INTERACTIONS BETWEEN ENVIRONMENTAL AND SAFETY PERFORMANCE IN VEHICLE DESIGN

By
Reza Tolouei

A thesis submitted to the University College London for the degree of Doctor of Philosophy

Centre for Transport Studies
University College London
June 2011

DECLARATION

I, *Reza Tolouei*, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature

Reza Tolouei

ACKNOWLEDGEMENT

This work was carried out in the Centre for Transport Studies at University College London. I am most grateful to Dr. Helena Titheridge and Professor Benjamin Heydecker who co-supervised me throughout this research. I express my gratitude to them for their constant guidance, advice, and support. I am also grateful to my examiners Professor Mike Maher and Dr. Arnab Majumdar who made a significant contribution to the quality of this work. Furthermore, I would like to thank Professor Richard Allsop, Professor Chris Lines, Professor Robert Noland, and Dr. Jeremy Broughton for their advice and comments on this work. I also wish to thank Department for Transport, especially David Robinson, as well as Daniel White in Experian for their cooperation and support in providing relevant data. Thanks are to Rees Jeffreys road fund for providing partial funding for this study. Finally, my special thanks go to the staff and students of the Centre for Transport Studies who were a constant source of support and joy throughout the period of this research.

ABSTRACT

One potential interaction between environmental and safety goals in transport is found within the vehicle fleet where fuel economy and safety impose conflicting requirements on vehicle design. Larger and heavier vehicles have a better secondary safety performance during a crash. On the other hand, they are associated with higher levels of fuel consumption and emissions. This issue has generated debate amongst researchers and policy makers when formulating policies to improve the environmental performance of the road transport system. An extensive review of literature reveals that arguments has often been based on either little research evidence, or research that has inadequacies in the applied methodologies.

This research investigates the safety consequences of changes in vehicles mass within the vehicle fleet aimed at increasing fleet fuel economy. The partial effects of mass on fuel consumption rate and secondary safety performance were estimated using a cross-sectional analysis of mass within the British passenger car fleet. Estimation results confirmed that fuel consumption increases as mass increases and were different for different fuel and transmission types. It was shown that vehicle mass has both protective and aggressive safety effects where vehicle size only tends to have protective effects; these were estimated using a novel methodology based on a detailed analysis of two-car crashes. The estimated relationships were used to investigate partial safety and environmental effects of changes in mass distribution within the fleet using an introduced incremental approach.

Results generally showed that the relationship between fuel economy and safety performance in vehicle design depends on the characteristics of the vehicle fleet, and in particular, mass distribution. It was shown that an informed change in the mass distribution not only imposes no trade-off between the fuel economy and safety goals, but also could lead to a desirable outcome in both aspects.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Environmental and safety goals in transport can interact in several different ways. One of these potential interactions can be found within the vehicle fleet where fuel economy[1] and safety impose requirements on vehicle design that conflict in some ways. Larger and heavier vehicles have a better safety performance in that they give a better protection to their occupants during a crash. On the other hand, these vehicles are associated with higher levels of fuel consumption and emissions. This issue has always generated debate amongst researchers and policy makers when policies to improve the environmental performance of the road transport system are to be formulated. One of the main concerns that is often raised is the consequent effect on traffic fatalities and injuries of a reduction in mass and size of vehicles in an attempt to improve the fuel economy of the vehicle fleet. One well-known example of these long-lasting debates has been in the US over the potential link between the increase in Corporate Average Fuel Economy (CAFE) standards since the 1970s, which changed the composition of US vehicle fleet with an increase in the proportion of smaller cars, and the increase in traffic fatalities and injuries in later years. CAFE standards are fleet average fuel economy standards for new cars and light trucks set by the US Department of Transport to be met by vehicle manufacturers (NRC, 2002). Similar arguments have been made in Europe more recently amongst researchers and policy makers over changes in mass and size of new cars as a result of environmental policies and emission targets (e.g. Buzeman et al., 1998; Broughton, 1999; Zachariadis, 2008).

1.1. Environmental objectives of road transport

The environmental impacts of road transport regarded as the ones of most concern are climate change, air quality, and fossil fuel consumption, which are all directly related to vehicle fuel consumption. Hence, decreasing vehicle fuel consumption is a desirable environmental policy leading to a reduction in the main adverse environmental impacts of road transport.

Carbon dioxide (CO_2), which is the most important greenhouse gas accounting for about two thirds of man-made global warming, results from the combustion of hydrocarbon

---

[1] The distance travelled per unit of fuel consumed

fuels in air. In the UK, for example, vehicle fuel consumption is responsible for about 22% of total CO_2 emissions, most of which comes from passenger cars (DfT, 2007). Passenger cars contribute to about 12% of overall EU emissions of CO_2 with their emissions increased by 26% over 1990-2004 period despite a reduction of 5% in overall EU greenhouse gas (GHG) emissions over the same period (EC, 2007).The European Union (EU) has set a target to cut overall greenhouse gas emissions by 20% by 2020 compared with the 1990 level (EC, 2007). EU notes that this can only be achieved if there is a significant carbon reduction in road transport sector. The UK Government has set a target to cut GHG emissions to at least 80% below the 1990 levels by 2050, with the transport sector having a significant role to play in achieving this target (DfT, 2010).

Air pollutants such as hydrocarbons (HC), carbon monoxide (CO), nitrogen oxides (NO_x), and particles (PM) are also emitted from vehicles; this is a result of the fact that combustion engines are not perfectly efficient. The European Union has regulated European emissions from vehicles to limit emission of these air pollutants through European Emission Standards (VCA, 2010). These are requirements that set specific limits for exhaust emissions of NO_x, HC, CO, and PM for the new vehicles sold in EU member states.

Introduction of more stringent emission standards and the European Union objective of further reducing carbon emissions through vehicle technology oblige manufacturers for the European car market to make changes to the design of their cars (Fontaras and Samaras, 2010). Amongst various design features, vehicle mass is a key variable having the potential to considerably affect the fuel consumption rate. Therefore, many polices formulated to help achieve the new targets and standards are likely to result in changes in mass and size of the vehicles in fleet.

1.2. Safety objectives of road transport

Although there has been a decreasing trend in the number of traffic injuries and fatalities in the last few years in most European countries (Noland, 2005), still many people are killed or seriously injured on the roads every year as a result of vehicle crashes. In Great Britain, for example, about 28,570 of road users were killed or seriously injured in 2008 of which about 12,000 (42%) were passenger car users (DfT, 2009).

Road safety policies aim to reduce the total number of road casualties in one of the two following ways:

- Reduction in the total number of traffic crashes of a given type and severity
- Reduction in the risk of occupant injury or fatality in the event of a crash of a given type and severity

There are a number of factors that contribute to traffic casualties including driver, vehicle, road, and other external factors. Vehicle design is one of the most important factors than can potentially affect both risk of crash involvement of the vehicle and risk of crash injury to the occupants when the vehicle is involved in a crash (Van Auken and Zellner, 2005).

Recent analyses suggest that vehicle safety-related factors play a more important role than other vehicle factors in the vehicle purchase process (Koppel et al., 2008). Various safety reports are published that assess safety performance of vehicles in different ways. For example, the European New Car Assessment Programme (Euro NCAP) compares safety performance of different vehicles in different groups of cars of the same type and size based on their performance in a variety of crash tests (Euro NCAP, 2010). Other studies compare safety performance of different vehicles based on real crash data. For example, the UK Department for Transport (DfT) estimates safety performance of popular makes and models in Great Britain, as driver injury risk, and reports the results for specific time periods (Broughton, 1996a).

1.3. Potential interaction in vehicle design

For a vehicle driven under controlled driving conditions, the fuel consumption rate varies substantially according to the vehicle design (Tenny and Lam, 1985; Biggs and Akcelik, 1987; Ross, 1997). Amongst various design features, vehicle mass is a key variable having the potential to considerably affect fuel consumption rate as it strongly affects forces resisting vehicle motion during a driving cycle (Gillespie, 1992; Redsell et al., 1993; Ross, 1994; Van den Brink and Van Wee, 2001; Burgess and Choi, 2003). It is generally known that an increase in vehicle mass increases fuel consumption.

Other vehicle design features such as fuel type, transmission type, engine size and other engine characteristics can also affect fuel consumption through influencing engine efficiency during a driving cycle (Ross, 1997; Van den Brink and Van Wee, 2001). Greater engine size is also linked with greater fuel consumption rate. However, because of the correlation between mass, engine size, and some other design features, the partial effects of these factors, where the effect is a result of a change in that factor holding all other factors constant, on vehicle fuel consumption is not clear.

Different types of vehicles also have different safety performances in the fleet depending on their design. In a two-vehicle crash, the injury risk of occupants in the lighter vehicle is higher than that of the heavier vehicle due to the greater velocity change during a collision (Evans, 2004). For example, in the case of a frontal collision between two vehicles with masses m_1 and m_2 travelling with speeds v_1 and v_2, it can be shown using Newtonian mechanics that the velocity change of the first vehicle during collision (Δv_1) depends on the proportion of the total mass contained by the other vehicle and the closing speed (Grime and Jones, 1970):

$$\Delta v_1 = \left(\frac{m_2}{m_1+m_2}\right)(v_1 + v_2). \tag{1.1}$$

As this relationship shows, injury risk to the occupants of a vehicle in a crash is influenced by the vehicle’s mass as well as the mass of the vehicle with which it collides. Although this relationship has been investigated in several studies (e.g. Grime and Hutchinson, 1979; Evans and Frick, 1993; Wood, 1997; Buzeman et al., 1998; Evans, 2004), there are major shortcomings in the methods used resulting in inconsistencies in the estimation results. However, it has been generally confirmed that heavier and larger vehicles give a better protection to their occupants in crashes while impose a greater risk of injury to the occupants of the other vehicles involved compared to lighter and smaller cars.

The results documented in the literature suggest that vehicle design (particularly mass) has the potential to cause conflict between environmental and safety goals within the vehicle fleet. The main concern is whether reducing vehicles’ mass and size across the fleet to improve fuel efficiency of fleet can have a detrimental effect on safety through increasing injury risk to the occupants of the vehicles.

1.4. Current research issues

An extensive review of relevant literature reveals that arguments over the issue of interaction between environmental and safety policies has often been made based on either little research evidence on one or both sides, or evidence based on research that has inadequacies in the applied methodologies.

A number of studies (e.g. Kahane, 1997; Buzeman et al., 1998; Broughton, 1999; Ross and Wenzel, 2001; Noland, 2004, Noland, 2005; Ahmad and Greene, 2005; Zachariadis, 2008) have addressed this issue using different methodologies and, as a consequence, there are inconsistencies in their results. Some of these studies are mainly empirical studies that have used aggregate time-series data (e.g. Noland, 2004, Noland, 2005; Ahmad and Greene, 2005). Other studies have made conclusions based on estimated effects of vehicle mass documented in the literature (e.g. Kahane, 1997; Buzeman et al., 1998; Broughton, 1999; Ross and Wenzel, 2001). However, various methodological issues are associated with these estimates making the conclusions inconsistent and in several cases, in opposite directions. An extensive and critical review of existing literature is presented and knowledge gaps in this topic are identified in the next chapter of this thesis.

Figure 1.1 shows the relationship between fuel consumption and safety in the vehicle fleet that is influenced by individual vehicles' mass. There are two distinct aspects of safety performance of a vehicle in fleet: primary safety performance, which is linked to the risk of crash involvement of the vehicle, and secondary safety performance, which is linked to the risk of occupant injury (to a specific level) when the vehicle is involved in a particular type of crash. In a two-car crash, risk of occupant injury in a vehicle is influenced not only by the level of protection that the vehicle offers to its occupants, but also by the level of aggressivity that is imposed to the occupants by the colliding vehicle. There are therefore two aspects of secondary safety performance for a vehicle that is involved in a two-car crash: A protective performance which is linked to the injury risk to the occupants of that vehicle, and an aggressive performance which is linked to the injury risk that the vehicle imposes to the occupants of the other vehicle with which it collides.

There is no evidence that suggests a direct effect of vehicle mass on primary safety performance of the vehicle; however, vehicle mass directly influences the secondary safety performance by affecting the velocity change of the vehicle in a crash, and hence the severity of the crash (see Equation 1.1). Therefore, for an individual vehicle in fleet, vehicle mass is a significant contributor to both fuel consumption and secondary safety performance of the vehicle with a potentially interacting effect. The effect of vehicle mass on secondary safety performance is more complicated than that on fuel consumption because in a crash between two vehicles, the secondary safety performance of a vehicle is influenced not only by mass of that vehicle, but also by mass of the other vehicle with which it collides (see Equation 1.1). Therefore, the secondary safety performance of a vehicle with a given mass in a fleet could vary depending on the mass distribution of vehicles in the fleet.

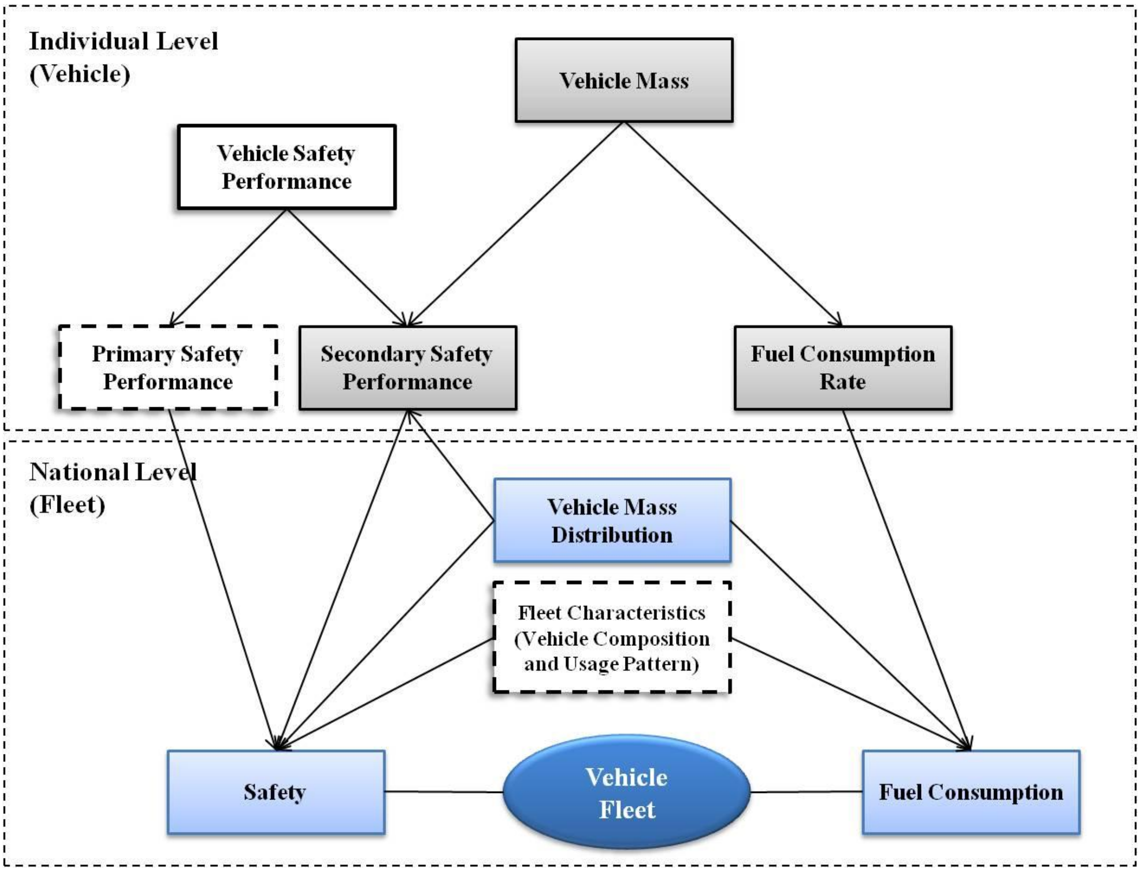


Figure 1.1: Relationship between fuel consumption and safety in the vehicle fleet that is influenced by individual vehicles' mass

Overall fuel consumption and safety of a vehicle fleet with a given vehicle composition and usage pattern are directly influenced by the fuel consumption rates and safety performances of individual vehicles in that fleet, respectively. Therefore, the key point in understanding whether vehicle mass imposes any conflict in safety and environmental goals as a whole in a vehicle fleet with given characteristics is to examine the issue of potential interaction between fuel consumption and safety performance in vehicle design within that vehicle fleet. Having investigated the detailed relationships between vehicle mass, vehicle fuel consumption rate, and vehicle secondary safety performance within the fleet, the environmental and safety consequences of changes in vehicles' mass in fleet can be investigated taking into account the characteristics of vehicle fleet and the usage pattern of vehicles. Depending on the characteristics of the vehicle fleet, a potential conflict in vehicle design could influence the relationship between the environmental and safety outcomes of the fleet. Little is known about the detailed relationships between vehicle design and each of safety or environmental goals within a vehicle fleet. Depending on the distribution of vehicles by specific design features and the usage pattern of different types of vehicles, conflict between safety and environmental goals in a given vehicle fleet is likely. A proper and reliable conclusion on the existence and the extent of any conflict between the goals can only be made when detailed relationships between vehicle mass and each side is known.

1.5. Research objectives

This research aims to investigate the likely safety consequences of changes in vehicles design (particularly mass) in the national vehicle fleet aimed at increasing fleet fuel economy. In particular, the following specific objectives are addressed:

1. To estimate the effects of vehicle design (particularly mass) on fuel consumption.
2. To estimate the protective and aggressive effects of vehicle mass in two-vehicle crashes separately.
3. To separate the effects of vehicle mass and size on secondary safety performance.
4. To examine whether there are specific design effects of different vehicle makes and models on their secondary safety performance beyond the effect of their mass.

5. To investigate partial safety and environmental consequences of differences in vehicle mass distribution in fleet.

Passenger cars in Great Britain with petrol and diesel fuel types were chosen as the vehicle fleet to study. Effect of vehicle usage pattern and its relationship with vehicle design is beyond the scope of this study. The primary safety performance of vehicles, which is unlikely to be influenced by vehicle mass, is also beyond the scope of this study.

1.6. Structure of the thesis

This thesis is split into seven chapters in a way to provide a systematic approach to investigating the potential interaction between fuel consumption and secondary safety performance in vehicle design within the vehicle fleet and its influence on overall safety and environmental objectives in fleet. Divided into sections and subsections, each chapter starts with a short introduction outlining what it covers and ends with a short summary of the findings. The rest of this thesis is organised as follows:

- *Chapter 2. Literature review*

This chapter presents a detailed review of the key studies that estimate the effect of vehicle mass on each of fuel consumption and safety, as well as studies that discuss the relationship between safety and environmental goals in road transport. A critical review of key literature is presented and knowledge gaps are identified.

- *Chapter 3. General methodology and study data*

In this chapter the general methodology used to investigate the interaction between fuel consumption and secondary safety performance in vehicle design is outlined, the data sources used in the study and their quality are discussed, and the process of development of different study datasets is explained.

- *Chapter 4. Vehicle mass and fuel consumption*

This chapter investigates the relationship between vehicle design and fuel consumption. The statistical modelling approach used to estimate the partial effects of mass on fuel consumption is explained followed by estimation results and interpretation of the developed statistical models.

- *Chapter 5. Vehicle mass and secondary safety*

Chapter 5 investigates the detailed relationship between vehicle mass and secondary safety performance where the protective and aggressive effects of mass are estimated. A novel methodology is introduced to analyse injury risk distribution in two-car crashes based on injury crash data and to estimate isolated effects of mass and size on driver injury risk.

- *Chapter 6. Safety and environmental consequences of changes in fleet mass distribution*

In Chapter 6, the likely safety and environmental consequences of changes in mass distribution within the vehicle fleet is investigated using the estimated effects of mass and other design features from Chapters 4 and 5. A number of hypothetical mass distribution scenarios for the British passenger car fleet are formulated and their partial effects on fleet fuel consumption and safety are estimated.

- *Chapter 7. Conclusions and discussion*

This is the final chapter of this thesis where the main study findings and contributions are outlined. This is followed by a discussion on policy implications, limitations of the study, and some suggestions on further research in this area.

CHAPTER 2. LITERATURE REVIEW

The objective of this chapter is to present a thorough review of the key literature relevant to the issue of interaction between environmental and safety performance in vehicle design. The chapter is organised as follows. The first section (2.1) provides a background on the issue while the next two sections (2.2 and 2.3) review the literature on the relationship between vehicle design and each of fuel consumption and secondary safety, respectively. Whilst the effects of different design factors are discussed, the emphasis is put on the effects of vehicle mass which is the key factor influencing both fuel consumption and safety performance with the potential to have interacting effects. The next section (2.4) discusses findings on the safety and environmental consequences of changes in vehicle design (in particular mass) within the fleet. The last section (2.5) summarises the knowledge gaps and limitations of the existing literature.

2.1. Background

Transport-related energy consumption has long been a challenge for policy makers with vehicle design, in particular mass, having played a key role in generating debates due to its disputed effects on overall safety outcome of a vehicle fleet. A particular case of such debates has been in the US over the safety consequences of the CAFE standards for new cars and light trucks since the 1970s. Figure 2.1 compares trends in traffic fatalities and average fuel economy of cars and light trucks (in terms of miles per gallon) in the US. Whilst the trends show a consistent increase in the fuel economy of vehicles since the 1970s because of the CAFE standards, an increase in fatalities in the mid and late 1970s as well as in 1980s are also apparent from the figure. Several studies have linked the changes in the US vehicle fleet towards more efficient cars, which normally tend to be lighter and smaller, as a result of the CAFE standards to the increase in the total number of fatalities in the mid and late 1970s (e.g. Crandall and Graham, 1989; NRC, 2002; Kahane, 1997; Kahane, 2003, Evans, 2004). This is on the grounds that lighter and smaller cars normally provide a lower level of protection to their occupants when they are involved in traffic crashes compared to larger and heavier cars.

Figure 2.2 compares trends in the average mass of new cars entering the fleet in the US and Europe. The trends show a greater average mass of new cars in the US compared to that in Europe as well as an increase in the average mass of new cars in both Europe and US in recent years. Although the average vehicle mass has been generally lower in Europe compared to that in the US, arguments over the safety consequences of a relatively greater proportion of light and small cars in European vehicle fleets are found less common in the literature compared to that of US. This could be partly due to the decreasing trends in the overall number of fatalities across the Europe (as shown in Figure 2.3), and partly due to the technological improvements in vehicle design leading to a reduction in fuel consumption rate per unit of vehicle mass (Schipper, 2008). However, in 1998, a voluntary agreement was signed between European Commission and all major vehicle manufacturers that aimed to limit the average CO_2 emission from the new passenger cars to 140 g/km by 2008-2009. Although this target was not achieved, this voluntary commitment was regarded as an important incentive and resulted in a new voluntary average CO_2 limit of 130 g/km for the new passenger car fleet by 2015 being agreed; this was introduced in 2009 (Fontaras and Samaras, 2010). With recent increases in the average mass of passenger cars in Europe specially since 2000 (Schipper, 2008) and the need to further reduce fuel consumption and carbon emissions through changes in vehicle design, concerns over the changes in fleet composition and its safety consequences have been raised and analyses have been performed to examine the safety consequences of fleet downsizing scenarios (e.g. Buzeman et al., 1998; Broughton, 1999; Noland, 2005; Zachariadis, 2008).

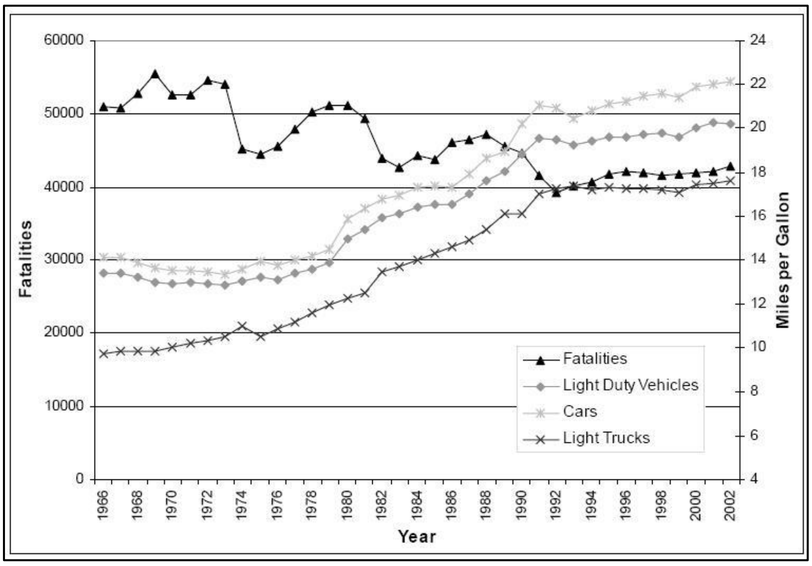


Figure 2.1: Trends in traffic fatalities and fuel economy in the US (Ahmed and Green, 2005)

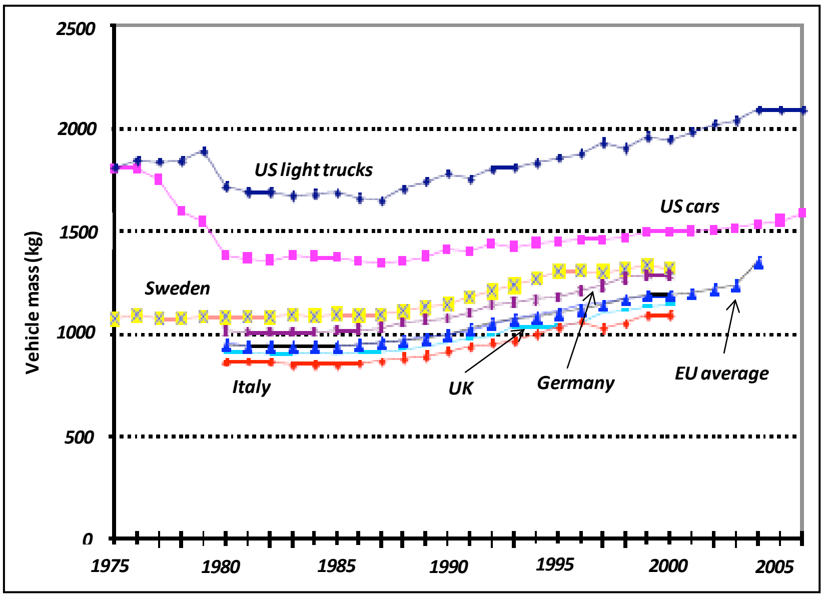


Figure 2.2: Trends in average mass for new cars in the US and Europe (Schipper, 2008)

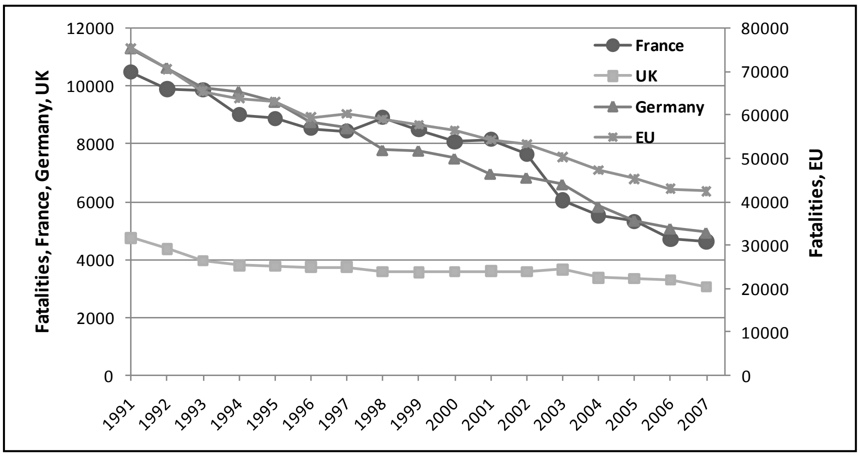


Figure 2.3: Trends in traffic fatalities in Europe (based on CARE, 2009)

2.2. Vehicle design and fuel consumption

2.2.1. Effect of vehicle mass

Amongst various design features, vehicle mass is a key variable having the potential to considerably affect the fuel consumption rate (Evans, 2004, Noland, 2005 and 2006). Depending on the engine efficiency of a vehicle and the energy required by vehicle accessories[1], a certain amount of fuel energy is consumed to overcome forces resisting

---

[1] Such as air conditioning, lights, audio systems, and heaters

vehicle motion during a driving cycle (Ross, 1997). In this thesis, this energy is referred to as the vehicle energy demand. Vehicle mass strongly affects vehicle energy demand, and hence, vehicle fuel consumption rate.

Vehicle energy demand during motion depends mainly on rolling, inertia, aerodynamic drag and gravitational losses; and vehicle mass in particular contributes directly to rolling, inertia, and gravitational losses (Gillespie, 1992; Redsell et al., 1993; Ross, 1994; Gyenes and Mitchell, 1994; Van den Brink and Van Wee, 2001; Burgess and Choi, 2003). A decrease in vehicle mass decreases the overall energy demand (through decreasing rolling resistance, inertia, and gravitational losses), and consequently, it decreases the fuel consumption rate.

Ross (1997) points out that the overall energy use depends on two factors, vehicle load (vehicle energy demand) and powertrain efficiency, where vehicle load is directly related to vehicle mass. Burgess and Choi (2003) performed a parametric study of the energy demands of passenger car on two competing inter-city routes in the UK that have different characteristics. They performed the analysis on different car categories defined according to the vehicle design factors (i.e. mass, size, body type) and represented by a typical make and model with a known mass. Based on simulations using a computer model, they found that total energy demand on both selected routes varied almost proportionally with changes in vehicle mass. For example, they calculated that decreasing vehicle mass by 10% could decrease total energy demand due to external resisting forces by 8.3% for an Audi A2 with a mass of about 900 kg. However, the effect of mass on fuel consumption remains unclear in their study.

While measuring the effect of mass on vehicle energy demand when the vehicle is driven under a known driving cycle is theoretically possible using available physical formulae related to these forces (for example, see Ross, 1997; Burgess and Choi, 2003), the effect of mass on fuel consumption of vehicle is not fully understood. This is mainly because the relationship between vehicle energy demand and its fuel consumption is not clear in the literature. Ross (1994) discussed the energy flow of US sales-weighted average 1993 model cars in the EPA[1] composite driving cycle. He showed that engine output is just 20% of fuel energy, while about 15% of fuel energy is used to overcome external forces resisting motion. Using an instantaneous model and collected on-road

---

[1] Environmental Protection Agency

second-by-second data, Biggs and Akcelik (1987) estimated that components related to vehicle energy demand that are needed to overcome forces resisting motion contribute to 30.5%, 51.7%, and 76.9% of total fuel consumption in a central business district, other urban, and non-urban areas, respectively.

A few studies have linked vehicle mass to its fuel consumption through its influence on vehicle energy demand. For example, Redsell et al. (1993) used measured fuel consumption together with various vehicle data collected based on experiments on different test routes to construct a fuel consumption prediction model by fitting the measured fuel consumption data to an expression that was derived theoretically. As well as vehicle performance and environmental parameters, the theoretical expression included forces resisting vehicle motion; therefore, it included vehicle mass. Based on their model, they argued that fuel consumption is related proportionally to the cube root of vehicle mass. While the data they used included a wide range of traffic, road, driver, and environmental factors, it was based on only three types of Vauxhall Cavalier car models (1300 cc, 1600 cc petrol, and 1600 cc diesel). There is a relatively high correlation between vehicle mass and engine size, both of which contribute to fuel consumption. Due to lack of variation between mass and engine size in the data they used (a sample of 3 cars), it is unlikely that their estimated relationship reflects the isolated effects of mass; it may contain the effects of engine size (as well as other vehicle design factors) as well. The fact that they found the engine size variable not to be statistically significant and hence removed it from the model also supports this argument. They did not validate their calibrated fuel consumption prediction model using other car models with various ranges of design features. In an older study, Biggs and Akcelik (1987) found a 10% increase in mass increases total fuel consumption in the central business district, other urban, and non-urban areas by 3.4%, 4.1%, and 3.2%, respectively, based on their estimated relationship between fuel consumption and vehicle energy demand in these areas as discussed earlier in this section. DeCicco and Ross (1996) estimated that a reduction of 10% and 20% in the mass of a 1300 kg passenger car would reduce its fuel consumption by about 4% and 10%, respectively, through reducing vehicle energy demand. The estimated effects of mass on fuel consumption in these studies are based on the estimated relationship between vehicle energy demand and fuel consumption, which is not consistent between different studies. Differences between these studies could be partly related to different collection methods and the different speed, acceleration, and deceleration patterns considered. In a driving cycle, traffic and road factors can strongly affect external forces that a vehicle must overcome (Burgess and Choi, 2003).

Other studies have estimated the effect of mass on fuel consumption or CO_2 emissions directly. It should be noted that due to the high correlation between fuel consumption and CO_2 emissions, a similar effect of mass on fuel consumption and CO_2 emission is expected. Van den Brink and Van Wee (2001) concluded that 100 kg of extra mass (based on 1000 kg vehicle mass) would lead to a 7-8% increase in fuel consumption as measured by the European fuel test cycle. However, as the authors state, this number also contains the effect of the larger engines that are generally associated with heavier cars. Discussing the carbon reduction benefits of diesel car penetration in Irish vehicle fleet, Zervas (2006) estimated two simple linear relationships between CO_2 emissions and vehicle mass for each of petrol and diesel cars, separately, under the New European Driving Cycle (NEDC) based on average mass of vehicles within a few defined car segments and their average CO_2 emissions. The estimated relationships generally show CO_2 emission increases as mass increases. Fontaras and Samaras (2010) discussed possible changes in vehicle characteristics for meeting European average CO_2 emission limit for passenger cars. Based on simulation results for 6 European car models typical of each of 6 broad car type categories under the NEDC, they estimated that a 5% and 10% reduction in vehicle mass would lead to a reduction of between 1.3-1.8% and 2.7-3.6% in CO_2 emissions, respectively, based on two estimated linear regression models between the change in mass and the simulated resulting change in CO_2 as shown in Figure 2.4. The limited number of car models they used in their study (6) on which the analysis was based makes it difficult to generalise the estimated effects of vehicle mass in this study for all car models which have a wide range of different design features.

While the studies discussed in this section generally show that fuel consumption and carbon emissions increase as mass increases, they do not reflect the partial effects of mass, where the effect is the result of a change in mass holding all other design factors constant. If other contributing factors are not fully controlled when estimating the effects of mass, the estimates may contain effects of other factors as well. One such case is to separate the effects of mass and engine size for different types of fuel and transmission. Larger engines are usually found in heavier cars; therefore the estimated effects of mass could contain effects of engine size as well.

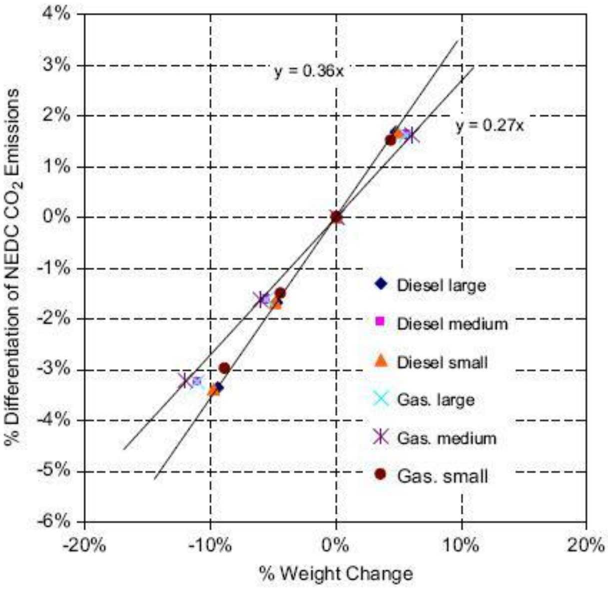


Figure 2.4: Change in CO_2 emission with respect to change in vehicle mass (Fontaras and Samaras, 2010)

2.2.2. Effects of other design factors

Vehicle design features such as fuel type, transmission type, engine size and other engine characteristics can potentially affect fuel consumption through influencing engine efficiency during a driving cycle. Ross (1997) discussed various technologies to reduce the energy consumption of automobiles through increasing engine efficiency based on physics of automobiles. In discussing how to improve mechanical efficiency through friction reduction by engine downsizing, he pointed out that all engine frictions are roughly proportional to engine displacement and downsizing the engine increases efficiency through increasing the ratio of maximum power to displacement, or specific power. Transmission is also regarded to play a critical energy role by determining the operating point of the engine (Stone, 1989 cited in Ross, 1997). Van den Brink and Van Wee (2001) listed the design factors determining engine efficiency as fuel type, engine size, and other engine characteristics (i.e. compression ratio, valve timing, injection/ignition timing). Reviewing the potential improvement in fuel economy through technological changes, DeCicco and Ross (1996) divided the technological improvements corresponding to the key engineering aspects of vehicle design in three categories: engine, transmission, and tractive load.

A number of other studies have found engine size as a main determinant of vehicle fuel consumption using statistical analysis of fuel consumption data (Tenny and Lam, 1985;

Biggs and Akcelik, 1987; Sorrell, 1992; Kirby et al., 2000; Leung and Williams, 2000; Kwon, 2006). However, there are inconsistencies in the estimation results as a result of different methodologies used and factors considered.

Tenny and Lam (1985) performed a statistical analysis of passenger car fuel data to estimate fuel consumption from travel survey data. They proposed two equations, one for urban driving conditions and one for rural and motorway driving conditions, to estimate fuel consumption as a function of average journey speed, engine size, and the payload. Arguing that the matching of the engine size to a car is largely dependent on the mass of the car by showing a high correlation between engine size and mass in the data, they stated that the effect of engine size shown by their introduced functions also included the effects of vehicle mass. They also raised the issue of whether the relationship between fuel consumption and engine size is different for manual transmission cars and automatic transmission cars. Using limited data on official fuel consumption tests in urban cycle, they concluded that such a relationship is not significantly different between the two types of transmissions.

Kirby et al. (2000) modelled new car fuel consumption as a function of a few variables including a lagged fuel consumption term, a time trend term and a fuel price term using data of officially certified fuel consumption rates. Introducing categorical variables for different engine sizes to their model, they found a significant effect of engine size on estimated fuel consumption with fuel consumption generally increasing as engine size increases. The estimated effects of engine size in their study are likely to include the effects of mass as well because vehicle mass is not included in their analysis. Leung and Williams (2000) estimated 0.85 mL/min increase in idle fuel flow rate by 100 cc increase in engine size using combined fuel consumption data from various earlier studies (Claffey, 1976; Watson, 1982; Post et al., 1982; Taylor and Young, 1996; all cited in Leung and Williams, 2000). Using European official fuel consumption rates for 1990 petrol cars, Sorrell (1992) estimated a linear regression model of fuel consumption and engine size and found that a 100 cc increase in engine size would increase fuel consumption by 0.31 L/100km. Kwon (2006) extended this analysis to use more recent data and a log-linear model. He estimated that a 10% increase in engine size increases fuel consumption of 2001 petrol cars by about 6%. These estimates do not reflect partial effects of engine size and may contain effects of other design features such as mass as well.

2.3. Vehicle design and secondary safety

As it was pointed out in Chapter 1, there is no strong evidence suggesting a direct effect of vehicle mass on primary safety (crash involvement) (Auken and Zellner, 2005); however, vehicle mass is directly related to secondary safety (risk of injury given a crash) through influencing the velocity change experienced by the vehicle occupants during the crash (see Equation 1.1). In order to draw conclusions on the safety effects of changes in vehicles' mass in fleet, it is also important to understand the relationships between secondary safety and other vehicle design factors that can change alongside with changes in vehicle mass (such as vehicle size).

2.3.1. Effect of vehicle mass

As mentioned earlier in this chapter, there are two aspects of the effects of mass of a subject vehicle on its safety performance in a crash with another vehicle: a protective effect related to the injury risk (injury probability) of occupants in the subject vehicle and an aggressive effect related to the injury risk that mass imposes to the occupants in the other vehicles in collision with the subject vehicle. Therefore, two-car crashes have been studied intensively in the literature to investigate the role of mass as they reflect both aspects of mass effects. Besides, they provide insight into crashes between any pair of vehicles and also into single-vehicle crashes (Evans, 2003).

In a two-car crash, Equation 1.1 implies that the relative mass of the two cars directly influences the velocity change (Δv). Δv has been regarded and used in vehicle safety research as the best measure of crash severity contributing to the injury risk of vehicle occupants (Evans, 1994). One difficulty in investigating the relationship between injury risk and Δv is the lack of information on the speed of the vehicles prior to a crash, which is required together with mass of the vehicles to calculate Δv. Information on Δv can only be made available when for a sample of two-car crashes, post-crash investigations are made to measure the deformation of the vehicles so that they can be used, together with structural parameters independently determined for the vehicles, to calculate Δv using physical equations (Evans, 1994). This is a costly procedure which is not normally done for all crashes; therefore, such data are usually limited in terms of the number of records and variability of mass.

However, Equation 1.1 implies that in a two-car crash between car 1 and car 2,

$$\frac{\Delta v_1}{\Delta v_2} = \frac{m_2}{m_1}\,. \qquad (2.1)$$

Equation 2.1 shows that in a two-car crash, the velocity change ratio is inversely related to the mass ratio of the cars. As a result of this equation and the lack of data on Δv, several studies have investigated the relative injury risk in two-car crashes as a function of mass ratio, the most important of which are reviewed in the next section. This is followed by a review of the studies that have linked the injury risk in a vehicle that is involved in a crash to its mass (either through its Δv or directly). It should be noted that driver injury risk has been studied extensively in the safety literature as a proxy for secondary safety performance of the vehicles (e.g. Evans and Wasielewski, 1987; Evans, 2004; Broughton, 1996a, 1996b, 1996c; DfT, 2006a). This is mainly because driver exposure is representative of vehicle exposure and that all vehicles involved in crashes have the same number of drivers but not the same number of passengers; hence, it is viewed as providing a consistent basis to assess the secondary safety performance of vehicles within a fleet.

2.3.1.1. Mass ratio and relative injury/fatality risk in two-car crashes

i) First law of two-car crashes (Evans' studies)

Evans (2004) has intensively studied the effect of vehicle mass in two-car crashes using 1978-1998 US crash fatality data and he has found a large effect for mass ratio in a two-car crash. Focusing on passenger cars, he has shown empirically that in a crash between two cars of different masses, the fatality risk ratio (R) of the lighter to heavier car increases as a power function of mass ratio ($\mu = {}^{m_2}/_{m_1}$) of the heavier to the lighter car (Evans and Frick, 1993):

$$R = \mu^u. \qquad (2.2)$$

This relationship, which is regarded by Evans (2004) as the "first law of two-car crashes", has been commonly accepted and used by the researchers and practitioners in the area of vehicle safety. Different values of exponent u for various sets of US crash data are estimated ranging from 2.70 (crashes in all directions) to 3.80 (frontal crashes) (Evans and Frick, 1992, Evans and Frick, 1993; Evans, 1994; Evans and Frick, 1994;

Evans, 2001, Evans, 2004). Table 2.1 shows a few examples of the estimated values of u for different sets of two-car crashes and Figure 2.5 shows the estimated relationship between R and μ for frontal crashes from 1975 to 1998 in the US (first row in Table 2.1) (Evans, 2004).

