in motorcycle- and automobile-related motor vehicle collisions. Accident Analysis and Prevention, Vol. 41, 2009, pp. 1075-1079.
92. Ibrahim, M. F.. Improvements and integration of a public transport system: the case of Singapore. Cities, Vol. 20, No. 3, 2003, pp. 205-216.
93. Ivan, J. N., C. Wang, and N. R. Bernardo. Explaining two-lane highway crash rates using land use and hourly exposure. Accident Analysis and Prevention, Vol. 32, 2000, pp. 787-795.
94. Jones, A.P., R. Haynes, V. Kennedy, I.M. Harvey, T. Jewell, and D. Lea. Geographical variations in mortality and morbidity from road traffic casualties in England and Wales. Health & Place, Vol.14, 2008, pp. 519-535.
95. Kluve, J., H. Schneider, A. Uhlendorff, and Z. Zhao. Evaluating continuous training programs using the generalized propensity score Journal of the Royal Statistical Society, Series A (Statistics in Society), Vol. 175 (2), 2012, pp. 587-617.

96. Kononov, J., B. Bailey, and B. K. Allery. Relationships between Safety and Both Congestion and Number of Lanes on Urban Freeways. Transportation Research Record: Journal of the Transportation Research Board, No. 2083, 2008, pp. 26-39.
97. Ladron de Guevara, F., S. P. Washington, and J. Oh. Forecasting Crashes at the Planning Level: Simultaneous Negative Binomial Crash Model Applied in Tucson, Arizona. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1897, 2004, pp. 191-199.
98. Lan, B., B. Persaud, C. Lyon, and R. Bhim. Validation of a full Bayes methodology for observational before–after road safety studies and application to evaluation of rural signal conversions. Accident Analysis and Prevention, Vol. 41 (3), 2009, pp. 574-580.
99. Lawson, R. A Review of the Effectiveness of Speed Cameras. The Association of British Drivers, 2011.
100. Lechner, M.. Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption. Econometric Evaluations of Active Labor Market Policies in Europe, Heidelberg, Physica, 2001.
101. Lee, C., and M. Abdel-Aty. Presence of passengers: does it increase or reduce driver's crash potential? Accident Analysis and Prevention, Vol. 40 (5), 2008, pp. 1703-1712.
102. Leigh, J. P., and E. M. Geraghty. High Gasoline Prices and Mortality from Motor Vehicle Crashes and Air Pollution. Journal of Occupational & Environmental Medicine, Vol. 50, 2008, pp. 249-254.
103. Leigh, J.P., and J. T. Wilkinson. The effect of gasoline taxes on highway fatalities. Journal of Policy Analysis and Management, Vol. 10, No. 3, 1991, pp. 474-481.
104. Li, H., D. J. Graham, and A. Majumdar. The effects of congestion charging on road traffic casualties: A causal analysis using difference-in-difference estimation. Accident Analysis and Prevention, Vol. 49, 2012, pp. 366-377
105. Li, W., A. Carriquiry, M. Pawlovich, and T. Welch. The choice of statistical models in road safety countermeasure effectiveness studies in Iowa. Accident Analysis and Prevention, Vol. 40, 2008, pp. 1531-1542.

106.Lord, D., A Manar, and A. Vizioli. Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments. Accident Analysis and Prevention, Vol. 37, 2005, pp. 185-199.

107.Lord, D., and P. Y. Park. Investigating the effects of the fixed and varying dispersion parameters of Poisson-gamma models on empirical Bayes estimates. Accident Analysis and Prevention, Vol. 40, 2008, pp. 1441-1457.

108.Lord, D., I. van Schalkwyk, S. Chrysler, and L. Staplin. A strategy to reduce older driver injuries at intersections using more accommodating roundabout design practices. Accident Analysis and Prevention, Vol. 39, 2007, pp. 427-432.

109.Lord, D., S. P. Washington, and J. N. Ivan. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accident Analysis and Prevention, Vol. 37, 2005, pp. 35-46.

110.Lord, D.. Modeling motor vehicle crashes using Poisson-gamma models: Examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. Accident Analysis and Prevention, Vol. 38, 2006, pp. 751-766.

111.Manski, C.. Statistical Treatment Rules for Heterogenous Populations. Econometrica, Vol. 72 (4), 2004, pp. 1221-1246.