Table 2.1: Estimated values of u in Equation 2.2 (Evans, 2004)

Description	$u \pm \Delta u$	Data period
Unbelted drivers, frontal crashes	3.58 ± 0.05	1975-1998
All drivers and crash directions	3.53 ± 0.03	1975-1989
Unbelted drivers	3.58 ± 0.04	1975-1989
Belted drivers	3.60 ± 0.13	1975-1989
Frontal crashes	3.74 ± 0.05	1975-1989
Drivers same gender, age within 5 years	3.80 ± 0.09	1975-1989
Rural crashes	3.45 ± 0.03	1975-1989
Urban crashes	3.63 ± 0.05	1975-1989

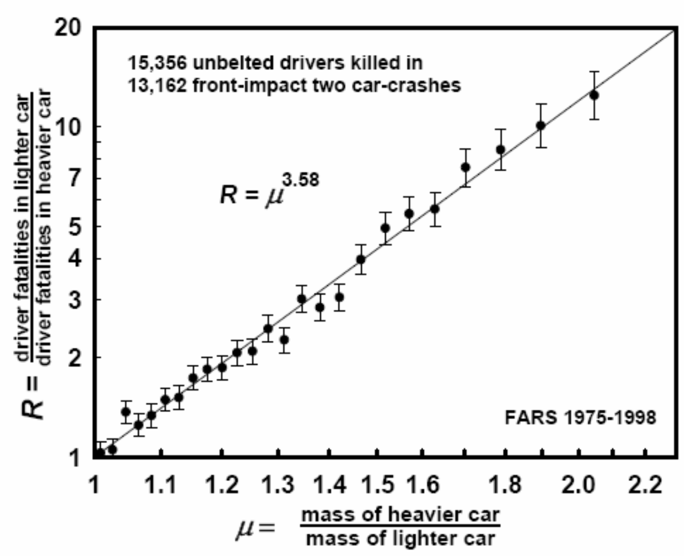


Figure 2.5: The estimated relationship between fatality risk ratio and mass ratio in frontal two-car crashes for unbelted drivers (Evans, 2001)

To estimate the value of u in this relationship, Evans and Frick (1993) confined the crash data to two-car crashes in which at least one of the drivers is killed. Then they defined a mass ratio, μ, for every crash between two cars of different masses as

$$\mu = \frac{m_H}{m_L} = \frac{\mathit{mass\ of\ the\ Heavier\ car}}{\mathit{mass\ of\ the\ Lighter\ car}}\,. \tag{2.3}$$

They aggregated crash data into categories associated with values of μ in given ranges, with the same symbol μ representing the average value in that range. From the crashes included in a given mass ratio range, an associated fatality ratio, R, was defined as

$$R = \frac{f_L}{f_H} = \frac{Number\ of\ driver\ fatalities\ in\ the\ Lighter\ car}{Number\ of\ driver\ fatalities\ in\ the\ Heavier\ car}\ . \tag{2.4}$$

The analysis that led to the Equation 2.2 was identified as “symmetric” case where the only distinguishing feature between the cars was their mass; therefore, the crashes in which the two cars have the same mass were excluded from the analysis. On the grounds that cars with identical mass and no other distinguishable feature must have equal fatality risks, the estimated relationship was constrained to the point $R=1$ and $\mu=1$.

The resulting dataset, which consisted of a few observations of R and μ calculated by aggregating crash data into certain number of categories, was least square fitted to $log(R) = u \times log(\mu)$ to estimate the value of exponent u in Equation 2.2. Each point was weighted in inverse proportion to the variance, σ^2. Evans and Frick (op.cit.) assumed a Poisson process for fatalities when the variance was given by $1/_{f_L} + 1/_{f_H}$.

In order to aggregate the data into a certain number of mass ratio categories for the analysis, the data was first divided as equally as possible into different number of categories. The chosen number of categories used to estimate the value of u, which depended on the number of available crash observations, would be the number that minimizes the value of $\Delta u/_{u}$, where Δu is the estimated standard error (Evans and Frick, 1994).

To support this empirical results, Evans (1994) attempted to explain Equation 2.2 by deriving this relationship using a combination of two sources of information: calculations based on Newtonian mechanics and US National Accident Sampling System (NASS) data.

He first showed that in a two-car crash between car 1 and car 2, the ratio of the speed changes for two cars is related to the mass ratio of the cars according to the Equation 2.1 This relationship can be easily derived from Equation 1.1. Then Evans (op.cit.) used NASS data for 1982-1991 period to empirically show the relationship between risk of driver injury or fatality (P) and the severity of a crash (measured by Δv, velocity change), which as first suggested by Joksch (1993), had the following form:

$$P = \left(\Delta v/_{\propto}\right)^k \tag{2.5}$$

where $\propto$ and k are parameters that are estimated from the crash data. Evans (1994) then used Equations 2.1 and 2.5 to derive Equation 2.2 subject to $k = u$.

While Evan's relationship provides a simple approach to estimate injury and fatality risk ratio as a function of mass ratio in two-car crashes, it is associated with some major disadvantages. As Evans (op.cit.) points out, the underlying assumption behind Equation 2.2 on the relationship between driver injury and fatality risk and vehicle velocity change (as appears in Equation 2.5) suffers from a major structural problem that gives values of risk greater than 1 when $\Delta v > \propto$.

The other disadvantage of Evans' methodology is its lack of flexibility in controlling for or estimating the effects of other contributing factors. Given the basis of the methodology which is aggregate analysis of crash data, a number of contributing factors that might vary within the aggregate categories of two-car crashes that are used to estimate the value of u in Equation 2.2 cannot be controlled. Therefore, the estimates are unlikely to reflect the partial effect of mass ratio on R. Excluding other effects from the estimates can be done by confining the data to more limited categories by placing restrictions on both vehicles accounting for different effects. However, this approach, which has been used in part by Evans (for examples, see Evans and Frick, 1993; Evans, 2001), can substantially reduce sample size leading to an increased uncertainty in the estimation results.

ii) Other studies

Wood (1997) derived the same kind of relationship as Equation 2.2 between relative injury risk and mass ratio in frontal two-car crashes using the fundamental relationships of Newtonian mechanics. In his calculations, relative injury risk is defined as the ratio of injury severity in the two cars when injury severity is assumed proportional to average body acceleration to the power of 2.5 as first suggested by Gadd (1966). However, on Wood's (op.cit.) calculations, the relationship between injury severity and injury risk (defined as the probability of injury in the event of the crash) remains unclear. Although he has compared his theoretical models with the results from the field data available in the literature (Ernst et al., 1991; Ernvall et al., 1992; Evans and Frick, 1993) and has found a high level of correlation between the two, the measure of relative risk used by him in his theoretical calculations (as defined above) is not necessarily equivalent to the one used in the literature based on field data (which is the ratio of injury risks, R). Based on his calculations, he concluded that in crashes between two cars of different size, the fundamental parameters contributing to the relative injury risk are masses of the cars and the structural energy absorption properties of the cars.

A few other studies have investigated a similar relationship between fatality and injury risk ratio and mass ratio in two-car crashes (e.g. Ernst et al., 1991; Ernvall et al., 1992; Joksch, 1998; Ross and Wenzel, 2001). These are all empirical studies based on aggregate analysis of crash data which have used a similar approach to that of Evans and Frick (1993) as it was explained earlier in this section.

2.3.1.2. Vehicles mass, velocity change, and injury/fatality risk

i) Velocity change (Δv) and risk of injury in two-car crashes

As mentioned briefly in the previous section, Evans (1994) used NASS data for 1982-1991 period to empirically investigate the relationship between risk of driver injury or fatality (P) and the severity of a crash (measured by Δv, velocity change). NASS data is a probability sample of crashes reported by the police in the US. Δv is coded for tow-away vehicles involved in these crashes, which is calculated using equations that include the measured amount of vehicle deformation as well as parameters related to the structure of the vehicle. More information about NASS data is available in (Evans, 1994). As explained in Section 2.3.1.1, Evans (op.cit.) used Equation 2.5 to show the relationship between P and Δv in different categories of crashes. His results are shown in Figure 2.6. In these relationships, P is set to 1.0 when $\Delta v > \propto$. This implies that, for example, unbelted and belted drivers have no chance of survival when Δv exceeds 70.6 mile/h (113.6 km/h) and 69.2 mile/h (111.4 km/h), respectively.

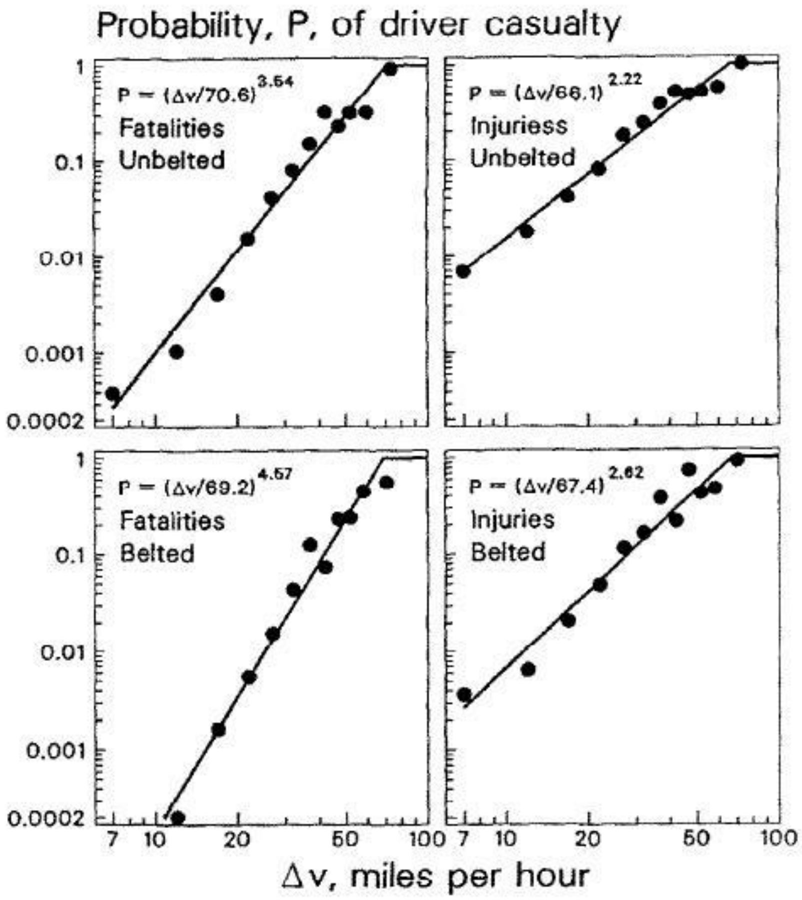


Figure 2.6: The estimated relationships between risk of injury and fatality versus Δv based on 1982-1991 NASS data (Evans, 1994)

Wood et al. (2007) discussed limits of survivability in frontal collisions where they compared the relationship between risk of injury and fatality and Δv based on two different functions. The first was a modified power function with a structure similar to that introduced by Joksch (1993) (Equation 2.5) but included a critical Δv (Δv_c) to avoid resulting in values greater than 1.0 for risk, as follows:

$$P = \left({}^{\Delta v}\!/_{\Delta v_c^{1+t\propto}} \right)^n \tag{2.6}$$

where $\propto$ is the transformed standard error and t is the normalised distance from the mean, both of which are estimated from the crash data (see Wood et al., 2007 for details). They used Monte Carlo simulation to estimate this relationship. Their estimated limits of survivability using this function (critical Δv) based on 1982-1991 NASS data for belted and unbelted drivers was 135 km/h and 145 km/h , respectively. Both of these values are higher than those estimated by Evans (1994) using Equation 2.5 and based on the same data. The second function they used was a logistic regression model which has the most appropriate functional form to estimate risk. This was also estimated using 1982-1991 NASS data. They compared their estimated relationships with the observations for driver fatalities and injuries. Figure 2.7 shows the results of this comparison for belted and unbelted drivers. Based on their estimated relationships, there is zero probability of survival for belted and unbelted drivers for velocity changes of 135-140 km/h and 145-150 km/h, respectively. Whilst they concluded that both functions are asymptotic to injury and fatality risk of 1.0 over the same range of *Δv*, they stated that to confirm any finding based on real crash data, further in-depth investigations were required. This was because the data they used included too few observations, specially for high-speed crashes, to result in statistically reliable estimates.

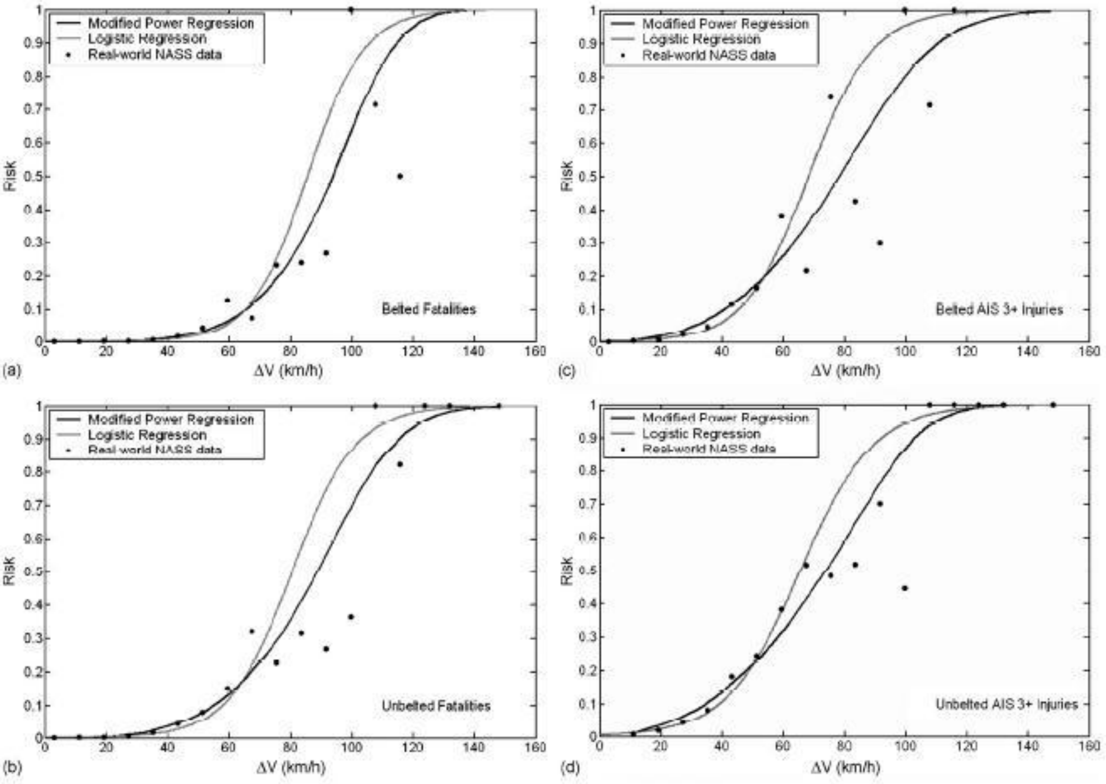


Figure 2.7: The estimated relationships between risk of injury and fatality versus *Δv:* Estimated logistic and modified power regression compared to NASS data (Wood et al., 2007)

Richards and Cuerden (2009) investigated the relationship between risk of injury and fatality and *Δv* in Britain using data from Co-operative Crash Injury Study (CCIS). CCIS is an ongoing project which collects in-depth real crash data in Great Britain according to a stratified sampling procedure which favours cars containing Killed or Seriously Injured (KSI) drivers. Vehicle examinations after the crash allow calculation of *Δv* for each vehicle. They used logistic regression to estimate driver fatality and KSI risk and as a function of Δv when risk is defined as the probability of fatality (or KSI) given that the driver is at least slightly injured. They found that for a belted driver in a frontal impact with another car, risk of fatality and KSI is 50% at a Δv of about 77 km/h and 39 km/h, respectively. For a belted driver on the struck side in side impact with another car, they found that risk of fatality and KSI is 50% at a Δv of about 50 km/h and 42 km/h, respectively. The confidence intervals of their estimations were relatively wide due to the small number of observations.

Toy and Hammitt (2003) used logistic regression models to separate the influence of different factors, including Δv, on risk of serious injury or death to driver in two-car crashes. The data they used was obtained from US Crashworthiness Data System (CDS), a computerized database claimed to be nationally a representative sample of police reported crashes in the US, which contains information on Δv. Including Δv to their estimated models as a proxy for crash severity to represent the effects of vehicles speed and vehicles mass, they found a significant effect of Δv on risk of serious injury or death to driver when other factors that contribute to risk are controlled in the model. Based on their results, a one unit change in Δv (1 km/hr) increases risk by 12%.

ii) Vehicle mass and risk of injury in two-car crashes

Other studies have discussed the relationship between vehicle mass and driver risk of injury when involved in crashes with different ranges of vehicles.

Broughton (1996a, 1996b) discussed the effect of vehicle mass on injury risk in two-car crashes based on British crash data where injury risk (D) is defined as the probability of driver injury when the vehicle is involved in a two-car crash in which at least one of the drivers is injured. Broughton (1996a) investigated the relationship between mass and driver injury risk using data of popular makes and models involved in two-car crashes in Great Britain from 1989 to 1992 (Figure 2.8). He found that driver risk falls steadily with increasing mass and that mass could explain a high proportion of variation in the casualty data. This generally reflected the greater protection of drivers in the heavier cars compared to that of drivers in the lighter cars in fleet; however, this relationship alone does not provide any information on the aggressive effect of vehicle mass in fleet.

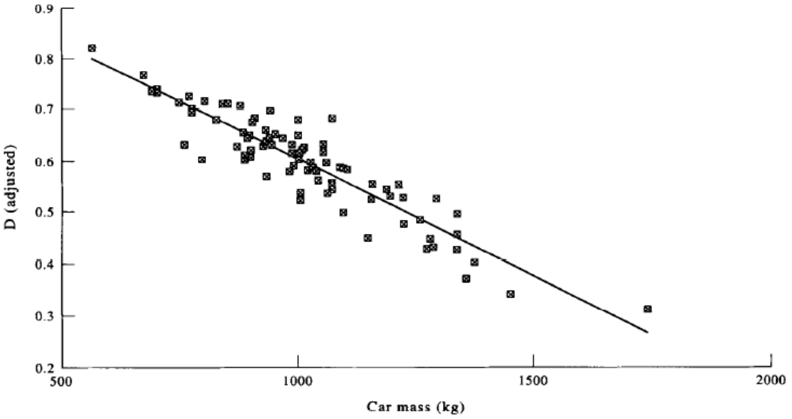


Figure 2.8: The relationship between average vehicle mass and its driver injury risk (D) in British fleet: 1989-1992 (Broughton, 1996a)

Estimating the proportion of drivers of cars in collision with different makes and models who are injured as a measure of aggressivity of these makes and models, Broughton (1996b) investigated the role of vehicle mass on the distribution of risk between the lighter and the heavier car in two-car crashes. He found that as the difference between the masses increases, the proportion of crashes in which the driver of the heavier car is injured diminishes significantly, while the proportion in which both drivers are injured diminishes only slightly. He concluded that the distribution of the risk of driver injury when two cars collide depends principally on the difference in mass. Although the measure of injury risk that he has used is not the ideal measure as it is not independent of risk of injury in the colliding car (Broughton, 1996c), he has compared different technical aspects of this measure (as will be discussed later in this chapter) with those of other alternative risk measures and has shown that it is the most satisfactory measure of risk estimated directly from the crash data that reflects secondary safety of vehicles in fleet (Broughton, 1996a, 1996c). This index, which was first defined and used by UK Department of Transport (DfT, 1993), is referred to as the British or DfT index (Broughton, 1996b). It should be noted that the ideal measure of driver injury risk in two-car crashes is the absolute injury risk defined as the probability of driver injury when the vehicle is involved in a crash, whether or not the driver in the colliding vehicle is injured. However, the major issue with this measure is that it cannot be directly estimated from the crash data because data on non-injury crashes (crashes in which neither of the drivers are injured) is not normally available.

On the other hand, when the measure of vehicle safety has been defined in a different way to include risk of crash involvement as well, there are inconsistencies in the findings on the effect of mass compared to those previously explained. Wenzel and Ross (2005) found that mass alone is only a modestly effective predictor of risk when risk is defined as driver deaths per year per million registered vehicles for a given car model and all types of crashes. For cars with roughly similar masses, they found that own-driver risk can vary greatly between manufacturers. The difference between their results and those from previous studies on the effect of mass could arise because they considered all types of crashes and also used a different measure of vehicle safety; one which is a measure of both primary safety (crash involvement) and secondary safety (injury risk). Therefore, care should be taken when findings of such studies on the effects of mass are interpreted and used for investigating safety outcome of changes in vehicles' mass in fleet.

2.3.2. Separate effects of mass and size

There is generally a high level of correlation between vehicle mass and size (vehicle length or wheelbase has been often used as a proxy for vehicle size in the literature). Many of the studies that have investigated the effect of mass on risk of injury and fatality have not controlled for the effect of vehicle size appropriately; therefore, their estimates could contain the effects of vehicle size as well. There are evidences in the literature suggesting different effects of mass and size on risk of injury and fatality given a crash; however, there are inconsistencies in the results of different studies. The main question, which has remained unclear in the literature, is whether there is any effect of vehicle size above and beyond that of mass ratio (Hutchinson and Anderson, 2009). The following presents a review of some of the key studies on the effects of vehicle size.

A few studies have attempted to explain the theoretical relationships between vehicle size and risk of injury and fatality in crashes. For example, Van Auken and Zellner (2005) investigated the independent effects of mass and size on crash worthiness (secondary safety performance), measured as the risk of fatality, based on theoretical models. They concluded that in collisions without local intrusion (when the passenger compartment within the vicinity of the vehicle occupant remains relatively intact during the impact) and for a similar vehicle mass, longer and wider vehicles could be expected to have less occupant fatality risk due to having a structure that allows more deformation. In the case of collisions with local intrusion, they concluded that for a given mass, vehicle size reduction tends to increase the likelihood and the amount of intrusion, resulting in an increased risk of fatality. They also showed that, as observed in the vehicle fleet, vehicle mass can vary independently of vehicle size through methods such as material substitutions or advanced structural design. Ross and Wenzel (2001) used basic physics to explain the relationship between vehicle mass, size, and risk of injury and fatality in a crash. They stated that not only the injury outcome is influenced by the velocity change, Δv, but also it is influenced by the time during which this velocity change occurs (the acceleration experienced). Then they argued based on physical formulae that an increase in crush distance, which is correlated with vehicle size, could increase the time, and hence, decrease the risk of injury by decreasing the acceleration experienced.

Several studies have attempted to investigate the separate effects of mass and size on risk of injury and fatality using statistical analysis of crash data. For example, Evans (2001) did so by analysing two-car crashes between cars of the same mass. Using various sets of data, he showed that when two cars of the same mass, M, crash into each other, the relative driver risk, R_{MM}, varies with the common mass of the cars according to the following equation:

$$R_{MM} = \frac{k}{M} \tag{2.7}$$

where k is a constant that is estimated from the crash data. Figure 2.9 shows fitted relationship between R_{MM} and M for five sets of crash data where the data is scaled to assign a risk of 1.0 to M = 1400 kg. He argued that although the relationship is given in terms of mass, it is in fact reflecting the effects of size because in crashes between cars of equal mass, Δv is always half of the closing speed irrespective of mass of the vehicles (see Equation 1.1). He has referred to Equation 2.7 as the “second law of two-car crashes” (Evans, 2004). This equation implies that, for a given mass, drivers in the larger vehicles have the lower risk of injury and fatality than those in the smaller vehicles, arguing that M in Equation 2.7 in fact reflects the effect of size. He used Equations 2.2 and 2.7 and introduced a model of absolute driver injury risk as a function of vehicle size (represented by vehicle length) and vehicle mass (Evans, 2004). This model of risk has not been tested with any real crash data.

Van Auken and Zellner (2005) investigated the safety effects of reductions in mass and size of passenger cars and light trucks using regression analysis of US crash data where they included mass, wheelbase, and track to their regression models. Based on the results for both risk of crash involvement and risk of fatality given a crash, and for different types of crashes, they generally concluded that vehicle mass reduction tends to decrease fatalities while vehicle wheelbase and track reduction tends to increase fatalities. As pointed out by Hutchinson and Anderson (2009), the estimated significant effects of mass and size that are in opposite directions for some types of crashes (i.e. “rollover” and “hit object”) in this study makes it difficult to draw a general conclusion on the separate effects of mass and size based on their results.

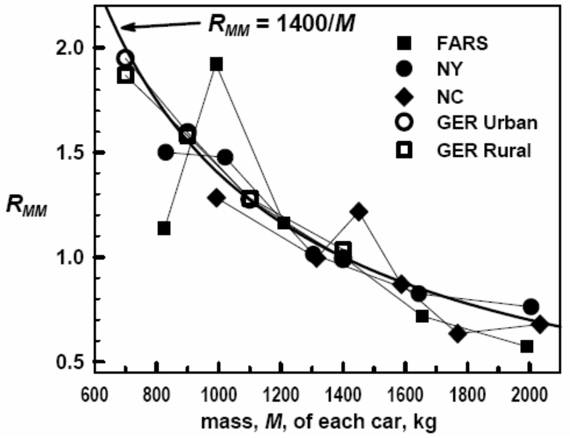


Figure 2.9: Relative risk (to that of cars with mass of 1400 kg) versus mass when cars of the same mass crash into each other (Evans, 2004)

A few other studies, all of which are based on the US crash data, have also found that an increase in vehicle size tends to result in a reduction in risk of injury to vehicle occupants (Evans and Wasielewski, 1987; Ross and Wenzel, 2001; Ross et al., 2006). However, other studies, based on different sets of data, have resulted in different conclusions (e.g. Grime and Hutchinson, 1979; Grime and Hutchinson, 1982; Broughton, 1999).

i) Vehicle mass as a proxy for vehicle size

According to Evans (2004), any estimated effect of mass in two-car crashes when mass ratio is controlled (e.g. in two-car crashes between cars having a similar mass) tends to reflect the effect of vehicle size because of the high correlation between mass and size. Analysing British crash data from 1969 to 1972, Grime and Hutchinson (1979, 1982) did not find a significant effect of mass on driver injury severity for the cars involved in two-car crashes with cars of similar mass as well as in single car crashes. This is directly in contrast with Evans' findings based on the US crash data (Evans, 2001). In a later study based on a more recent British crash data (1991 to 1994), Broughton (1999) modelled risk of driver fatality in two-car crashes, when risk is defined as the number of driver fatalities where at least one of the drivers is injured (of any severity), using logistic regression models. He included mass ratio, as well as absolute mass of the subject vehicle whose driver risk was being estimated, in the models and surprisingly found a significant and positive effect of mass above the effect of mass ratio. This translated as the risk of a driver fatality increases with car mass in two car crashes when all other factors including mass ratio are unchanged. These results, which are in contrast with other theoretical and empirical findings, suggests possible existence of collinearity in the estimated models due to a high correlation of absolute mass and mass ratio in the data, both of which were included in the models as explanatory variables. As suggested by Hutchinson and Anderson (2009), differences on the findings of these studies on the effects of absolute mass or size in two-car crashes, when mass ratio is controlled, is partly due to the different data used; as Figure 2.2 suggests, American cars have tended to be larger than European cars.

2.3.3. Specific design effects of makes and models

As mentioned earlier, there are two distinct measures of safety performance for a vehicle that is involved in a two-car crash: "Secondary Safety Performance" which is linked to the injury risk to the occupants of that vehicle, and "Aggressivity Performance" which is linked to the injury risk that the vehicle imposes to the occupants of the other vehicle. Different levels of secondary safety and aggressivity performance are associated with different types of vehicles in fleet depending on their mass and other design features. Studies have suggested protective and aggressive effects for vehicle design features other than mass (i.e. stiffness, geometry, body structure) in two-car crashes (Buzeman et al., 1998; Toy and Hammitt, 2003; Van Auken and Zellner, 2005). In vehicle safety research, the effects of such factors have often been investigated as represented by specific secondary safety and aggressivity performance of makes and models in fleet (e.g. Broughton, 1996a, 1996c; Newstead et al. 2000; Wenzel and Ross, 2005; DfT, 2006a), the most important of which are reviewed as follows.

i) Secondary safety performance

Broughton (2007) analysed the influence of type of car and registration year on the number of car driver casualties in Great Britain and found that the mean risk of death for the driver of the smallest types of cars (minis and superminis) is four times the risk for the largest type (4x4s and people carriers) when the types are arranged in order of increasing mass and physical dimensions. He also found that compared to older cars, newer more modern cars are safer for their occupants and more aggressive to occupants of cars with which they collide. He argued that these effects are partly because of an increase in the mass of new cars. Other studies have also found a greater risk of injury for older cars compared to newer cars in fleet (Blows et al., 2003; Frampton et al, 2002; Broughton, 2007). This could be partly related to the effects of increased average mass of newer cars, which is not controlled in these studies, and partly related to a better design of newer car models compared to older ones.

Wenzel and Ross (2005) estimated a combined risk for each make and model in 1997-2001 US fleet. The combined risk was the sum of the risk to the drivers in all kinds of crashes and the risk to the drivers of the other vehicles in two-car crashes when risk was defined as the driver deaths per year per million registered vehicles. While this measure gives an indication of risk of being involved in fatal crashes per ownership for different car models, it does not take into account the influence of vehicle usage. It seems possible that some makes and models have a significantly different usage than others and hence significantly different exposure to the risk of being involved in crashes. Besides, the defined measure is not an appropriate index to compare secondary safety performance of makes and models as it is influenced by the primary safety performance (risk of crash involvement) of vehicles as well.

As it was mentioned briefly in Section 2.3.1.2, DfT estimates secondary safety performance of popular makes and models in Great Britain, which have different design features, as the driver injury risk and report the results for specific time periods. The latest DfT report on secondary safety of vehicles is available for 186 models of cars involved in traffic crashes during 2000 to 2004 in Great Britain (DfT, 2006a). The DfT estimates of secondary safety performance are calculated for cars using data from the two-car crashes where at least one driver was injured. The DfT safety index for car model m (D_m) is defined as the following.

> D_m is the proportion of drivers of a car model m who are injured when involved in two-car crashes where at least one driver is injured.

Adjusted D for all makes and models are calculated using logistic regression models to allow for speed limit (proxy for accident[1] severity), first point of impact, driver sex, and driver age.

Broughton (1996a, 1996c), in two papers, discussed the DfT method for estimating safety indices as a measure of secondary safety performance of vehicles. In the first paper, he concluded from theoretical considerations that DfT indices provide the most satisfactory means of comparing the secondary safety performance of different models of cars compared to alternative available indices (Broughton, 1996c). Broughton (1996a) suggested that it is sensible to concentrate on the “all casualties” index (D_{all}) as it is shown to be highly correlated with “ksi” index (D_{ksi}) and also it is more discriminating because of the much larger number of accidents used in its calculation. Broughton (1996c) discussed practical aspects of the indices in the second paper. He showed that the DfT indices are not biased by ignoring the differences in the distribution of “other” cars for different car models involved in accidents. He found that indices calculated from individual years of data are consistent with the indices calculated for the grouped data from 1989-92 and argued that it is justified to accumulate data over several years to provide more reliable results. On the other hand, the fact that the index is a relative measure which compares the safety of different models at the same time limits the number of years over which the index should be calculated. This is because the design of vehicles in fleet changes over time. He also discussed that the indices are closely clustered when calculated for different model variants within makes and models; therefore, it is justifiable to calculate aggregate indices for each make and model to provide more reliable results.

ii) Aggressivity performance

To complement the DfT secondary safety index, Broughton (1996c) defined an aggressivity index for car model m (A_m) as the following.

> A_m is the proportion of drivers of cars who are injured when involved in collision with car model m where at least one driver is injured.

---

[1] The words “accident” and “crash” have been used interchangeably throughout this thesis

As he stated, the defined *D* and *A* indices for makes and models are not independent of injury risk in the "other" cars in collision with them. This is due to the correlation between the defined risk measures (*D* and *A*). For example, when estimating secondary safety performance of a given make and model, if the "other" cars in collisions with this make and model are hypothetically replaced with less physically vulnerable drivers leading to a reduction in driver injury risk in the "other" car, this would result in an increase in the estimated secondary safety performance of this make and model. In an ideal situation, the estimated secondary safety and aggressivity performance of makes and models should be independent of risk of injury in the "other" cars in collision with them. This is the main disadvantage associated with DfT's defined secondary safety index (*D*) as well as Broughton's defined aggressivity index (*A*) which is because non-injury crash data is not available in Great Britain. The other disadvantage of the DfT methodology is the fact that the estimated effects do not reflect partial effects of vehicle mass and other specific design features of makes and models.

Les and Fildes (2001) reviewed different methods of estimating vehicle aggressivity and proposed and compared two aggressivity rating methods where the "subject" car aggressivity is estimated based on injury outcome to the driver of the "other" vehicle involved in a two-vehicle crash using logistic regression techniques. They correctly concluded that one major disadvantages of all available methods that estimate vehicle secondary safety and aggressivity is use of the concept of "relative risk" instead of the ideal measure of "absolute risk" where risk of injury in one vehicle in a two-vehicle crash is independent of risk of injury in the colliding vehicle. As mentioned before, absolute risk of injury cannot be calculated directly from the crash data due to lack of data on crashes in which neither of drivers are injured. As a result of using "relative risk" in estimating secondary safety and aggressivity, there is the potential issue of correlation between them and its influence on the estimates.

Newstead et al. (2000) used a preferred method of estimating secondary safety and aggressivity performance to compare safety performance of makes and models in Australian fleet. They used Police reported two-car crash data from three states (Victoria, New South Wales, and Queensland) during 1987-98; these data included crashes that resulted in death or injury or a vehicle being towed away. The indices that they introduced are the products of two probabilities. The first is the probability of being injured when involved in a crash where a vehicle is towed-away, and the second is the probability of an injured driver being killed or hospitalised when the vehicle is involved in a crash. While this results in two independent indices for secondary safety and aggressivity, their estimation is solely dependent on the availability of non-injury crash data. Besides this, the effect of the mass of the colliding vehicle (which contributes to crash severity) is not controlled in their study.

2.4. Safety and environmental effects of changes in vehicle design within the fleet

The effects of mass on fuel consumption and safety risk that are documented in the literature suggest that vehicle mass has the potential to cause a conflict between environmental and safety goals. The main concern is whether reducing vehicle mass within a fleet to improve fuel efficiency can have a detrimental effect on safety through changes in injury and fatality risk. The following reviews the key literatures that have examined this issue.

Ross and Wenzel (2001) examined a safety-fuel economy scenario in the US that focused on changes in vehicle design (mainly mass) within the fleet with a priority of reducing traffic fatalities with a view of increasing fuel economy as well. The ultimate goal of the scenario was to narrow the range of masses while maintaining or increasing selected spaces. As a result, the masses of heavier and larger vehicle types would be reduced while masses of lighter vehicle types would remain the same. Discussing possible mass-reduction technologies, they estimated a reduction of about 2000 fatalities (based on 1999 fatality rates) as well as an improvement in fleet fuel economy. They only made rough estimates of the effects, which were all based on the estimated effects of mass and size available in the literature as discussed earlier in this chapter. They have also used a number of assumptions in estimating these effects which are not fully described (such as changes in vehicle ownership and vehicle usage pattern, or changes in vehicles distribution by engine size, fuel type, and other design features that could influence fuel consumption).

Broughton (1999) investigated the likely effects of uniform fleet downsizing (reducing mass and size of all passenger cars in fleet proportionally), which leads to a reduction in energy consumption and atmospheric pollutions, on road safety in Great Britain. Using a series of statistical models estimating driver injury and fatality risk in different types of crashes, he concluded that a uniform reduction in mass and size would lead to fewer injuries and fatalities subject to a number of assumptions that are discussed fully in the study. This was mainly based on a model that unexpectedly estimated a significant and positive effect of absolute mass above that of mass ratio on risk of injury and fatality when a vehicle is involved in two-car crashes. As discussed in Section 2.3.2, there are potential issues in the methodology used in this study that are not addressed properly.

Buzeman et al. (1998) investigated changes in crash injuries and fatalities as a result of a number of hypothetical changes in mass distribution of vehicles in the Swedish passenger car fleet and concluded that a uniform mass reduction of 20% increases fatalities by 5.4% while a reduction of mass range by 20% reduces fatalities by 3%. They used the relationship between risk of injury and Δv estimated by Evans (1994) as the base of their calculations to estimate the injury risks. Therefore, their methodology is dependent on the availability of information on the impact speed (closing speed) of two-car crashes (the second term in Equation 1.1) for which they used the observed distribution of impact speeds from the 1982-1999 US NASS data as it was given by Evans (op.cit.). To calculate the total number of injuries, they multiplied the risk of injury by the probability of crash involvement for different categories of Δv where the probability of crash involvement assumed to be a function of vehicle mass. This implied that when mass of a vehicle changes, the probability of that vehicle to be involved in a crash of a given Δv changes as well. As it was mentioned earlier, there is no strong evidence of a direct effect of vehicle mass on risk of crash involvement in the literature. They also made a number of assumptions in their analysis; for example, they assumed that the distribution of closing speed is equivalent to that of Δv (mass distribution is similar in each Δv category). Besides, in their analysis, which is solely based on frontal two-car crashes, they do not discuss the effects of vehicle size.

In addressing the safety implications of improving vehicle fuel economy, Zachariadis (2008) investigated the relationship between fuel consumption and secondary safety performance of recent model European motor cars using a collective source of independent data on vehicle mass and fuel consumption together with the Euro NCAP crash test performance records for different vehicles. He used statistical modelling to investigate various relationships between mass, safety performance, rate of CO_2 emissions and vehicle design features. From this, he concludes that additional safety performance increases mass slightly and does not necessarily increase fuel consumption when other design features are taken into account. There are some important issues associated with his study. Whilst the use of the Euro NCAP crash test results as an indicator of vehicle safety was by itself an innovation, as he noted, Euro NCAP safety crash test safety ratings are not comparable between different groups of vehicles. However, this is not addressed adequately in the study: he estimated a single model for all groups and then a separate model for each group. The former is inappropriate because the safety ratings are not all measured on the same scale, whilst the latter is inappropriate because the greater effects of differences between vehicle classes are not represented. The other issue associated with Euro NCAP tests is that they do not reflect the effect of relative mass on injury outcome in two-car crashes properly whilst data based on real crashes can do so because they include crashes between cars with different ranges of mass. Therefore, the main conclusion of this paper that “there is no trade-off between better car safety and CO_2 emission reduction” is not supported appropriately by the analyses reported in the paper. The trade-off between fuel economy and safety performance depends on the mass distribution within the fleet.

A few studies have addressed the issue of trade-off between fuel economy and safety goals as a whole in fleet using different methodologies and, as a consequence, there are inconsistencies in the results. These studies are mainly empirical studies that have used aggregate time-series data.

A report by the US National Research Council (NRC, 2002) concluded that changes in masses of cars and light trucks in the US since the 1970s, some of which was due to Corporate Average Fuel Economy (CAFE) standards, could have resulted in 1300 to 2600 additional fatalities in 1993. This conclusion was based on an earlier analysis by Kahane (1997) where he estimated the effect of mass reduction in passenger cars, light trucks and vans on fatalities. Findings from this report were later superseded by applying different analytical techniques to more recent crash data where Kahane (2003) estimated a larger fatality increase as mass is reduced for all crash modes. Crandall and Graham (1989) analysed US time-series data from 1947-1981 and found that additional fatalities occurred as a result of CAFE standards through estimating an increase in fatalities by a decrease in vehicle mass and by linking higher fuel efficiencies to a decrease in mass of new cars.

The methodology they used is questioned by other studies in terms of the type of data and modelling approach used (Noland, 2004) and the time-series period selected (Ahmad and Greene, 2005). Noland (2004, 2005) used count data methods and accounted for heterogeneity and other contributing factors to analyse the effects of average fuel economy of vehicles on traffic-related fatalities. He examined two different aggregate datasets. Using US state-level time-series data, he found that improvements in fuel efficiency were associated with increased fatalities in the 1970s, but this effect had largely disappeared after the mid 1980s (Noland, 2004). He also analysed country-level time-series data from 13 countries and found that changes in vehicle efficiency are not associated with changes in traffic fatalities (Noland, 2005). Using cointegration analysis and time-series data on US light duty vehicle fuel economy and highway fatalities, Ahmad and Greene (2005) found the unexpected result that the stationary linear relationship between the average fuel economy of passenger cars and light trucks, and highway fatalities is negative meaning that reduced fuel consumption is linked to fewer fatalities. The inconsistencies in the results of these studies linking average fuel consumption to the number of fatalities are partly due to the different vehicle fleets and different time periods studied. Besides, since the effect of vehicle mass is not controlled for, it is not clear to what extent the changes in average fuel consumption are related to the changes in vehicle mass.

2.5. Summary

The key studies relevant to the issue of trade-off between fuel consumption and secondary safety performance in vehicle design imposed by vehicle mass were reviewed when the studies were grouped under three main topics: the relationship between vehicle design and vehicle fuel consumption, the relationship between vehicle design and vehicle safety performance, and the safety and environmental consequences of changes in vehicle design within the fleet. Whilst the investigated effects of different vehicle design features were reviewed, the special emphasis was placed on the estimated effects of mass which is the key design variable that potentially imposes the trade-off. Table 2.2 gives a summary of the most relevant reviewed studies. The table includes the name of the authors, year of publication, main analysis method, data, main findings, and major shortcomings associated with each study.

A group of studies have linked vehicle mass to its fuel consumption through the way mass influences vehicle energy demand. The estimated effects in these studies, which are based on the estimated relationship between vehicle energy demand and fuel consumption, vary among the different studies and so are not consistent with each other. As discussed earlier, differences could be partly related to different collection methods and the different speed, acceleration, and deceleration patterns considered. Other studies have estimated the effect of mass on fuel consumption directly using simple linear or log-linear regression models; however, the estimates do not reflect reliably the partial effects of mass and include effects of other factors which have not been controlled properly as a result of the type of data used or the analysis methods applied. In summary, the review of literature on the relationship between vehicle mass and its fuel consumption revealed that there are key questions on the effects of mass for which the literature does not provide any reliable answer. These include:

- What is the effect of a reduction in vehicle mass, holding all other design factors including engine size constant, on its fuel consumption rate?
- Is partial effect of vehicle mass on its fuel consumption rate different between cars with different design features (different fuel types or transmission types)?
- How different the relationship between vehicle mass and its fuel consumption rate is when the vehicle is driven under different driving conditions (e.g. urban versus rural driving conditions)?

Whilst it is generally agreed that a reduction in vehicle mass in a vehicle fleet results in a reduction in overall fuel consumption and carbon emissions, all other factors remaining constant, there are disagreements on the effects of such changes on safety. This is due to the fact that while vehicle mass has a protective effect on the injury outcome of its occupants in crashes, it also has an aggressive effect on the injury outcome of occupants of the other vehicles with which it collides (see Equation 1.1). Therefore, in order to be able to predict the likely safety outcome of a change in vehicles' mass in fleet reliably, detailed relationship between mass and secondary safety performance of vehicles in the fleet should be clear.

The statistical modelling techniques that have been used in the area of accident modelling depend on the measure of safety analysed. The typical models used to estimate injury outcome of crashes (secondary safety) are fundamentally different from those used to estimate crash frequencies (primary safety). Since the number of crashes occurring at an entity, which is a measure of primary safety, is non-negative, discrete, and random, count data models such as Poisson and negative binomial regression models (and their extensions) have been used commonly in the literature to model crash occurrence[1]. On the other hand, injury outcome in the event of a crash, which is a measure of secondary safety, can be described as a binary variable (e.g. injury / no injury). In this case, logistic regression model has been the most frequently used model in the literature to estimate the probability of injury in the event of a crash. Jones and Jorgensen (2003) state that the popularity of logistic regression is mainly because of its ease of interpretation, widespread acceptability, and provision of appropriate estimation routines in the majority of statistical packages. Alternatively, injury outcome can be described as an ordinal variable (e.g. no injury, slight injury, serious injury, fatality). To model injury severity as an ordinal measure, ordered probability models such as order logit and probit have been used in the literature.

It was discussed in Section 2.3.1.1 that the relationship between fatality and injury risk ratio and mass ratio in two-car crashes introduced by Evans and Frick (1993) (Equation 2.2) is the most well-known relationship; however, certain disadvantages are associated with the methodology behind this relationship (discussed in Section 2.3.1.1) which suggests that there is a scope for more investigations on this relationship as well as further research for an alternative relationship.

Review of available literature on the effects of mass in two-car crashes revealed that the partial effect of the relative mass of the cars on absolute driver injury risk is not clear. A key question is that, for example, if in a two-car crash with a given mass ratio and absolute injury risk to the driver in each car, the mass ratio changes while all other factors remain constant, how does the absolute driver injury risk in each car change? The answer to this question is the key to an estimate of the change in the total number of crash injuries and fatalities as a result of a change in mass distribution within the vehicle fleet.

---

[1] For examples, see Mountain, et al., 1998; Kumara and Chin, 2005; Miaou, et al., 2005; Kim and Washington, 2006; Lord and Mannering, 2010.

The other important issue which has not been addressed properly in the literature is the isolated effects of vehicle mass and size. As discussed in Section 2.3.2, the key question is whether there is any effect of vehicle mass or size beyond and above that of mass ratio in two-car crashes. This is in particular important because there is the potential to reduce vehicle mass while maintaining its size through various mass-reduction technologies (Wenzel and Ross, 2001).

One important issue regarding the studies that estimate fatality and injury risk is the definition of risk, and in particular, the choice of denominator (Wenzel and Ross, 2001). It was discussed in Section 2.3.1.2 that the ideal measure of driver injury risk for a subject vehicle is the absolute risk defined as the proportion of driver injuries in the subject vehicle when involved in two-car crashes irrespective of driver injury outcome in the colliding vehicles. As discussed, this measure cannot be directly calculated from the crash data which include no information on non-injury crashes. Instead, many studies have focused on relative measures of injury risk which can be directly calculated from the injury crash data (crashes in which there is at least one driver injury) (Broughton, 1996c).

As a result of the knowledge gaps and the uncertainties in the underlying relationships (summarised above), there are differences and sometimes conflicts amongst the results of the studies that have investigated the issue of potential interaction between fuel economy and safety performance in vehicle design within the fleet. Such gaps could limit the creditability of research findings on the existence of any trade-off. These gaps will be addressed in detail in the following chapters and a methodology will be introduced to investigate the partial effects of a given change in mass distribution within the fleet on each of fleet fuel economy and crash injuries and fatalities based on the estimated underlying relationships.

Table 2.2: Summary of the most relevant reviewed studies

Subject	Study	Main analysis method / major assumptions	Data source	Main findings	Major shortcomings
Vehicle mass and fuel consumption	Burgess and Choi (2003)	Parametric study of vehicle energy demand using computer simulation	On-road measurements	Total energy demand of vehicles varies almost proportionally with changes in vehicle mass.	The relationship between energy demand and fuel consumption is not discussed; hence, the effect of mass on fuel consumption is unclear.
	Biggs and Akcelik (1987)	Assuming an energy-related instantaneous model for fuel consumption rate	Collected on-road second-by-second data	A 10% increase in mass increases total fuel consumption by 3.2- 4.1%.	The effect of engine size on fuel consumption is not controlled properly.
	Redsell et al. (1993)	Fitting the fuel consumption data to an expression derived theoretically based on physical formulae	Measurements from designed experiments	Fuel consumption is related proportionally to the cube root of vehicle mass.	The effects of mass and engine size are not isolated properly because of the correlation between mass and engine size and lack of sufficient variation in the data.
	Van den Brink and Van Wee (2001)	Regression analysis Linear relationship between fuel consumption rate and vehicle factors	Fuel consumption measurements under European test cycles	100 kg of extra mass would lead to 7-8% increase in fuel consumption.	The effect of engine size is not controlled; hence, the estimates contain the effects of engine size as well.
	Zervas (2006)	Regression analysis Linear relationship between vehicle mass and CO_2 emission rate	CO_2 emission measurements under European test cycles	CO_2 emission increases as vehicle mass increases.	The effects of other vehicle design factors (e.g. engine characteristics) are not controlled.
	Fontaras and Samaras (2010)	Simulation of CO_2 emission rates for a few car models under defined driving cycles Regression analysis Linear relationship between the change in mass and the simulated change in CO_2	Computer simulations	A 10% reduction in vehicle mass would lead to a reduction of 2.7-3.6% in CO_2 emissions.	Too few car models (6) with limited ranges of variation in design characteristics were used.

Table 2.2 (continued)

Vehicle mass and driver injury risk	Evans and Frick (1992), Evans and Frick (1993), Evans (1994), Evans and Frick (1994), Evans (2001), Evans (2004)	Regression analysis Linear regression of aggregate two-car crash data: $\left(log(R) = u \times log(\mu)\right)$	US two-car crash data where at least one of the drivers is killed from Fatality Analysis Reporting System (FARS) (1975-1998)	In a crash between two cars of different masses, the fatality risk ratio (R) of the lighter to heavier car increases as a power function of mass ratio (μ) of the heavier to the lighter car ($R = \mu^u$) where u is between 2.70 to 3.80 depending on crash characteristics.	The underlying assumed relationship between injury risk and velocity change suffers from a major structural problem. The methodology lacks flexibility in controlling for or estimating the effects of other contributing factors. The estimated effects are unlikely to reflect the partial effects of mass.
	Wood (1997)	Theoretical calculations based on fundamental relationships of Newtonian mechanics	Various crash data from the literature	In two-car crashes, the ratio of injury severity increases as a power function of mass ratio	The relationship between injury severity and injury risk (the probability of injury in the event of the crash) is not discussed. The measure of relative risk used in the theoretical calculations is not necessarily equivalent to the one used in the literature based on field data and hence, not comparable.
	Wood et al. (2007)	Fitting a power function between injury risk and velocity change using Monte Carlo simulation Logistic regression models to estimate injury risk as a function of velocity change	Crash data from US National Accident Sampling System (NASS) (1982-1991)	In two-car crashes, driver injury risk increases by velocity change where there is zero probability of survival for belted and unbelted drivers for velocity changes of 135-140 km/h and 145-150 km/h, respectively.	The data included too few observations, especially for high-speed crashes, to result in statistically reliable estimates. The effect of mass on injury risk is not known.
	Richards and Cuerden (2009)	Regression analysis Logistic regression models to estimate driver injury risk as a function of velocity change	UK crash data based on Co-operative Crash Injury Study (CCIS)	In two-car crashes, driver injury risk increases by velocity change.	The confidence intervals of the estimations are wide due to the small number of observations. The effect of mass on injury risk is not known.

Table 2.2 (continued)

Vehicle mass and driver injury risk	Broughton (1996a, 1996b)	Regression analysis A linear relationship between relative driver injury risk and mass of makes and models Relative risk defined as the probability of driver injury when involved in a two-car injury accident	Police reported two-car crash data in Great Britain (1989-1992)	Driver injury risk falls steadily with increasing mass and mass can explain a high proportion of variation in the casualty data. The distribution of the risk of driver injury when two cars collide depends principally on the difference in mass.	The measure of injury risk used is not the ideal measure as it is not independent of risk of injury in the colliding car. The estimated relationship does not provide any information on the aggressive effect of vehicle mass. The effects of mass and size are not isolated.
	Wenzel and Ross (2005)	Descriptive statistical analysis of crash data Risk defined as driver fatalities per year per million registered vehicles for a given car model and all types of crashes	US fatality crash data from FARS (1997-2001)	Mass alone is only a modestly effective predictor of risk.	The safety measure used is a measure of both primary safety (crash involvement) and secondary safety (injury risk); therefore, it cannot be used to reflect the effects of mass fully.
Vehicle size and driver injury risk	Ross and Wenzel (2001)	Theoretical calculations based on basic physical formulae	-	Injury outcome in two-car crashes is influenced by the velocity change. Increase in crush distance could increase the time, and hence, decrease the risk of injury by decreasing the acceleration experienced.	Separate effects of mass and size on injury risk are not quantified.
	Evans (2001)	Linear regression of aggregate crash data Theoretical calculations	US two-car crash data between cars of the same mass from FARS (1975-1998)	For a given mass, drivers in the larger vehicles have the lower risk of injury and fatality than those in the smaller vehicles. A model of absolute driver injury risk as a function of vehicle size and mass is derived.	The introduced model of absolute risk is derived theoretically and has not been tested with any real crash data.

Table 2.2 (continued)

Vehicle size and driver injury risk	Van Auken and Zellner (2005)	Regression analysis A three-stage weighted logistic regression	US crash data from FARS (1995-1999) Accident data files from 7 US states (1995-1999)	Vehicle mass reduction tends to decrease fatalities while vehicle wheelbase and track reduction tends to increase fatalities.	The unexpected estimated significant effects of mass and size that are in opposite directions for some types of crashes in this study makes it difficult to draw a general conclusion on the separate effects of mass and size based on their results.
	Broughton (1999)	Regression analysis Logistic regression models of relative injury risk Relative risk defined as the number of driver fatalities where at least one of the drivers is injured	Police reported two-car crash data in Great Britain (1991-1994)	The risk of driver fatality increases with car mass in two car crashes when all other factors including mass ratio are unchanged.	The measure of injury risk in each vehicle is not independent of injury outcome in the colliding vehicle. Collinearity is likely in the estimated models due to a high correlation of absolute mass and mass ratio in the data
Make and model and driver injury risk	Broughton (2007)	Regression analysis of casualty rates by type of car Casualty rates defined as the ratio of number of driver fatalities to the number of registered vehicles	Police reported two-car crash data in Great Britain (2001-2005)	The mean risk of fatality for the driver of the smallest types of cars (minis and superminis) is 4 times the risk for the largest type (4x4s and people carriers).	The effect of vehicle mass is not controlled.
	Broughton (1996c), DfT (2006a)	Regression analysis Logistic regression models of relative driver injury risk and make and model Relative risk defined as the number of driver injuries where at least one of the drivers is injured	Police reported two-car crash data in Great Britain (1989-1992 and1990-1994)	New secondary safety and aggressivity indices for each make and model involved in two-car crashes are introduced when the effects of driver age, driver gender, speed limit, and point of impact is controlled.	The effect of vehicle mass is not controlled. The measure of injury risk in each vehicle is not independent of injury outcome in the colliding vehicle.

Table 2.2 (continued)

Make and model and driver injury risk	Newstead et al. (2000)	Regression analysis Logistic regression models of crash involvement probability and driver injury probability	Police reported two-car crash data from 3 Australian states (1987-1998)	Secondary safety and aggressivity indices for makes and models are introduced as the product of two probabilities: crash involvement and driver injury given a crash.	Estimation of indices is solely dependent on the availability of non-injury crash data. The effect of vehicle mass is not controlled.
Safety and environmental effects of changes in vehicles' mass	Ross and Wenzel (2001)	Estimating the effects of a safety-fuel economy scenario based on the estimated effects of mass and size available in the literature	US crash fatality data from FARS (1999)	Narrowing the range of masses in fleet while maintaining or increasing vehicle dimensions results in a reduction of about 2000 fatalities (based on 1999 US fatality rates) as well as an improvement in fleet fuel economy.	Changes in vehicle ownership and vehicle usage pattern are not considered. Changes in vehicles distribution by engine size, fuel type, and other design features that could influence fuel consumption are not discussed.
	Broughton (1999)	Estimating the effects of a uniform fleet downsizing scenario based on the estimated effects of mass and size using logistic regression models The number of accidents were assumed to remain constant	Police reported two-car crash data in Great Britain (1991-1994)	A uniform reduction in mass and size of vehicles would lead to fewer injuries and fatalities.	Existence of collinearity is likely in the estimated models used to estimate the effects due to a high correlation of absolute mass and mass ratio in the data.
	Buzeman et al. (1998)	Estimating the effects of a number of mass distribution scenarios using the relationship between risk of injury and velocity change estimated by Evans (1994) The probability of crash involvement assumed to be a function of vehicle mass	Swedish two-car crashes (1995)	A uniform mass reduction of 20% in Swedish fleet increases fatalities by 5.4% while a reduction of mass range by 20% reduces fatalities by 3%.	Methodology is dependent on the availability of information on the impact speed of two-car crashes. There is no strong evidence of a direct effect of vehicle mass on risk of crash involvement in the literature. The effect of vehicle size is not discussed.

Table 2.2 (continued)

Safety and environmental effects of changes in vehicles’ mass	Zachariadis (2008)	Regression analysis Linear relationships between safety performance, mass, rate of CO_2 emissions, and vehicle design features	A collective source of independent data on vehicle mass and fuel consumption Euro NCAP crash test performance records (model years 2000-2007)	Additional safety performance increases mass slightly and does not necessarily increase fuel consumption when other design features are taken into account.	Euro NCAP safety crash test ratings are not comparable between different groups of vehicles. The safety ratings are not all measured on the same scale in the modelling process. The greater effects of differences between vehicle classes are not represented in the models.
	Crandall and Graham (1989)	Observational study Regression analysis of aggregate time-series data	US time-series crash fatality and fuel consumption data (1947-1981)	Additional fatalities occurred as a result of Corporate Average Fuel Economy (CAFE) standards	The effect of mass is not controlled properly. The methodology used has certain disadvantages.
	Noland (2004, 2005)	Observational study Regression analysis Fixed-effect negative binomial models to analyse the effects of average fuel economy of vehicles on traffic-related fatalities	US state-level time-series data on traffic fatalities and average fuel economy Country-level time-series data on traffic fatalities and average fuel economy from 13 countries (1970-1996)	Improvements in fuel efficiency in the US were associated with increased fatalities in the 1970s Changes in vehicle efficiency in Europe during 1970s and 1980s are not associated with changes in traffic fatalities	The effects of vehicle mass are not controlled; hence, it is not clear to what extent the changes in average fuel consumption are related to the changes in vehicle mass.
	Ahmad and Greene (2005)	Observational study Regression analysis Linear relationship between the average fuel economy of vehicles and traffic fatalities	US time-series crash fatality and fuel consumption data (1966-2002)	Reduced fuel consumption is linked to fewer fatalities	The effects of vehicle mass are not controlled; hence, it is not clear to what extent the changes in average fuel consumption are related to the changes in vehicle mass

CHAPTER 3. GENERAL METHODOLOGY AND STUDY DATA

The previous chapter presented a thorough review of the key literature relevant to the issue of interaction between environmental and safety performance in vehicle design. This revealed that the most promising approach to address this issue is to investigate the relationship between vehicle mass and each of fuel consumption and safety performance of vehicles within a vehicle fleet, separately, using cross-sectional data that belongs to a specific period of time. Therefore, the main objectives of this chapter are to outline the general study methodology, to introduce the data sources, and to explain the process of development of the final study datasets. The first section (3.1) summarises the methodology and discusses the data requirements. Possible primary and secondary data sources[1], their availability, advantages, and disadvantages are discussed in the second section (3.2). The process of data quality checks and development of the final study datasets are explained in the third section (3.3).

3.1. Methodology outline and data requirements

As discussed in detail in Chapter 2, the majority of studies that have addressed the issue of conflict between fuel economy and safety in the vehicle fleet caused by mass are empirical studies that have used aggregate time-series data. In such studies, the characteristics of the vehicle fleet and the time period to which the data belongs could influence the results. This partly contributes to the inconsistencies in the results of these studies. Besides, since the effect of vehicle mass is not controlled for, it is not clear to what extent the changes in fuel consumption are related to the changes in vehicle mass rather than other contributing factors that tend to change over time.

This study uses a different approach; it examines the issue of potential interaction between environmental and safety performance in vehicle design, within a vehicle fleet, caused by vehicle mass. Addressing this issue is the key point in understanding whether there is any conflict in safety and environmental goals as a whole in a vehicle fleet with given characteristics. Partial effects of vehicle mass on fuel consumption and secondary safety performance (safety performance in the event of a crash) of vehicles within the

---

[1] Primary data is observed or collected directly from the event or experience while secondary data is collected from the external sources that collect, process, or analyse the primary data.

fleet are investigated using disaggregate cross-sectional analysis of mass within a national vehicle fleet at a specific period of time. The estimated effects are then used to investigate likely safety and environmental consequences of changes in vehicle design, in particular mass, within the vehicle fleet. The scope of this study is the passenger car fleet in Great Britain.

It was discussed in Chapter 2 that several vehicle design factors could affect vehicle fuel consumption. These include mass, engine characteristics, fuel type, and year of manufacture. Apart from vehicle design, other factors including vehicle condition, environmental condition, driver, and road factors could also influence vehicle fuel consumption. As discussed in detail in Chapter 2, current estimates of the effect of mass on fuel consumption available in the literature do not reflect partial effects of mass where the effect is the result of a change in mass, holding all other factors constant. This is partly due to the methods applied, and partly due to insufficient data used. If other contributing factors are not fully controlled when estimating the effects of mass, the estimates may contain effects of other factors as well. In this study, statistical modelling techniques are used to estimate partial effects of mass on fuel consumption where the effects of other contributing factors are controlled. The ideal data required for such an analysis is a cross-sectional dataset of fuel consumption that includes information on various design features of vehicles (including vehicle mass) that contribute to vehicle fuel consumption. Such a dataset should ideally include a wide range of vehicle design and fuel consumption rates to reflect variation of fuel consumption rate by vehicle design variables. Besides, the data collection method should ensure that the effects of other contributing factors on fuel consumption rate are controlled as much as possible.

Vehicle mass is linked to the risk of injury and fatality of vehicle occupants during a crash as well. The literature review in Chapter 2 revealed that occupants of lighter cars in fleet have a greater risk of injury and fatality in crashes compared to the occupants of heavier cars. This is mainly due to a greater velocity change that the lighter vehicles undergo when involved in crashes with the other vehicles as suggested by Equation 1.1. However, Equation 1.1 also shows that the velocity change of a vehicle in two-car crashes is influenced not only by mass of that vehicle, but also by mass of the other vehicle in the crash. Therefore, a detailed analysis of two-car crashes is required to investigate the partial protective and aggressive effects of mass. Although several studies have analysed two-car crashes to investigate the effects of mass, there are important flaws in the methodologies and knowledge gaps in the estimated effects (see Chapter 2, Section 2.3.1.2 for details). In this study, a novel methodology is introduced to estimate protective and aggressive effects of vehicle mass when the effects of vehicle size and other factors that can contribute to injury risk are controlled. The ideal data required for this analysis is a sample of two-car crashes where information on mass and size of both vehicles involved in the crashes is available. This sample should ideally include crashes between pairs of vehicles with wide ranges of mass and size.

Having estimated partial effects of vehicle mass on fuel consumption and secondary safety performance (both protective and aggressive effects of mass), the results are used to investigate the partial effects of different hypothetical scenarios of fleet mass distribution on overall fleet fuel consumption and the total number of casualties as relative changes with respect to a base fleet with a base mass distribution. Therefore, detailed vehicle registration data for the base fleet is required where mass and other design features of vehicles are known. Based on the estimated outcome of these mass distribution scenarios, different policy options are discussed, recommendations are made, and direction of further work in this area is suggested.

3.2. Possible data sources

The following sections introduce possible sources of data (primary and secondary) and discuss availability, advantages, and disadvantages associated with each of them with regards to the study requirements. Having compared various aspects of possible sources for each type of data, the source that best suits the study requirements is identified for further quality assessments (checking the reliability and accuracy of data) and potential use as the data source for the main analyses.

3.2.1. Two-car crash data

As mentioned earlier, ideally a sample of two-car crashes is required where information on mass and size of both vehicles involved in the crashes is available. A sequential search was conducted to select from different available sources the data that best meet these requirements. There are three possible sources of accident data in Great Britain: data from National Health Service (NHS), insurance company data, and police-reported accident records.

3.2.1.1. NHS health data

One potential source of accident data is the NHS injury data which provides information on the injuries sustained in traffic accidents. There are various databases within NHS including Ambulance Service data, Accident and Emergency (A&E) department data, hospital inpatient data, and specialised health databases (e.g. trauma audit and research network data, General Practitioners' data). These are believed to vary substantially within the UK in terms of both availability and specific details due to lack of a unified data recording system (Ward, et al., 2006).

The main advantage of this data, when compared to Police records, is that they usually include more details about the nature of the casualty and hence provide a better assessment of the injury severity. However, the main issue regarding this data is lack of sufficient information on the accident itself, and more importantly, on the details of the vehicles involved in the accident. Besides, health data includes no information on slight injuries and it only includes those fatalities that died in the hospital; these are only around 20% of the total number of fatalities in road accidents in Great Britain (DfT, 2009). It is also difficult to link the hospital injury data to other datasets such as Police records mainly because hospital records do not have Ordnance Survey grid references as do the Police records. Furthermore, the English language description of the location within the database is not precise (Ward, et al., 2002). These major shortcomings together with the difficulty in collecting all the NHS data from different units and hospitals across Great Britain and processing them led to the rejection of NHS injury data as an appropriate source of crash data for use in the statistical modelling of this study.

3.2.1.2. Insurance company data

Another potential source of accident data is insurance claims data collected and held by various insurance companies. These are based on accident injury claims of the occupants as well as property damage claims of the owners of insured cars. The main advantage of insurance company data compared to the alternative sources is the availability of information on damage-only accidents (accidents in which there are no casualties; these are not usually included in the data from police reports). Investigations were made to assess the availability and suitability of insurance data for this study.

These included approaching insurance companies to discuss the possibility of obtaining data from them as well as a thorough review of the literature.

Although there are studies in the literature that have investigated the secondary safety of vehicles using data from insurance companies (e.g. Gustafsson, et. al., 1989; Newstead et. al., 2000), none of them has used a recent UK based insurance data. Broughton (1996c) claimed that in Great Britain, accident claim data from insurance companies are completed less consistently than police accident reports and hence, it does not form a complete database for a detailed analysis of secondary safety.

Correspondences with a few insurance companies[1] revealed the substantial difficulty that could arise in providing data because of confidentiality. Even if negotiations were successful with a certain number of companies to provide the data, the resulting database would be unlikely to be a complete dataset representing accidents in a national scale with a sufficient number of records to perform a reliable analysis of two-car crashes. Therefore, investigations were continued for an alternative source of data which has the desired properties to be used in this research.

3.2.1.3. Police reported accident data

The database compiled from the police reports of road accidents that result in injury or fatality in Great Britain is called STATS19. This database has been the main data source for safety research as well as the basis for setting and monitoring casualty reduction targets in Great Britain. STATS19 includes accident, vehicle, and casualty datasets for all road accidents involving personal injury (slight or serious) and death which are notified to the Police within 30 days of occurrence (DfT, 2004). Basic STATS19 data is publicly available for use as in annual basis[2].

Although the STATS19 vehicle record does not include make and model, mass, dimension, and other design information for the vehicles involved in crashes, it includes the Vehicle Registration Mark (VRM) of the vehicles. This provides the opportunity to enhance the STATS19 vehicle details with the information available from the UK

---

[1] Including Aviva (former Norwich Union) and Automobile Association

[2] Department for Transport, Local Government and the Regions, *Road Accident Data, 2000-2006* [computer file]. Colchester, Essex: UK Data Archive [distributor], http://www.data-archive.ac.uk.

Driver and Vehicle Licensing Agency (DVLA) such as make and model, model variant, and some other technical design information. This unique opportunity as well as availability, popularity, and completeness of STATS19 data compared to the other alternative sources suggested this database as the preferred source of data to perform further investigations to verify its quality and accuracy; these will be explained later in this chapter.

3.2.2. Vehicle fuel consumption data

As mentioned in Section 3.1, the ideal data required for the analysis of vehicle mass and fuel consumption is a cross-sectional dataset of vehicle fuel consumption rates that includes information on a wide range of various design features for vehicles (including vehicle mass) that contribute to fuel consumption rate. Besides, the data collection method should ensure that the effects of other contributing factors on fuel consumption rate are controlled as much as possible. There are generally two sources of fuel consumption data: experimental data and laboratory data.

3.2.2.1. Experimental data

One potential source of fuel consumption data is the on-board fuel consumption measurements collected during actual on-the-road designed experiments. The data that are already available in the literature[1] are based on the experiments that are designed to address specific objectives other than those of this study and they generally relate to one or a limited number of car models. Therefore, they lack sufficient variety in terms of vehicle makes and models and design features such as mass and engine size that could influence vehicle fuel consumption in different ways. On the other hand, design and implementation of an experiment specifically for the purpose of this study, which would involve a number of car models with various design features driven under specific driving cycles, was found to be costly and beyond the available resources for this research.

---

[1] Examples include Redsell et al., 1993; De Vlieger, 1997; Leung and Williams, 2000; North et. al., 2006.

3.2.2.2. Laboratory data

Officially certified fuel consumption rates for specific makes and models are measured under controlled driving cycles, vehicle conditions, and ambient temperature (VCA, 2007a). A fuel consumption and emission databank is available from the UK Vehicle Certification Agency (VCA), which is one of the executive agencies of DfT[1], for the new car models that have entered the market in the United Kingdom since 2000. The data is available in the form of two fuel consumption rates (urban and extra-urban) for different variants of each make and model stratified by the year on the market, fuel type, engine size, transmission type and the Euro emission standard that the vehicle satisfies. The main advantages of VCA fuel consumption data compared to different available experimental data are the large variability in makes and models and their design features and consistency in the way fuel consumption is measured for these make and model variants. The fact that measurements are conducted under controlled driving cycles, vehicle condition, and ambient temperature ensures that the effects of the majority of factors that contribute to vehicle fuel consumption besides the vehicle design are controlled. These desirable properties suggested that VCA car fuel consumption data can be used as the primary source of data for this study. Further investigations were made to verify the quality and reliability of VCA data; these will be explained later in this chapter.

3.2.3. Vehicle mass and dimension data

The general issue associated with available sources of data for two-car crashes and fuel consumption is lack of information on vehicle mass and dimensions within the data. However, availability of detailed design information for different variants of makes and models in both STATS19 two-car crash data and VCA fuel consumption data provides the opportunity to extract mass and dimension data for these makes and models from other external sources. One potential option to collect data on vehicle technical information including mass and dimensions is to obtain the data directly from individual manufacturers. Alternatively, such data can be obtained from a single secondary source

---

[1] Source: DfT webpage http://www.dft.gov.uk/about/howthedftworks/aboutthedftexecutiveagencies, Accessed December 2010.

that includes technical design figures for various model variants of different manufacturers in a unified and consistent format.

3.2.3.1. Individual manufacturers’ technical data

The most reliable source of information on vehicle mass and dimension is vehicle manufacturers’ brochures or tables. These usually include detailed technical data for various model variants. Investigations revealed that such data is available in the manufacturers’ official web pages only for the most recently produced models. Approaching a few manufacturers and requesting the data for all produced model variants in the past few years suggested that this would be a substantially difficult and time-consuming task. The difficulties faced included their unwillingness to provide the data, the incompleteness of the provided data (e.g. not covering all models or all the production years requested), and offering to sell the data or claiming the cost of data processing. These limitations led to a search for an alternative and secondary source which could provide the official manufacturers’ technical data for all the produced models in various years in a consistent and appropriate format.

3.2.3.2. Secondary sources of manufacturers’ technical data

After a thorough investigation, the online edition of CAR magazine[1], which is the oldest monthly motoring magazine in the UK being launched first in 1962, was identified as a potential source of vehicle design data for all makes and models. CAR holds a web-based databank of vehicle technical data for various model variants of the majority of manufacturers. CAR claims these data are official figures coming directly from vehicle manufacturers. The databank is updated from time to time with new information added any time a new model is produced by a manufacturer. Therefore, the biggest advantage of CAR databank, apart from its availability, is the fact that it has a unified and consistent format for all makes and models and it includes data for all model years produced (as early as 1980s). These favourable properties suggested that CAR magazine’s web-based databank of vehicle technical data can be used potentially as the primary source of vehicle design data in this study to assign vehicle mass and dimension figures to the makes and models in both fuel consumption and two-car crash

---

[1] http://data.carmagazine.co.uk/cars/specs/

data. However, the quality and accuracy of CAR magazine's data needed to be examined carefully to confirm it as the final source of design data for this research; this will be explained in the next section of this chapter.

3.2.4. Vehicle registration data

It was mentioned in Section 3.1 that in order to estimate safety and environmental consequences of changes in vehicle design in a fleet (in particular, fleet mass distribution), ideally, a detailed dataset of all the vehicles driven on the roads is required; this is not normally available. However, a disaggregate cross-sectional dataset that includes data on vehicle registration by make and model and other design features could be used instead as a reasonable proxy. Detailed vehicle registration data is available; however, it is not in the public realm. Correspondence with Driver and Vehicle Licensing Agency, which is one of the executive agencies of DfT[1] that maintains a record of all vehicles registered in the UK, revealed that non-personal anonymised vehicle registration data extracted from the DVLA vehicle register is available through certain commercial selling agents.

3.3. Study data

As discussed in the previous section, a thorough investigation of all possible data sources and their availability, advantages, and disadvantages led to selecting police reported accident data (STATS19), VCA fuel consumption data, CAR magazine's vehicle technical data, and DVLA vehicle registration data as the potential sources of data for this study. Table 3.1 summarises the sources of data that were investigated, and compares their advantages and disadvantages. The selected data sources were studied in greater detail to verify their quality and accuracy before being used to develop the final study datasets. The following sections explain these steps in detail.

---

[1] Source: DfT webpage http://www.dft.gov.uk/about/howthedftworks/aboutthedftexecutiveagencies, Accessed December 2010.

Table 3.1: Comparison of different possible data sources

Type of Data	Data source	Main advantages	Main disadvantages
Two-car crash data	NHS	- Including details about the nature of the casualty	- Lack of information on the details of the vehicles involved in the accidents - No data on slight injuries - Only including those fatalities that died in the hospital - Difficulty in collecting all the NHS data from different units and hospitals across Great Britain
	Insurance companies	- Availability of information on damage-only accidents	- Lack of consistency between different datasets - Substantial difficulty in obtaining the data because of confidentiality
	Police reports (STATS19)	- Possibility of adding data on technical details of vehicles to the accident records - Availability - Being widely used within the safety research - Completeness	- Lack of information on non-injury crashes - Lack of information on speed of impact - Lack of data on vehicle mass
Vehicle fuel consumption data	Experimental data	- Reflecting actual on-the-road fuel consumption and vehicle performance	- Lack of sufficient variety in vehicle makes and models and design features in the data that are currently available - Lack of sufficient resources to design and implement an experiment specifically for the purpose of this study
	Laboratory data (VCA)	- Large variability in makes and models and their design features - Consistency in the way fuel consumption is measured for different model variants - Availability of data for urban and extra-urban driving cycles	- Lack of data on vehicle mass
Vehicle mass and dimension data	Individual manufacturers	- The most reliable source of information on vehicle mass and dimension	- Publicly available only for the most recently produced models - Substantial difficulties in obtaining the data from manufacturers for all produced makes and models in the past few years
	CAR magazine	- Availability - Completeness (data for all model years produced) - A unified and consistent format for all makes and models	
Vehicle registration data	DVLA	- Availability - Completeness	- Lack of data on vehicle mass

3.3.1. Quality of data

3.3.1.1. STATS19 accident database

As was mentioned earlier, the database compiled from the police reports of road accidents that result in injury in Great Britain is called STATS19. STATS19 data is analysed nationally based on a great variety of characteristics it contains and the results are used extensively for research work and the improvement of road safety in relation to roads, road users, vehicles and traffic movement. The data also form the basis for annual statistics on road accidents and casualties published by DfT, the Scottish Executive (SE) and the National Assembly for Wales (NAfW) (DfT, 2004).

A set of information has to be collected by a Police Officer when an injury road accident is reported to them. The information is collected and processed in a designed agreed format, details of which are explained in a separate document published by DfT referred to as STATS20 (DfT, 2004). The collected data is checked locally and validated by the Police or local councils before transmission to the DfT. A number of validity checks are then applied to the data by DfT to ensure the data is consistent (Scottish Executive, 2005). These includes, for each variable, checks on character positions, variable format, and acceptable range, as well as various consistency checks (for example, accident reference number must be unique within the dataset, or if in the casualty records, the type of casualty is driver and sex of casualty is female, sex of driver in the vehicle record must be coded as female). The validation system will identify errors and missing data. Depending on the type of error, the data is either corrected manually or sent back to the originator for correction or confirmation. The details of the validity checks and the error procedures that are carried out on STATS19 data are given in a separate document by DfT, referred to as STATS21 (DfT, 2004).

One generally-accepted[1] issue regarding the STATS19 data is the problem of under-reporting; that is, it is claimed that while very few, if any, fatal accidents are not known to the police, a large number of less serious accidents are not reported. The police do not attend all accidents and there is no legal requirement to report accidents if details are exchanged by those involved at the scene. By matching STATS19 data with hospital A&E department data in London, Ward et al. (2002) roughly estimated an overall

---

[1] See, for example, Ward et al., 2002; Ward et al., 2006; DfT, 2009.

reporting rate of about 70%. Ward et al. (2006) state that the issue of under-reporting influences use of STATS19 data as the single dataset to investigate casualty trends and use of multiple datasets (e.g. STATS19 and health data) can provide a better platform to monitor trends in road traffic casualties.

However, the issue of under-reporting is unlikely to introduce any bias to the analysis of two-car crashes in this study. The main analysis in this study is performed on a sample of two-car crashes where at least one of the drivers is Killed or Seriously Injured (KSI); the majority of these types of crashes are believed to have been reported to the Police. Besides, there is no evidence in the literature suggesting that under-reporting in two-car crashes is systematically related to the type of crash or the characteristics of the vehicles or drivers involved in the crash. In an attempt to compare the serious injury data in STATS19 with that from hospital inpatient data, DfT (2009) found that despite the fact that the number of injuries are not directly comparable between the two dataset due to different definitions used, the two dataset show a similar pattern in terms of sex and age group of casualties as well as the time of accident, especially for car occupants.

DfT (2009) states that despite the issue of under-reporting, STATS19 remains the most detailed, complete and reliable single source of information on road casualties covering the whole of Great Britain.

3.3.1.2. VCA fuel consumption database

It was explained in Section 3.2.2 that officially certified fuel consumption rates for specific makes and models that are measured under controlled driving cycles, vehicle conditions, and ambient temperature are held and published by the UK Vehicle Certification Agency. The VCA is an executive agency of the UK Department for Transport and is the designated UK vehicle type approval authority (DfT, 2010). VCA is responsible for the creation and management of the new car fuel consumption and exhaust emission figures. VCA's car fuel database[1] is the official UK source for car fuel consumption and exhaust emission figures (VCA, 2007a).

The fuel consumption testing is carried out either by independent test organisations, or by the manufacturers or importers themselves at their own test facilities. Before the

---

[1] www.vcacarfueldata.org.uk

results are officially recognised, DfT inspects the test laboratories and witnesses some tests being carried out; or checks that the figures have been certified by a European government under the agreed arrangements for mutual recognition of test results (VCA, 2007b).

3.3.1.3. CAR magazine's vehicle design database

As was mentioned in Section 3.2.3, CAR magazine holds a web-based databank of vehicle technical data for various model variants of the majority of manufacturers. There is a separate web page for each car model within a given make, and there is a separate web page for each model variant for a given make and model. Figure 3.1 outlines the structure of these web pages for a given make of car where n represents the number of car models within that make and m_n represents the number of model variants for car model n. As an example, the technical data web page for a variant of BMW 3-series (Saloon 318i 4d) is shown in Figure 3.2.

As mentioned earlier in this chapter, CAR magazine claims that its database includes official figures that have been received directly from manufacturers whenever new models have been produced. In order to verify this claim and examine the accuracy of CAR magazine's data, a recent car model from each manufacturer was chosen at random and mass and dimension data for all variants of the chosen model were downloaded from the tables or brochures available in the manufacture's website. These were cross-compared with the corresponding data for that manufacturer in the CAR magazine's database. This process was performed for the 33 most popular manufacturers in Great Britain that accounted for about 98% of the British passenger car fleet in 2007. The results of this comparison are reflected in Table 3.2.

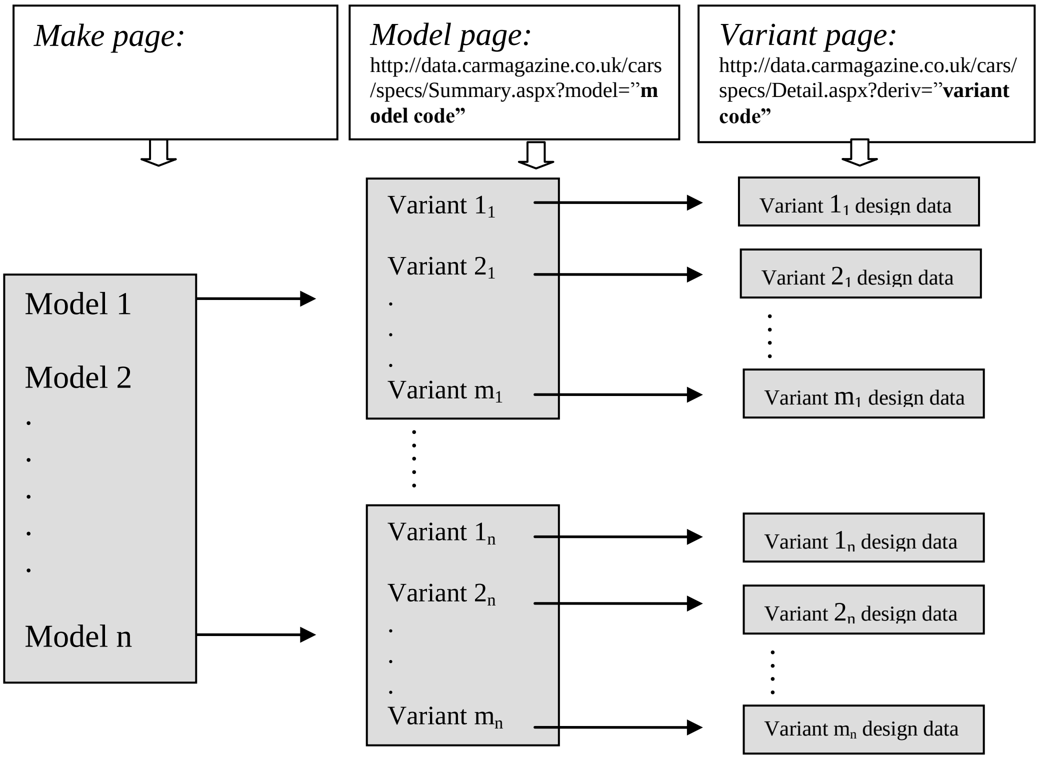
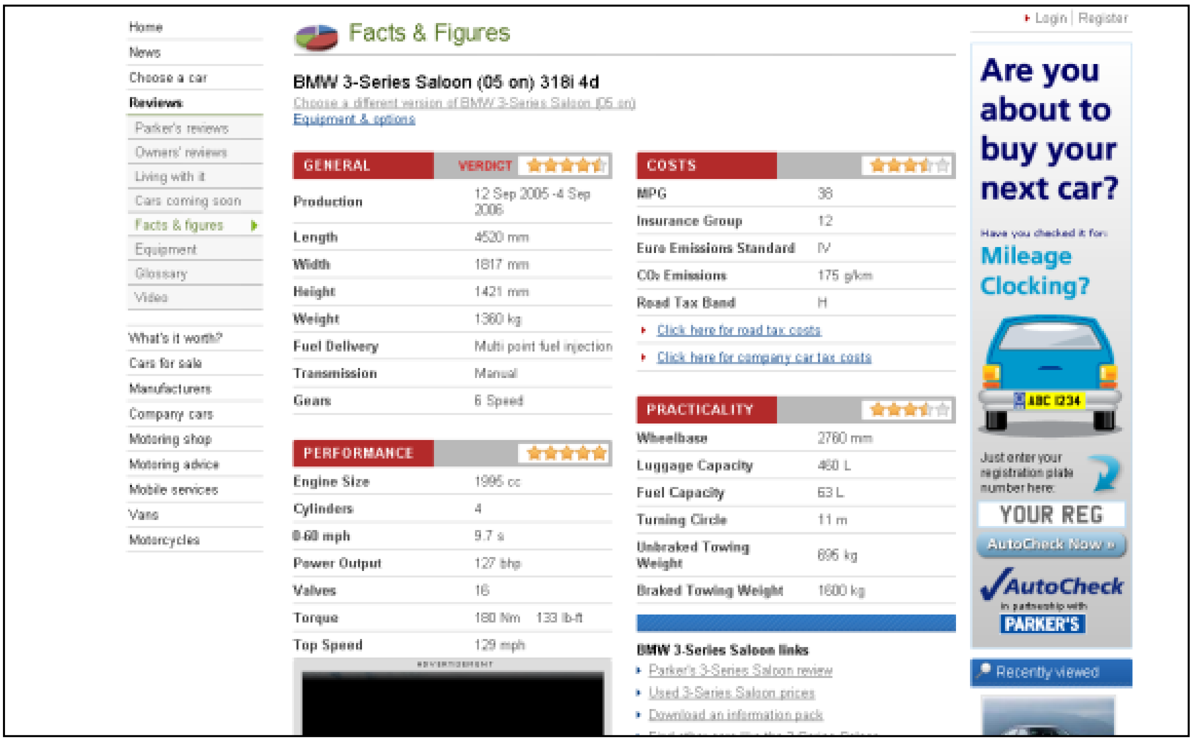


Figure 3.1: Structure of CAR magazine’s web pages of vehicle technical data for a given vehicle make



BMW 3-Series Saloon (05 on) 318i 4d

Choose a different version of BMW 3-Series Saloon (05 on)

Equipment & options

GENERAL	VERDICT
Production	12 Sep 2005 -4 Sep 2006
Length	4520 mm
Width	1817 mm
Height	1421 mm
Weight	1360 kg
Fuel Delivery	Multi point fuel injection
Transmission	Manual
Gears	6 Speed

PERFORMANCE	
Engine Size	1995 cc
Cylinders	4
0-60 mph	9.7 s
Power Output	127 bhp
Valves	16
Torque	180 Nm 133 lb-ft
Top Speed	129 mph

COSTS	
MPG	38
Insurance Group	12
Euro Emissions Standard	IV
C0₂ Emissions	175 g/km
Road Tax Band	H

- Click here for road tax costs
- Click here for company car tax costs

PRACTICALITY	
Wheelbase	2760 mm
Luggage Capacity	460 L
Fuel Capacity	63 L
Turning Circle	11 m
Unbraked Towing Weight	695 kg
Braked Towing Weight	1600 kg

BMW 3-Series Saloon links

- Parker's 3-Series Saloon review
- Used 3-Series Saloon prices
- Download an information pack

Figure 3.2: CAR magazine’s web page of technical data for a variant of BMW 3-series

Table 3.2: Comparison of CAR magazine's data with manufacturers' data

Manufacturer	Fleet registrations in 2007[1]			Cross-comparison with CAR's database			
	Number	Percentage	Cumulative percentage	The selected model	Number of variants	Percent matched	
						Mass	Dimensions
FORD	4,821,550	15.8	16.0	Focus Hatchback	52	100	100
VAUXHALL	3,822,811	12.5	28.0	Astra Hatchback	51	100	100
VOLKSWAGEN	2,174,618	7.1	35.0	Passat Saloon	22	100	100
PEUGEOT	2,153,043	7.1	42.0	3010	27	100	100
RENAULT	1,959,447	6.4	49.0	Scenic	46	100	100
ROVER[2]	1,456,008	4.8	54.0	-			
TOYOTA	1,373,030	4.5	58.0	Land Cruiser	6	100	100
NISSAN	1,277,047	4.2	62.0	Micra Hatchback	9	100	100
BMW	1,188,337	3.9	66.0	Z4 Roadster	13	100	100
CITROEN	1,141,240	3.7	70.0	C3	16	100	100
HONDA	983,427	3.2	73.0	Accord Tourer	26	100	100
MERCEDES-BENZ	945,552	3.1	76.0	E-Class Saloon	37	100	100
FIAT	881,838	2.9	79.0	Bravo	17	100	100
AUDI	753,679	2.5	82.0	A4 Saloon	150	100	100
LAND ROVER	697,531	2.3	84.0	Freelander	43	100	100
VOLVO	571,028	1.9	86.0	S60	19	100	100
MAZDA	461,726	1.5	87.0	5	5	100	100
MITSUBISHI	341,130	1.1	88.0	Lancer Sportback	11	100	100
SKODA	340,802	1.1	90.0	Fabia Hatchback	25	100	100
SUZUKI	327,658	1.1	91.0	Swift Hatchback	8	100	100
HYUNDAI	319,498	1.0	92.0	i30 Hatchback	22	100	100
JAGUAR	294,263	1.0	93.0	XF Saloon	10	100	100
SEAT	283,332	0.9	94.0	Alhambra	13	100	100
SAAB	250,095	0.8	94.0	9-5 Saloon	25	100	100
MINI	248,052	0.8	95.0	Cooper	40	100	100
KIA	202,820	0.7	96.0	Rio	28	100	100
DAEWOO[3]	164,644	0.5	96.0	-			
SUBARU	122,868	0.4	97.0	Legacy Tourer	5	100	100
ALFA ROMEO	117,746	0.4	97.0	MiTo	24	100	100
LEXUS	104,995	0.3	98.0	RX	5	100	100
PORSCHE	104,876	0.3	98.0	Cayenne	6	100	100
CHRYSLER	97,905	0.3	98.0	Grand Voyager	5	100	100
DAIHATSU	87,998	0.3	98.0	Sirion	19	100	100
OTHERS	465,630	1.5	100.0				
Total	30,536,224	100.0			785	100	100

1. Source: vehicle registration data from Driver and Vehicle Licensing Agency (DVLA)
2. Production of the Rover models ceased in 2005 when manufacturer MG Rover Group entered administration.
3. Daewoo, as of 2001, was taken over by GM Group rebranding most of its popular models as Chevrolet

Cross-comparison of mass and dimension data for 785 different model variants shows an exact match for each of these models between the figures from CAR magazine's online database and those published by manufacturers. Whilst this comparison is only possible for the most recent models for which data is publicly available from manufacturers, the results support CAR magazine's claim that the technical figures it holds are manufacturers' official figures and hence, they are accurate.