112.Marchesini, P., and W. Weijermars. The relationship between road safety and congestion on motorways. Working paper, SWOV Institute for Road Safety Research, 2010.

113.Marshall, W. E., and N. W. Garrick. Does street network design affect traffic safety? Accident Analysis and Prevention, Vol. 43, 2011, 769-781.

114.Martin, J.. Relationship between crash rate and hourly traffic flow on interurban motorways. Accident Analysis and Prevention, Vol. 34, 2002, 619-629.

115.McDonald, T. L., W. P. Erickson, and L. L. McDonald. Analysis of Count Data from Before-after Control-Impact Studies. Journal of Agricultural, Biological, and Environmental Statistics, Vol. 5, No. 3, 2000, pp. 262-279.

116.Miaou, S. P., and D. Lord. Modelling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and Bayes versus empirical Bayes. Transportation Research Record, Vol. 1840, 2003, pp. 31-40.

117.Miguel, E., and M. Kremer. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. Econometrica, Vol. 72(1), 2004,159-217.

118.Mitra, S., and S. Washington. On the nature of over-dispersion in motor vehicle crash prediction models. Accident Analysis and Prevention, Vol. 39, 2007, pp. 459-468.

119.Neyman, J.. On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9, translated in Statistical Science, (with discussion), Vol. 5, No 4, 1990, pp. 465-480.

120.Noland, R. B., and L. Oh. The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. Accident Analysis and Prevention, Vol. 36, 2004, pp. 525-532.

121.Noland, R. B., and M. A. Quddus. A spatially disaggregate analysis of road casualties in England. Accident Analysis and Prevention, Vol. 36, 2004, pp. 973-984.

122.Noland, R. B., and M. G. Karlaftis. Sensitivity of crash models to alternative specifications. Transportation Research Part E, Vol. 41, 2005, pp. 439-458.

123.Noland, R. B., M. A. Quddus, and W. Y. Ochieng. The effect of the London congestion charge on road casualties: an intervention analysis. Transportation, Vol. 35, 2008, pp. 73-91.

124.Noland, R. B.. Traffic fatalities and injuries: the effect of changes in infrastructure and other trends. Accident Analysis and Prevention, Vol. 35, 2003, pp. 599-611.

125.Noland, R.. Traffic fatalities and injuries: are reductions the result of 'improvements' in highway design standards? In: Transportation Research Board 80th Annual Meeting, Washington, D.C., 2000.

126.Olszewski, P., and L. Xie. Modelling the effects of road pricing on traffic in Singapore. Transportation Research Part A, Vol. 39, 2005, pp. 755-772.

127.Ossiander, E. M., and P. Cummings. Freeway speed limits and traffic fatalities in Washington State. Accident Analysis and Prevention, Vol. 34, 2002, pp. 13-18.

128.Paleti, R., N. Eluru, and C. Bhat. Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. Accident Analysis and Prevention, Vol. 42 (6), 2010, pp. 1839-1854.

129.Park, E.S., and D. Lord. Multivariate Poisson–lognormal models for jointly modelling crash frequency by severity. Transportation Research Record, Vol. 2019, 2007, pp. 1-6.

130.Park, E.S., J. Park, and T.J. Lomax. A fully Bayesian multivariate approach to before-after safety evaluation. Accident Analysis and Prevention, Volume 42, 2010, pp. 1118-1127.

131.Pau, M., and S. Angius. Do speed bumps really decrease traffic speed? An Italian experience. Accident Analysis and Prevention, Vol. 33, 2001, 585-597.

132.Pearl, J. Causality: Models, Reasoning, and Inference, Cambridge University Press, 2000.

133.Persaud, B., and C. Lyon. Empirical Bayes before-after safety studies: lessons learned from two decades of experience and future directions. Accident Analysis and Prevention, Volume 39, 2007, pp. 546–555.

134.Persaud, B., B. Lan, C. Lyon, and R. Bhim. Comparison of empirical Bayes and full Bayes approaches for before–after road safety evaluations. Accident Analysis and Prevention, Vol. 42, 2010, pp. 38-43.

135.Persaud, B., R. Retting, and C. Lyon. Crash reductions following installation of centre-line rumble strips. Accident Analysis and Prevention, Vol. 36, 2004, pp. 1073-1079.