3.3.1.4. DVLA vehicle registration database

As discussed briefly in Section 3.2.4, vehicle registration data by make and model and other design features is available from DVLA. DVLA is an executive agency of the UK Department for Transport that maintains detailed record of all vehicles registered in the UK (DfT, 2010). It is also responsible for maintaining an up-to-date record of all those who are entitled to drive various types of vehicles as well as contributing to different Government policies.

All new and imported vehicles are legally required to be registered by DVLA if they are to be used on public roads. During this process, the vehicle details (make, model, year of manufacture, engine capacity, etc) as well as current keeper's details are recorded. DVLA has to be notified whenever any details of the vehicle or the keeper are changed. Information on registered vehicles is held by DVLA in a database of vehicle registration. The non-personal anonymised vehicle registration data from DVLA for different quarters of UK vehicle fleet is available for purchase.

3.3.2. Final study datasets

Using the sources of data explained in Section 3.3.1, specific datasets were developed to address the objectives of this research. A fuel consumption dataset was developed that included disaggregate cross-sectional data on different variants of makes and models, their design characteristics, and their fuel consumption rates for the cars that were available in the British market between 2000 and 2007. This dataset will be used to estimate the effects of vehicle design features on the fuel consumption rate. A two-car crash dataset was developed that included crashes that occurred in 2000-2006 period. This dataset will be used for a detailed cross-sectional analysis of mass and driver injury risk. A vehicle design dataset was developed that included technical information for different variants of makes and models. This will be used to assign mass and size data to the vehicles in each of fuel consumption and two-car crash datasets based on the available information on makes and models and their design features. Finally, a dataset of registered vehicles that included vehicle design data was obtained from Driver and Vehicle Licensing Agency reflecting the British passenger car fleet in the last quarter of 2007. This dataset will be used to investigate the safety and environmental consequences of different mass distribution scenarios using the estimated effects of mass. To maintain consistency between different datasets, all data were collected so that they belong to a similar period of vehicle fleet as much as the availability of data permitted. A summary of these datasets together with their source, the period to which they belong, and a brief description of the data is given in Table 3.3. The following sections describe different characteristics of these datasets in more details. The process of data manipulation, data augmentation, and dataset development for each of these are also explained.

Table 3.3: Final study datasets

Type of data	Source	Period	Description
Two-car crashes	Department for Transport (STATS19)	2000-2006	Personal injury road accidents in Great Britain
Vehicle fuel consumption	Vehicle Certification Agency	2000-2007	Official UK fuel consumption and exhaust emissions figures
Vehicle design data	CAR Magazine	1980-2007	Technical specifications of different make and model variants
Vehicle registration	Driver and Vehicle Licensing Agency	Quarter 4 of 2007	Number of vehicles registered in Great Britain by make and model and design factors

3.3.2.1. Fuel consumption dataset

In order to reliably estimate the partial effects of mass on fuel consumption, the effects of all other contributing factors should be controlled. The fuel consumption of a vehicle driven on a road is determined by vehicle design (i.e., fuel type, year of manufacture, engine characteristics, mass, technological features), vehicle condition (i.e., vehicle age, vehicle maintenance, use of air conditioning, engine operating temperature), driving cycle (i.e., driving style, traffic factors, road factors), and ambient temperature. Use of officially certified fuel consumption rates for specific makes and models, which are measured under controlled driving cycles, vehicle condition, and ambient temperature, makes it possible to control for the effect of the majority of factors that can affect vehicle fuel consumption. As explained earlier in this chapter, these data are available from the VCA for new car models on the market in the United Kingdom since 2000 (VCA, 2007a). The data are available in the form of two fuel consumption rates (urban and extra-urban) for different variants of each make and model stratified by the year on the market, fuel type, engine size, transmission type and the Euro emission standard[1] that the vehicle satisfies. However, vehicle mass and size data is not available in this database.

Given that the study objective is to find the relationship between vehicle mass and its fuel consumption, a cross-sectional dataset of makes and models that includes both fuel consumption and vehicle mass, as well as other design factors that contribute to fuel consumption, is required. Availability of detailed design information for different variants of makes and models in the VCA fuel consumption data provides the opportunity to extract mass data for these makes and models from external sources.

As highlighted earlier, vehicle technical data from CAR magazine was used to develop a vehicle design dataset in order to assign mass data to the make and model variants in VCA fuel consumption database. Creating a complete vehicle design dataset by exploring the relevant web pages and saving the required data for all variants of makes and models proved to be a very time consuming process. Therefore, a computer program was written in Visual Basic (VB) and used to download all the information from the web and store them as a single dataset. The inputs to the developed computer program are makes of the cars. For a given car make, the program loads the web page related to that make and correspondingly loads the web pages related to all model variants within that make and downloads the relevant design data for them in a text file. A simplified processing flowchart of this program is shown in Figure 3.3.

---

[1] These are requirements that set specific limits for exhaust emissions of NO_x, HC, CO, and PM for new vehicles sold in European Union (EU) member states.

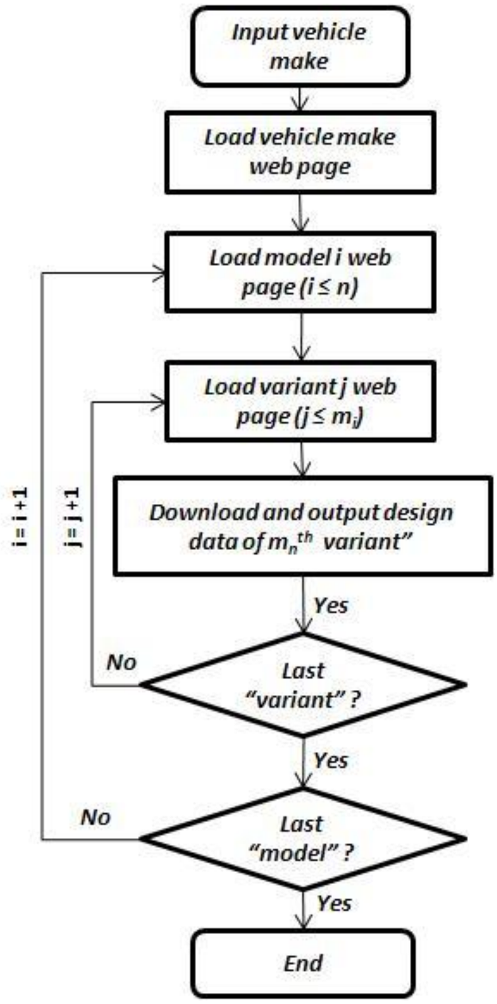


Figure 3.3: Simplified processing flowchart of the computer program that downloads vehicle design data of all model variants of a given vehicle make

This computer program was used to create a dataset of vehicle design that included information on vehicle mass and dimensions as well as other design features for different variants of makes and models. This dataset will be referred to as the design dataset. The total number of model variants for which design data were downloaded was about 27,000. Unfortunately, mass data was not available for all the model variants; it was only available for about 89% of them. The majority of CAR magazine's missing data on mass and dimension relate to the relatively older makes and models. Checking with a few manufacturers revealed that when the data is missing for a model variant in CAR magazine's database, it is also not available in manufacturers' tables or brochures; hence, it cannot be obtained from other sources.

Table 3.4 gives a descriptive summary of makes and models in the developed design dataset. Data was downloaded for 37 vehicle manufacturers. The vehicle makes with minimum and maximum number of models in the datasets were Isuzu (2 models) and

Volkswagen (20 models), respectively. However, Vauxhall has the maximum number of model variants (3031) in the dataset. As Figure 3.2 showed, the other technical data downloaded for each model variant included vehicle dimensions (length, width, height, and wheelbase), transmission type, engine size, and number of doors.

Table 3.4: Descriptive summary of makes and models in the design dataset

ID	Make	No. of models	No. of variants	Percentage of mass availability
1	Ford	14	2029	75
2	Vauxhall	15	3031	89
3	Peugeot	18	1787	88
4	Rover	10	740	78
5	Renault	15	2114	90
6	Volkswagen	20	1542	81
7	Nissan	14	879	92
8	Citroen	17	670	90
9	Toyota	17	880	94
10	Fiat	17	445	95
11	Honda	14	710	96
12	BMW	12	1692	94
13	Volvo	10	1514	99
14	Mercedes	17	2193	76
15	Audi	10	1555	93
16	Mazda	10	292	98
17	Land rover	4	359	85
18	Hyundai	13	238	97
19	Suzuki	11	165	85
20	Skoda	4	377	92
21	Seat	7	359	88
22	Daewoo	5	45	100
23	Mitsubishi	14	345	95
24	Saab	2	745	86
25	MG	6	173	92
26	Proton	6	103	88
27	Jaguar	6	421	99
28	Subaru	4	209	90
29	Kia	11	251	98
30	Daihatsu	6	77	100
31	Alfa Romeo	8	380	98
32	Jeep	5	105	89
33	Mini	5	76	100
34	Lexus	5	224	98
35	Isuzu	2	52	100
36	Porsche	4	113	95
37	Chrysler	7	103	98
Total		365	26,993	89

The developed design dataset was used to assign vehicle mass and dimension data to the makes and models in the VCA fuel consumption dataset based on different design features of makes and models. A cross-sectional dataset of makes and models that includes data on fuel consumption, mass, and other vehicle design factors was thus developed. A separate computer program was written in VB to perform this task as performing this manually would be very time consuming. A simplified processing flowchart of this program is shown in Figure 3.4.

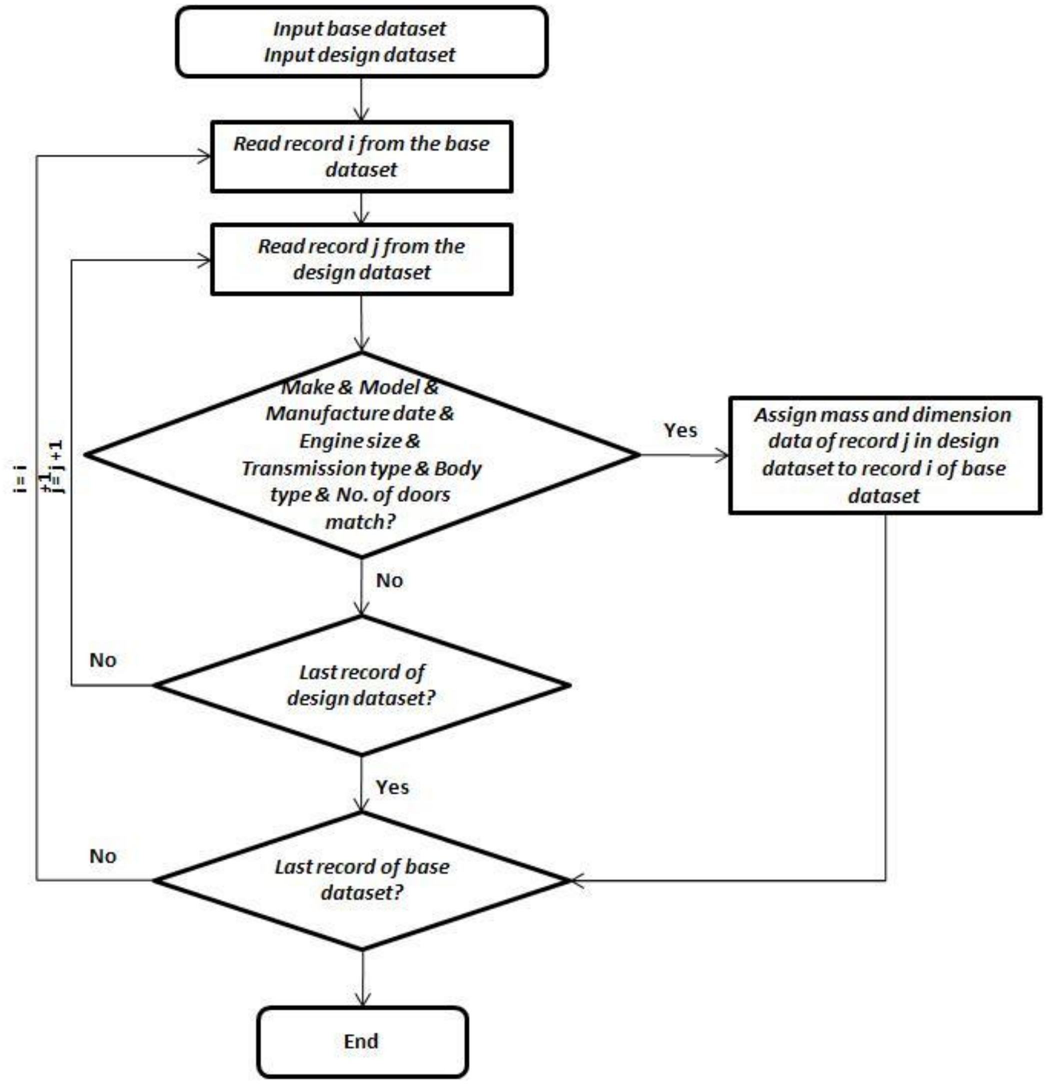


Figure 3.4: Processing flowchart of the computer program written to assign mass and dimension data to the makes and models in a defined base dataset

The general inputs to this program are a fuel consumption dataset (or in general, a base dataset of make and model variants with their design characteristics) and a vehicle design dataset (that includes mass and dimension data). As the processing flowchart shows, all the records from the base dataset that include make and model and their design information are read, a match for each vehicle from the design dataset is found (if available) based on the information on make and model and other design features, and the relevant mass and dimension data from the matched model variant in the design dataset are assigned to the vehicle record in the base dataset. This program was used to add mass and dimension data to the cross-sectional fuel consumption dataset of makes and models available from VCA. Table 3.5 shows the number of records for the base dataset and final dataset, which includes vehicle mass, by year. This is the year in which the vehicle is available on the market; it is used as a proxy for vehicle's year of manufacture. As the table shows, vehicle mass could be assigned to about 94% of makes and models in the fuel consumption dataset. The other 6% were excluded from the final dataset as mass data was not available for them neither in CAR magazine's database nor from manufacturers.

Table 3.5: Number of records for the fuel consumption dataset before and after assigning vehicle mass

Year on the market	Number of records		Percentage of mass availability
	Base dataset	Final dataset (sample)	
2000	934	828	89%
2001	2417	2194	91%
2002	1156	1106	96%
2003	1691	1617	96%
2004	1450	1398	96%
2005	2155	2062	96%
2006	1760	1638	93%
2007	1863	1812	97%
Total	13426	12655	94%

One option was to make some assumptions for the mass of those makes and models for which data was unavailable. However, due to the high variability of mass for different variants within makes and models, this would introduce some uncertainty to the fuel consumption dataset. If the developed sample has a sufficient number of records and it is shown to have similar characteristics to the base dataset reflecting the same range of variability in vehicle design, it can be used reliably for the analysis of vehicle fuel consumption. In order to examine this, the distribution of vehicles by engine size, fuel type, and transmission type (the design variables that contribute to fuel consumption) were compared between the base and the final (sample) dataset. The comparison results are reflected in Figures 3.5 to 3.7 and Tables 3.6 to 3.8. For each of engine size, fuel type, and transmission type, the results show a close match between sample and base datasets. In particular, Tables 3.6 to 3.8 show that the maximum relative difference in proportions between sample and full dataset is 0.06. These comparison results suggest that the sample dataset has similar characteristics to the base dataset; therefore, it will be used for the analysis of vehicle fuel consumption which will be covered fully in Chapter 4.

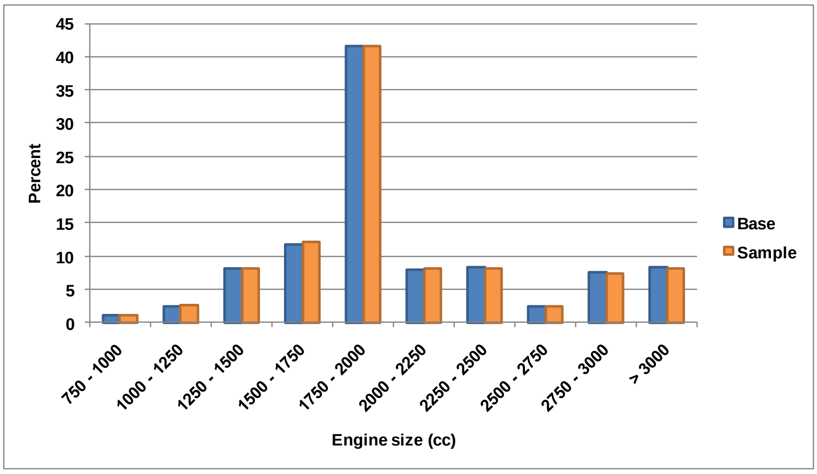


Figure 3.5: Distribution of vehicles by engine size: sample versus base dataset

Table 3.6: Distribution of vehicles by engine size: sample versus base dataset

Engine size category	Population proportion (%)	Sample proportion (%)	Absolute difference (%)	Relative difference[1]
750 - 1000	1.2	1.2	0.0	0.00
1000 - 1250	2.4	2.6	0.1	0.04
1250 - 1500	8.2	8.2	0.0	0.00
1500 - 1750	11.8	12.1	0.2	0.02
1750 - 2000	41.6	41.5	0.1	0.00
2000 - 2250	7.9	8.2	0.3	0.04
2250 - 2500	8.4	8.2	0.2	0.02
2500 - 2750	2.5	2.5	0.0	0.00
2750 - 3000	7.6	7.5	0.1	0.01
> 3000	8.4	8.1	0.3	0.04
Total	100.0	100.0	0.0	0.00

1. The ratio of absolute difference to population proportion

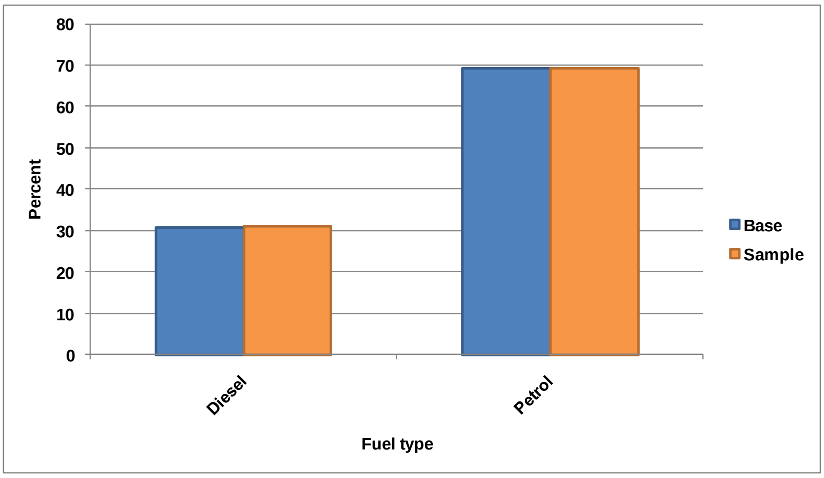


Figure 3.6: Distribution of vehicles by fuel type: sample versus base dataset

Table 3.7: Distribution of vehicles by fuel type: sample versus base dataset

Engine size category	Population proportion (%)	Sample proportion (%)	Absolute difference (%)	Relative difference[1]
Diesel	30.7	30.9	0.2	0.01
Petrol	69.3	69.1	0.2	0.00
Total	100.0	100.0	0.0	0.00

1. The ratio of absolute difference to population proportion

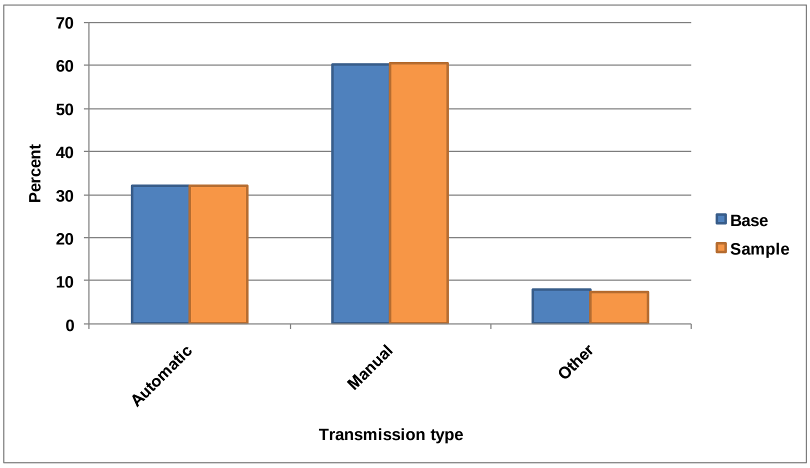


Figure 3.7: Distribution of vehicles by transmission type: sample versus base dataset

Table 3.8: Distribution of vehicles by transmission type: sample versus base dataset

Engine size category	Population proportion (%)	Sample proportion (%)	Absolute difference (%)	Relative difference[1]
Automatic	32.0	32.0	0.1	0.00
Manual	60.1	60.5	0.4	0.01
Other	8.0	7.5	0.5	0.06
Total	100.0	100.0	0.0	0.00

1. The ratio of absolute difference to population proportion

3.3.2.2. Two-car crash dataset

As discussed earlier in this chapter, the ideal dataset required to analyse the effects of vehicle mass on secondary safety performance of vehicles in fleet should include two-car crashes where detailed design data (including vehicle mass) is available for both of the vehicles involved in the crash. Such a dataset is not readily available in Great Britain. However, various datasets from different sources were collected, reconfigured, and linked together to develop a sample dataset of two-car crashes for this study.

i) Individual datasets

Table 3.9 shows the individual datasets that were used to develop the final two-car crash dataset together with their sources and brief descriptions.

Table 3.9: Individual datasets used to develop the final two-car crash dataset

Data source	Dataset	Description
STATS19	Accidents	Injury accidents
	Vehicles	Vehicles involved in injury accidents
	Casualties	Casualties resulted from injury accidents
DfT	VRM	Make and model variants of vehicles involved in injury accidents
	MM	Manufactured makes and models and their unique codes
CAR magazine	Design	Design data for manufactured make and model variants

As pointed earlier, STATS19 is a large database that includes data on all road accidents involving personal injury or fatality. In 2006, for example, it included information on about 190,000 different types of accidents involving about 350,000 different types of vehicles and resulting in about 260,000 casualties of different severities. As it is shown in Table 3.9, there are three separate datasets within STATS19: accident, vehicle, and casualty.

The accident dataset is composed of a number of records related to the injury accidents each of which include information on different aspects of the accident such as accident severity, number of vehicles and casualties involved, time and location of the accident, weather condition and road characteristics. Each accident record has a unique accident reference number and the number of records in the data is the same as the number of injury accidents reported.

The vehicle dataset contains information on all vehicles involved in injury accidents. Each record in the data includes information on different aspects of the vehicle including vehicle type, vehicle position in the time of accident, age and sex of the driver. The number of records is the same as the number of vehicles involved in injury accidents where each record includes data on accident reference number as well as a unique reference number for that vehicle involved in that particular accident so that vehicle records can be linked to particular accidents.

The casualty dataset includes various information regarding personal injuries and fatalities. Such information includes injury severity, age and sex of the casualty, and the type of casualty (driver, passenger, or pedestrian). There are three levels of casualty severity in the STATS19 data: killed (within 30 days), seriously injured, and slightly injured. Each record in the casualty data includes the accident and vehicle (if available) reference numbers as well as a unique reference number for that casualty linked with the particular vehicle (if available) and particular accident. Therefore, records in the casualty data can be linked to particular vehicles in particular accidents using these unique reference numbers. It should be noted that STATS19 does not include any information on non-injured passengers or pedestrians involved in injury accidents.

While STATS19 provides the opportunity to develop a two-car crash dataset which is required to undertake a disaggregate cross-sectional analysis of mass, it lacks detailed mass and size data for the vehicles involved in crashes which are key variables to the analysis. However, since 1989, the Vehicle Registration Mark (VRM) of the vehicles involved in accidents has been recorded by the Police (Broughton, 2007). This allows the basic vehicle dataset of STATS19 to be augmented with data from DVLA on make and model and design characteristics of those vehicles for which VRM is recorded. The DVLA dataset of vehicles involved in injury accidents for which VRM data is recorded is referred to as the Vehicle Registration Mark (VRM) dataset and is developed annually and held by Department for Transport (DfT). The VRM records include vehicle design data such as year of manufacture, engine size, fuel type, and propulsion type, as well as the accident and vehicle reference number. It also includes two unique codes for the make and model of the vehicle.

Another dataset held by DfT, called the Make and Model dataset (or MM dataset), includes information on registered makes and models and their unique codes. The codes are unique for all variants of different makes and models stratified by engine specification, engine capacity, body structure, and number of doors. As a result, the make and model dataset contains as many as 38,000 records. This data can be linked to the VRM records to provide information on makes and models and model variants of vehicles. These data are not publicly available to everyone. After a series of correspondences, DfT agreed to provide both VRM and MM datasets for the vehicles involved in crashes during 2000-2006 to be used specifically for this study.

Unfortunately, neither the VRM dataset nor the MM dataset include data on mass and size which are the key variables of interest. However, VRM includes make and model and detailed design information for the vehicles for which the registration mark is recorded. This provides the opportunity to assign mass and size data to the vehicles, based on their make and model and design features, using the design dataset as described in Section 3.3.2.1.

ii) Developing the final two-car crash dataset from individual datasets

The individual datasets listed in Table 3.9 were used to develop the final dataset of two-car crashes where the design data on both of the vehicles is available. The data was limited to two-car crashes during 2000 to 2006 in Great Britain (for which DfT provided the relevant VRM data). The development of the final two-car crash dataset included the following three main steps:

1. A dataset of all two-car crashes in which at least one of the drivers is injured (slight, serious, or fatal) was developed from the STATS19 database.
2. This dataset was linked with the VRM and MM datasets to add make and model information to the vehicles in the two-car crash dataset where the data was available.
3. The resulting two-car crash dataset was then linked with the design dataset to assign mass and dimension data to the vehicles where such data was available.

These steps are shown in Figure 3.8. All the process of data reconfiguration and augmentation shown in this figure was performed using Microsoft Access package.

In the first step, a dataset of total two-car crashes during 2000 to 2006 in which at least one of the drivers was injured (D1 dataset) was developed. This was done in a number of stages. First two-car crashes were extracted from all types of crashes, hence excluding single-vehicle crashes, crashes involving other types of vehicles (coaches, lorries, etc), and crashes involving other road users (pedestrians, motorcyclists, etc). The relevant data on the both cars involved in the crash were then added to each record in the data. An independent dataset of makes and models involved in injury crashes during 2000 to 2006 was also developed by linking the DfT's VRM dataset to the make and model (MM) dataset.

In the second step, the developed two-car crash dataset in the first step was linked to the developed make and model dataset (called VRM-MM) to create a sample dataset of two-car crashes where detailed make and model data is available for both of the cars involved in the crash (D2 dataset). This was done using the unique accident and vehicle references available in both datasets. A considerable number of two-car crashes were excluded from the dataset due to lack of make and model data on one of the cars involved in the crash. This exclusion will be discussed later in this section.

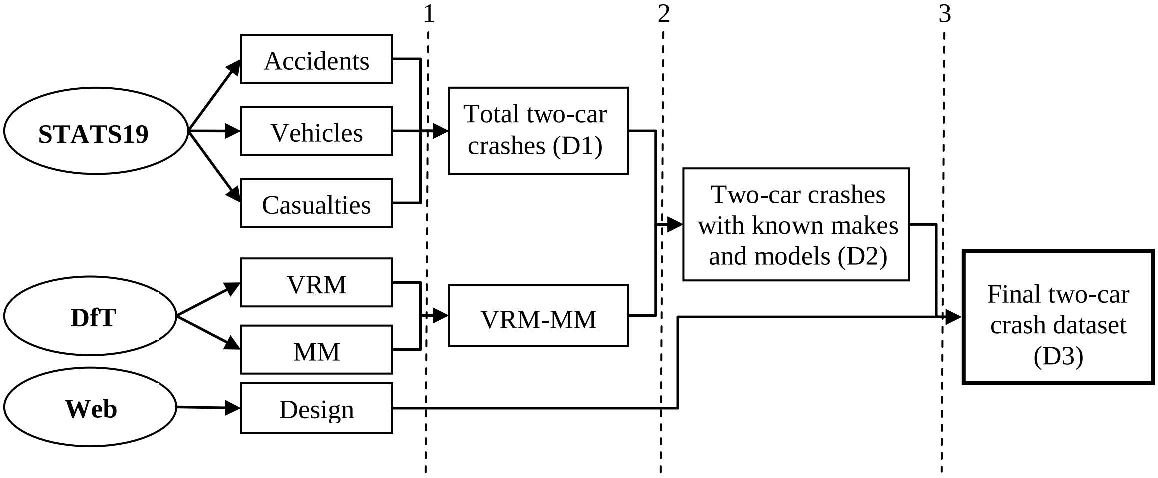


Figure 3.8: Process of developing final two-car crash dataset

In the third step, vehicle design data available in the design dataset were assigned to the cars. This could be done with great precision as detailed information on the variants of different makes and models such as year of manufacture, engine specification, engine capacity, body structure, and number of doors was available in both the VRM-MM and the design datasets. The computer program that was used to add vehicle mass to fuel consumption dataset (see Section 3.3.2.1 and Figure 3.4) was used to assign mass and dimension data to the vehicles in the D2 dataset. Due to a detailed level of vehicle design information available in both D2 and design datasets, a one-to-one match between variants of makes and models was possible when data was available. The resulting dataset was a final dataset of two-car crashes where mass and dimension data was available for both cars involved in the crash (D3 dataset).

Unfortunately, due to the fact that VRM data is not recorded for all the vehicles in STATS19 data and vehicle mass and dimension data is not available for all the variants of makes and models in the VRM dataset, a considerable proportion of two-car crashes were excluded from the data in each stage. As a result, the final two-car crash dataset only included about 21% of overall two-car crashes occurred in 2000-2006 period. Table 3.10 shows the number of two-car crashes available at the end of each step of dataset development by different years of data (years of accidents). As this table suggests, the proportion of missing data decreases with the more recent year of crash. This is partly because more technical data is available for newer cars, and partly because the Police recorded a higher proportion of VRM data in years that are more recent. The final sample of two-car crashes included about 85,000 two-car crashes in which at least one of the drivers was injured.

Table 3.10: Number of records by year of crash during the process of two-car crash dataset development

Year	Total two-car crashes (D1)	Two-car crashes with known makes and models (D2) (percent of total)	Two-car crashes with available vehicle mass data (D3) (percent of total)
2000	63,184	46,586 (74%)	6493 (10%)
2001	61,438	43,999 (72%)	8157 (13%)
2002	60,296	44,628 (74%)	10325 (17%)
2003	57,650	40,791 (71%)	12365 (21%)
2004	56,360	38,528 (68%)	12149 (23%)
2005	54,503	30,764 (56%)	14393 (26%)
2006	52,318	24,109 (46%)	21088 (39%)
Total	405,749	269,405 (66%)	84970 (21%)

The distribution of crashes by crash and driver characteristics (factors that potentially contribute to driver injury outcome in two-car crashes) were compared between the final sample dataset (D3) and the full two-car crash dataset (D1). The results summarised in Figure 3.9 and Table 3.11 show that the distribution of speed limit, direction of impact, driver age, and driver gender in the developed sample dataset is similar to that in the full two-car crash data. As Table 3.11 shows, the relative difference of proportions between sample and full data for all the crash categories compared remains below 0.2. The developed sample dataset will be used for a detailed analysis of vehicle mass and driver injury risk, which will be covered fully in Chapter 5.

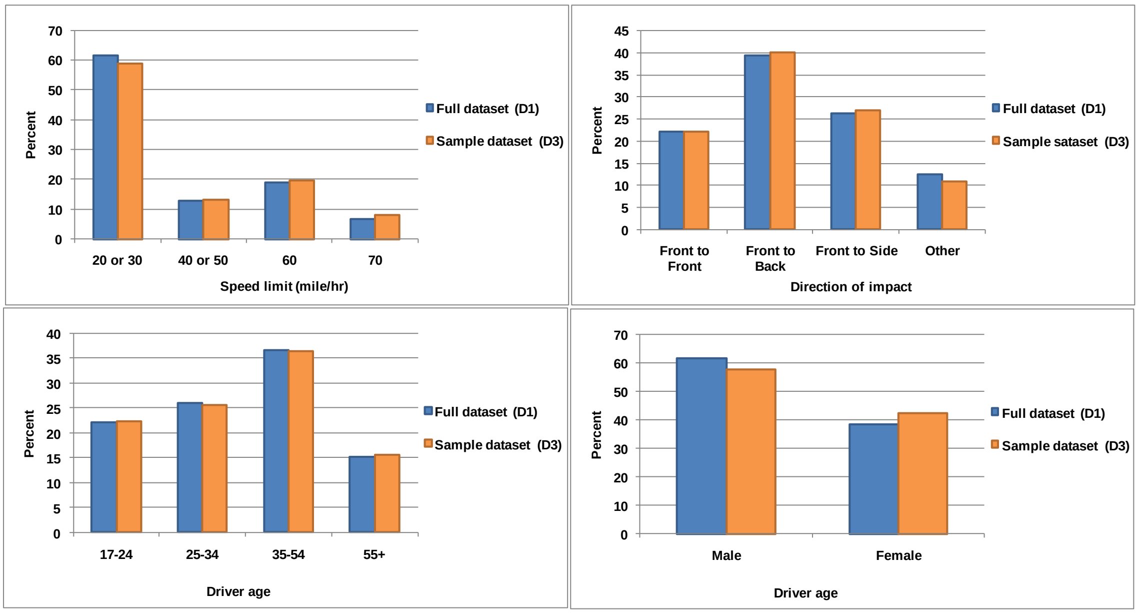


Figure 3.9: Distribution of two-car crashes by crash and driver characteristics: Full dataset (D1) and sample dataset (D3)

Table 3.11: Distribution of two-car crashes by crash and driver characteristics: Full dataset (D1) and sample dataset (D3)

Variable	Population proportion (%)	Sample proportion (%)	Absolute difference (%)	Relative difference[1]
Speed Limit				
20 or 30	61.5	59.0	2.5	0.04
40 or 50	12.7	13.2	0.5	0.04
60	19.0	19.7	0.7	0.04
70	6.8	8.1	1.3	0.19
Total	100.0	100.0	0.0	0.00
Direction of Impact				
Front to Front	22.1	22.2	0.1	0.00
Front to Back	39.3	40.1	0.8	0.02
Front to Side	26.2	26.9	0.7	0.03
Other	12.4	10.8	1.6	0.13
Total	100.0	100.0	0.0	0.00
Driver age				
17-24	22.1	22.3	0.2	0.01
25-34	26.1	25.7	0.4	0.02
35-54	36.6	36.5	0.2	0.01
55+	15.2	15.5	0.3	0.02
Total	100.0	100.0	0.0	0.00
Driver Sex				
Male	61.6	57.8	3.8	0.06
Female	38.4	42.2	3.8	0.10
Total	100.0	100.0	0.0	0.00

1. The ratio of absolute difference to population proportion

3.3.2.3. Vehicle registration dataset

In order to estimate safety and environmental consequences of changes in vehicle design in a fleet (in particular, fleet mass distribution), ideally, a detailed dataset of all the vehicles driven on the roads is required; this is not normally available. However, a disaggregate cross-sectional dataset that includes data on vehicle registration by make and model and other design features could be used instead as a reasonable proxy. Detailed vehicle registration data is available; however, it is not in the public realm. Correspondence with DVLA revealed that non-personal anonymised vehicle data extracted from the DVLA vehicle register is available through five commercial selling agents. The data includes information on various variables, but does not include the name and address of vehicle keepers, vehicle registration mark, vehicle identification or engine number. This information has been omitted to protect and ensure the anonymity of individual vehicle keepers. This data was purchased to be used for this study[1].

The vehicle registration data in Great Britain, which belonged to the last quarter of 2007, included number of vehicles registered by make and model and various other design and ownership factors and could be extracted at any desired level of aggregation. The data was extracted at a disaggregate level for all variants of registered makes and models in the fleet to give the number of registered vehicles in the fleet when they are stratified by the following variables:

- Make and model
- Model trim
- Year of manufacture
- Body type
- Number of doors
- Fuel type
- Engine size
- Transmission type

The final dataset of registered vehicles included about 102,200 records belonging to different variants of 740 makes and models from 63 manufacturers. The total number of registered makes and models in Great Britain in the last quarter of 2007 was

---

[1] Data was purchased from Experian plc.

30,536,224. Table 3.12 shows the 20 most popular makes and models in the fleet together with their overall number and their share of total registered vehicles in fleet. The most popular make and model in the British fleet in 2007 was Ford Fiesta followed by Vauxhall Astra and Ford Focus. These 20 makes and models form about 50% of all registered makes and models in the fleet.

Table 3.12: The top 20 most popular makes and models in 2007 British vehicle fleet by total number of registrations

Rank	Manufacturer	Make and Model	Total registered	Percentage of overall
1	Ford	Ford Fiesta	1,320,261	4.32
2	Vauxhall	Vauxhall Astra	1,228,558	4.02
3	Ford	Ford Focus	1,160,813	3.80
4	Vauxhall	Vauxhall Corsa	1,099,649	3.60
5	Volkswagen	Volkswagen Golf	908,022	2.97
6	Ford	Ford Mondeo	868,560	2.84
7	Renault	Renault Clio	778,060	2.55
8	BMW	BMW 3 Series	713,816	2.34
9	Vauxhall	Vauxhall Vectra	649,451	2.13
10	Peugeot	Peugeot 206	611,762	2.00
11	Nissan	Nissan Micra	595,148	1.95
12	Volkswagen	Volkswagen Polo	564,530	1.85
13	Ford	Ford Escort	533,609	1.75
14	Fiat	Fiat Punto	514,105	1.68
15	Ford	Ford Ka	462,512	1.51
16	Honda	Honda Civic	460,176	1.51
17	Renault	Renault Megane	388,823	1.27
18	Volkswagen	Volkswagen Passat	341,760	1.12
19	Renault	Renault Scenic	340,467	1.11
20	Peugeot	Peugeot 306	331,806	1.09

This dataset, referred to as vehicle registration dataset, did not include mass and dimension data for the makes and models; therefore, a version of the computer program explained in Section 3.3.2.1 was used to assign mass and dimension data to different variants of makes and models in the vehicle registration dataset from the design dataset (see Section 3.3.2.1). Mass and dimension data could be assigned to about 22,350,000 registered vehicles in fleet (about 73% of overall registered vehicles); the data was not available for the rest of the vehicles. This dataset will be used to define a base fleet when a number of mass distribution scenarios were formulated to investigate their likely effects on fleet fuel consumption and crash injuries. This will be covered in Chapter 6 of this thesis.

3.4. Summary and conclusions

As discussed in Section 3.1, the aim of this study is to examine the issue of potential interaction between environmental and safety performance in vehicle design, within a vehicle fleet, associated with vehicle mass. As noted, addressing this issue is key in understanding whether there is any conflict in safety and environmental goals as a whole in a vehicle fleet with given characteristics. To address this, partial effects of vehicle mass on each of fuel consumption and secondary safety performance of vehicles within the fleet will be investigated in this study (in Chapters 4 and 5) using disaggregate cross-sectional analysis of mass within a national vehicle fleet at a specific period of time. Having estimated these effects, the results can be used to investigate the partial effects of different hypothetical scenarios of fleet mass distribution on overall fleet fuel consumption and the total number of casualties.

The data required for the analysis of mass and fuel consumption was identified to be a cross-sectional dataset of vehicle fuel consumption that includes information on various design features of vehicles (including vehicle mass) that are associated with vehicle fuel consumption. Such a dataset should ideally include a wide range of vehicle design and fuel consumption rates to reflect the association between these variables. Besides this, the data collection method should allow that the effects of other contributing factors on fuel consumption rate to be controlled as far as possible. VCA fuel consumption database was identified as the data source that best meets these requirements (see Section 3.2.2 and Table 3.1); however, it lacks the information on mass and size of the vehicles. After verifying its reliability and accuracy, vehicle technical data from CAR magazine was used to develop a vehicle design dataset for assigning mass and dimension data to the make and model variants in VCA fuel consumption database (see Section 3.3.2.1 for details). The resulting fuel consumption dataset will be used to estimate the effect of mass on fuel consumption rate; this will be explained in detail in Chapter 4 of this thesis.

It was discussed that the data required for the analysis of mass and secondary safety is a sample of two-car crashes where information on mass and size of both vehicles involved in the crashes is available. This sample should ideally include crashes between pairs of vehicles with wide ranges of mass and size. Such a dataset is not readily available in Great Britain. However, it was concluded that police reported accident data has distinct advantages over the alternative sources of data (see Section 3.2.1 and Table 3.1) and can be used as the basis to develop the two-car crash dataset for this study. Various datasets from different sources including STATS19, DfT, and CAR magazine were collected, reconfigured and linked together to develop a sample dataset of two-car crashes that included information on mass and size of both vehicles (see Section 3.3.2.2. for details). This dataset will be used to model the effects of vehicle mass and other contributing factors in two-car collisions; this will be investigated in Chapter 5 of this thesis.

Finally, in order to estimate safety and environmental consequences of changes in mass distribution within the fleet, a detailed dataset of all the vehicles driven on the roads is required; this is not normally available. However, it was discussed that a disaggregate cross-sectional dataset that includes data on vehicle registration by make and model and other design features could be used instead as a reasonable proxy. Such a dataset was obtained from DVLA. Since the data did not include mass and dimension information for the makes and models, vehicle technical data from CAR magazine was used to add information on vehicle mass and dimensions to the DVLA vehicle registration dataset. The developed dataset was used to define a base fleet when the likely effects of a number of mass distribution scenarios on fleet fuel consumption and crash injuries were investigated; this will be covered in Chapter 6 of this thesis.

CHAPTER 4. VEHICLE MASS AND FUEL CONSUMPTION

The main objective of this chapter is to estimate the partial effects of vehicle mass on fuel consumption. The chapter is organised as follows. The first section (4.1) introduces factors that affect vehicle fuel consumption, discusses the relationship between vehicle mass and fuel consumption, and describes official fuel consumption measurements as an appropriate database for this analysis. The development and characteristics of the fuel consumption dataset is explained in detail in the second section (4.2). The third section (4.3) explains details of the applied methodology and reports the estimation results. The modelling results are interpreted and discussed in the next section (4.4). The final section (4.5) summarises and discusses the findings.

4.1. Background

Vehicle fuel consumption contributes to global warming (through the emission of CO_2), air quality (through the emission of toxic air pollutants), and fossil fuel consumption. These factors are regarded as the main environmental impacts of road transport. Therefore, reducing fuel consumption of the vehicle fleet is a desirable policy that improves the general environmental performance of the road transport system through a reduction in greenhouse gas emissions, toxic exhaust emissions, and oil consumption.

4.1.1. Factors affecting vehicle fuel consumption[1]

The determinants of vehicle fuel consumption are classified into four main groups: vehicle design factors, vehicle condition factors, driving cycle, and climatic conditions. This is illustrated in detail in Figure 4.1.

Fuel consumption varies substantially according to the vehicle design. The main design features of the vehicle that influence fuel consumption are fuel type, mass, engine size and characteristics, transmission type, year of manufacture, and other technological features (see Chapter 2, Sections 2.2.1 and 2.2.2).

---

[1] In this chapter, the term “fuel consumption” refers to the fuel consumption rate in l/100km

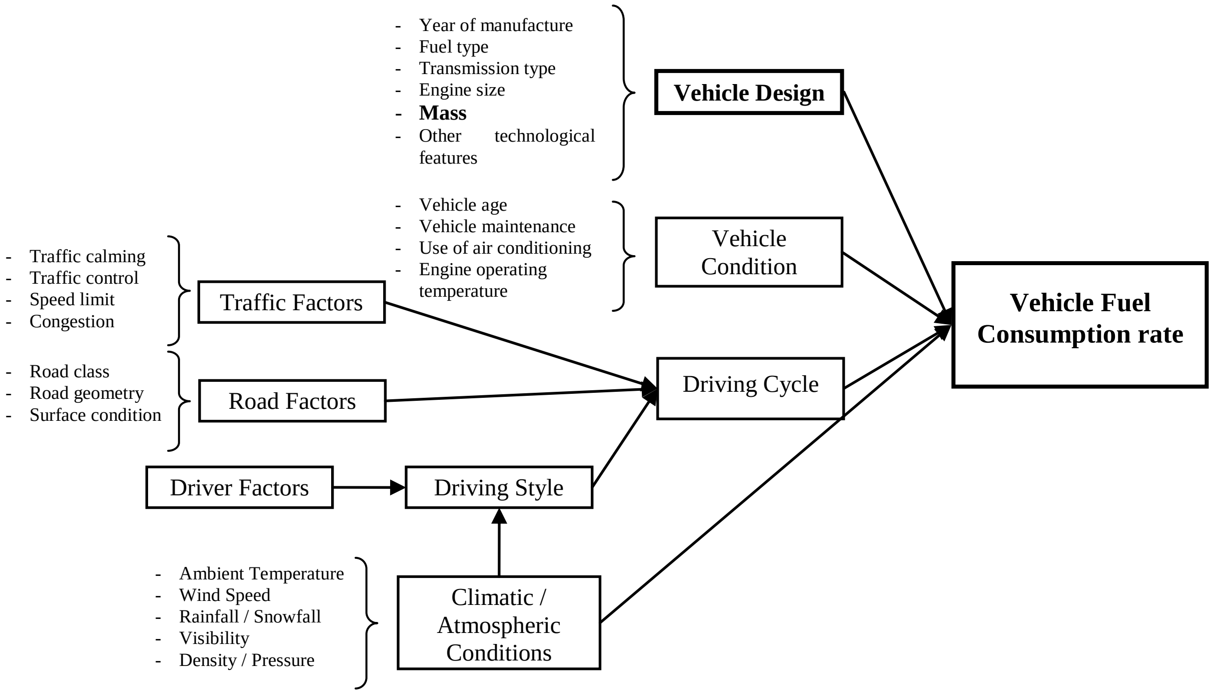


Figure 4.1: Determinants of vehicle fuel consumption rate

For a given vehicle design, fuel consumption can vary depending on various vehicle condition factors. Some of these include vehicle age, vehicle maintenance, use of air conditioning inside the vehicle, and operating temperature of the vehicle engine (Ross, 1994; Highway Agency, 2007). Driving cycle represents the speed profile of the vehicle over a given period of time or distance travelled. It includes a series of acceleration, deceleration, steady speed, and idling periods. Therefore, it can substantially influence the fuel consumption of a vehicle of a given design and condition. Driving cycle is influenced by driving style, traffic and road factors. Driving style, which is driver behaviour whilst driving, varies amongst different drivers depending on how quickly they accelerate, decelerate, change gears and make other decisions. In fact, driving cycle reflects traffic and road conditions and the driver’s response to them. Fuel consumption rate could also vary by different climatic and atmospheric conditions. For example, higher air density, which is proportional to ambient pressure at constant temperature, is associated with higher aerodynamic drag, and hence, higher fuel consumption rate (Redsell et al., 1993). An increase in temperature is linked to a decrease in fuel consumption and emissions due to a reduction in vehicle drag; besides, combustion is more efficient in warmer weather (Redsell et al., 1993). Other weather conditions (e.g. rain, wind, visibility) could influence fuel consumption through affecting driving style.

4.1.2. Relationship between mass and fuel consumption

Amongst various design features, vehicle mass is a key variable having the potential to considerably affect fuel consumption rate. Depending on the engine efficiency of a vehicle and the energy required by vehicle accessories[1], a certain amount of fuel energy is consumed to overcome forces resisting vehicle motion during a driving cycle and vehicle mass substantially contributes to most of these resistances including rolling, acceleration and gravitational losses (see Chapter 2, Section 2.2.1). Therefore, two vehicles with similar engine characteristics but a different mass have a different fuel consumption rate when they are driven under similar driving and road conditions. Current estimates of the effect of vehicle mass on fuel consumption that are available in the literature do not reflect the partial effects of mass where the effect is the result of a change in mass, holding all other factors constant. This is partly because not all the factors that contribute to fuel consumption are controlled. If other contributing factors are not fully controlled when estimating the effect of mass, the estimates may contain effects of other factors as well. One such case is to separate the effects of mass and engine size for vehicles using different types of fuel and with different transmission systems. Larger engines are usually found in heavier cars and they also tend to weigh more; therefore the estimated effects of mass could contain the effects of engine size as well.

In order to reliably estimate the partial effect of mass on fuel consumption, the effects of all other contributing factors should be controlled. Use of officially certified fuel consumption rates for specific makes and models, which are measured under controlled driving cycles, vehicle condition and ambient temperature, makes it possible to control for the effect of the majority of factors that can affect vehicle fuel consumption besides the vehicle design. As explained in Chapter 3, these data are available from the UK Vehicle Certification Agency (VCA) for the new car models on the market in the United Kingdom since 2000 (VCA, 2007a). This data was used as a basis to estimate the partial effects of vehicle design features, particularly mass, on fuel consumption rate.

---

[1] Such as air conditioning, lights, audio systems and heaters

4.1.3. Official fuel consumption measurements

All new cars that are approved for sale in Europe are required to meet certain emission standards. They have therefore passed through a fuel consumption test, which was first described by European Union Directive 80/1268/EEC (EC, 1980) and was further amended several times (EC, 1993; EC, 1999; EC, 2004). These measurements take place using controlled driving cycles, vehicle condition and ambient temperature and are available for different makes and models on the market.

In the first fuel consumption tests, described by European Union Directive 80/1268/EEC, measurements were taken for three driving cycles: a constant speed of 90 km/h, a constant speed of 120 km/h, and an urban cycle (EC, 1980). The urban cycle, starting from when the engine is warmed-up, consisted of a series of accelerations, steady speeds, decelerations and idling. This Directive was further amended by Directive 93/116/EEC as a result of which, the three-part tests were replaced by a two-part test: an urban driving cycle and an extra-urban driving cycle (EC, 1993). The extra-urban cycle was conducted immediately following the urban cycle and consisted of roughly half steady-speed driving and the remainder accelerations, decelerations, and some idling. The pattern of speed changes in the urban cycle were not changed from the previous urban cycle, but the conditions under which the fuel consumption was measured were changed; similar to emission measurements, fuel consumption measurement started 40 seconds after start-up instead of starting when the engine is warmed up. Directive 93/116/EEC was amended by Directive 99/100/EC in which the allowance of a 40 second warm-up period for measuring emissions for urban cycle was removed so that measurements begin immediately on start-up (EC, 1999). Directive 80/1268/EEC was then amended in 2004 (EC, 2004). In the last amendment, the driving cycles and measurements remained unchanged from the previous one. Table 4.1 shows a summary of changes in fuel consumption measurement tests since 1980.

Table 4.1: Summary of standards and changes to fuel consumption measurement tests

Year	EU Directive	Driving cycles	Start of measurements in urban cycle	Fuel consumption measurement method
1980	80/1268/EEC	3 cycles: 90 & 120 km/h, urban	Warmed-up mode	Direct measurement
1993	93/116/EEC	2 cycles: urban, extra-urban	40 s after start up	Carbon balance method
1999	99/100/EC	2 cycles: urban, extra-urban	Start up	Carbon balance method
2004	2004/3/EC	2 cycles: urban, extra-urban	Start up	Carbon balance method

According to the European Union Directive 93/116/EEC, fuel consumption is calculated based on carbon balance method using the emissions of CO_2 and other carbon related emissions (CO and HC) which are measured directly during the tests. Carbon balance method, which relates the carbon measured in the exhaust gas to the carbon content of the fuel consumed based on the law of conservation of mass, is explained in annex I to Directive 93/116/EEC (EC, 1993). Carbon balance has been a standard and accurate method within the automotive industry for calculating vehicle fuel consumption that is also recognised and used by US Environmental Protection Agency (EPA) to calculate and report official vehicle fuel consumption rates (Ensfield et al., 2006). In this method, fuel consumption is calculated using the following formulae (EC, 1993):

- for petrol vehicles

$$FC = \frac{0.1154}{D}[(0.866 \times HC) + (0.429 \times CO) + (0.273 \times CO_2)] \tag{4.1}$$

- for diesel vehicles

$$FC = \frac{0.1155}{D}[(0.866 \times HC) + (0.429 \times CO) + (0.273 \times CO_2)] \tag{4.2}$$

where,

FC is fuel consumption in litre per 100 km;
HC is measured emission of hydrocarbon in g/km;
CO is measured emission of carbon monoxide in g/km;
CO_2 is measured emission of carbon dioxide in g/km; and
D is density of the test fuel at 15°C.

It is noted that since the proportion of CO_2 emissions is substantially greater than that of HC and CO emissions for a given distance travelled, CO_2 is the main determinant of fuel consumption rate.

Details of the speed profiles of the urban and extra-urban driving cycles are given by Pelkmans and Debal (2006). Figure 4.2 reflects the speed profile of urban and extra-urban driving cycles. The average speed in urban cycle is 19 km/h and the distance travelled is 4 km while the average speed in extra-urban cycle is 63 km/h and the distance travelled is 7 km. The test is carried out at an ambient temperature of 20°C to 30°C (Pelkmans and Debal, 2006). The cars tested have to be run-in and therefore must have been driven for at least 3000 kilometres before testing (VCA, 2007b). Pelkmans and Debal (2006) estimated a gap of 10% to 20% between test and actual fuel consumption rates. They argued that this is because European test cycles have too smooth an acceleration profile to be realistic.

The tests are carried out in the laboratory on a chassis dynamometer, which is a rolling bed that simulates driving conditions. The chassis dynamometer simulates main resistance forces in real driving condition including acceleration and rolling resistance for the vehicle being tested. Roller drums that are in contact with wheels load the vehicle drive train. The system is connected to a computer that logs speed and power from the wheels and calibrates the system using the input weight of the vehicle so that correct loads can be applied to the wheels to simulate real driving condition for a specific driving cycle. However, since the vehicle body is stationary in these tests, they do not simulate aerodynamic drag which exists in real driving condition. A slight difference of less than 4 percent in energy consumption between chassis dynamometer and driving on road has been measured for the same driving cycles (Wang, et. al., 1999). This suggests that chassis dynamometer simulates forces resisting vehicle motion reasonably well.

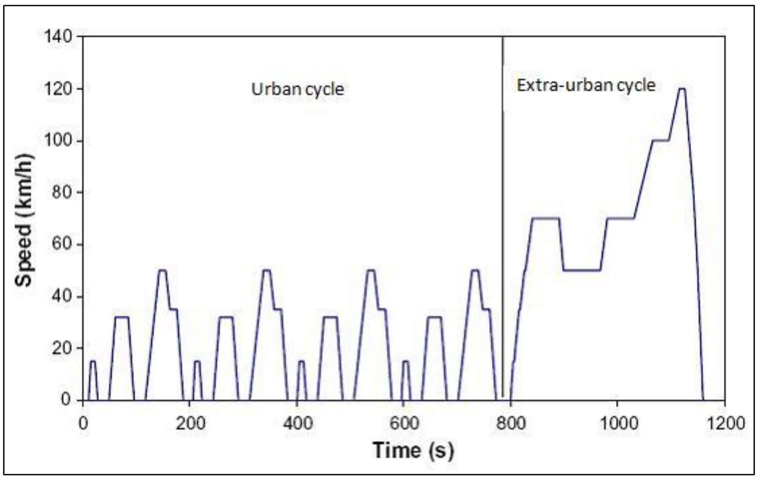


Figure 4.2: Speed profile of urban and extra-urban driving cycles (Pelkmans and Debal, 2006)

4.2. Fuel consumption dataset

As mentioned earlier in this chapter, VCA fuel consumption data was used to estimate the partial effects of vehicle mass on fuel consumption. The process of developing the fuel consumption dataset was explained in detail in Chapter 3, Section 3.3.2.1. The dataset consisted of 12655 records for 112 popular makes and models from 34 manufacturers.

Table 4.2 gives a summary of the variables in the VCA fuel consumption dataset with a brief description for each. VCA defines variants of a make and model and includes them in its database when they are different in at least one of the variables shown in this table. As an example, all variants of Ford Fiesta (as defined by VCA) in the fuel consumption dataset are given in Appendix 1. In several cases, the only difference between variants is their trim (which describes the detailed version of a model). Table 4.3 shows a few examples of two make and model variants where they only differ in trim. This shows that there are model variants in the data that are identical in year, fuel type, transmission type, mass, engine size, Euro emission standard, urban and extra-urban fuel consumption (all of which will be included as variables in the analysis of fuel consumption) and only differ in a detailed technical characteristic as described by trim. The table also shows that the content and format of trim is not consistent between different makes and models.

On the grounds that the variables shown in Table 4.2, excluding “Trim”, are the main design variables that account for the majority of variation in vehicle fuel consumption, records in the data that were identical in all of these variables and only differed in trim (see Table 4.3 for examples) were defined as duplicate and hence excluded from the original dataset; the resulting final dataset included a total of 9737 model variants from 112 makes and models.

Table 4.2: Available variables in the fuel consumption dataset

Variable	Type	Description
Make and model	Categorical	Make and model of the vehicle
Trim	String	Detailed version of the model
Urban fuel consumption	Continuous	Fuel consumption in l/100km
Extra-urban fuel consumption	Continuous	Fuel consumption in l/100km
Fuel Type	Categorical	Petrol and diesel
Engine size	Continuous	Engine displacement in cc
Transmission type	Categorical	Manual, automatic, semi-automatic, continuously variable , and sequential shift gearbox
Euro emission standard	Categorical	The Euro emission standard that the vehicle satisfies (I, II, III)
Year	Nominal	The calendar year in which the vehicle is on the market (2000-2007)

Table 4.3: A few example of variants having only a slight difference in trim

Make and Model	Variant	Trim
Audi A6	1	3.0 TDI V6 quattro Tiptronic (225 PS)
	2	3.0 TDI V6 quattro Tiptronic (225 PS) (DPF)
BMW 300	1	318i Saloon - 01 January 2001 to 30 June 2001
	2	318i Saloon - 01 July 2001 to 30 April 2002
Ford Fiesta	1	1.8 TDdi E-Diesel
	2	1.8 Turbo E-Diesel
Ford Fiesta	1	1.6 Duratec (4.06 FDR)
	2	1.6 Duratec (4.25 FDR)
Honda CRV	1	2.0 i-VTEC Executive
	2	2.0 i-VTEC SE
Mercedes A Class	1	A 170 CDI
	2	A 170 CDI LWB
Mercedes C Class	1	C200 CDI no DPF 245 Tyres at rear
	2	C200 CDI with DPF 245 Tyres at rear
Renault Scenic	1	2.0 VVT (JM1N06)
	2	2.0 VVT 136
Vauxhall Astra	1	2.0i 16v Turbo 2 Door Convertible
	2	2.0i 16v Turbo 2 Door Convertible From VIN: W0L0
Vauxhall Corsa	1	1.3CDTI SXI 5 Door Hatchback From VIN: W0L0XCF68
	2	1.3CDTI Design 5 Door Hatchback From VIN: W0L0XC

Table 4.4 shows, for different manufacturers, makes and models by number of variants in the fuel consumption dataset before and after the exclusion of duplicate records. Depending on the design varieties within these makes and models during the study period (2000-2007), the number of variants in the dataset for each make and model differs and ranges from 1 (Rover Mini) to 735 (Vauxhall Vectra). The existence of more records from a specific make and model in the dataset does not necessarily mean that there are also the same proportion of registered vehicles of that make and model on the roads.

Descriptive statistics of the two fuel consumption rates (urban cycle, extra-urban cycle), engine size, mass, and frontal area for two types of fuel[1] are compared in Table 4.5. In the dataset, the average fuel consumption of petrol cars is greater than diesel cars in both driving cycles while the average engine size is almost the same for both types of fuel (though the range of petrol engine sizes is almost twice that of diesel ones) and diesel cars are, on average, more massive than petrol cars.

---

[1] The two types of fuel considered in this study are petrol and diesel which account for more than 99% of the VCA fuel consumption data

Table 4.4: Makes and models and the number of their variants in the fuel consumption dataset

Make	Make and Model	No. of variants		Make	Make and Model	No. of variants	
		Original data[1]	Final data[2]			Original data	Final data
Alfa	Alfa 156	142	106	Mitsubishi	Mitsubishi Carisma	60	34
Audi	Audi A3	249	133		Mitsubishi Shogun	187	118
	Audi A4	98	91		Mitsubishi Space wagon	21	14
	Audi A6	406	298	Nissan	Nissan Almera	63	60
	Audi TT	35	35		Nissan Micra	54	51
BMW	BMW 300	604	490		Nissan Primera	75	67
	BMW 500	305	267		Nissan Serena	5	4
	BMW 700	79	72		Nissan Terrano	40	36
Citroen	Citroen C3	57	57	Peugeot	Peugeot 106	20	13
	Citroen C5	107	107		Peugeot Saxo	47	34
	Citroen Picasso	55	49		Peugeot 205	105	87
	Citroen Synergie	19	14		Peugeot 306	53	37
	Citroen Xantia	34	34		Peugeot 307	100	100
	Citroen Xsara	141	108		Peugeot 406	169	133
Daewoo	Daewoo Lanos	24	12	Proton	Proton Persona	63	40
	Daewoo Matiz	26	10	Renault	Renault Clio	90	79
	Daewoo Nubira	42	32		Renault Espace	81	69
Daihatsu	Daihatsu Charade	16	14		Renault Laguna	251	218
Fiat	Fiat Bravo	52	34		Renault Megane	218	187
	Fiat Punto	29	29		Renault Scenic	113	99
	Fiat Seicento	19	14	Rover	Rover 200/400	17	17
Ford	Ford Escort	7	5		Rover 25/45	19	17
	Ford Fiesta	162	137		Rover 75	138	108
	Ford Focus	330	232		Rover Mini	1	1
	Ford Galaxy	108	91	Saab	Saab 9-3	200	164
	Ford Ka	43	40		Saab 9-5	163	148
	Ford Maverick	4	4	Seat	Seat Ibiz/Co	26	23
	Ford Mondeto	384	324		Seat Leon	76	62
	Ford Puma	10	7		Seat Toledo	50	46
Honda	Honda Accord	252	169	Skoda	Skoda Fabia	201	155
	Honda CIVIC	159	125		Skoda Felicia	10	6
	Honda CRV	60	23		Skoda Octavia	281	201
Hyundai	Hyundai Accent	43	38	Subaru	Subaru Impreza	61	51
	Hyundai Atoz	4	4		Subaru Legacy	190	168
	Hyundai Coupe	46	39	Suzuki	Suzuki Baleno	15	14
	Hyundai Lantra	4	4		Suzuki Swift	43	27
Isuzu	Isuzu Trooper	68	60		Suzuki Vitara	3	2
Jaguar	Jaguar S Type	76	40	Toyota	Toyota Avensis	118	98
	Jaguar X Type	46	46		Toyota Celica	17	15
	Jaguar XJ Type	50	30		Toyota Corolla	119	103
Chrysler jeep	Jeep Cherokee	41	34		Toyota Land Cruiser	45	45
	Jeep Grand Cherokee	22	19		Toyota MR2	8	7
Landrover	Landrover Defender	16	14		Toyota Previa	23	21
	Landrover Discovery	39	32		Toyota RAV-4	27	26
	Landrover Freelander	47	37		Toyota Yaris	73	66
	Landrover Rangerover	38	33	Vauxhall	Vauxhall Astra	540	340
Lexus	Lexus IS200	13	13		Vauxhall Corsa	376	259
Mazda	Mazda 626	37	22		Vauxhall Frontera	65	47
	Mazda MX-5	36	33		Vauxhall Omega	114	100
Mercedes Benz	Mercedes A class	365	233		Vauxhall Vectra	735	394
	Mercedes C Class	313	245	Volkswagen	Volkswagen Beetle	79	77
	Mercedes E Class	545	388		Volkswagen Golf/Jet	361	284
	Mercedes ML Class	71	58		Volkswagen Passat	214	204
	Mercedes S Class	141	117		Volkswagen Polo	87	86
MG	MG MGF	6	6	Volvo	Volvo SV40	79	50
Mini	Mini	65	56		Volvo V70	76	61
					TOTAL	12655	9737

1. Fuel consumption data before the exclusion of duplicate model variants.
2. Fuel consumption data after the exclusion of duplicate model variants.

Table 4.5: Descriptive statistics of continuous design variables in the fuel consumption dataset

Vehicle variable	Fuel type	Descriptive statistics				
		Mean	Min	Max	Std. Deviati	Obs.
Urban fuel consumption (l/100km)	Petrol	12.24	6.00	24.10	2.92	6798
	Diesel	8.52	4.90	15.80	2.00	2939
Extra-urban fuel consumption (l/100km)	Petrol	6.80	3.90	14.00	1.29	6798
	Diesel	5.28	3.50	11.00	1.20	2939
Engine size (cc)	Petrol	2142	796	6209	804	6798
	Diesel	2118	1248	4164	471	2939
Mass (kg)	Petrol	1383	720	2687	255	6798
	Diesel	1490	875	2717	263	2939
Frontal area (cm^2)	Petrol	25838	19035	38982	2432	6648
	Diesel	26630	21516	42486	2911	2875

For petrol cars in the dataset, the minimum urban fuel consumption rate (6 litre/100km) and extra-urban fuel consumption rate (3.9 litre/100km) respectively belong to a variant of Daihatsu Charade with engine size of 989 cc and kerb mass of 720 kg, and a variant of Vauxhall Corsa with engine size of 998 cc and kerb mass of 975 kg. The maximum urban fuel consumption rate (24.1 litre/100km) and extra-urban fuel consumption rate (14 litre/100km) respectively belong to an Estate Mercedes M-class with engine size of 6208 cc and kerb mass of 2310 kg, and a Jeep Grand Cherokee with engine size of 4700 cc and kerb mass of 2073 kg.

For diesel cars, the minimum urban fuel consumption (4.9 litre/100km) and extra-urban fuel consumption (3.5 litre/100km) respectively belong to a variant of Citroen C3 with engine size of 1398 cc and kerb mass of 1022 kg, and to a Vauxhall Astra with engine size of 1686 cc and mass of 1225 kg. Both maximum urban fuel consumption (15.8 litre/100km) and extra-urban fuel consumption (11 litre/100km) belong to a Toyota Land Cruiser with engine size of 4164 cc and kerb mass of 2520 kg.

Table 4.6 shows the distribution of model variants in the final dataset within different categories of three design variables: fuel type, transmission type, and the emission standard that the model variant satisfies. The category with the most records is manual transmission petrol cars meeting Euro IV emission standards. The Euro II, III, and IV emission standards are specific limits for exhaust emissions of NO_x, HC, CO, and PM (which are progressively more stringent). They apply to all passenger cars sold in the European Union member states that are manufactured from January 1996 to January 2000, from January 2000 to January 2005, and from January 2005 to mid 2008, respectively (Highway Agency, 2007).

Table 4.6: Distribution of model variants within categorical design variables in the fuel consumption dataset

Transmission type	Fuel type	Euro emission standard			Total
		II	III	IV	
Manual	Petrol	443	1597	1759	3799
	Diesel	158	1158	692	2008
Automatic	Petrol	220	1299	1035	2554
	Diesel	32	427	241	700
Other	Petrol	17	105	323	445
	Diesel	0	73	158	231
Total		870	4659	4208	9737

4.3. Fuel consumption modelling

4.3.1. Relationship between variables

A correlation analysis was performed to investigate the linear relationship between the variables in the dataset. The continuous variables in the dataset (urban and extra-urban fuel consumption, engine size, mass, frontal area) were included in the analysis. Table 4.7 shows the estimated correlation coefficients. There is a positive and significant correlation between each of the three design variables (engine size, mass, frontal area) and the two fuel consumption rates in the dataset. As the table shows, there is also a positive correlation between engine size, mass, and frontal area. As it was expected, the two fuel consumption rates are highly and significantly correlated.

Table 4.7: Correlation coefficients between variables (all significant at $\alpha = 0.001$)

	Urban fuel cons.	Extra-urban fuel cons.	Engine size	Mass	Frontal area
Urban fuel cons.	1.000	0.917	0.760	0.617	0.414
Extra-urban fuel cons.		1.000	0.686	0.666	0.598
Engine size			1.000	0.751	0.409
Mass				1.000	0.670
Frontal area					1.000

The relationships between engine size and each of urban and extra-urban fuel consumption is shown in Figures 4.3 and 4.4, respectively. The relationships are shown separately for petrol and diesel cars. As it was expected, the fuel consumption generally increases as engine size increases. Comparison of these two figures suggests that engine size has a different effect on each of urban and extra-urban fuel consumption (there is a greater effect on urban fuel consumption as suggested by the steeper gradient). For each driving cycle, the relationships show that for a given engine size, a petrol car tends to have a higher consumption than a diesel car. Besides, the relationship between engine size and fuel consumption could be different for petrol and diesel cars as suggested by these figures.

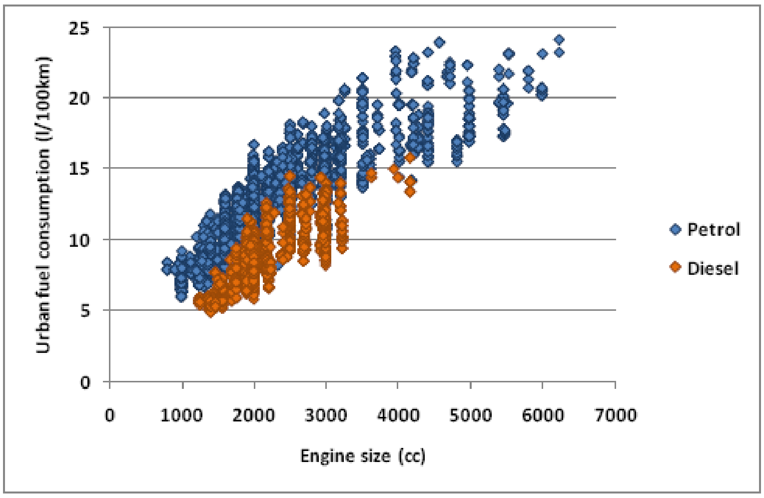


Figure 4.3: Relationship between engine size and urban fuel consumption

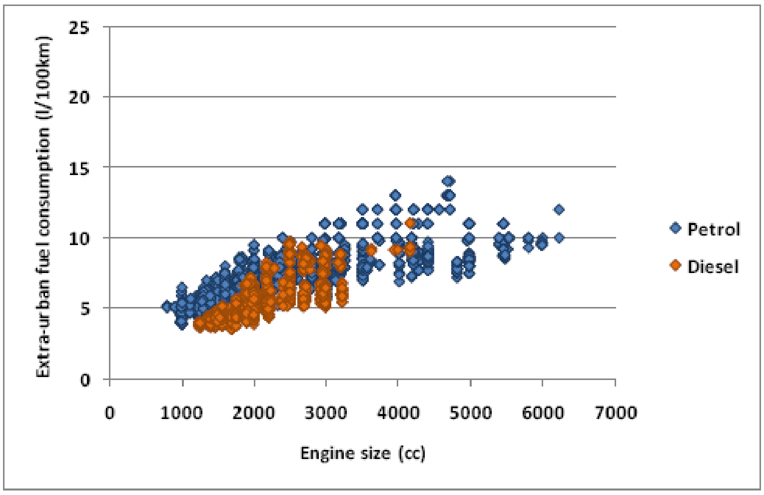


Figure 4.4: Relationship between engine size and extra-urban fuel consumption

Figures 4.5 and 4.6 show the relationship between mass and each of urban and extra-urban fuel consumption. Comparison of the two figures suggests that the relationship is different for urban and extra-urban driving cycles (there is a greater effect of mass for the urban cycle, this is suggested by the steeper gradient). The relationships suggest that in both driving cycles, a petrol car tends to have higher consumption rates than a diesel car of the same engine size. The figures also suggest a different effect of mass on fuel consumption for petrol and diesel cars.

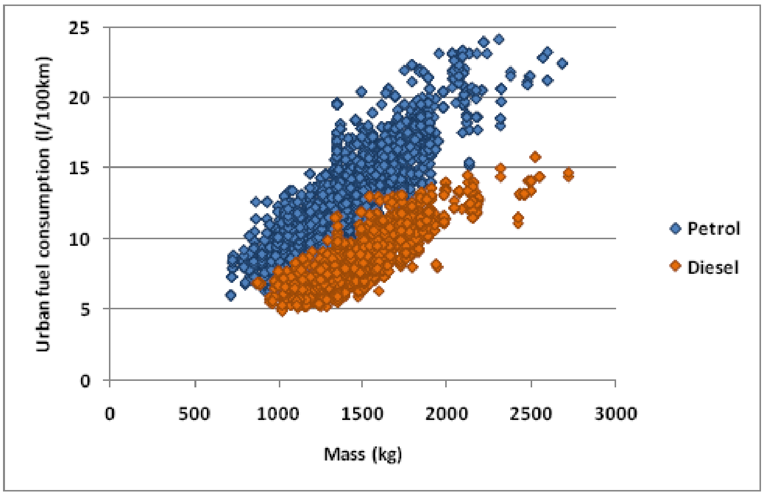


Figure 4.5: Relationship between mass and urban fuel consumption

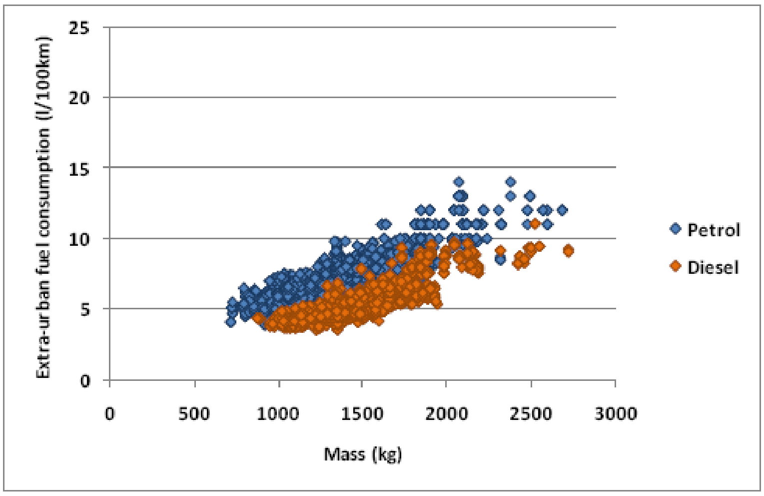


Figure 4.6: Relationship between mass and extra-urban fuel consumption

The objective of the analysis is to separate the effects of different design factors on fuel consumption. Therefore, a statistical modelling approach was used to investigate the partial effects of mass and other vehicle-related variables on fuel consumption. Use of official fuel consumption rates, which are measured under controlled conditions, automatically controls for driving cycle, vehicle condition and ambient temperature. Design factors were accounted for by defining relevant variables in the dataset for each driving cycle. This dataset provided an opportunity to undertake cross-sectional analysis of mass to estimate its partial effects on fuel consumption. Dependent variables were urban fuel consumption and extra-urban fuel consumption.

4.3.2. Choice of statistical model

The choice of the statistical model depends on a number of factors including the nature of the dependent variable, the relationship between dependent variable and explanatory variables, and the nature of stochastic variability. Basic linear regression model is the simplest case of regression models that assumes a linear relationship between a dependent variable and a set of explanatory variables under certain assumptions. For the simple case of a single explanatory variable x,

$$y_i = \propto + \beta x_i + \epsilon_i \tag{4.3}$$

under the assumption that $\epsilon_1, \epsilon_2, \dots \epsilon_i$ ... are independent errors, with

$$E(\epsilon_i) = 0 \tag{4.4}$$

$$Var(\epsilon_i) = \sigma^2 \tag{4.5}$$

where, y_i denotes observation *i* of the dependent variable *y*,

x_i denotes observarion *i* of the explanatory variable *x*, and

ε_i denotes the error associated with observation i.

Parameters $\propto$ and β can be estimated using different methods including least squares and maximum likelihood. If the errors are assumed to be normal, then t tests can be performed on the parameter estimates and least squares and maximum likelihood estimation provide identical results; details of these are available in many statistical textbooks including (Maddala, 2001) and (Greene, 1993)

The assumptions of simple linear regression (relationships 4.4 and 4.5) imply that the errors have a zero mean and a common variance σ^2. Preliminary examinations of the data which included estimation of linear regression models for each of urban and extra-urban fuel consumption rates suggested that these assumptions are violated, therefore least square regression technique is not appropriate for analyzing this data.

The class of Generalized Linear Models (GLM) includes models useful for analysis of data that has non-normal error distributions from the exponential family (McCullagh and Nelder, op.cit.). This class of statistical models are characterised by three elements:

1. A probability distribution for the dependent variable y depending on the mean μ and variance σ^2.
2. A linear predictor of the explanatory variables:

$$\eta = \sum \beta_j x_j \tag{4.6}$$

3. A function linking the linear predictor η to the mean μ:

$$\eta = g(\mu) \tag{4.7}$$

In the generalised linear models, normality and constant variance are no longer a requirement for the error component of the model (McCullagh and Nelder, op.cit.). Given the nature of the dependent variables, which are positive continuous quantities with a positively skewed distribution, generalized linear model with gamma error structure was found appropriate to model the fuel consumption variation. A logarithmic link function was used in the estimated models. Therefore, the estimated fuel consumption models will have the following form:

$$\mu = \exp\left(\beta_0 + \sum \beta_j X_j\right) \tag{4.8}$$

where, μ is mean fuel consumption rate,

$\{\beta_j\}$ are a set of coefficients whose values are to be estimated, and

$\{X_j\}$ are a set of explanatory variables.

In generalised linear models, goodness of fit of an estimated model is assessed based on model deviance. Deviance is a measure of discrepancy of the model and is defined as:

$$D = -2(LL_{model} - LL_{max}) \tag{4.9}$$

where LL_{model} is the log likelihood achieved by the model under investigation and LL_{max} is the maximum log likelihood achievable in a full model. Therefore, deviance is a measure of the distance between the model under investigation and the full model. The form of the deviance function varies depending on the distribution. For a generalised linear model with gamma distribution, deviance has the following form (McCullagh and Nelder, op.cit.):

$$D = 2\sum\left[-ln\left({}^{y}/_{\hat{\mu}}\right) + {}^{(y-\hat{\mu})}/_{\hat{\mu}}\right]. \tag{4.10}$$

Another measure of goodness of fit used to compare the performance of different generalised linear models from the same dataset and with the same specification but different number of parameters is Akaike Information Criterion (AIC). It is a relative measure of model performance based on the value of log likelihood. AIC is calculated using the following formula:

$$AIC = 2\mathrm{k} - 2\mathrm{LL} \tag{4.11}$$

where k is the number of parameters in the model and LL is the log likelihood values of the estimated model. The model with a lower value of AIC has a better performance.

4.3.3. Explanatory variables

Explanatory variables extracted and defined for the study, along with their definitions, are listed in Table 4.8. “Frontal area” was calculated and included as a variable because it affects aerodynamic drag during vehicle motion. The variable “Time” is the number of years after 2000 when the vehicle is available on the market for the first time, therefore it takes ordinal values ranging from 0 (for year 2000) to 7 (for year 2007). “Time” is used as a proxy for vehicle year of manufacture. This variable was introduced to account for possible technological improvements to vehicle design. The difference between this variable and the “Euro” variable is that the “Euro” variable only relates to changes made in vehicle design to achieve compliance with certain emissions standards (either through improving fuel efficiency or technologies such as exhaust catalysts) while the variable “Time” is a proxy for all other technological changes in vehicle design that have not been controlled by other variables. Variables “(Make)i” and “(Make & Model)j” represent make of the vehicle and make and model of the vehicle, respectively. These are introduced to examine whether the estimated mean fuel consumption rates are different for different manufacturers or makes and models when other factors are controlled.

Table 4.8: Explanatory variables used in the fuel consumption models

Variable	Type	Definition
Engine size	Continuous	Engine displacement in cc
Mass	Continuous	Kerb mass in kg
Frontal area	Continuous	(Width x Height) in m^2
Time	Ordinal	Number of years after 2000 to which the dataset belongs
Categorical variables for fuel type		
Diesel	Binary	1 if the fuel type is diesel; 0 if the fuel type is petrol
Categorical variables for Euro standard		
Euro II	Binary	1 if vehicle meets Euro II; 0 otherwise
Euro III	Binary	1 if vehicle meets Euro III; 0 otherwise
Categorical variables for transmission type		
Automatic transmission	Binary	1 if the vehicle has an automatic transmission; 0
Other transmission	Binary	1 if the vehicle has a transmission other than manual or Automatic; 0 otherwise
Categorical variables for vehicle make		
$(\text{Make})_i$	Nominal	1 if vehicle is Make i; -1 if vehicle is Make n; 0 otherwise (i = 1,...,n)
Categorical variables for vehicle make and model		
$(\text{Make \& Model})_j$	Nominal	1 if vehicle is make and model j; -1 if vehicle is make and model m; 0 otherwise (j = 1,...,m)

As Table 4.8 shows, the widely-used dummy coding method was used to code the three categorical variables fuel type, Euro standard, and transmission type. In coding each categorical variable, the category with the most number of observations was taken as the reference category. In the case of transmission type, for example, manual transmission type was taken as the reference because almost 60% of cars in the dataset have manual transmission. Examples of “Other transmission” include semi automatic, continuously variable and sequential shift gearbox. Based on these definitions, the reference category is a manual petrol car meeting Euro IV emission standards which is available in the 2000 UK market. On the other hand, effect coding method was used to code variables representing vehicle make and vehicle make and model as defined in Table 4.8. This method of coding for vehicle makes (makes and models) avoids an arbitrary choice of a single vehicle make (make and model) as the reference group and allows estimation of relative effects of different makes (makes and models). In this method, the reference is the grand mean of fuel consumption rate across vehicle makes[1] (makes and models) which is represented by the constant in the model. The estimated coefficient of each vehicle make (make and model) in the regression model represents the difference

---

[1] This is mean of the mean fuel consumption rate of all vehicle make (or make and model) categories.

between the fuel consumption rate of that vehicle make (make and model) and the grand mean.

4.3.4. Model estimation results

As it was mentioned before, generalised linear models with gamma error structures were estimated for each of fuel consumption rates. The parameters of the statistical models were estimated based on maximum likelihood method using statistical package “R”. Residual plots were used to diagnose whether the relationship between continuous explanatory variables and the dependent variables are linear or nonlinear. Wonnacott and Wonnacott (1990) discuss diagnosis of model adequacy using plots of model residuals against the explanatory variables and note that a bow-shaped residual plot for an explanatory variable suggests a non-linear effect of that variable. The plot of model residuals against “Engine size” was bow-shaped suggesting a non-linear effect of engine size and the logarithmic transformation was found to be the best relationship when the performance of competing models with “Engine size” and different transformations of “Engine size” were compared. Therefore, a logarithmic transformation of “Engine size” is used in the models. No clear sign of a non-linear effect of “Mass” was observed in the data.

A relatively high correlation between the variables “Mass”, “Engine size” and “Frontal area” in the dataset raised the concern of collinearity between them and its consequence on model estimates. To examine this, the three variables were added to the models one by one and in separate steps. Changes in model performance (measured by log likelihood values through the Akaike Information Criterion (AIC)) as well as the sign and significance of estimated coefficients were compared for all possible estimated models. Lattices shown in Figures 4.7 and 4.8 compare AIC values for these estimated models for urban and extra-urban fuel consumption, respectively. In these lattices, Model 4.1 is the reference model that includes fuel type, transmission type, time, and Euro emission standard as the explanatory variables and the three other variables (“Engine size”, “Mass”, and “Frontal area”) were added to this model one by one.

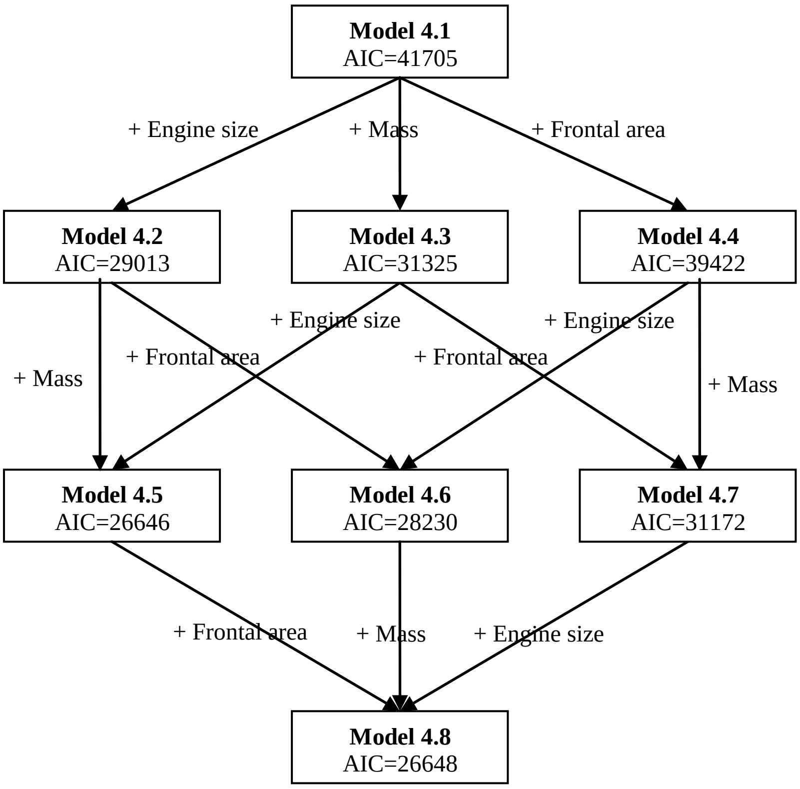


Figure 4.7: Effect of adding design variables on model performance (urban fuel consumption models)

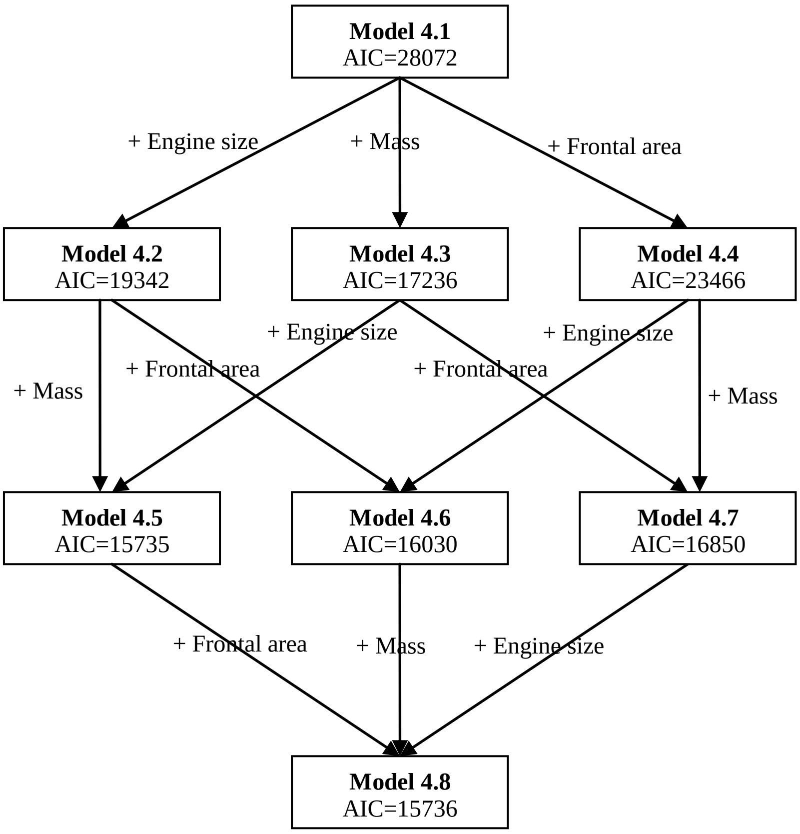


Figure 4.8: Effect of adding design variables on model performance (extra-urban fuel consumption models)

For both urban and extra-urban fuel consumption, the estimation results showed that adding “Frontal area” to the model which already included “Mass” and “Engine size” (Model 4.5 to Model 4.8) did not improve model performance and the variable “Frontal area” in model 4.8 was not statistically significant[1]. On the other hand, adding “Mass” to a model which already included “Engine size” (Model 4.2 to Model 4.5) considerably improved model performance and no unexpected change was found in the sign and influence of other variables. Based on these results, only the variable “Frontal area” was excluded from the models. The detailed estimation results of these models for both urban and extra-urban fuel consumption are given in Appendix 2.