136.Persaud, B.N., E. Hauer, R. Retting, R. Vallurapalli, and K. Mucsi. Crash reductions following traffic signal removal in Philadelphia. Accident Analysis and Prevention, Vol. 29, 1997, pp. 803-810.

137.Personal Travel Factsheet. Department for Transport, 2011.

138.Phillips, K. B. and J. A. McCutchen. Economic Regulation vs Safety Regulation of the Trucking Industry - Which More Effectively Promotes Safety? Transportation Quarterly, Vol. 45, 1991, pp. 323-340.

139.Pirdavani, A., T. Brijs, T. Bellemans, B. Kochan, and G. Wets. Evaluating the road safety effects of a fuel cost increase measure by means of zonal crash prediction modelling. Accident Analysis and Prevention, Vol. 50, 2013, pp. 186-95.

140.Quddus, M. A., M. G. H. Bell, J. Schmocker, and A. Fonzone. The impact of the congestion charge on the retail business in London: An econometric analysis. Transport Policy, Vol. 14, 2007, pp. 433-444.

141.Quddus, M. A.. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. Accident Analysis and Prevention, Vol. 40, 2008, pp. 1486-1497.

142.Quddus, M. A.. Time series count data models: An empirical application to traffic accidents. Accident Analysis and Prevention, Vol. 40, 2008, pp. 1732-1741.

143.Quddus, M., C. Wang, and S. Ison. Road traffic congestion and crash severity: an econometric analysis using ordered response models. ASCE Journal of Transportation Engineering, Vol. 136(5), 2010, pp. 424-435.

144.Rifaat, S. M., R.Tay, and A.Barros. Effect of street pattern on the severity of crashes involving vulnerable road users. Accident Analysis and Prevention, Vol. 43, 2011, pp. 276-283.

145.Rubin, D. B.. Bayesian inference for causal effects: The Role of Randomization. Annals of Statistics, Vol. 6, 1978, pp. 34-58.

146.Rubin, D.. Assignment to Treatment Group on the Basis of a Covariate. Journal of Educational Statistics, Vol. 2(1), 1977, pp. 1-26.

147.Rubin, D.. Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies. Journal of Educational Psychology, Vol. 66, 1974, pp. 688-701.

148.Rubin, D.. Matching to Remove Bias in Observational Studies. Biometrics, Vol. 29, 1973, pp. 159-183.

149.Rubin, D.. The Use of Matched Sampling and Regression Adjustments to Remove Bias in Observational Studies. Biometrics, Vol. 29, 1973, pp. 185-203.

150.Santos, G., and J. Bhakar. The impact of the London congestion charging scheme on the generalised cost of car commuters to the city of London from a value of travel time savings perspective. Transport Policy, Vol. 13, 2006, pp. 22-33.

151.Savolainen, P., F. Mannering, D. Lord, and M. Quddus. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis and Prevention, Vol. 43, 2011, pp. 1666-1676.

152.Sawalha, Z., and T. Sayed. Evaluating safety of urban arterial roadways. Journal of Transportation Engineering, Vol. 127 (2), 2001, pp. 373-378.

153.Schwarz, G. E.. Estimating the dimension of a model. Annals of Statistics, Vol. 6, 1978, pp. 461-464.

154. Shefer, D., and P. Rietveld. Congestion and Safety on Highways: Towards an Analytical Model. Urban Studies, Vol. 34, No. 4, 1997, 679-692.
155. Siddiqui, C., M. Abdel-Aty, and K. Choi. Macroscopic spatial analysis of pedestrian and bicycle crashes. Accident Analysis and Prevention, Vol. 45, 2012, pp. 382-391.
156. Smiley, A. et al.. Accidents, Convictions, and Demerit Points: An Ontario Driver Records Study. Transportation Research Record, Vol. 1238, 1989, pp. 53-64.
157. Southworth, M., and B. Parthasarathy. The suburban public realm I: its emergence, growth and transformation in the American metropolis. Journal of Urban Design, Vol. 1 (3), 1996, pp. 245-263.
158. Spiegelhalter, D., A. Thomas, N. Best, and D. Lunn. WinBUGS User Manual. MRC Biostatistics Unit, Cambridge, 2005. Available from http://www.mrc-cam.ac.uk/bugs.
159. Swift, P., and D. Painter. Residential street typology and injury accident frequency. Congress for the New Urbanism VI, Denver, CO, 1998.
160. Talen, E., 2006. Design for diversity: evaluating the context of socially mixed neighbourhoods. Journal of Urban Design 11 (1), 1–32.
161. Taylor, M. C., D. A. Lyman, and A. Baruya. The effects of drivers' speed on the frequency of road accidents. TRL Report 421, Crowthorne, Berkshire: Transport Research Laboratory, 2000.
162. Tuerk, A, and D. J. Graham. Traffic impacts of the central London congestion charging scheme. Working paper, Imperial College London, 2010.
163. Vuong, Q.H.. Likelihood Ratio Tests for Model Selection and Non-Nested Hypothesis. Econometrica, Volume 57, 1989, pp. 307-333.
164. Wakefield, J.C., N.G. Best, and L. Waller. Bayesian Approaches to Disease Mapping, on Spatial Epidemiology: Methods and Applications. Oxford University Press, 2000.
165. Walle, S. V., and T. Steenberghen. Space and time related determinants of public transport use in trip chains. Transportation Research Part A, Vol. 40, 2006, pp. 151-162.
166. Wang, C., M. A. Quddus, and S. G. Ison. Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England. Accident Analysis and Prevention, Vol. 41, 2009, pp. 798-808.

167.Wang, X., and M. Abdel-Aty. Temporal and Spatial Analyses of Rear-end Crashes at Signalized Intersections. Accident Analysis and Prevention, Vol. 38, 2006, pp. 1137-1150.

168.Ward, H., R, Lyons, and R. Thoreau. Under-reporting of Road Casualties-Phase 1. Road Safety Research Report No.69, Department for Transport ; London, 2006.

169.Wardrop, J. G.. Some theoretical aspects of road traffic research. Proceedings, Institution of Civil Engineers, PART II, Vol.1, 1952, pp. 325-378.

170.Wedagama, D. M., R. N. Bird, and A. V. Metcalfe. The influence of urban land-use on non-motorised transport casualties. Accident Analysis and Prevention, Vol. 38, 2006, pp. 1049-1057.

171.West, R., J. Elander, and D. French. Decision making, personality and driving style as correlates of individual accident risk. Contractor Report 309, Crowthorne, Berkshire, Transport Research Laboratory, 1992.

172.White, M. J.. The "Arms Race" on American Roads: The Effect of Heavy Vehicles on Traffic Safety and the Failure of Liability Rules. Journal of Law and Economics, Vol. 47, No. 2, 2004, pp. 333-355.

173.White, M.J.. The “arms race” on American roads: the effect of heavy vehicles on traffic safety and failure of liability rules. Journal of Law and Economics, Vol. 47, No. 2, 2004, pp. 333-356.

174.Wichiensin, M., M. G. H. Bell, and H. Yang. Impact of congestion charging on the transit market: An inter-modal equilibrium model. Transportation Research Part A, Vol. 41, 2007, pp. 703-713.

175.Wier, M., J. Weintraub, E. H. Humphreys, E. Seto, and R. Bhatia. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis and Prevention, Vol. 41, 2009, pp. 137-145.

176.Winston, C., V. Maheshri, and F. Mannering. An exploration of the offset hypoth-esis using disaggregate data: the case of airbags and antilock brakes. Journal of Risk and Uncertainty, Vol. 32 (2), 2006, pp. 83-99.

177.Xie, Y., Y. Zhang, and F. Liang. Crash injury severity analysis using Bayesian ordered probit models. Journal of Transportation Engineering, Vol. 135 (1), 2009, pp. 18-25.

178. Yamamoto, T., J. Hashiji, and V. Shankar. Underreporting in traffic accident data, bias in parameters and the structure of injury severity models. Accident Analysis and Prevention, Vol. 40 (4), 2008, pp. 1320-1329.
179. Yanmaz-Tuzel, O., and K. Ozbay. A comparative Full Bayesian before-and-after analysis and application to urban road safety countermeasures in New Jersey, Accident Analysis and Prevention, Vol. 42, 2010, pp. 2099-2107.
180. Ye, F., and D. Lord. Investigating the effects of underreporting of crash data on three commonly used traffic crash severity models: multinomial logit, ordered probit and mixed logit models. Transportation Research Record, forthcoming.
181. Zajac, S. S., and J. N. Ivan. Factors influencing injury severity of motor vehicle–crossing pedestrian crashes in rural Connecticut. Accident Analysis and Prevention, Vol. 35, 2003, pp. 369-379.