These results suggest that it is possible to isolate the separate influence of “Mass” and “Engine size” on fuel consumption using statistical models. The relationship between “Mass” and “Engine size” shown in Figure 4.9 also shows that there is the potential to isolate the effect of these two variables in the models. The relationship is closer to a power function than a linear one and in many cases, considerable variation in mass value is observed for a given engine size.

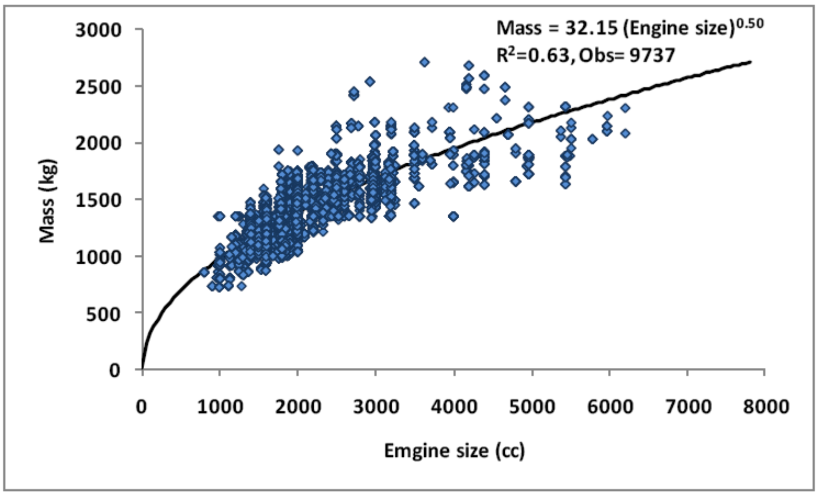


Figure 4.9: Relationship between vehicle mass and engine size

Four hypotheses were made on the relationship between mass and fuel consumption, and engine size and fuel consumption as suggested by Figures 4.4 to 4.6:

1. The relationship between engine size and fuel consumption is different for different fuel types.

---

[1] The statistical significance is defined at $\alpha=0.05$ throughout this thesis unless otherwise stated

2. The relationship between engine size and fuel consumption is different for different transmission types.
3. The relationship between mass and fuel consumption is different for different fuel types.
4. The relationship between mass and fuel consumption is different for different transmission types.

To examine these hypotheses, interaction terms between these variables were included in the models.

Tables 4.9 and 4.10 show final model estimation results for urban and extra-urban driving cycles, respectively. For each driving cycle three different models were estimated. Model A includes vehicle design factors (i.e. engine size, mass, fuel type, transmission type, year on the market, and Euro emission standard that the vehicle meets) as explanatory variables. The potential weakness associated with this model is the expected high correlation between the error terms of many variants of a vehicle make or a vehicle make and model, and hence violating the regression assumption of independent errors. Models B and C include vehicle design factors included in Model A as well as fixed effects for vehicle makes and vehicle makes and models, respectively, as explanatory variables. This can eliminate the potential issue of correlation of errors in Model A as explained.

In general, the significant improvement in deviance from the null model in all models shows a good performance of the estimated models. Note that the deviance of the perfect model is zero. Therefore, in the urban fuel consumption model A, for example, inclusion of the explanatory variables has reduced the deviance from 778 (deviance of the model with no explanatory variable) to 71. In both urban and extra-urban driving cycles, the goodness of fit of the models (measured through the AIC) is substantially improved from Model A to Model C.

Table 4.9: Model estimation results for urban driving cycle (dependent variable: log (µ) where µ is mean fuel consumption rate in l/100km)

Variable		Model A				Model B				Model C			
		Coef.	Std. Error	t-stat	p-value	Coef.	Std. Error	t-stat	p-value	Coef.	Std. Error	t-stat	p-value
Constant		-2.08	0.05	-40.28	0.000	-2.14	0.05	-43.20	0.000	-1.86	0.05	-38.78	0.000
Ln (Engine size)		0.56	0.01	69.75	0.000	0.57	0.01	74.62	0.000	0.55	0.01	77.06	0.000
Mass		0.00022	0.00001	23.57	0.000	0.00021	0.00001	23.47	0.000	0.00009	0.00001	8.33	0.000
Diesel		-0.46	0.09	-5.27	0.000	-0.71	0.08	-8.57	0.000	-0.56	0.08	-7.28	0.000
Euro II		0.054	0.004	12.93	0.000	0.041	0.004	9.64	0.000	0.020	0.004	4.89	0.000
Euro III		0.016	0.002	7.36	0.000	0.018	0.002	7.72	0.000	0.010	0.002	4.51	0.000
Time		-0.009	0.001	-16.08	0.000	-0.008	0.001	-15.63	0.000	-0.008	0.000	-15.95	0.000
Automatic transmission		1.92	0.07	28.17	0.000	1.86	0.06	29.08	0.000	1.71	0.06	28.20	0.000
Other transmission		-0.32	0.12	-2.68	0.007	0.00	0.11	-0.01	0.991	-0.09	0.10	-0.91	0.362
Diesel x Ln (engine size)		-0.02	0.01	-1.38	0.168	0.02	0.01	1.75	0.085	0.00	0.01	0.37	0.712
Automatic x Ln (engine size)		-0.26	0.01	-24.48	0.000	-0.24	0.01	-24.94	0.000	-0.23	0.01	-24.82	0.000
Other x Ln (engine size)		0.04	0.02	2.15	0.032	-0.01	0.02	-0.70	0.485	0.00	0.02	0.28	0.783
Diesel x mass		0.00014	0.00001	12.08	0.000	0.00011	0.00001	9.56	0.000	0.00010	0.00001	9.30	0.000
Automatic x mass		0.00009	0.00001	7.72	0.000	0.00007	0.00001	6.64	0.000	0.00010	0.00001	9.47	0.000
Other x mass		0.00003	0.00002	1.46	0.144	0.00009	0.00002	4.05	0.000	0.00007	0.00002	3.21	0.001
Manufacturer / Make and model fixed effects													
ALFA	156	-	-	-	-	0.091	0.008	0.000	-1.86	0.096	0.007	13.701	0.000
AUDI	A3	-	-	-	-	0.002	0.004	0.549	0.583	-0.049	0.006	-7.617	0.000
	A4	-	-	-	-					0.010	0.008	1.268	0.205
	A6	-	-	-	-					0.046	0.005	8.987	0.000
	TT	-	-	-	-					0.083	0.012	6.793	0.000
BMW	300	-	-	-	-	-0.072	0.003	-21.069	0.000	-0.052	0.004	-13.685	0.000
	500	-	-	-	-					-0.033	0.005	-6.379	0.000
	700	-	-	-	-					-0.055	0.010	-5.662	0.000
CITROEN	C3	-	-	-	-	-0.024	0.004	-5.629	0.000	-0.164	0.010	-16.759	0.000
	C5	-	-	-	-					-0.010	0.007	-1.423	0.155
	PICASSO	-	-	-	-					-0.039	0.010	-3.862	0.000
	SAXO	-	-	-	-					0.013	0.013	1.035	0.301
	SYNERGIE/	-	-	-	-					0.057	0.019	3.030	0.002
	XANTIA	-	-	-	-					0.043	0.012	3.540	0.000
	XSARA	-	-	-	-					-0.044	0.007	-6.316	0.000

Table 4.9 (continued)

DAEWOO	LANOS	-	-	-	-	0.089	0.011	8.349	0.000	0.095	0.020	4.679	0.000
	MATIZ	-	-	-	-					0.142	0.023	6.227	0.000
	NUBIRA	-	-	-	-					0.030	0.013	2.432	0.015
DAIHATSU	CHARADE	-	-	-	-	-0.083	0.021	-4.012	0.000	-0.150	0.019	-7.744	0.000
FIAT	BRAVO	-	-	-	-	0.040	0.009	4.451	0.000	0.031	0.012	2.526	0.012
	PUNTO	-	-	-	-					-0.069	0.013	-5.154	0.000
	SEICENTO	-	-	-	-					0.057	0.019	2.946	0.003
FORD	ESCORT	-	-	-	-	-0.014	0.003	-4.221	0.000	0.039	0.031	1.244	0.214
	FIESTA	-	-	-	-					-0.083	0.007	-12.685	0.000
	FOCUS	-	-	-	-					-0.054	0.005	-10.979	0.000
	GALAXY	-	-	-	-					0.043	0.008	5.296	0.000
	KA	-	-	-	-					-0.007	0.012	-0.557	0.578
	MAVERICK	-	-	-	-					0.144	0.035	4.100	0.000
	MONDETO	-	-	-	-					0.025	0.004	5.795	0.000
	PUMA	-	-	-	-					-0.064	0.027	-2.409	0.016
HONDA	ACCORD	-	-	-	-	-0.072	0.005	-15.056	0.000	-0.011	0.006	-1.881	0.060
	CIVIC	-	-	-	-					-0.159	0.007	-24.046	0.000
	CRV	-	-	-	-					-0.041	0.015	-2.785	0.005
HYUNDAI	ACCENT	-	-	-	-	-0.014	0.009	-1.672	0.095	-0.053	0.012	-4.521	0.000
	ATOZ	-	-	-	-					-0.024	0.035	-0.692	0.489
	COUPE	-	-	-	-					-0.011	0.011	-1.004	0.315
	LANTRA	-	-	-	-					0.002	0.035	0.070	0.944
ISUZU	TROOPER	-	-	-	-	0.072	0.010	6.965	0.000	0.154	0.010	14.997	0.000
JAGUAR	S TYPE	-	-	-	-	-0.019	0.008	-2.565	0.010	0.000	0.012	-0.034	0.973
	X TYPE	-	-	-	-					0.017	0.011	1.617	0.106
	XJ	-	-	-	-					0.017	0.013	1.262	0.207
CHRYSLER	CHEROKEE	-	-	-	-	0.031	0.011	2.806	0.005	0.090	0.013	7.071	0.000
	GRAND	-	-	-	-					0.112	0.017	6.554	0.000
LANDROVER	DEFENDER	-	-	-	-	0.049	0.008	6.081	0.000	0.189	0.020	9.678	0.000
	DISCOVERY	-	-	-	-					0.101	0.014	7.159	0.000
	FREELANDE	-	-	-	-					0.157	0.012	13.404	0.000
	RANGEROV	-	-	-	-					0.090	0.015	5.947	0.000
LEXUS	IS200	-	-	-	-	0.091	0.021	4.264	0.000	0.102	0.019	5.254	0.000
MAZDA	626	-	-	-	-	0.000	0.011	0.039	0.969	-0.106	0.015	-7.037	0.000
	MX-5	-	-	-	-					0.037	0.013	2.959	0.003

Table 4.9 (continued)

MERCEDES	A CLASS	-	-	-	-	-0.048	0.003	-15.063	0.000	-0.087	0.005	-17.071	0.000
	C CLASS	-	-	-	-					-0.003	0.005	-0.497	0.619
	E CLASS	-	-	-	-					-0.006	0.005	-1.341	0.180
	ML CLASS	-	-	-	-					0.014	0.011	1.251	0.211
	S CLASS	-	-	-	-					-0.019	0.008	-2.310	0.021
MG	MGF	-	-	-	-	0.012	0.029	0.413	0.221	0.012	0.029	0.413	0.680
MINI	MINI	-	-	-	-	0.022	0.010	2.270	0.000	0.022	0.010	2.270	0.000
MITSUBISHI	CARISMA	-	-	-	-	-0.009	0.006	-1.500	0.134	-0.061	0.012	-5.020	0.000
	SHOGUN	-	-	-	-					0.062	0.007	8.524	0.000
	SPACEWAG	-	-	-	-					-0.026	0.019	-1.358	0.174
NISSAN	ALMERA	-	-	-	-	-0.046	0.006	-8.358	0.000	-0.065	0.009	-6.988	0.000
	MICRA	-	-	-	-					-0.145	0.010	-14.243	0.000
	PRIMERA	-	-	-	-					-0.061	0.009	-7.012	0.000
	SERENA	-	-	-	-					0.198	0.035	5.652	0.000
	TERRANO	-	-	-	-					0.125	0.012	10.155	0.000
PEUGEOT	106	-	-	-	-	-0.039	0.004	-8.851	0.000	0.035	0.020	1.737	0.082
	205	-	-	-	-					-0.069	0.008	-8.523	0.000
	306	-	-	-	-					0.003	0.012	0.295	0.768
	307	-	-	-	-					-0.055	0.007	-7.645	0.000
	406	-	-	-	-					-0.048	0.006	-7.534	0.000
PROTON	PERSONA	-	-	-	-	0.058	0.013	4.647	0.000	0.042	0.012	3.630	0.000
RENAULT	CLIO	-	-	-	-	-0.022	0.004	-6.110	0.000	-0.070	0.008	-8.430	0.000
	ESPACE	-	-	-	-					0.090	0.009	10.104	0.000
	LAGUNA	-	-	-	-					-0.009	0.005	-1.701	0.089
	MEGANE	-	-	-	-					-0.071	0.006	-12.779	0.000
	SCENIC	-	-	-	-					-0.028	0.007	-3.888	0.000
ROVER	200/400	-	-	-	-	0.013	0.007	1.870	0.061	-0.025	0.017	-1.478	0.139
	25/45	-	-	-	-					-0.072	0.017	-4.219	0.000
	75	-	-	-	-					0.046	0.007	6.454	0.000
SAAB	9-3	-	-	-	-	0.028	0.005	5.844	0.000	0.048	0.006	8.255	0.000
	9-5	-	-	-	-					0.053	0.006	8.315	0.000
SEAT	LBIZ/CO	-	-	-	-	-0.020	0.007	-2.921	0.004	-0.049	0.015	-3.309	0.001
	LEON	-	-	-	-					-0.020	0.009	-2.227	0.026
	TOLEDO	-	-	-	-					-0.031	0.010	-2.963	0.003

Table 4.9 (continued)

SKODA	FABIA	-	-	-	-	-0.053	0.005	-11.680	0.000	-0.072	0.006	-11.775	0.000
	FELICIA	-	-	-	-					-0.002	0.029	-0.066	0.947
	OCTAVIA	-	-	-	-					-0.055	0.005	-10.266	0.000
SUBARU	IMPREZA	-	-	-	-	-0.019	0.006	-3.390	0.001	0.111	0.010	11.152	0.000
	LEGACY	-	-	-	-					-0.045	0.006	-7.828	0.000
SUZUKI	BALENO	-	-	-	-	-0.047	0.012	-3.958	0.000	-0.065	0.019	-3.428	0.001
	SWIFT	-	-	-	-					-0.105	0.014	-7.497	0.000
	VITARA	-	-	-	-					0.032	0.050	0.642	0.521
TOYOTA	AVENSIS	-	-	-	-	-0.064	0.004	-14.659	0.000	-0.069	0.007	-9.408	0.000
	CELICA	-	-	-	-					0.005	0.018	0.256	0.798
	COROLLA	-	-	-	-					-0.073	0.007	-10.192	0.000
	LAND	-	-	-	-					0.009	0.012	0.770	0.441
	MR2	-	-	-	-					-0.061	0.027	-2.302	0.021
	PREVIA	-	-	-	-					-0.004	0.015	-0.258	0.796
	RAV-4	-	-	-	-					-0.028	0.014	-1.991	0.047
	YARIS	-	-	-	-					-0.149	0.009	-16.204	0.000
VAUXHALL	ASTRA	-	-	-	-	-0.004	0.003	-1.255	0.210	-0.040	0.004	-9.138	0.000
	CORSA	-	-	-	-					-0.055	0.005	-10.041	0.000
	FRONTERA	-	-	-	-					0.183	0.011	17.026	0.000
	OMEGA	-	-	-	-					0.072	0.007	9.697	0.000
	VECTRA	-	-	-	-					-0.002	0.004	-0.585	0.559
VOLKSWAGEN	BEETLE	-	-	-	-	-0.006	0.004	-1.760	0.078	0.015	0.008	1.872	0.061
	GOLF/JE	-	-	-	-					-0.033	0.005	-7.279	0.000
	PASSAT	-	-	-	-					0.051	0.005	9.440	0.000
	POLO	-	-	-	-					-0.081	0.008	-10.216	0.000
VOLVO	SV40	-	-	-	-	-	-	-	-	0.014	0.010	1.417	0.156
	V70	-	-	-	-					0.021	0.009	2.250	0.024
Model statistics													
Observations		9737				9737				9737			
Null deviance		778				778				778			
Residual Deviance		71				60				47			
Log L value		-12900				-12081				-10941			
AIC		25832				24260				22137			

Table 4.10: Model estimation results for extra-urban driving cycle (dependent variable: log (μ) where μ is mean fuel consumption rate in l/100km)

Variable		Model A				Model B				Model C			
		Coef.	Std. Error	t-stat	p-value	Coef.	Std. Error	t-stat	p-value	Coef.	Std. Error	t-stat	p-value
Constant		-1.04	0.05	-20.17	0.000	-1.05	0.05	-22.89	0.000	-0.85	0.04	-20.83	0.000
Ln (Engine size)		0.34	0.01	42.56	0.000	0.35	0.01	48.96	0.000	0.34	0.01	55.93	0.000
Mass		0.00025	0.00001	27.19	0.000	0.00025	0.00001	30.35	0.000	0.00013	0.00001	13.99	0.000
Diesel		0.04	0.09	0.49	0.627	-0.22	0.08	-2.84	0.004	-0.11	0.07	-1.62	0.106
Euro II		0.07	0.00	17.40	0.000	0.036	0.004	9.21	0.000	0.012	0.003	3.42	0.001
Euro III		0.02	0.00	8.93	0.000	0.006	0.002	3.00	0.003	-0.004	0.002	-1.96	0.050
Time		-0.01	0.00	-12.87	0.000	-0.008	0.000	-15.58	0.000	-0.008	0.000	-18.09	0.000
Automatic transmission		1.31	0.07	19.20	0.000	1.19	0.06	20.07	0.000	1.02	0.05	19.72	0.000
Other transmission		-0.61	0.12	-5.13	0.000	-0.40	0.10	-3.85	0.000	-0.37	0.09	-4.30	0.000
Diesel x Ln (engine size)		-0.09	0.01	-6.93	0.000	-0.04	0.01	-3.20	0.000	-0.04	0.01	-4.46	0.000
Automatic x Ln (engine size)		-0.18	0.01	-17.48	0.000	-0.16	0.01	-17.46	0.000	-0.14	0.01	-17.59	0.000
Other x Ln (engine size)		0.11	0.02	5.62	0.000	0.06	0.02	3.97	0.000	0.05	0.01	3.82	0.000
Diesel x mass		0.00025	0.00001	20.62	0.000	0.00015	0.00001	13.87	0.000	0.00011	0.00001	11.90	0.000
Automatic x mass		0.00010	0.00001	8.51	0.000	0.00006	0.00001	5.57	0.000	0.00007	0.00001	8.10	0.000
Other x mass		-0.00011	0.00002	-4.74	0.000	-0.00004	0.00002	-1.92	0.055	0.00001	0.00002	0.73	0.463
Manufacturer / Make and model fixed effects													
ALFA	156	-	-	-	-	0.026	0.007	3.662	0.000	0.024	0.006	3.930	0.000
AUDI	A3	-	-	-	-	-0.042	0.004	-10.852	0.000	-0.075	0.006	-13.464	0.000
	A4	-	-	-	-					-0.054	0.007	-8.231	0.000
	A6	-	-	-	-					-0.021	0.004	-4.774	0.000
	TT	-	-	-	-					0.037	0.010	3.534	0.000
BMW	300	-	-	-	-	-0.093	0.003	-29.258	0.000	-0.084	0.003	-25.987	0.000
	500	-	-	-	-					-0.060	0.004	-13.560	0.000
	700	-	-	-	-					-0.084	0.008	-10.112	0.000
CITROEN	C3	-	-	-	-	-0.054	0.004	-13.524	0.000	-0.128	0.008	-15.209	0.000
	C5	-	-	-	-					-0.074	0.006	-12.385	0.000
	PICASSO	-	-	-	-					-0.055	0.009	-6.326	0.000
	SAXO	-	-	-	-					-0.021	0.011	-1.948	0.051
	SYNERGIE/	-	-	-	-					0.071	0.016	4.401	0.000
	XANTIA	-	-	-	-					-0.023	0.011	-2.156	0.031
	XSARA	-	-	-	-					-0.096	0.006	-15.925	0.000

Table 4.10 (continued)

DAEWOO	LANOS	-	-	-	-	0.033	0.010	3.296	0.001	-0.032	0.017	-1.813	0.070
	MATIZ	-	-	-	-					0.087	0.019	4.470	0.000
	NUBIRA	-	-	-	-					-0.007	0.011	-0.634	0.526
DAIHATSU	CHARADE	-	-	-	-	-0.019	0.019	-0.965	0.335	-0.089	0.017	-5.372	0.000
FIAT	BRAVO	-	-	-	-	0.008	0.008	0.936	0.349	-0.024	0.011	-2.241	0.025
	PUNTO	-	-	-	-					-0.059	0.012	-5.100	0.000
	SEICENTO	-	-	-	-					-0.023	0.017	-1.370	0.171
FORD	ESCORT	-	-	-	-	-0.039	0.003	-12.946	0.000	0.045	0.027	1.689	0.091
	FIESTA	-	-	-	-					-0.077	0.006	-13.729	0.000
	FOCUS	-	-	-	-					-0.078	0.004	-18.533	0.000
	GALAXY	-	-	-	-					0.050	0.007	7.242	0.000
	KA	-	-	-	-					-0.040	0.010	-3.973	0.000
	MAVERICK	-	-	-	-					0.124	0.030	4.133	0.000
	MONDETO	-	-	-	-					-0.036	0.004	-9.533	0.000
	PUMA	-	-	-	-					-0.018	0.023	-0.775	0.438
HONDA	ACCORD	-	-	-	-	-0.043	0.004	-9.850	0.000	-0.031	0.005	-6.212	0.000
	CIVIC	-	-	-	-					-0.101	0.006	-17.892	0.000
	CRV	-	-	-	-					0.093	0.013	7.349	0.000
HYUNDAI	ACCENT	-	-	-	-	-0.001	0.008	-0.139	0.890	-0.049	0.010	-4.863	0.000
	ATOZ	-	-	-	-					0.038	0.030	1.268	0.205
	COUPE	-	-	-	-					0.001	0.010	0.151	0.880
	LANTRA	-	-	-	-					-0.075	0.030	-2.502	0.012
ISUZU	TROOPER	-	-	-	-	0.149	0.010	15.443	0.000	0.229	0.009	25.949	0.000
JAGUAR	S TYPE	-	-	-	-	-0.054	0.007	-7.791	0.000	-0.040	0.010	-4.069	0.000
	X TYPE	-	-	-	-					-0.044	0.009	-4.891	0.000
	XJ	-	-	-	-					-0.011	0.011	-0.991	0.322
CHRYSLER	CHEROKEE	-	-	-	-	0.123	0.010	12.107	0.000	0.162	0.011	14.891	0.000
	GRAND	-	-	-	-					0.225	0.015	15.317	0.000
LANDROVER	DEFENDER	-	-	-	-	0.134	0.007	18.014	0.000	0.392	0.017	23.414	0.000
	DISCOVERY	-	-	-	-					0.218	0.012	18.029	0.000
	FREELANDE	-	-	-	-					0.214	0.010	21.298	0.000
	RANGEROV	-	-	-	-					0.133	0.013	10.215	0.000
LEXUS	IS200	-	-	-	-	0.104	0.020	5.247	0.000	0.110	0.017	6.567	0.000
MAZDA	626	-	-	-	-	0.053	0.010	5.431	0.000	-0.048	0.013	-3.762	0.000
	MX-5	-	-	-	-					0.078	0.011	7.222	0.000

Table 4.10 (continued)

MERCEDES	A CLASS	-	-	-	-	-0.070	0.003	-23.629	0.000	-0.077	0.004	-17.520	0.000
	C CLASS	-	-	-	-					-0.035	0.004	-8.091	0.000
	E CLASS	-	-	-	-					-0.061	0.004	-15.634	0.000
	ML CLASS	-	-	-	-					0.084	0.010	8.701	0.000
	S CLASS	-	-	-	-					-0.093	0.007	-13.101	0.000
MG	MGF	-	-	-	-	-0.051	0.029	-1.744	0.081	-0.076	0.025	-3.085	0.002
MINI	MINI	-	-	-	-	0.010	0.010	0.995	0.320	-0.029	0.008	-3.460	0.000
MITSUBISHI	CARISMA	-	-	-	-	0.098	0.006	16.695	0.000	-0.077	0.010	-7.444	0.000
	SHOGUN	-	-	-	-					0.197	0.006	31.400	0.000
	SPACEWAG	-	-	-	-					0.066	0.016	4.113	0.000
NISSAN	ALMERA	-	-	-	-	0.021	0.005	4.071	0.000	-0.053	0.008	-6.740	0.000
	MICRA	-	-	-	-					-0.057	0.009	-6.511	0.000
	PRIMERA	-	-	-	-					-0.028	0.007	-3.718	0.000
	SERENA	-	-	-	-					0.258	0.030	8.593	0.000
	TERRANO	-	-	-	-					0.274	0.011	25.870	0.000
PEUGEOT	106	-	-	-	-	-0.061	0.004	-14.824	0.000	-0.024	0.017	-1.384	0.166
	205	-	-	-	-					-0.084	0.007	-12.118	0.000
	306	-	-	-	-					-0.050	0.010	-4.944	0.000
	307	-	-	-	-					-0.072	0.006	-11.538	0.000
	406	-	-	-	-					-0.084	0.005	-15.531	0.000
PROTON	PERSONA	-	-	-	-	0.091	0.012	7.754	0.000	0.068	0.010	6.820	0.000
RENAULT	CLIO	-	-	-	-	-0.019	0.003	-5.848	0.000	-0.052	0.007	-7.348	0.000
	ESPACE	-	-	-	-					0.106	0.008	13.864	0.000
	LAGUNA	-	-	-	-					-0.041	0.004	-9.521	0.000
	MEGANE	-	-	-	-					-0.075	0.005	-15.735	0.000
	SCENIC	-	-	-	-					0.002	0.006	0.345	0.730
ROVER	200/400	-	-	-	-	-0.054	0.006	-8.736	0.000	-0.101	0.015	-6.832	0.000
	25/45	-	-	-	-					-0.104	0.015	-7.047	0.000
	75	-	-	-	-					-0.033	0.006	-5.394	0.000
SAAB	9-3	-	-	-	-	-0.053	0.005	-11.658	0.000	-0.029	0.005	-5.689	0.000
	9-5	-	-	-	-					-0.046	0.006	-8.427	0.000
SEAT	LBIZ/CO	-	-	-	-	-0.040	0.006	-6.261	0.000	-0.079	0.013	-6.194	0.000
	LEON	-	-	-	-					-0.050	0.008	-6.477	0.000
	TOLEDO	-	-	-	-					-0.052	0.009	-5.793	0.000

Table 4.10 (continued)

SKODA	FABIA	-	-	-	-	-0.078	0.004	-18.438	0.000	-0.093	0.005	-17.837	0.000
	FELICIA	-	-	-	-					-0.084	0.025	-3.413	0.001
	OCTAVIA	-	-	-	-					-0.091	0.005	-20.011	0.000
SUBARU	IMPREZA	-	-	-	-	-0.008	0.005	-1.520	0.129	0.088	0.009	10.373	0.000
	LEGACY	-	-	-	-					-0.034	0.005	-6.877	0.000
SUZUKI	BALENO	-	-	-	-	0.017	0.011	1.525	0.127	0.017	0.016	1.015	0.310
	SWIFT	-	-	-	-					-0.061	0.012	-5.062	0.000
	VITARA	-	-	-	-					0.175	0.042	4.113	0.000
TOYOTA	AVENSIS	-	-	-	-	0.012	0.004	2.866	0.004	-0.043	0.006	-6.919	0.000
	CELICA	-	-	-	-					0.011	0.016	0.720	0.471
	COROLLA	-	-	-	-					-0.013	0.006	-2.102	0.036
	LAND	-	-	-	-					0.175	0.010	16.785	0.000
	MR2	-	-	-	-					-0.056	0.023	-2.458	0.014
	PREVIA	-	-	-	-					0.094	0.013	7.080	0.000
	RAV-4	-	-	-	-					0.068	0.012	5.756	0.000
	YARIS	-	-	-	-					-0.057	0.008	-7.308	0.000
VAUXHALL	ASTRA	-	-	-	-	-0.046	0.003	-16.211	0.000	-0.079	0.004	-21.150	0.000
	CORSA	-	-	-	-					-0.103	0.005	-22.047	0.000
	FRONTERA	-	-	-	-					0.219	0.009	23.812	0.000
	OMEGA	-	-	-	-					0.021	0.006	3.327	0.001
	VECTRA	-	-	-	-					-0.070	0.003	-20.188	0.000
VOLKSWAGEN	BEETLE	-	-	-	-	-0.039	0.003	-11.416	0.000	-0.014	0.007	-2.055	0.040
	GOLF/JE	-	-	-	-					-0.060	0.004	-15.330	0.000
	PASSAT	-	-	-	-					-0.014	0.005	-2.994	0.003
	POLO	-	-	-	-					-0.104	0.007	-15.346	0.000
VOLVO	SV40	-	-	-	-	-	-	-	-	-0.076	0.009	-8.813	0.000
	V70	-	-	-	-					0.027	0.008	3.420	0.001
Model statistics													
Observations		9737				9737				9737			
Null deviance		478				478				478			
Residual Deviance		68				50				35			
Log L value		-7422				-5972				-4110			
AIC		14876				12042				8473			

Comparison of Model B and Model C for both urban and extra-urban fuel consumption models shows that the fixed effects are different for different makes and models within manufacturers. This is confirmed by the substantially better performance of Model C compared to that of Model B (measured by log likelihood through the AIC) and the fact that the majority of the estimated fixed effects for makes and models in Model C (about 75% in urban fuel consumption model and 90% in extra-urban fuel consumption model) are statistically significant at 5% level as shown by the p-values in Tables 4.9 and 4.10. In several cases, the estimated fixed effects for different car models of a manufacturer are in different directions.

For each of the urban and extra-urban models, comparison of the estimated coefficients of vehicle design variables between Model A and C shows that most of the coefficients have similar values except the coefficient of “Mass” which is reduced substantially by adding fixed effects for makes and models. This suggests that the estimated fixed effects for makes and models are substantially related to the effect of vehicle mass. In fact, whilst the estimated coefficient of “Mass” in Model A is related to both between and within make and model effects, this coefficient in Model C reflects within make and model effects of mass only. This is consistent with the main intended use of the estimated models which is to predict the effect on fuel consumption of a change in design by manufacturers to reduce mass in any make and model. The comparisons made between the estimated models clearly show that, for both urban and extra-urban models, Model C is the best model reflecting variation in fuel consumption rate by vehicle design factors.

One assumption that is made in the model estimation process is that the effects of mass and engine size on fuel consumption are similar for different makes and models. To check this assumption, the interactions between each of mass and engine size, and make and model variables were added to Model C. The results showed that the estimated coefficients for the majority of makes and models were not statistically significant at 5% level. In addition to this, a few makes and models with a reasonable number of variants in the data were chosen at random and three different fuel consumption models (for urban driving cycle) were fitted to each of them to include each of “Mass” and “Engine size” separately as well as “Mass” and “Engine size” together as the explanatory variables. The estimation results, summarised in Appendix 3 of this thesis, showed that while the estimated coefficients are not stable in many cases due to lack of sufficient sample sizes, the overall results suggested that the observed variation in the estimated effects of mass and engine size between different makes and models are relatively small and randomly distributed rather than being systematically related to a design factor such as mass or engine size. These results support the approach taken of estimating one grand model for all makes and models in the dataset.

The plot of observed versus predicted values based on Model C are shown in Figures 4.10 and 4.11 for urban and extra-urban fuel consumption rate, respectively. Note that the scale is different for the two plots depending on the range of fuel consumption values. A good fit between measured values of fuel consumption and estimated values by the models shows that the developed models can estimate an average fuel consumption for a make and model with a given mass, engine size, fuel type, transmission type, and year of manufacture reasonably well in the defined urban and extra-urban driving cycles.

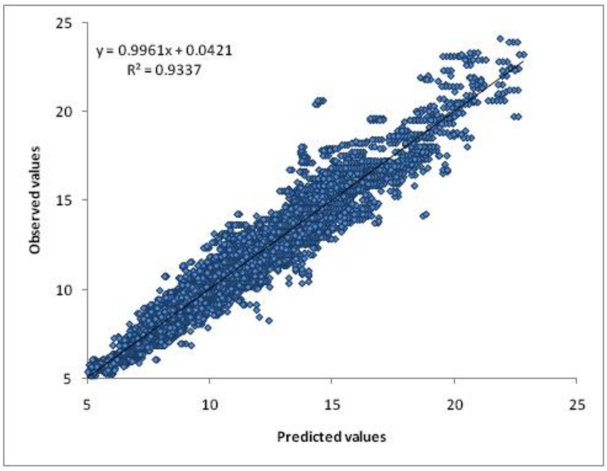


Figure 4.10: Observed values versus estimated values – urban fuel consumption model

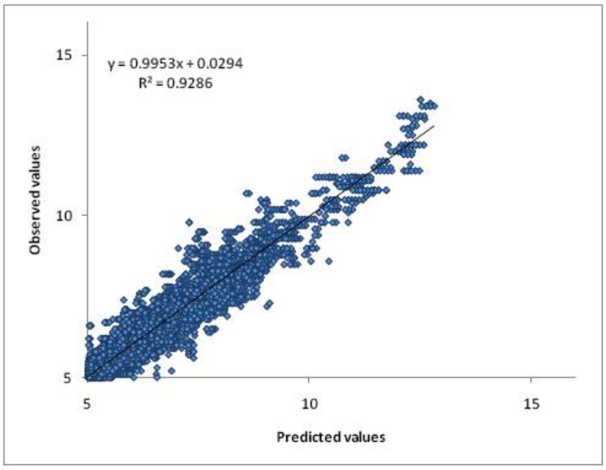


Figure 4.11: Observed values versus estimated values – extra-urban fuel consumption model

4.4. Interpretation of modelling results

Two sets of models were developed, one for each of urban and extra-urban cycles, to investigate the effects of car design on fuel consumption, while vehicle mass is the variable of interest in this study. The signs of coefficient estimates are directly related to their influence on fuel consumption. Since the link function in the estimated models is a log function, an increase of b in the value of an explanatory variable with a coefficient β, while all other variables are constant, increases the dependent variable by $\left(e^{\beta.b} - 1\right) \times 100$ percent. All the estimated effects in the following sections are based on Model C. The effects of mass are discussed first, followed by a discussion on the effects of other design variables.

4.4.1. Effect of vehicle mass

The effect of vehicle mass on fuel consumption was found to be statistically significant at 5% level. As explained earlier, the estimated coefficients of “Mass” in Models C reflect the effect of mass within makes and models. As expected, a positive coefficient of “Mass” in both urban and extra-urban models confirms that fuel consumption increases as mass increases. In most cases, interaction terms between mass and each of fuel and transmission type are significant at 5% level showing that the partial effect of mass is different for different fuel and transmission types. In both models, the interaction term between mass and diesel fuel type is positive and statistically significant at 5% level which shows that the effect of mass on fuel consumption of diesel cars is greater than that of petrol cars of the same design in both driving cycles. This could be partly a consequence of different characteristics of Diesel and petrol engines and their relationship with vehicle mass. This result suggests that the potential for saving fuel by switching from petrol to diesel is greater in lighter cars and decreases as vehicle mass increases. Results for the interaction terms between mass and transmission type show that in both driving cycles, mass in automatic cars has a greater influence on fuel consumption compared to manual cars. The effect of mass on cars with transmission types other than manual or automatic (see Section 4.3) is not significantly different from that of manual cars for the extra-urban cycle; however, this effect is greater in the case of the urban cycle.

Table 4.11 shows the partial effects of mass on fuel consumption for different combinations of fuel and transmission. These effects are calculated for both driving cycles as the percent change in fuel consumption within makes and models caused by a 100 kg increase in mass of model variants when vehicle design and all other contributing factors are held constant. The greatest partial effect of mass was found for automatic diesel cars when, for a make and model, a 100 kg increase in mass of its model variants would increase their typical urban and extra-urban fuel consumption by 2.9% and 3.1%, respectively. The results also show that the effect of mass on fuel consumption is greater in the extra-urban driving cycle than, as might be expected, in the urban cycle. Vehicle mass directly contributes to rolling resistance; this increases with speed due to an increase in the work being done in deforming the tyre over a given time (Heisler, 2002). This can result in a greater influence of mass on fuel consumption in extra-urban driving which includes a higher average speed than does urban driving.

Table 4.11: Partial effects of mass - percent change in fuel consumption with 100 kg increase in mass

Fuel type	Transmission type	Percent change in fuel consumption	
		Urban cycle	Extra-urban cycle
Petrol	Manual	0.90	1.29
	Automatic	1.88	2.02
	Other	1.57	NS[1]
Diesel	Manual	1.89	2.39
	Automatic	2.89	3.12
	Other	2.57	NS

1. The effect is not significantly different from the reference case (manual petrol car)

4.4.2. Effects of other design factors

Since a logarithmic transformation of “Engine size” is used in the models as an explanatory variable, the coefficient of this variable shows the elasticity of fuel consumption with respect to engine size. Positive values for this variable confirm that fuel consumption increases as engine size increases in both driving cycles. Table 4.12 shows the estimated elasticity of fuel consumption with respect to engine size for different combinations of fuel and transmission type when all other design factors are held constant. The effect of engine size on fuel consumption is not significantly different (at 5% level) for diesel cars from that of petrol cars for the urban cycle, while this effect is less for diesel cars for the extra-urban cycle. In both driving cycles, the effect of engine size is greater for manual petrol cars compared to automatic petrol cars.

In a manual petrol car, increasing engine size by 10% while holding all other design features constant, increases fuel consumption in urban and extra-urban driving cycle by 5.5% and 3.4%, respectively. Unlike vehicle mass, the effect of engine size on fuel consumption was found to be greater in the urban cycle than in the extra-urban cycle. The reason for this could be the better optimised performance of engines at higher speeds as well as the presence of less acceleration and idling in the extra-urban cycle compared to the urban driving cycle.

Table 4.12: Estimated elasticities of fuel consumption with respect to engine size by fuel and transmission type for different driving cycles

Fuel type	Transmission type	Estimated elasticities of fuel consumption	
		Urban cycle	Extra-urban cycle
Petrol	Manual	0.55	0.34
	Automatic	0.33	0.21
	Other	NS[1]	0.40
Diesel	Manual	NS	0.30
	Automatic	NS	0.16
	Other	NS	0.35

1. The effect is not significantly different from the reference case (manual petrol car)

As the interaction terms between fuel and mass suggest (Tables 4.9 and 4.10), the effect of fuel type on fuel consumption is greater for lighter vehicles in both cycles and this effect decreases as mass increases. It is notable that variables related to Time and the Euro emission standard that a vehicle satisfies affect fuel consumption in the same direction (Tables 4.9 and 4.10). Cars that meet Euro II and Euro III emission standards consume more fuel than cars meeting Euro IV emission standards. This is expected as many design technologies aim to reduce emissions by reducing vehicle fuel consumption. The negative coefficients for the “Time” variable show that even within each Euro standard, fuel consumption decreases for new vehicles coming to the market each year compared to the previous year. This is the result of technological improvements to vehicle design. Findings on the effects of transmission type on fuel consumption show that a car with automatic transmission consumes more fuel compared to a car with the same design but manual transmission for the both driving cycles. The size of this effect depends on the values of mass and engine size for a given fuel type. The fuel consumed by manual cars in practice might be slightly greater than that measured during the test as drivers under test conditions are likely to be more efficient in changing gears in order to optimize fuel consumption.

4.4.3. Effects of makes and models

Estimation results of Models C in Tables 4. 9 and 4.10 showed fixed effects of makes and models on fuel consumption. It was also explained that these estimated effects include the effects of mass between makes and models. Figure 4.12 shows, for both urban and extra-urban driving cycles, plots of the estimated fixed effects of makes and models that are statistically significant at 5% level (in terms of the percent change in fuel consumption from the mean fuel consumption of makes and models) against the average mass of the particular makes and models in the dataset. For both cases, the plots generally show a positive correlation between mass and the estimated effects; that is, the greater the mass of a make and model, the higher its fuel consumption rate. This supports the conclusion that the estimated effects for makes and models shown in Tables 4.9 and 4.10 are largely related to the effect of mass between makes and models while the estimated coefficient of “Mass” in Model C reflects the effect of mass within makes and models.

Despite the general positive correlation between the estimated fixed effects of makes and models and their average mass reflected in Figure 4.12, the plots show that a few makes and models have a considerably higher or lower fuel consumption compared to other makes and models in a similar range of average mass. Generally, the discrepancies in the estimated effects of makes and models that have a similar average mass is partly explained by the fact that the distribution of mass within makes and models are different for different makes and models in the fuel consumption data. For example, examination of data shows that Mercedes S Class and BMW 700 (both of which show a lower fuel consumption than that of models with a similar average mass) have a mass distribution that is positively skewed. Therefore, the average mass of these makes and models is greater than the predominant mass of the variants of these makes and models in the data. In fact, the estimated effects for these makes and models are based on the observations (model variants) the majority of which have a mass less than the average mass of these makes and models in the data. On the other hand, models showing a considerably higher fuel consumption than other models with a similar mass (such as Daewoo Matiz and Nissan Serena) show to have a mass distribution that is negatively skewed (i.e. the majority of their model variants in the data have a mass that is greater than their average mass).

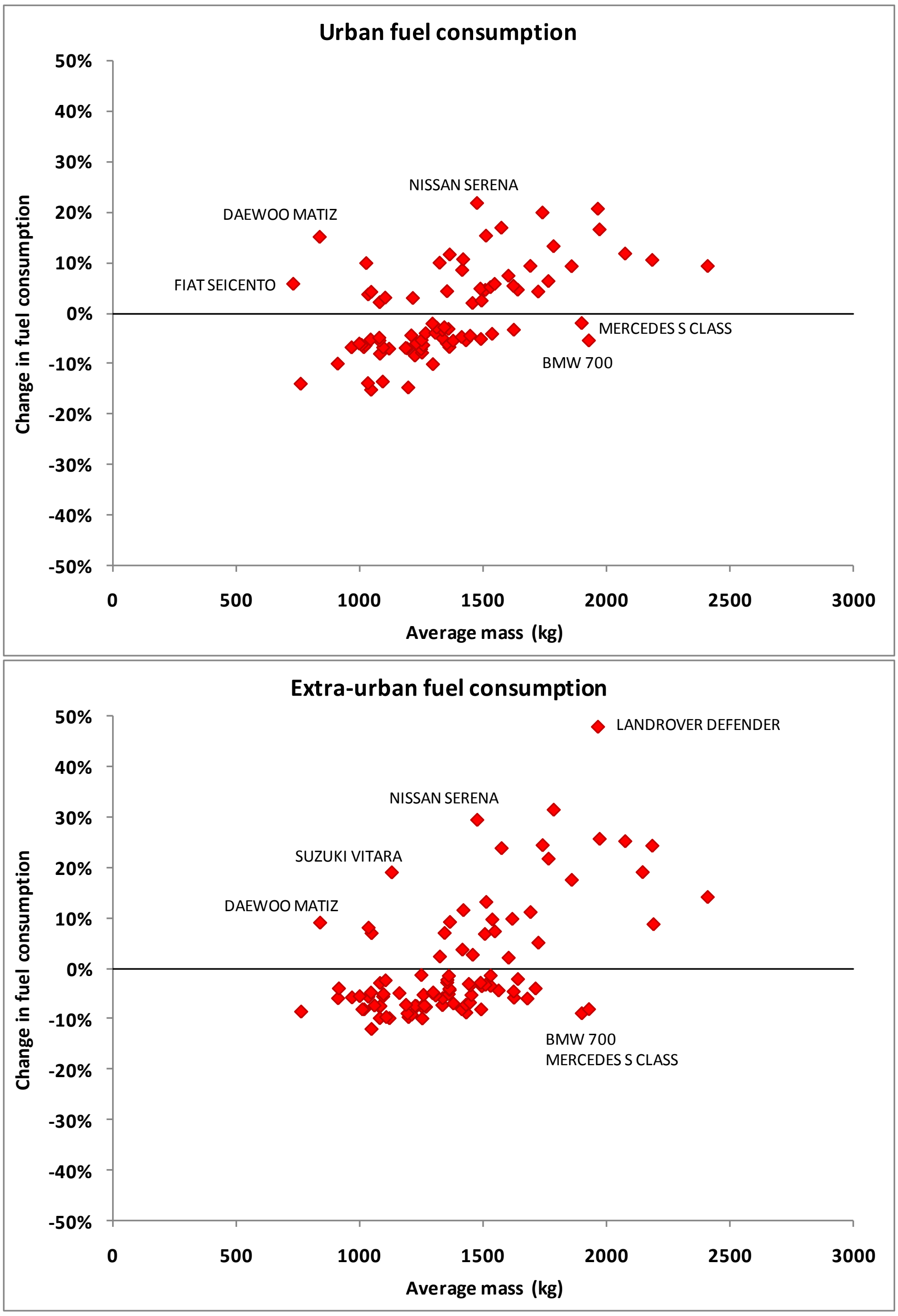


Figure 4.12: Estimated fixed effects of makes and models on fuel consumption versus average mass

4.5. Summary and conclusions

The effects of vehicle design on fuel consumption were investigated using cross-sectional data of a sample of popular makes and models in Great Britain. Partial effects of mass and engine size were estimated for different fuel and transmission types in both urban and extra-urban driving cycles while the effects of other contributing factors were controlled.

Comparison of different sets of estimated models showed the necessity to include not just vehicle make, but vehicle make and model on top of other design variables in the regression models. The results showed that by adding fixed effects of vehicle makes and models, the estimated coefficients of design variables remain broadly the same except the coefficient of “Mass” which is reduced substantially. This shows that in the final models (Models C in Tables 4.9 and 4.10), the estimated coefficient of “Mass” reflects the within make and model effects of mass while the between make and model effects of mass are reflected in the estimated coefficients of makes and models. The fact that only within make and model effects of mass could be isolated by the estimated models limits the range of fleet mass distribution scenarios that can be examined in this study; this will be explained in more detail in Chapter 6.

The effect of mass on fuel consumption within makes and models in both urban and extra-urban driving cycles was found to be statistically significant. Based on the modelling results, the partial effect of mass on fuel consumption is different for different combinations of fuel and transmission types. This effect is significantly greater for diesel cars compared to petrol cars of the same design in both European driving cycles. This could be partly a consequence of different characteristics of diesel and petrol engines and their relationship with vehicle mass. It was also found that in automatic transmission cars, mass has a significantly greater influence on fuel consumption in both driving cycles, compared to manual transmission cars of the same design. It was found that a 100 kg increase in mass of model variants within a make and model could increase their fuel consumption by 0.9% to 3.1% depending on fuel type, transmission type, and the driving cycle. The greatest partial effect of mass was found for automatic diesel cars when a 100 kg increase in mass would increase typical urban and extra-urban fuel consumption by 2.9% and 3.1%, respectively. The results also showed that the effect of mass on fuel consumption is greater in the extra-urban driving cycle than that in the urban cycle. Vehicle mass directly contributes to rolling resistance; this increases with speed due to an increase in the work being done in deforming the tyre over a given time (Heisler, 2002). This can result in a greater influence of mass on fuel consumption in extra-urban driving which includes a higher average speed than does urban driving.

These estimated effects of mass are generally lower than those estimated by Van den Brink and Van Wee (2001) whose estimation contains the effects of engine size as well (see Chapter 2, Section 2.2.1). The effects found here are not directly comparable with other estimated effects of mass discussed in Chapter 2, Section 2.2.1 due to different types of relationships considered. Considering the relatively low gap between test and actual fuel consumption rates (see Section 4.1.3) and the share of each component of vehicle energy demand in determining fuel consumption, as discussed in Chapter 2, no considerable change is expected in the partial effects of mass in practical urban and extra-urban conditions from those estimated here. However, more research is needed to investigate changes in these mass effects in different driving cycles involving more accelerations/decelerations and factors such as wind, hills and corners. In fuel consumption tests, the vehicle is assumed to be driven on a flat road and driving in a straight direction.

Modelling results confirmed that fuel consumption increases as engine size increases in both driving cycles. Based on modelling results, the partial effect of engine size on fuel consumption is not significantly different for diesel cars from that of petrol cars for the urban cycle, while this effect is less for diesel cars for the extra-urban cycle. In both driving cycles, the effect of engine size is greater for manual petrol cars compared to automatic petrol cars. Unlike vehicle mass, the effect of engine size on fuel consumption was found to be greater in the urban cycle than in the extra-urban cycle. The reason for this could be the better optimised performance of engines at higher speeds as well as the presence of less acceleration and idling in the extra-urban cycle compared to the urban driving cycle. Depending on driving cycle, fuel type, and transmission type, results suggest that 10% increase in engine size when all other design factors are held constant, could increase fuel consumption by 1.6% to 5.5%. These effects are lower than those estimated by Kwon (2006) whose estimate does not reflect the partial effects of engine size (see Chapter 2, Section 2.2.2). The effects are not directly comparable with other estimates discussed in Chapter 2, Section 2.2.2 due to different types of relationships considered.

It was found that cars meeting more stringent emission standards also consume less fuel and within each Euro standard, fuel consumption decreases for the new vehicles coming to the market each year compared to the previous year. It was also found that a car with automatic transmission consumes more fuel compared to a car with the same design features but manual transmission in both driving cycles.

The modelling results showed that while there is clearly a relationship between fuel consumption and each of mass and engine size, this dataset provides no evidence of an association between fuel consumption and the frontal area of a vehicle. This might be due to the fact that the aerodynamic drag is not properly simulated as a function of vehicle size in the fuel consumption tests. More research is needed using data based on real driving condition to investigate this potential association more fully.

CHAPTER 5. VEHICLE MASS AND SECONDARY SAFETY

The main objective of this chapter is to estimate the protective and aggressive effects of vehicle mass in two-car crashes by isolating the effects of vehicle mass and size. The chapter is organised as follows. The first section (5.1) provides a brief theoretical background on the effect of vehicle mass and other factors on occupant injury risk. The second section (5.2) investigates the relationship between vehicle mass and its relative driver injury risk in the British vehicle fleet at a given period of time. The third section (5.3) extends this to a disaggregate cross-sectional analysis of injury risk in two-car crashes where a novel methodology is introduced to estimate the partial effects of mass and size on absolute driver injury risk in both vehicles involved in the crash. The findings are summarised and discussed in the final section (5.4).

5.1. Background

As mentioned in Chapter 1, Section 1.4, there are two distinct aspects for the safety performance of a vehicle in fleet: primary safety performance, which is linked to the risk of crash involvement of the vehicle, and secondary safety performance, which is linked to the risk of occupant injury (to a specific level) when the vehicle is involved in a particular type of crash. While there is no evidence of any direct effect of vehicle mass on the primary safety performance of a vehicle, mass is directly related to the secondary safety performance of the vehicle. Research has shown that heavier and larger vehicles generally provide a higher level of safety for their occupants when involved in crashes compared to smaller and lighter vehicles (see Chapter 2, Section 2.2.1). The best available measure of crash severity which has been used in vehicle safety research is the prompt velocity change that the vehicle undergoes during a crash (Evans, 2004; Toy and Hammitt, 2003). Although the time during which this velocity change occurs is also important in the injury outcome of a crash, it is usually assumed to be identical for all crashes as data on this measure is not usually available.

In order to investigate the relationship between vehicle mass and secondary safety performance, two-car crashes have been studied intensively in vehicle safety research. This is because they form a case for vehicle crashes where both protective and aggressive effects of mass are best represented where the closing speed is identical for both drivers in the crash. Two-car crashes can also provide insight into crashes between any pair of vehicles and also into single-vehicle crashes (Evans, 2003). However, there are certain disadvantages or shortcomings associated with the methodologies used; these were discussed extensively in Chapter 2, Section 2.2.1.

As discussed in Chapter 1, Section 1.3, in a two-vehicle crash, the injury risk of occupants in the lighter vehicle is higher than that in the heavier vehicle due to the greater velocity change during the collision. For example, in the case of a frontal collision between two vehicles with masses m_1 and m_2 travelling with speeds v_1 and v_2, it can be easily shown using Newtonian mechanics that the velocity change of the first vehicle during the collision (Δv_1) depends on the proportion of the total mass contained by the other vehicle $\left(\frac{m_2}{m_1+m_2}\right)$ and the closing speed $(v_1 + v_2)$:

$$\Delta v_1 = \left(\frac{m_2}{m_1+m_2}\right)(v_1 + v_2)\ . \tag{5.1}$$

A consistent basis for the analysis of vehicle secondary safety performance is achieved by estimating crash injury risk to the driver because driver exposure is the representative of vehicle exposure. Besides, all the vehicles involved in crashes have the same number of drivers but not necessarily the same number of passengers and data on uninjured passengers is not usually available. However, when driver injury risk is used to represent secondary safety performance, it is important that the effect of factors contributing to the risk of injury, which might be different for different drivers, are controlled.

The main factors that contribute to driver injury risk in a two-car crash are outlined in Figure 5.1. Apart from driver factors and velocity change, vehicle size has also the potential to affect driver injury risk. It was discussed in Chapter 2, Section 2.3.2 that for a given Δv, a larger vehicle can give a better protection to its occupants by providing more crush space in the event of a crash. Isolating the effects of mass and size has long been an important issue in vehicle safety research. In many studies, the estimated effects of mass contain the effects of size as well because of a relatively high correlation between mass and size factors. However, there is theoretically a fundamental difference between the effects of mass and size. Vehicle mass has both protective and aggressive effects while vehicle size only tends to have a protective effect. Separate effects of mass and size on driver injury risk in two-car crashes will be estimated in this chapter.

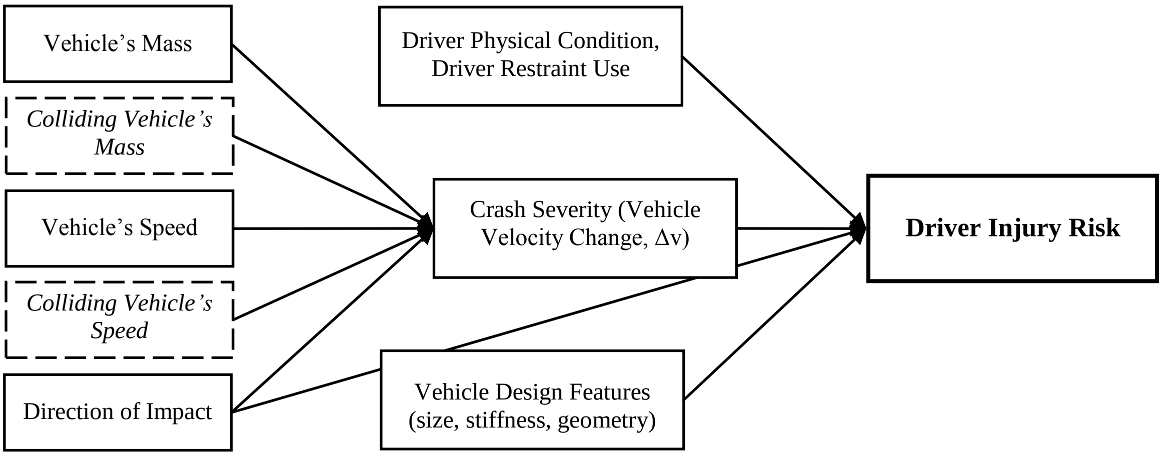


Figure 5.1: Main determinants of driver injury risk in a two-car crash

The analysis of driver injury risk in this chapter is divided into two main parts. First, the relationship between vehicle mass and its relative driver injury risk in the British fleet at a given period of time is shown. Then the analysis is extended to a detailed disaggregate cross-sectional analysis of two-car crashes where a novel methodology is introduced to estimate partial effects of mass and size on absolute driver injury risk.

5.2. The relationship between vehicle mass and its relative driver injury risk in fleet

This section describes the analyses performed to investigate the relationship between a measure of driver injury risk defined and estimated by UK Department for Transport (DfT) for popular makes and models in British fleet, and average mass of these makes and models.

5.2.1. Data

DfT estimates secondary safety performance of popular makes and models in Great Britain, as relative driver injury risk, and reports the results for specific time periods. The DfT risk estimates are calculated for cars using data from two-car crashes where at least one driver was injured. The two DfT safety indices are defined as:

D_{all} = Proportion of drivers of a given car model who are injured when involved in two-car accidents where at least one driver is injured.

D_{ksi} = Proportion of drivers of a given car model who are killed or seriously injured (ksi) when involved in two-car accidents where at least one driver is injured.

These are relative measures of injury risks where D in one vehicle is influenced by the injury risk in the other vehicle in the two-car crash. DfT used logistic regression models to estimate adjusted D_{all} and D_{ksi} for those makes and models registered on or after the 1 January 1995 which were involved in two-car crashes during 2000-2004. DfT adopted a threshold of 150 crash involvements to include makes and models to the analysis in order to achieve reliable estimates. In addition to make and model, variables related to speed limit (proxy for accident severity), first point of impact, driver sex and driver age are included in the DfT's model estimation process as explanatory variables to control for the effects of these factors. A brief explanatory note to this modelling process is available in appendix 3 to the DfT report (DfT, 2006a).

The latest DfT report on secondary safety of vehicles is available for 186 models of cars involved in accidents during 2000 to 2004 in Great Britain (DfT, 2006a); this provides an opportunity sample of popular makes and models with estimated relative driver injury risks in the fleet. In the DfT report, two estimates of risk (D_{all} and D_{ksi}) are available for each make and model as well as registration dates of that make and model. The estimates reflect relative secondary safety performance of makes and models in the British fleet. Table 5.1 shows a sample part of the DfT tables of risk estimates for the first 10 makes and models (out of 186) in the report.

Mass data for these makes and models was extracted from the developed design dataset explained in Section 3.3.2 of Chapter 3. DfT's driver injury estimates are aggregated for makes and models while mass varies within each make and model by various design factors; therefore, typical masses were assigned to the makes and models according to the following process. Each make and model was disaggregated to a number of design categories based on its period of manufacture (only for those included in the registration years of that make and model within the DfT database), body structure (estate, saloon, hatchback, coupe, etc), and engine size to minimize mass variation within different design categories of each make and model. For each make and model, a typical mass was then assigned to each design category according to the following rules:

- If mass varies for a design category by transmission type, mass for the manual transmission type is taken as the typical mass (any difference is usually slight).
- If mass varies for a design category by number of doors, mass for the one with the maximum number of doors is taken as the typical mass (any difference is usually slight).
- If mass for a design category varies by other engine specifications:
    - When variations are slight, the predominant mass (based on number of specifications with the same mass) is assigned to that design category.
    - When variations are substantial, separate masses are assigned to different specifications of that design category; hence defining new design categories.

Table 5.1: A sample of the DfT data on risk estimates (DfT, 2006a)

Car model	Registration dates	Percentage of drivers injured when involved in an injury crash[a]					
		Fatal or serious injuries (D_{ksi})			All injuries (D_{all})		
		Adjusted estimate	Confidence interval		Adjusted estimate	Confidence interval	
Audi TT	1999 to 2001	5	2	11	64	55	71
BMW Z3	1996 to 2001	4	2	8	62	54	69
Ford Puma	1997 to 2001	4	3	7	72	68	76
Hyundai Coupe	1995 to 2001	7	5	12	73	67	78
Mazda MX-5	1990 to 2004	6	4	10	72	68	77
MG MGF	1995 to 2003	6	4	9	77	73	81
Toyota Celica	1990 to 2004	6	4	9	62	57	67
Toyota MR2	1990 to 1996	7	4	12	67	60	73
Citroen AX	1990 to 1995	8	7	10	80	79	82
Citroen C3	2002 to 2004	3	1	7	72	64	78

[a] Injury crash is defined in this case as a crash in which at least one driver is injured.

Only those design categories for which mass data was available were included in the dataset. The total number of design categories in the dataset for which mass data was available was 757 for a total of 111 makes and models. Therefore, the makes and models in the dataset had different ranges of variation in mass, length, and engine size. The mass range changed from zero for makes and models with only one defined design category to 503 kg for the Volkswagen Passat with 29 defined categories. The plot of D_{all} associated with the makes and models in the dataset against mass for different design categories of these makes and models is shown in Figure 5.2. Each estimated value of D_{all}, which belongs to a particular make and model, corresponds to a range of masses belonging to different design categories of that particular make and model.

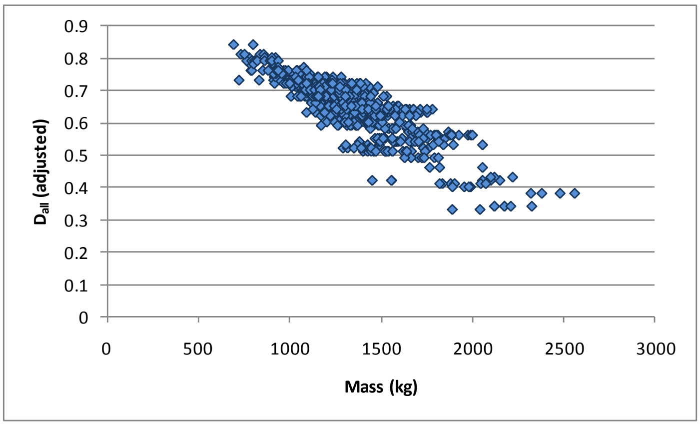


Figure 5.2: Plot of the estimated relative driver injury risk (D_{all}) against mass of different design categories of makes and models

Table 5.2 shows descriptive statistics of the two safety indices, average mass and mass range, as well as average engine size, average length, engine size range, and length range in the driver injury risk dataset. Within the dataset, Landrover Defender has the lowest “all casualties” index (0.33), and Hyundai Atoz and Rover Mini have the highest “all casualties” index (0.84). Average mass in the dataset is 1330 kg; it ranges from 690 kg (Rover Mini) to 2435 kg (Toyota Land Cruiser). The average mass of design categories within each make and model was used as the typical mass of that make and model in the final dataset. This dataset, which includes an overall of 111 records, was used to investigate the relationship between vehicle mass and its relative driver injury risk in the 2000-2004 British vehicle fleet.

Table 5.2: Descriptive statistics of the driver injury risk dataset (sample size: 111)

Variable	Descriptive statistics			
	Mean	Min	Max	Std. Deviation
“All casualties” index (D_{all}) of make and model	0.65	0.33	0.84	0.10
“ksi” index (D_{ksi}) of make and model	0.05	0.01	0.14	0.02
Average mass of make and model (kg)	1330	690	2435	335
Mass range within make and model[1] (kg)	175	0	503	113
Average engine size of make and model (cc)	2025	896	4382	665
Engine size range within make and model[1] (cc)	744	0	3510	636
Average length of make and model (cm)	431	305	510	40
Length range within make and model[1] (cm)	20	0	110	21

[1] Makes and models in the dataset have different ranges of variation in mass, length, and engine size

5.2.2. Analysis of driver injury risk

As explained in Chapter 2, Section 2.3.3, Broughton (1996a, 1996c) discussed different aspects of the DfT method for estimating safety indices as a measure of secondary safety of vehicles. Broughton (1996c) concluded that the DfT indices provide the most satisfactory means of comparing relative secondary safety of different models of cars. He also suggested that it is more sensible to concentrate on the "all casualties" index (D_{all}) as it is shown to be highly correlated with the "ksi" index (D_{ksi}) and it is more discriminating because of the much larger number of accidents used in its calculation (see Table 5.1 for examples of these indices and Table 5.2 for descriptive statistics of them).

Adjusted relative crash injury risk to drivers for all injuries (D_{all}), which is available as an aggregate measure of secondary safety performance for each make and model in the dataset, was used as the dependent variable to estimate the effect of mass on safety performance. A linear model was estimated using Ordinary Least Squares (OLS) to investigate the effect of vehicle mass on its adjusted relative crash injury risk to the driver. The model included average mass of variants for each make and model as the explanatory variable. The estimation result of this model is shown in Table 5.3. To examine the effect of uncertainty in the value of mass for makes and models (having variants with different ranges of mass) on the estimation results, a Weighted Least Squares (WLS) regression, where the weights were assigned to the records based on their mass range to give a greater weight to the records with a lower mass range, was also applied followed by a limited sensitivity analysis of model estimation results. Use of WLS provided the opportunity to give more weights to more reliable observations (those having lower mass range) in estimating the model. The results from the WLS estimation were similar to those from the OLS estimation (shown in Table 5.3) suggesting no considerable effect of variation in the values of mass within makes and models on the values of estimated coefficients (Tolouei, 2007; Tolouei and Titheridge, 2009).

Table 5.3: Model estimation results (dependent variable: adjusted relative crash injury risk to driver)

	Model 5.1		
	Coefficient	Std. Error	t-stat
(Constant)	1.036	0.017	60.134
Mass (kg)	-0.00030	0.00001	-23.532
Observations	111		
R^2	0.836		

Modelling results show that the adjusted "all casualties" index (D_{all}) decreases steadily with increasing mass and mass can explain about 84% of variation in this index. Based on the estimation results, a 100 kg increase in mass decreases relative risk of injury to the driver in a two-car injury accident by 3%. Based on the limited sensitivity analysis which was carried out, this estimated effect can change between 2.6% and 3.2% according to the typical values of mass assigned to makes and models in the dataset (Tolouei, 2007). These effects are estimated using injury accidents in Great Britain during 2000 to 2004. It is notable that this is estimated as a partial effect of mass on driver injury risk as the influences of driver age and sex, speed limit (proxy for accident severity) and first point of impact are already controlled for in the estimation of adjusted injury risks (DfT, 2006a).

It has been argued that higher engine performance and power could be associated with greater speeds and greater injury risk. This was examined, using the available data, by estimating the effect of engine size (as a proxy for engine power) on crash injury risk to the driver when the effect of mass is controlled for. Average engine size of variants for each make and model was used as the typical engine size of that make and model. The ratio of engine size to mass was also calculated and averaged over all variants for each make and model in the dataset. High variation in the values of engine size within makes and models, which is the result of design varieties, introduces a high level of uncertainty to this analysis. On average, the range of variation in engine size values within makes and models in the dataset was about 750 cc; the standard deviation of typical engine size values between makes and models was about 660 cc (see Table 5.2).

The variables "Mass" and "Engine size" in the dataset were highly correlated (correlation coefficient of 0.93) while there was less correlation between mass and the ratio of engine size to mass (correlation coefficient of 0.34). When "Mass" was replaced by "Ratio of engine size to mass" (Model 5.2 in Table 5.4) the coefficient was negative and significant, with the model explaining only 10% of variation in crash injury risk (suggested by the estimated value of R^2). This would imply that for a given mass, cars with larger engines had lower crash injury risk, which does not support the hypothesis made on the effect of higher engine performance and power. Replacing "Mass" with "Engine size" (Model 5.3 in Table 5.4) also reduced model performance (measured by R^2). These results show that vehicle mass is the best single explanator of driver injury risk (Model 5.1 in Table 5.3). To examine a possible effect of engine size, the variables "Engine size" and "Ratio of engine size to mass" were included separately in the model which had vehicle mass as the explanatory variable (Model 5.1 in Table 5.3). The results of these two estimated models are shown in Table 5.5. These variables do not have statistically significant coefficients in the models and their inclusion in the model does not improve model performance; the R^2 value is not improved measurably beyond 0.836 and the fitted coefficients of the new variables have t values (0.266 and -0.043 respectively) not significantly different from zero.

Table 5.4: Model estimation results (dependent variable: adjusted relative crash injury risk to driver)

	Model 5.2			Model 5.3		
	Coefficient	Std. Error	t-stat	Coefficient	Std. Error	t-stat
(Constant)	0.941	0.087	10.76	0.920	0.019	49.46
Engine size (cc)	-	-	-	-0.00014	0.00001	-15.66
Engine size / mass (cc/kg)	-1.98	0.058	-3.43	-	-	-
Observations	111			111		
R^2	0.097			0.708		

Table 5.5: Model estimation results (dependent variable: adjusted relative crash injury risk to driver)

	Model 5.4			Model 5.5		
	Coefficient	Std. Error	t-stat	Coefficient	Std. Error	t-stat
(Constant)	1.038	0.019	55.66	1.038	0.038	55.664
Mass (kg)	-0.00030	0.00003	-9.14	-0.00030	0.00001	-9.145
Engine size (cc)	0.000004	0.00002	0.27	-	-	-
Engine size / mass (cc/kg)	-	-	-	-0.0011	0.026	-0.043
Observations	111			111		
R^2	0.836			0.836		

It has been also argued that for a given mass, larger vehicles are associated with lower risk of injury as they provide more crumple room in crashes. This was also examined using the available data by adding the average length of each make and model, as the typical length, to the regression models. The variables added to the regression models in separate steps were “Length” and “Ratio of length to mass”. The results, reflected in Tables 5.6 and 5.7, were generally similar to those of adding the effects of engine size (Tables 5.4 and 5.5); they showed no effect of length over and above mass that is statistically significant and they confirmed that vehicle mass is the best single explanator of driver injury risk.

These results generally show that while there is clearly a relationship between crash injury risk and vehicle mass, this dataset provides no evidence of an association between crash injury risk and either engine power or performance (as represented by engine size) or vehicle size (as represented by length). However, because of a high correlation between these variables and the high level of uncertainty in the available data, as discussed earlier, more research is needed using disaggregate cross-sectional data in order to address this issue more fully.

Table 5.6: Model estimation results (dependent variable: adjusted relative crash injury risk to driver)

	Model 5.6			Model 5.7		
	Coefficient	Std. Error	t-stat	Coefficient	Std. Error	t-stat
(Constant)	0.078	0.030	2.611	1.458	0.079	18.413
Length (cm)	-	-	-	-0.0019	0.0002	-10.335
Length / mass (cm/kg)	1.674	0.087	19.209	-	-	-
Observations	111			111		
R^2	0.772			0.495		

Table 5.7: Model estimation results (dependent variable: adjusted relative crash injury risk to driver)

	Model 5.8			Model 5.9		
	Coefficient	Std. Error	t-stat	Coefficient	Std. Error	t-stat
(Constant)	0.988	0.055	17.985	0.968	0.140	6.937
Mass (kg)	-0.00031	0.00002	-15.043	-0.00028	0.00004	-6.488
Length (cm)	0.00016	0.00017	0.927	-	-	-
Length / mass (cm/kg)	-	-	-	0.123	0.250	0.492
Observations	111			111		
R^2	0.836			0.836		

5.2.3. Comparison with an earlier study

Broughton (1996a) estimated the effect of mass on adjusted D_{all} using data on 91 popular makes and models in Great Britain involved in accidents from 1989 to 1992. He established approximate mass for 87 makes and models and estimated a 4.5% increase in adjusted D_{all} for a 100 kg increase in mass. Figure 5.3 compares his results with my results based on the estimated Model 5.1 (in Table 5.3). The mean prediction intervals at 95% confidence level for the makes and models in the dataset are also shown for the new estimated model in this plot.

Figure 5.2 reveals an important finding: a car with a given mass would have a different secondary safety performance in the British fleet in the two periods of time (i.e. a higher injury risk in the 2000-2004 fleet compared to the 1989-1992 fleet). On the other hand, a 100 kg increase in mass had a greater effect on injury risk (4.5% decrease) in the 1989-1992 fleet compared to that in the 2000-2004 fleet (3% decrease).

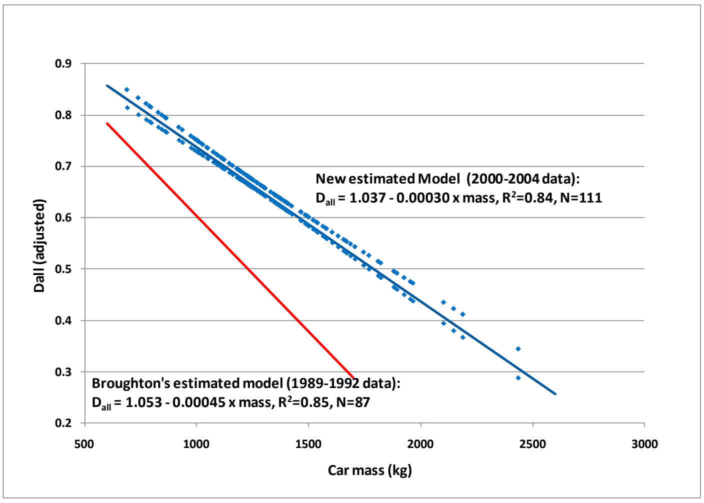


Figure 5.3: Comparing the results of this study with the results from Broughton's study (1996b) on the effect of vehicle mass on driver injury risk

Broughton (2007) discussed trends in changes in the car fleet in Great Britain over the years 1997 to 2003 and showed that the numbers of the smallest and largest cars have grown, while the numbers of cars of intermediate size have been stable. He also argued that the mass of the typical new car had increased by about 20% between 1990 and 2001, while the mass of the typical large/luxury saloon (largest cars) was about 80% greater than the mass of the typical Mini/Supermini (smallest cars) throughout. These findings suggest that the effect of mass on relative driver injury risk (D) can significantly change over time based on fleet characteristics in which mass distribution is an important factor. This is the result of the effect of vehicle mass on injury risk to the driver of that vehicle as well as to the driver of the other vehicle in a two-vehicle crash as represented by Equation 5.1. A detailed analysis of two-car crashes is needed to understand the protective and aggressive effects of vehicle mass. Ideally, the effect of mass on absolute driver injury risk, where the injury risk is independent from the driver injury risk in the other vehicle in crash, should be investigated. However, this measure of risk cannot be directly calculated from the injury crash data due to lack of data on non-injury crashes. In the next section, a methodology is introduced to analyse the relationship between mass and absolute driver injury risk in two-car crashes based on injury crash data.

5.3. Analysis of injury risk in two-car crashes

Although the analysis presented in Section 5.2 confirmed that secondary safety performance of a vehicle in fleet generally increases as mass decreases, it also showed that the relationship between mass and driver injury risk (probability of driver injury in the event of a crash) changes over time depending on the mass distribution of vehicles in fleet. In order to fully understand the likely changes in overall injury outcome of a fleet as a result of a change in mass distribution within the fleet, a detailed analysis of two-car crashes is required to estimate the partial effects of vehicle mass on injury risk to its driver as well as to that of the other driver involved in the crash. Besides, it is ideally preferred to represent the vehicle secondary safety performance by absolute driver injury risk where the injury risk in one vehicle is independent from the injury risk in the other vehicle in two-car crashes. A methodology is introduced based on disaggregate cross-sectional analysis of two-car crash data to estimate the effects of vehicle mass, as well as other factors, on absolute driver injury risk.

5.3.1. Methodology

As was discussed in Section 5.1 of this chapter, in the case of a frontal collision between two vehicles with masses m_1 and m_2 travelling with speeds v_1 and v_2 immediately before the collision, the velocity change of the vehicles during the collision (Δv_1 and Δv_2) are given by (Grime and Jones, 1970):

$$\Delta v_1 = \left(\frac{m_2}{m_1+m_2}\right)(v_1 + v_2) \tag{5.2}$$

$$\Delta v_2 = \left(\frac{m_1}{m_1+m_2}\right)(v_1 + v_2)\ , \tag{5.3}$$

where $v = v_1 + v_2$ is the closing speed of the vehicles. Similarly, Δv_1 and Δv_2 in the case of a front to back collision where $v_1 > v_2$ are given by:

$$\Delta v_1 = \left(\frac{m_2}{m_1+m_2}\right)(v_1 - v_2) \tag{5.4}$$

$$\Delta v_2 = \left(\frac{m_1}{m_1+m_2}\right)(v_1 - v_2)\ , \tag{5.5}$$

where, in this case, $v = v_1 - v_2$ is the closing speed of the vehicles. Finally, in the case of a front to side collision at right angle, Δv for each vehicle has two components in x and y directions. It can be shown that the magnitude of Δv_1 and Δv_2 are given by:

$$\Delta v_1 = \left(\frac{m_2}{m_1+m_2}\right)\left(\sqrt{{v_1}^2 + {v_2}^2}\right) \tag{5.6}$$

$$\Delta v_2 = \left(\frac{m_1}{m_1+m_2}\right)\left(\sqrt{{v_1}^2 + {v_2}^2}\right). \tag{5.7}$$

In this case, $v = \sqrt{{v_1}^2 + {v_2}^2}$. The mass ratio of the vehicles in crash is defined as

$$\mu = \frac{m_2}{m_1}. \tag{5.8}$$

Therefore, Δv_1 and Δv_2in Equations 5.2 to 5.7 can be rearranged as the following:

$$\Delta v_1 = v\left(\frac{\mu}{\mu+1}\right) \tag{5.9}$$

$$\Delta v_2 = v\left(\frac{1}{\mu+1}\right). \tag{5.10}$$

In a two-vehicle collision, the probability of injury of the driver of vehicle 1, $p_1(v)$, increases with closing speed v and with increasing the value of mass ratio μ while the probability of injury of the driver of vehicle 2, $p_2(v)$, increases with closing speed v and with decreasing the value of mass ratio μ. One of the functional forms having the appropriate properties to describe $p_1(v)$ and $p_2(v)$, both of which range between zero and one, is the logistic function. Logistic function has a widespread acceptability amongst researchers and has been used commonly in the literature to model driver injury risk (see Chapter 2 for examples). Therefore, the logistic function is chosen in this study to describe $p_1(v)$ and $p_2(v)$ as

$$p_1(v) = \frac{exp[C_1+\beta\Delta v_1]}{1+exp[C_1+\beta\Delta v_1]} = \frac{exp\left[C_1+\beta v\left(\frac{\mu}{\mu+1}\right)\right]}{1+exp\left[C_1+\beta v\left(\frac{\mu}{\mu+1}\right)\right]} \tag{5.11}$$

and

$$p_2(v) = \frac{exp[C_2+\beta\Delta v_2]}{1+exp[C_2+\beta\Delta v_2]} = \frac{exp\left[C_2+\beta v\left(\frac{1}{\mu+1}\right)\right]}{1+exp\left[C_2+\beta v\left(\frac{1}{\mu+1}\right)\right]}. \tag{5.12}$$

In these equations, C_1 and C_2 represent the characteristics of the driver (age, gender, etc) and the vehicle (dimensions, make, model, etc) that could contribute to the driver injury risk as outlined in Figure 5.1. C_1 and C_2 can be expressed as the following:

$$C_j = \sum_i \alpha_{ji} x_{ji}\ ;\ j = 1,2\ \&\ i = 0,1,2,\ldots\ \&\ x_{j0} = 1 \tag{5.13}$$

where, j denotes the driver number (1 or 2),

$\{x_{ji}\}$ are a set of driver and vehicle characteristics for vehicle j, and

$\{a_{ji}\}$ are a set of parameters to be estimated in the model fitting process.

As can be seen from Equations 5.11 and 5.12, probability of driver injury in each vehicle depends on the closing speed v. The main difficulty associated with the analysis of injury risk in two-car crashes arises because vehicles' speed immediately prior to the crash (v_1 and v_2) are not usually observed; therefore, closing speed v is rarely known.

Suppose $f(v)$ represents the probability distribution of closing speed v that is generally characterised by a mean m and a vector ϕ of constant parameters. Then the overall probabilities of any collision resulting in injury of the drivers of vehicles 1 and 2 are:

$$P_1 = \int_v p_1(v) f(v) dv \tag{5.14}$$

and

$$P_2 = \int_v p_2(v) f(v) dv \,. \tag{5.15}$$

There are four possible driver injury outcomes of any two-vehicle collision depending on the driver injury outcome of each vehicle. Since the probability of injury depends on the closing speed v, and v is common to the two vehicles in the collision, the two events of driver 1 and driver 2 being injured are dependent to each other. Therefore the four possible injury outcomes are in fact joint injury probabilities as shown in Table 5.8. In this case, we expect that if the driver of vehicle 1 is injured, it is more likely that the driver of vehicle 2 is injured too.

Table 5.8: Possible joint injury outcomes of a two-vehicle collision

	Driver 1 not injured	Driver 1 injured
Driver 2 not injured	$\pi_{00} = \int_v \big(1 - p_1(v)\big)\big(1 - p_2(v)\big) f(v) dv$	$\pi_{10} = \int_v p_1(v)\big(1 - p_2(v)\big) f(v) dv$
Driver 2 injured	$\pi_{01} = \int_v \big(1 - p_1(v)\big) p_2(v) f(v) dv$	$\pi_{11} = \int_v p_1(v) p_2(v) f(v) dv$

According to the Equations 5.11 to 5.13, the π_{ij} are functions of the parameter β, the parameters $\boldsymbol{\alpha}$ relating to the vehicle and driver characteristics, and the parameters m and $\boldsymbol{\phi}$ characterising the distribution of closing speed $f(v)$. Since speeds of the vehicles prior to the collision are not observed, m and $\boldsymbol{\phi}$ are nuisance parameters that will be estimated in the model estimation process.

The other difficulty in estimating the absolute driver injury risk (P_1 and P_2) is the fact that no observation is available if there are no injuries (π_{00} is unknown). However, conditional driver injury risk defined as the probability of driver injury in a collision in which there is at least one driver injury can be calculated directly from the observed data. Conditional joint injury probabilities are defined as below:

$$q_{ij} = Prob(Z_1 = i, Z_2 = j \mid i = 0{,}1 \;\&\; j = 0{,}1 \;\&\; i + j \geq 1) = \frac{\pi_{ij}}{1-\pi_{00}} \tag{5.16}$$

where Z for each driver represents the binary injury outcome (0=no injury, 1=injury). Closing speed v determines the "severity" of the collision (measured by $i + j$), whilst mass ratio μ determines the "imbalance" between the injuries of the two drivers (measured by $|i - j|$). The three possible conditional joint injury outcomes as defined by Equation 5.16 are shown in Table 5.9. The observed values of these are available from the crash data.

Table 5.9: Conditional joint injury outcomes of a two-vehicle collision

	Driver 1 not injured	Driver 1 injured
Driver 2 not injured	-	$q_{10} = {}^{\pi_{10}}/1 - \pi_{00}$
Driver 2 injured	$q_{01} = {}^{\pi_{01}}/1 - \pi_{00}$	$q_{11} = {}^{\pi_{11}}/1 - \pi_{00}$

i) Maximum likelihood estimation of parameters

The joint injury probabilities π_{ij} shown earlier in Table 5.8 can be formed using Equations 5.11 to 5.15 as a function of unknown parameters (β, $\boldsymbol{\alpha}$, m, $\boldsymbol{\phi}$) where a probability density function is assumed for closing speed ($f(v)$). Having formed π_{ij}, the three conditional joint injury probabilities (q_{ij}) shown in Table 5.9 can be described as functions of these unknown parameters:

$$q_{10} = \frac{\int_v \frac{exp\left[C_1+\beta v\left(\frac{\mu}{\mu+1}\right)\right] f(v)\, dv}{\left[1+exp\left[C_1+\beta v\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\beta v\left(\frac{1}{\mu+1}\right)\right]\right]}}{1-\int_v \frac{f(v)\, dv}{\left[1+exp\left[C_1+\beta v\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\beta v\left(\frac{1}{\mu+1}\right)\right]\right]}} \tag{5.17}$$

$$q_{01} = \frac{\int_v \frac{exp\left[C_2+\beta v\left(\frac{1}{\mu+1}\right)\right] f(v)\, dv}{\left[1+exp\left[C_1+\beta v\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\beta v\left(\frac{1}{\mu+1}\right)\right]\right]}}{1-\int_v \frac{f(v)\, dv}{\left[1+exp\left[C_1+\beta v\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\beta v\left(\frac{1}{\mu+1}\right)\right]\right]}} \tag{5.18}$$

$$q_{11} = \frac{\int_v \frac{exp\left[C_1+\beta v\left(\frac{\mu}{\mu+1}\right)\right] exp\left[C_2+\beta v\left(\frac{1}{\mu+1}\right)\right] f(v)\, dv}{\left[1+exp\left[C_1+\beta v\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\beta v\left(\frac{1}{\mu+1}\right)\right]\right]}}{1-\int_v \frac{f(v)\, dv}{\left[1+exp\left[C_1+\beta v\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\beta v\left(\frac{1}{\mu+1}\right)\right]\right]}} . \quad (5.19)$$

The unit of observation will be two-car collisions with three possible conditional joint injury outcomes. For any values of the parameters the probabilities of the observed conditional joint injury outcomes can be calculated for each collision. By combining these over the whole dataset, the likelihood function can be calculated as the following. For each observation, define,

$$y_{ij} = \begin{cases} 1, & if\ Z_1 = i\ and\ Z_2 = j\,; \\ 0, & otherwise; \end{cases} \quad (5.20)$$

where i and j show, respectively, the binary injury outcome for the driver 1 and 2 (i.e. $i = 0,1,\ j = 0,1,\ i + j \geq 1$). The likelihood function over the whole dataset can be calculated using the following:

$$L(\boldsymbol{\alpha}, \beta, m, \boldsymbol{\phi}) = \prod_{n=1}^{N} [({q_{10}}^{y_{10}})_n ({q_{01}}^{y_{01}})_n ({q_{11}}^{y_{11}})_n] \quad (5.21)$$

where N denotes the total number of records in the dataset. An optimisation algorithm can then be applied to find the values of the parameters that maximise the logarithm of the likelihood function[1] (log-likelihood function) shown below:

$$l(\boldsymbol{\alpha}, \beta, m, \boldsymbol{\phi}) = \sum_{n=1}^{N} \left[\left(y_{10}\log_e(q_{10})\right)_n + \left(y_{01}\log_e(q_{01})\right)_n + \left(y_{11}\log_e(q_{11})\right)_n\right]. \quad (5.22)$$

ii) Probability distribution of closing speed

As mentioned earlier, a distribution form is required for the closing speed with a given probability density function $f(v)$ the parameters of which $(m, \boldsymbol{\phi})$ will be estimated in the model estimation process. In this study, two continuous probability distributions are investigated separately to describe the distribution of closing speed. A normal distribution is investigated first because it is a simple well-described distribution which is defined with only two parameters (mean and standard deviation). However, the disadvantage of normal distribution in this case is that it is an unbounded distribution; hence specific constraints are required on the distribution parameters to ensure the

---

[1] Since the logarithm is a monotonic function, parameters that maximise the likelihood function also maximise the log likelihood function.

values of closing speed v remains positive during the model fitting process. Therefore, a log-normal distribution is also investigated for v. This has the advantageous property that it is bounded below by 0, therefore it is free from any constraint required to ensure positive values for closing speed v during the model fitting process.

Evidence suggests that vehicles have different average speeds in different types of roads as classified according to their speed limit (DfT, 2007). Therefore, it can be argued that the distribution of closing speed is different in different types of roads where the speed limit varies. This will be investigated by including the variable speed limit, which is observed for each collision, in the probability density function of the distributions being investigated. The model estimation results for the different density functions formulated will be compared to determine the best distribution form that describes the closing speed for the two-car collisions being studied.

D1: Normal distribution

A normal distribution with parameters m (mean) and σ (standard deviation) is assumed for the closing speed v. Two different probability density functions are characterised depending on whether m is constant or varies by the speed limit of the road.

D1.A: Same distribution for all speed limits

It is assumed that, for all the collisions, v is normally distributed with mean m and standard deviation σ. Transforming the normal random variable v to the standardised normal variable z with mean 0 and standard deviation 1 results in the following:

$$v = m + z\sigma = m(1 + zC) \tag{5.23}$$

where C is the coefficient of variation of closing speed. Therefore, the following probabilities are equivalent:

$$P(v_1 < v < v_2) = P(z_1 < z < z_2)\ . \tag{5.24}$$

This will provide the opportunity to use the unit normal density function ($\phi(z)$) to calculate the probabilities described in Equations 5.17 to 5.19 for any given values of m and C. Replacing the closing speed v in Equations 5.17 to 5.19 from Equation 5.23 results in an expression that includes the product of two unknown parameters m and β

($m' = \beta m$). A new variable u is therefore defined as the product of the parameter β and the closing speed v as

$$u = \beta v = \beta m(1 + zC) = m'(1 + zC)\ . \tag{5.25}$$

This new variable u is then normally distributed with mean m' and coefficient of variation C. The probability density function for u can be expressed based on the unit normal density function for z according to the following:

$$f(u) = \mathrm{g}(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} exp[-0.5(z)^2] \tag{5.26}$$

where $z = \frac{u-m\prime}{m\prime C}$. The conditional joint injury probabilities, required to form the log-likelihood function in Equation 5.22, are then calculated according to the following:

$$q_{10} = \frac{\int_z \frac{exp\left[C_1+\mathrm{u}(z)\left(\frac{\mu}{\mu+1}\right)\right] \mathrm{g}(z)\ dz}{\left[1+exp\left[C_1+\mathrm{u}(z)\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\mathrm{u}(z)\left(\frac{1}{\mu+1}\right)\right]\right]}}{1-\int_z \frac{\mathrm{g}(z)\ dz}{\left[1+exp\left[C_1+u\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+u\left(\frac{1}{\mu+1}\right)\right]\right]}} \tag{5.27}$$

$$q_{01} = \frac{\int_z \frac{exp\left[C_2+\mathrm{u}(z)\left(\frac{1}{\mu+1}\right)\right] \mathrm{g}(z)\ dz}{\left[1+exp\left[C_1+\mathrm{u}(z)\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\mathrm{u}(z)\left(\frac{1}{\mu+1}\right)\right]\right]}}{1-\int_z \frac{\mathrm{g}(z)\ dz}{\left[1+exp\left[C_1+\mathrm{u}(z)\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\mathrm{u}(z)\left(\frac{1}{\mu+1}\right)\right]\right]}} \tag{5.28}$$

$$q_{11} = \frac{\int_z \frac{exp\left[C_1+\mathrm{u}(z)\left(\frac{\mu}{\mu+1}\right)\right] exp\left[C_2+\mathrm{u}(z)\left(\frac{1}{\mu+1}\right)\right] \mathrm{g}(z)\ dz}{\left[1+exp\left[C_1+\mathrm{u}(z)\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\mathrm{u}(z)\left(\frac{1}{\mu+1}\right)\right]\right]}}{1-\int_z \frac{\mathrm{g}(z)\ dz}{\left[1+exp\left[C_1+\mathrm{u}(z)\left(\frac{\mu}{\mu+1}\right)\right]\right]\left[1+exp\left[C_2+\mathrm{u}(z)\left(\frac{1}{\mu+1}\right)\right]\right]}} \tag{5.29}$$

where $u(z)$ and $g(z)$ are given by Equations 5.25 and 5.26, respectively. The integration is calculated numerically over the standardised normal variable z. In maximising the log-likelihood function over the dataset (Equation 5.22), the following constraints should be applied:

1. $C \geq 0\ ,$ (5.30)
2. $m' \geq 0\ ,$ (5.31)
3. $1 - z^n C \geq 0 \xrightarrow{yields} C \leq {}^{1}\!/_{z^n}\ .$ (5.32)

Therefore, the parameters $\boldsymbol{\alpha}$ (representing driver and vehicle factors as shown in Equation 5.13) and m' and C (characterising the distribution of $u = \beta v$) that maximise Equation 5.22 are estimated. The estimated values of these parameters are then used to predict the values of absolute driver injury risk for vehicle 1 and 2 for any given value of mass ratio μ using the following equations:

$$P_1 = \int_v p_1(v) f(v) dv = \int_z \frac{exp\left[C_1 + \mathrm{u}(z)\left(\frac{\mu}{\mu+1}\right)\right] \mathrm{g}(z)\, dz}{1 + exp\left[C_1 + \mathrm{u}(z)\left(\frac{\mu}{\mu+1}\right)\right]} \tag{5.33}$$

$$P_2 = \int_v p_2(v) f(v) dv = \int_z \frac{exp\left[C_2 + \mathrm{u}(z)\left(\frac{1}{\mu+1}\right)\right] \mathrm{g}(z)\, dz}{1 + exp\left[C_2 + \mathrm{u}(z)\left(\frac{1}{\mu+1}\right)\right]}\,. \tag{5.34}$$

D1.B: Different distributions for different speed limits

As was pointed earlier, it is hypothesised that the distribution mean varies proportionally with the speed limit ($m = \omega_1 \times$ speed limit). Therefore:

$$m' = \beta m = \beta(\omega_1 v^l) = s v^l \tag{5.35}$$

where v^l denotes the speed limit and $s = \beta\omega_1$ (both β and ω_1 are unknown parameters). Therefore, the only difference between this case and the formulation shown in *D1.A* is that instead of parameter m', the parameter s is estimated which is used together with the observed value of speed limit to determine the distribution mean m' for each collision.

D2: Log-normal distribution

As was discussed earlier in this section, a log-normal distribution with parameters m and σ is also investigated for the closing speed v where m and σ are, respectively, mean and standard deviation of the associated normal distribution (i.e. m and σ are mean and standard deviation of $\log_e v$). It was pointed earlier that log-normal distribution has the advantageous property over normal distribution that it is bounded below by zero; hence, it is free from the constraints used in the case of normal distribution to ensure positive values for v (Equations 5.30 to 5.32). Two cases are investigated as described below:

D2.A: Same distribution for all speed limits

In this case, similar to that of normal distribution (*D1.A*), the same log-normal distribution for v is assumed for all the collisions irrespective of the speed limit of the road. v has a log-normal distribution with parameters m and σ, therefore $u = \beta v$ has a log-normal distribution with parameters $m' = m + \log_e \beta$ and σ (i.e. m' and σ are mean and standard deviation of $\log_e u$). Transforming $\log_e u$ which has a normal distribution to $\log_e z$ which has a standard normal distribution with mean 0 and standard deviation 1 results in the following relationship:

$$u = exp[m' + \sigma \log_e z] = \beta exp(m) z^{\sigma} \ . \tag{5.36}$$

In this equation, both β and m, in addition to the standard deviation σ, are unknown parameters. A new variable θ is therefore defined as $\theta = \beta exp(m)$. Therefore,

$$u = \theta z^{\sigma} \ . \tag{5.37}$$

The probability density function $f(u)$ can be expressed based on the unit normal density function for $\log_e z$ according to the following:

$$f(u) = \mathrm{g}(z) = \frac{1}{\mathrm{z}} \phi(\log_e z) = \frac{1}{z\sqrt{2\pi}} exp[-0.5(\log_e z)^2] \tag{5.38}$$

where $\log_e z = \frac{\log_e u - m\prime}{\sigma}$. The conditional joint injury probabilities that are required to form the likelihood function in Equation 5.22 are then approximated according to the Equations 5.27 to 5.29 where $u(z)$ and $g(z)$ are given by Equations 5.37 and 5.38, respectively. Therefore, the parameters $\boldsymbol{\alpha}$ (representing driver and vehicle factors as shown in Equation 5.13), and θ and σ that maximise Equation 5.22 are estimated subject only to the following constraint:

$$\sigma \geq 0 \ . \tag{5.39}$$

The resulting values of these parameters are then used to estimate the values of absolute driver injury risk for vehicle 1 and 2 for any given value of mass ratio μ using Equations 5.33 and 5.34.

D2.B: Different distributions for different speed limits

We now consider the possibility that the distribution mean for $\log_e v$ is related to the speed limit according to the following:

$$m = log_e(\omega_2 s^l) \ . \tag{5.40}$$

Replacing this in Equation 5.36 gives

$$u = \beta \omega_2 z^{\sigma} s^{l} \,. \tag{5.41}$$

In this equations, both β and ω_2, as well as the standard deviation σ, are unknown parameters. A new variable t is therefore defined as $t = \beta\omega_2$. Equation 5.41 can be rewritten as

$$u = t z^{\sigma} s^{l} \,. \tag{5.42}$$

Therefore, the only difference between this case and the formulation shown in *D2.A* is that instead of parameter θ, the parameter t is estimated which is used together with the estimated value of σ as well as the observed value of speed limit s^{l} to characterise the distribution of closing speed.

5.3.2. Two-car crash dataset

The data used to analyse driver injury risk in two-car crashes in this study is based on STATS19 Police reported data which includes road accidents that involve personal injury or death. Data from 2000 to 2006 was used to extract two-car crashes in which at least one of the drivers was injured. The process of developing a sample of injury two-car crashes that included data on the design variables of the colliding cars was detailed in Chapter 3, Section 3.3.2.2. As explained, this sample, which included about 21% of the total injury two-car crashes during 2000-2006, was used for the analysis of two-car crashes.

There are three levels of casualty severity in STATS19 data: killed (within 30 days as a result of sustained injury), seriously injured[1], and slightly injured[2]. The main analysis of two-car crashes was performed for serious or fatal injuries only due to the greater importance of these injuries. This is consistent with the similar studies based on STATS19 data (for example see Broughton, 1996b; Broughton, 2007). Therefore, the final sample dataset included two-car crashes where at least one of the drivers was either Killed or Seriously Injured (KSI); this included a total of 5,795 two-car crashes.

---

[1] Examples include fracture, internal injury, severe cuts, crushing, burns, concussion, sever general shock requiring hospital treatment, and detention in hospital as an in-patient (DfT, 2004)

[2] Examples include sprains not necessarily requiring hospital treatment, neck whiplash injury, bruises, slight cuts, and slight shock requiring roadside attention (DfT, 2004)

Descriptive statistics of vehicle design variables in the dataset (mass, length, width, height, and wheelbase) are shown in Table 5.10 (the maximum number of observations in this table is twice the total number of collisions because there are two vehicles per collision). The average vehicle mass in the dataset is 1135 kg; it ranges from 690 kg (for a variant of Citroen AX) to about 2600 kg (for a variant of Land Rover Range Rover). Vehicle length varies from 270 cm (for a variant of Volkswagen Polo) to 516 cm (for a variant of Mercedes S class) with the average of 413 cm. As the statistics for “Length” and “Wheelbase” suggest, there is more variation in “Length” compared to “Wheelbase” in the dataset. This suggests that “Length” is preferred to “Wheelbase” as a variable that represents the vehicle size. The average width of vehicles in the dataset is 178 cm where the minimum and maximum widths belong to variants of Renault Clio and Ford Mondeo, respectively.

Table 5.10: Descriptive statistics of vehicle design variables in the dataset (injury level: KSI)

Vehicle variable	Descriptive statistics				
	Min	Mean	Max	Std. Deviation	Obs.
Mass (kg)	690	1135	2599	250	11590
Length (cm)	270	413	516	36	11590
Width (cm)	142	178	223	15	11590
Height (cm)	122	143	194	10	10158
Wheelbase (cm)	142	255	448	13	10440

Table 5.11 shows the distribution of two-car crashes in the dataset by type of impact and speed limit of the road. Frontal crashes alone constitute about 43% of all crashes in the dataset. This is probably because these are high severity crashes resulting in a greater number of KSI drivers. The most common crash category in the dataset is frontal crashes on roads with a speed limit of 60 mile/hr.

Table 5.11: Distribution of crashes by type of impact and speed limit in the dataset (injury level: KSI)

Type of impact	Speed limit (mile/hr)				Total
	20 or 30	40 or 50	60	70	
Frontal	798	283	1324	80	2485
Front to back	401	131	174	171	877
Front to side	693	241	746	140	1820
Unknown/Other	263	57	197	96	613
Total	2155	712	2441	487	5795

Driver factors in a vehicle that is involved in a two-vehicle crash can potentially contribute to the risk of injury to the driver of that vehicle (through correlation with the physical strength of the driver) as well as to that to the driver of the colliding vehicle (through influencing driving style and aggressivity). Distribution of drivers involved in injury crashes by age and gender are reflected in Table 5.12. The category with the largest number of records in the dataset is that of male drivers aged 35-54.

Table 5.12: Distribution of drivers by age and gender in the dataset (injury level: KSI)

Driver age	Driver gender			Total
	Male	Female	Unknown	
17-24	1639	806	1	2446
25-34	1588	900	3	2491
35-54	2465	1474	13	3952
+55	1522	748	1	2271
Unknown	196	81	153	430
Total	7410	4009	171	11590

5.3.3. Injury[1] risk modelling

The methodology explained in Section 5.3.1 was used to analyse driver injury risk in two-car crashes. The effects of different factors on driver injury probability (Equations 5.33 and 5.34) were estimated by forming and maximising the log-likelihood function, described by Equation 5.22, over the two-car crash dataset. Four different distributional assumptions, defined and formulated in detail in Section 5.3.1, are investigated for the closing speed v; these are summarised in Table 5.13.

Table 5.13: Summary of the defined distributions for closing speed v

Distribution of v	Type A (distribution is independent of speed limit v^l)	Type B (distribution is dependent on speed limit v^l)
Normal (D1)	$v \sim N(m,\sigma)$ $m' = \beta m$ $u = \beta v \sim N(m',\sigma)$ $u = m'(1 + zC)$ $f(u) = \phi(z)$ Unknown parameters: m', C	$v \sim N(m,\sigma)$ $m' = \beta m$ $u = \beta v \sim N(m',\sigma)$ $m' = sv^l$ $u = sv^l(1 + zC)$ $f(u) = \phi(z)$ Unknown parameters: s, C
Log-normal (D2)	$v \sim LogN(m,\sigma)$ $u = \beta v \sim LogN(m + \log_e \beta, \sigma)$ $u = \beta exp(m) z^\sigma$ $u = \theta z^\sigma$ $f(u) = 1/z\, \phi(\log_e z)$ Unknown parameters: θ, σ	$v \sim LogN(m,\sigma)$ $u = \beta v \sim LogN(m + \log_e \beta, \sigma)$ $m = log_e(\omega_2 s^l)$ $u = \beta \omega_2 z^\sigma s^l$ $u = tz^\sigma s^l$ $f(u) = 1/z\, \phi(\log_e z)$ Unknown parameters: t, σ

v^l: speed limit

---

[1] In this section, "injury" refers to fatality or serious injury (KSI); therefore, slight injury is treated as "no injury".

The maximum likelihood estimation was performed using the "R" software (different optimisation algorithms are available all of which were shown to give almost identical results). In calculating the log-likelihood function value (Equation 5.22), the integration in Equations 5.27 to 5.29 was performed numerically using Simpson's rule. In the case of the normal distribution, the integration was carried over z in the interval $[-4, +4]$, where the area under the curve $\phi(z)$ (see Equation 5.26) is almost equal to 1, with the increments of 0.05. In the case of the log-normal distribution, the integration is carried over z in the interval $[0{,}40]$, where the area under the curve $\frac{1}{z}\phi(z)$ (see Equation 5.38) is almost equal to 1, with the increments in z of 0.01. The analysis was performed separately for three different collision types: front to front collisions, front to back collisions, and front to side collisions.

5.3.3.1. Front to front collisions

i) Maximum likelihood estimation results

In the model estimation process for front to front collisions, vehicles 1 and 2 had the same labels (vehicle 1, vehicle 2) as those in the original STATS19 data. In the first step, the simplest model form that includes no driver or vehicle effects except mass ratio μ (i.e. $C_1 = C_2 = \propto_0$) was estimated for different closing speed distributions to find the distribution form that led to the best description of the injury severity distribution. Therefore there were three parameters to estimate for each distribution form: the constant $\propto_0$ that represents C_1 and C_2 in Equations 5.33 and 5.34, and two parameters that describe the closing speed distribution as summarised in Table 5.13. The maximum likelihood estimation results for normal distribution of v and log-normal distribution of v are shown in Table 5.14 and 5.15, respectively. In these tables, the models have the same labels as their assumed distribution equivalents. Figures 5.4 and 5.5 show how the maximised log-likelihood value varies by the coefficient of variation of closing speed (C) (in the normal distribution) and the standard deviation of logarithm of closing speed (σ) (in the log-normal distribution), respectively. As expected, the plots show that, for each distribution form, the maximised log-likelihood values are identical when $C=\sigma=0$.

Table 5.14: Maximum likelihood estimation results: Normal distribution of *v*

Parameters	Model D1.A0				Model D1.B0			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
$\propto_0$	-3.82	0.43	-4.66	-2.98	-3.07	0.23	-3.40	-2.56
m'	6.31	0.74	4.86	7.76	-	-	-	-
s	-	-	-	-	0.10	0.01	0.09	0.11
C	0.25	0.27	-0.27	0.77	0.19	0.17	-0.15	0.53
LL value	-2507.18				-2465.54			
AIC	2513				2472			
Obs	2485				2485			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

Table 5.15: Maximum likelihood estimation results: Log-normal distribution of *v*

Parameters	Model D2.A0				Model D2.B0			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
$\propto_0$	-4.39	0.60	-5.58	-3.21	-5.18	0.28	-5.73	-4.63
θ	5.00	1.46	2.13	7.87	-	-	-	-
t	-	-	-	-	0.07	0.01	0.05	0.09
σ	0.47	0.22	0.05	0.90	0.73	0.06	0.62	0.85
LL value	-2507.84				-2443.42			
AIC	2514				2449			
Obs	2485				2485			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

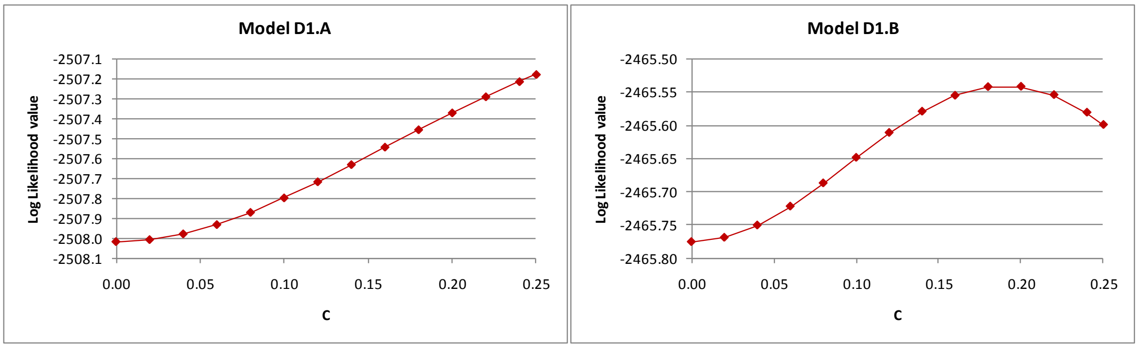


Figure 5.4: Maximised log-likelihood versus coefficient of variation of closing speed (*C*): normal distribution

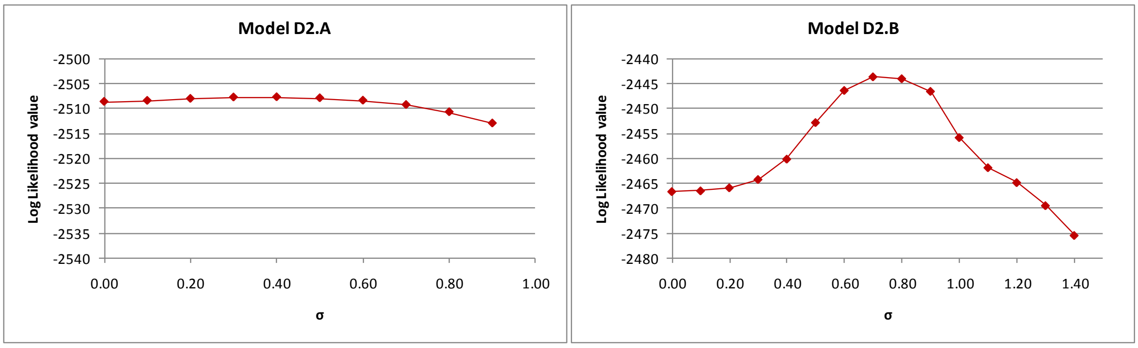


Figure 5.5: Maximised log-likelihood versus standard deviation of logarithm of closing speed (σ): log-normal distribution

The results, reflected in Tables 5.14 and 5.15 and Figures 5.4 and 5.5, show that the best model with highest log-likelihood value is Model D2.B0, in which all the estimated parameters are statistically significant and the log-likelihood is substantially better than the next best model. According to this model, the closing speed has a log-normal distribution with a mean value that depends on the speed limit ($m = log_e(\omega s^l)$). The variance of the distribution of closing speed has the single parameter σ, which in the lognormal distribution gives the relationship to the mean speed: $Var(v) = [E(v)]^2[\exp(\sigma^2) - 1]$. This model fitted the data substantially better than the corresponding one that did not use speed limit in the model for the distribution of closing speed. Therefore, Model D2.B0 was expanded to include variables related to driver and vehicle characteristics.

The variables related to driver age and driver gender were added to Model D2.B0 to investigate the effects of these factors. These contribute to C_1 and C_2 in Equations 5.33 and 5.34 as shown by Equation 5.13. The first hypothesised was that injury risk to the driver in each vehicle is influenced by the physical condition of the driver as represented by driver age and gender. These variables were added to Model D2.B0 to test this hypothesis. C_1 and C_2 for this model, labelled as Model D2.B1, are described as the following:

$$C_1 = \sum_i \alpha_{1i} x_{1i}\ ;\ i = 0,1,2, \dots \&\ x_{10} = 1 \tag{5.43}$$

$$C_2 = \sum_i \alpha_{2i} x_{2i}\ ;\ i = 0,1,2, \dots \&\ x_{20} = 1. \tag{5.44}$$

The maximum likelihood estimation results for Model D2.B1 are reflected in Table 5.16. The widely-used dummy coding method has been used to code the categorical variables related to driver age and gender that were shown in Table 5.12 where male driver aged 35-54 is taken as the reference category.

Table 5.16: Maximum likelihood estimation results: adding the effects of driver characteristics

Parameters	Model D2.B1			
	Est.	Std. Error	CI-	CI+
$\alpha_{10} = \alpha_{20}$	-5.178	0.268	-5.703	-4.653
$\alpha_{11} = \alpha_{21}$ (Female)	-0.417	0.080	-0.574	-0.260
$\alpha_{12} = \alpha_{22}$(Age 17-24)	-0.241	0.102	-0.441	-0.040
$\alpha_{13} = \alpha_{23}$(Age 25-34)	-0.190	0.100	-0.385	0.006
$\alpha_{14} = \alpha_{24}$ (Age +55)	0.646	0.100	0.449	0.843
t	0.060	0.010	0.041	0.079
σ	0.797	0.061	0.677	0.917
LL value	-2393.72			
AIC	2408			
Obs	2485			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

It was then hypothesised that injury risk to the driver in each vehicle is influenced not only by the physical vulnerability of its driver, but also by the driving style of the driver of the colliding vehicle as represented by variables age and gender (e.g. the effect of more aggressive driving). Therefore in the new model, labelled as Model D2.B2, C_1 and C_2 are described as the following:

$$C_1 = \sum_i \alpha_{1i} x_{1i} + \alpha_{2i} x_{2i}\,;\ i = 0,1,2,\ldots \,\&\, x_{10} = x_{20} = 1 \tag{5.45}$$

$$C_2 = \sum_i \alpha_{1i} x_{2i} + \alpha_{2i} x_{1i}\,;\ i = 0,1,2,\ldots \,\&\, x_{10} = x_{20} = 1. \tag{5.46}$$

The estimation results for Model D2.B2 are shown in Table 5.17.

Table 5.17: Maximum likelihood estimation results: adding the effects of driver characteristics and driving style

Parameters	Model D2.B2			
	Est.	Std. Error	CI-	CI+
$\alpha_{10} = \alpha_{20}$	-5.185	0.581	-6.324	-4.046
α_{11} (Female)	0.320	0.320	-0.307	0.946
α_{12}(Age 17-24)	-0.001	0.378	-0.742	0.740
α_{13}(Age 25-34)	-0.150	0.391	-0.916	0.616
α_{14} (Age +55)	-0.410	0.418	-1.230	0.410
α_{21} (Female)	0.718	0.316	0.099	1.337
α_{22}(Age 17-24)	0.189	0.366	-0.528	0.906
α_{23}(Age 25-34)	0.023	0.394	-0.751	0.796
α_{24} (Age +55)	-1.063	0.425	-1.895	-0.230
t	0.086	0.009	0.069	0.103
σ	0.629	0.051	0.528	0.729
LL value	-2393.724			
AIC	2416			
Obs	2485			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

Comparison of the maximum likelihood estimation results for Models D2.B0, D2.B1, and D2.B2 reflected in Tables 5.15 to 5.17 shows that Model D2.B1 has the best goodness of fit (measured by log-likelihood through the AIC) as well as the estimated parameters that are statistically significant (except for $\alpha_{13} = \alpha_{23}$ which shows that there is no difference between age range 25-34 and the reference category of 35-54). Therefore it is the best model that represents the effects of drivers' age and gender on driver injury probability in two-car crashes. Thus we cannot detect through this modelling approach any age and gender-specific effect on driving style that influence injury risk in the colliding vehicles.

One of the fundamental questions in the analysis of injury risk in two-car crashes which has remained unclear is whether there is any effect of vehicles' size beyond the effect of mass ratio. In order to examine this, the variables related to vehicle size were added to Model D2.B1 as explanatory variables (see Equations 5.43 and 5.44). Two models were estimated: Model D2.B3 in which vehicle size is represented by “vehicle length” (m), and Model D2.B4 in which vehicle size is represented by “vehicle length × vehicle width” (m^2)[1]. The results, reflected in Table 5.18, shows that both models have a better goodness of fit than that of Model D2.B1 (measured by log-likelihood through the AIC) as well as statistically significant estimated coefficients for the variable “Size”. This confirms that there is an effect of vehicle size above and beyond that of mass ratio in frontal collisions. The negative coefficient of size in these models, which are statistically significant at 5% level, shows that vehicle size is protective. The goodness of fit of Model D2.B4 is significantly better than that of Model D2.B3. Therefore Model D2.B4, in which vehicle size is represented by “vehicle length × vehicle width”, is the best model that reflects the partial effects of different contributing factors on driver injury probability in frontal crashes where injury is defined by either fatality or serious injury.

Table 5.18: Maximum likelihood estimation results: adding the effects of vehicle size

Parameters	Model D2.B3				Model D2.B4			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
$\alpha_{10} = \alpha_{20}$	-3.002	1.006	-4.973	-1.031	-2.982	0.691	-4.336	-1.627
$\alpha_{11} = \alpha_{21}$ (Female)	-0.388	0.081	-0.547	-0.230	-0.401	0.081	-0.559	-0.243
$\alpha_{12} = \alpha_{22}$(Age 17-24)	-0.283	0.104	-0.486	-0.080	-0.274	0.103	-0.477	-0.072
$\alpha_{13} = \alpha_{23}$(Age 25-34)	-0.206	0.100	-0.403	-0.010	-0.194	0.101	-0.391	0.003
$\alpha_{14} = \alpha_{24}$ (Age +55)	0.625	0.101	0.427	0.822	0.635	0.101	0.438	0.833
$\alpha_{15} = \alpha_{25}$ (Size)	-0.401	0.174	-0.741	-0.060	-0.204	0.056	-0.314	-0.094
t	0.047	0.011	0.025	0.068	0.038	0.011	0.016	0.059
σ	0.899	0.092	0.719	1.078	0.983	0.115	0.758	1.209
LL value	-2389.86				-2383.36			
AIC	2406				2399			
Obs	2485				2485			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

The estimated values of parameters in Model D2.B4 were used to predict driver injury probabilities for different values of the explanatory variables using Equations 5.33 and 5.34 when in these equations, C_1 and C_2 are given by Equations 5.43 and 5.44, and $\mathrm{u}(z)$ and $\mathrm{g}(z)$ are given by Equations 5.37 and 5.38, respectively. As was mentioned earlier,

---

[1] Due to a substantial number of missing data on vehicle height, the effect of volume is not examined in these models.

the integration is calculated numerically using Simpson's rule[1] over the values of z in the interval [0,40] with increments of 0.01.

ii) Effects of vehicle mass

The estimated injury probabilities for a few examples of two-car crashes are shown in Table 5.19 where examples are defined depending on the values of mass ratio μ and speed limit; these are for crashes with drivers in the reference category (male drivers aged 35-54). As the model estimation results showed, the driver injury probabilities (P_1 and P_2) are influenced not only by mass ratio, but also by "Size" of the vehicles in crash. Therefore, the estimated values of P_1 and P_2 can be different for a given value of μ depending on the dimensions of the vehicles. The relationship between vehicle mass and "Size" in the dataset is shown in Figure 5.6 when "Size" is defined by "vehicle length × vehicle width". The trend in the data is closer to an exponential function than a linear one. In the two-car crash examples in Table 5.19, an average value of "Size" is calculated for the given values of mass using the relationship shown in Figure 5.6; these are used in estimating P_1 and P_2.

Table 5.19: The effect of mass ratio (μ) on injury probabilities (P_1 and P_2) in frontal collisions

Crash		μ (m_2/m_1)	Speed limit	P_1	P_2	$R=P_1/P_2$
1	$m_1=1000$ $m_2=1000$	1.0	40	0.079	0.079	1.00
2	$m_1=1000$ $m_2=1500$	1.5	40	0.101	0.043	2.38
3	$m_1=1000$ $m_2=2000$	2.0	40	0.116	0.025	4.56
4	$m_1=1000$ $m_2=1000$	1.0	60	0.135	0.135	1.00
5	$m_1=1000$ $m_2=1500$	1.5	60	0.170	0.078	2.19
6	$m_1=1000$ $m_2=2000$	2.0	60	0.194	0.048	4.00

The results suggest that, for example, if two cars with a similar mass (1000 kg) crash into each other in a road where the speed limit is 60 mile/hr, the probability of each driver being killed or seriously injured is about 13.5%. However, if car 2 had a mass twice that of car 1 (1000 kg compared to 2000 kg), the probability of driver of car 1 (lighter car) being killed or seriously injured would increase to about 19.4% while the probability of driver of car 2 (heavier car) being killed or seriously injured would decrease to about 4.8%. These results are consistent with vehicle mass having both

---

[1] For example, see (Moin, 2001) for description

protective and aggressive effects in two-car crashes. The results also show that, in general, the probability of injury increases with speed limit; this represents the effect of the closing speed of the vehicles involved in the collision on driver injury probability.

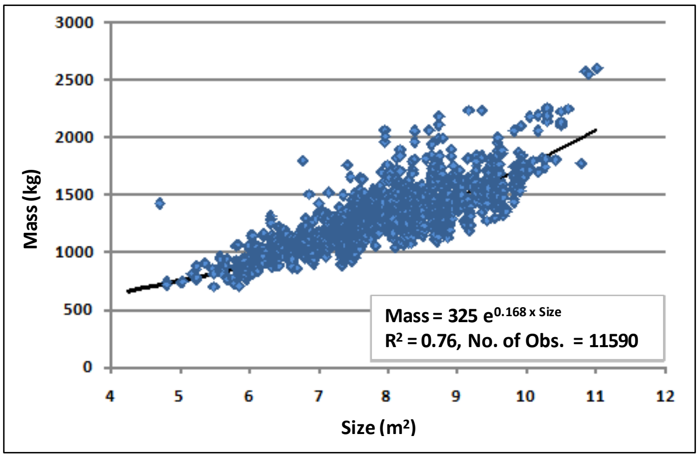


Figure 5.6: Relationship between vehicle mass and size (Length × Width)

iii) Effects of vehicle size

It was shown that vehicle size has a protective effect above the effect of mass ratio in frontal two-car crashes. It was also shown that the best variable representing the effect of vehicle size is the product of vehicle length and vehicle width. The estimated effects of vehicle size, based on the estimated Model D2.B4, are shown for a few examples of frontal two-car crashes in Table 5.20; these are for crashes with drivers in the reference category (male drivers aged 35-54).

Table 5.20: The effects of vehicle mass (kg) and vehicle size (Length × Width (m^2)) on injury probabilities (P_1 and P_2) in frontal collisions

Crash		μ (m_2/m_1)	Speed limit	P_1	P_2	$R=P_1/P_2$
1	m_1=1000 $Size_1$=6 m_2=1000 $Size_2$=6	1.0	60	0.145	0.145	1.00
2	m_1=1000 $Size_1$=6 m_2=1000 $Size_2$=7	1.0	60	0.145	0.131	1.10
3	m_1=1000 $Size_1$=6 m_2=1000 $Size_2$=8	1.0	60	0.145	0.119	1.21
4	m_1=1000 $Size_1$=6 m_2=1500 $Size_2$=9	1.5	60	0.181	0.079	2.30
5	m_1=1000 $Size_1$=7 m_2=1500 $Size_2$=9	1.5	60	0.166	0.079	2.11
6	m_1=1000 $Size_1$=8 m_2=1500 $Size_2$=9	1.5	60	0.152	0.079	1.94

The results are shown for two sets of mass ratios (1.0 and 1.5) where the size of one of the cars varies while all other factors including mass of the two cars are kept constant. Comparison of the estimated values of P_1 and P_2 for crashes 1 to 3 shows that increasing “Size” for car 2 from 6 m^2 to 8 m^2, when its mass is constant (1000 kg), decreases the probability of its driver being killed or seriously injured from about 14.5% to about 11.9%. On the other hand, in a frontal crash where mass ratio is 1.5 (1000 kg compared to 1500 kg), increasing “Size” for the lighter car (car 1) from 6 m^2 to 8 m^2 decreases its driver injury probability from about 18.1% to about 15.2% without affecting the driver injury probability of car 2.

The findings on the effects of vehicle size is important from the policy point of view because the relationship between mass and size reflected in Figure 5.6 suggests that there is the potential to make changes to vehicle design to increase the size of vehicles while vehicle mass is maintained. This could increase the safety performance of a vehicle without any adverse impact on the safety performance of the other vehicles in the fleet.

iv) Effects of driver factors

The estimated coefficients of driver age and driver gender variables were used to estimate their partial effects on driver injury risks in both vehicles using the same methodology that was used to estimate the partial effects of mass ratio and vehicle size. The results for a few examples of frontal collisions where they are different only in a driver factor are reflected in Table 5.21.

Table 5.21: The effects of driver age and gender on injury probabilities (P_1 and P_2) in frontal collisions

Crash		μ (m_2/m_1)	Speed limit	P_1	P_2	$R=P_1/P_2$
1	m_1=1000 Driver$_1$=male aged 35-54 m_2=1000 Driver$_2$=female aged 35-54	1.0	60	0.135	0.112	1.21
2	m_1=1000 Driver$_1$=male aged 35-54 m_2=1000 Driver$_2$=male aged 17-24	1.0	60	0.135	0.119	1.14
3	m_1=1000 Driver$_1$=male aged 35-54 m_2=1000 Driver$_2$=male aged +55	1.0	60	0.135	0.183	0.74

A negative coefficient for female driver in Model D2.B4 shows a lower injury probability for female drivers than male drivers. The results show that, for example, in a frontal collision between two cars with the same mass (1000 kg) but different driver genders, the probability of injury for the male driver is about 13.5% while the probability of injury for the female driver is about 11.2%. This effect is not in accordance with the general expectation that female drivers are generally more vulnerable than male drivers when involved in similar crashes due to a relatively less physical strength. One possible explanation might be given by the type of cars female drivers tend to driver compared to male drivers. For example, they might tend to drive model variants that are newer or have better secondary safety features. Examination of the available two-car crash data (crashes between 2000-2006) shows that, for example, about 59% of the vehicles that were driven by male drivers were registered for the first time before 2000 while, for female drivers, this figure is about 50%. This suggests there is a general tendency for female drivers to drive vehicles that are newer compared to male drivers. It should be noted that in the analysis of two-car crashes in Great Britain during 2000-2004, DfT (2006) found the consistent results that female drivers are less likely to be killed than men drivers when involved in the crashes.

On the other hand, the estimated effects for driver age show that a younger driver has a lower risk of injury than an older driver of a similar vehicle when involved in crashes (P_2=0.119 in crash 2 compared to P_2=0.183 in crash 3); this is in accordance with the prevailing wisdom.

5.3.3.2. Front to side collisions

i) Maximum likelihood estimation results

In analysing front to side collisions, vehicles 1 and 2 are labelled so that vehicle 1's first point of impact is front while vehicle 2's first point of impact is side (nearside or offside). Similar to the analysis of frontal collisions explained in Section 5.3.3.1, in the first step, the simplest model form that includes no driver or vehicle effects except mass ratio μ was estimated for different closing speed distributions to find the distribution form that best describes the closing speed v in this type of collisions. Unlike the symmetric case of frontal collisions, different constants are assumed for the vehicles involved in front to side collisions (i.e. $C_1 = \alpha_{10}$, $C_2 = \alpha_{20}$). The maximum likelihood estimation results for normal distribution of v and log-normal distribution of v are shown in Tables 5.22 and 5.23, respectively. In these tables, the models have the same labels as their assumed distribution equivalents. Figures 5.7 and 5.8 show how maximised log-likelihood varies by coefficient of variation of closing speed (C) (in the normal distribution) and standard deviation of logarithm of closing speed (σ) (in the log-normal distribution), respectively.

Table 5.22: Maximum likelihood estimation results: Normal distribution of *v*

Parameters	Model D1.A0				Model D1.B0			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
$\propto_{10}$	-3.990	0.558	-5.083	-2.897	-3.638	0.223	-4.075	-3.200
$\propto_{20}$	-3.513	0.556	-4.602	-2.423	-3.160	0.225	-3.601	-2.718
m'	5.031	1.158	2.761	7.301	-	-	-	-
s	-	-	-	-	0.093	0.008	0.077	0.110
C	0.250	0.460	-0.652	1.152	0.199	0.166	-0.127	0.525
LL value	-1693.87				-1663.73			
AIC	1702				1672			
Obs	1820				1820			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

Table 5.23: Maximum likelihood estimation results: Log-normal distribution of *v*

Parameters	Model D2.A0				Model D2.B0			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
$\propto_{10}$	-4.373	0.514	-5.380	-3.365	-5.218	0.32	-	-4.591
$\propto_{20}$	-3.894	0.513	-4.899	-2.889	-4.729	0.31	-	-4.104
θ	4.182	1.057	2.110	6.255	-	-	-	-
t	-	-	-	-	0.052	0.01	0.03	0.072
σ	0.406	0.187	0.038	0.773	0.727	0.07	0.57	0.876
LL value	-1694.20				-1662.88			
AIC	1702				1671			
Obs	1820				1820			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

Similar to that in frontal collisions, the best model with highest likelihood value is Model D2.B0, in which all the estimated parameters are statistically significant at 5% level and the log-likelihood of -1662.9 is slightly better than the next best model (D1.B0 with log-likelihood of -1663.7). This model fitted the data substantially better than the corresponding one (D2.A0, log-likelihood -1694.2) that did not use speed limit in the model for the distribution of closing speed. Therefore, Model D2.B0 was expanded to include variables related to driver and vehicle characteristics.

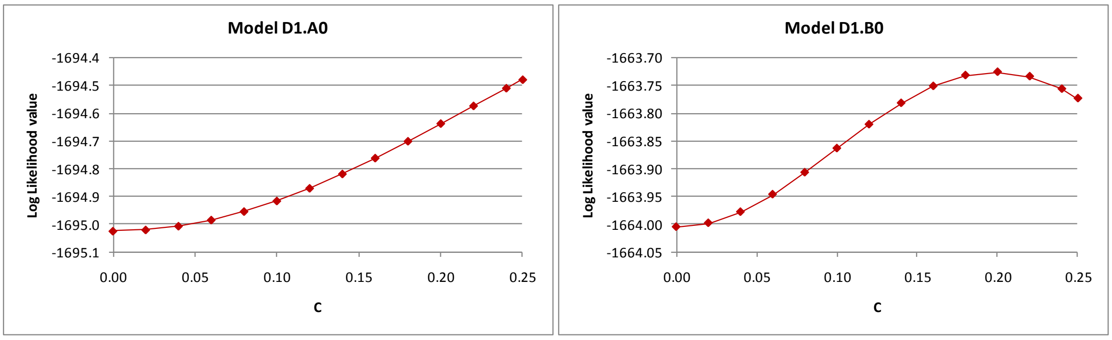


Figure 5.7: Maximised log-likelihood versus coefficient of variation of closing speed (*C*): normal distribution

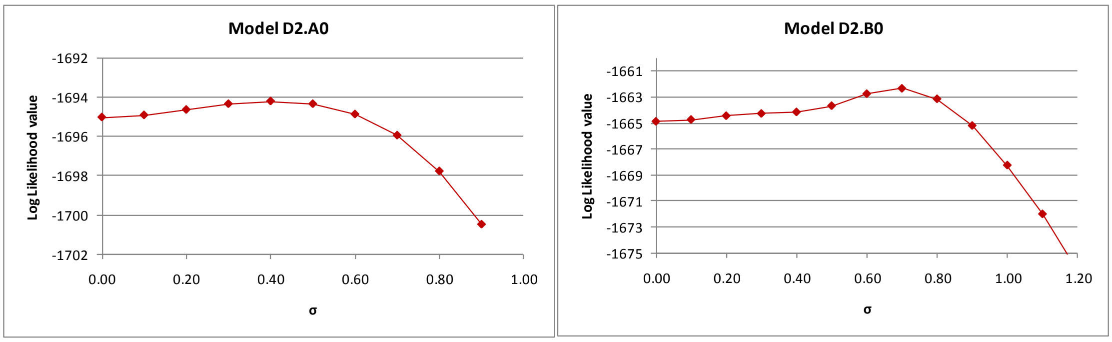


Figure 5.8: Maximised log-likelihood versus standard deviation of logarithm of closing speed (σ): log-normal distribution

The variables related to driver age and driver gender were added to Model D2.B0 to investigate the effects of these factors. These contribute to C_1 and C_2 in Equations 5.33 and 5.34 as shown by Equation 5.13. A similar modelling approach as that explained for frontal collisions was taken. The maximum likelihood estimation results for Models D2.B1 and D2.B2 are shown in Tables 5.24 and 5.25, respectively (see Section 5.3.3.1 for the definition of these models).

Similar to the case of frontal collisions, comparison of the maximum likelihood estimation results for Models D2.B0, D2.B1, and D2.B2 reflected in Tables 5.23 to 5.25 shows that Model D2.B1 has the best goodness of fit (measured by log-likelihood through the AIC). Therefore it is the best model that represents the effects of drivers' age and gender on driver injury probability in front to side crashes. The sign and significance of the estimated parameters are similar to that of frontal collisions; the only difference is in the estimated effect of driver age 17-24 which is not statistically significant in the case of front to side collisions so that there is no statistically significant difference between the injury risk of driver in the 3 age bands (17-24, 25-34, 35-54) when involved in front to side collisions.

Table 5.24: Maximum likelihood estimation results: adding the effects of driver characteristics

Parameters	Model D2.B1			
	Est.	Std. Error	CI-	CI+
α_{10}	-4.495	0.393	-5.265	-3.725
α_{20}	-4.032	0.393	-4.802	-3.263
$\alpha_{11} = \alpha_{21}$ (Female)	-0.338	0.084	-0.504	-0.173
$\alpha_{12} = \alpha_{22}$ (Age 17-24)	0.054	0.108	-0.158	0.265
$\alpha_{13} = \alpha_{23}$ (Age 25-34)	-0.094	0.166	-0.418	0.231
$\alpha_{14} = \alpha_{24}$ (Age +55)	0.566	0.109	0.353	0.779
t	0.051	0.011	0.031	0.072
σ	0.712	0.098	0.521	0.903
LL value	-1640.165			
AIC	1656			
Obs	1820			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

Table 5.25: Maximum likelihood estimation results: adding the effects of driver characteristics and driving style

Parameters	Model D2.B2			
	Est.	Std. Error	CI-	CI+
α_{10}	-5.349	0.609	-6.543	-4.155
α_{20}	-4.885	0.609	-6.079	-3.690
α_{11} (Female)	-0.029	0.320	-0.656	0.597
α_{12}(Age 17-24)	1.106	0.395	0.332	1.880
α_{13}(Age 25-34)	0.277	0.353	-0.415	0.968
α_{14} (Age +55)	-0.166	0.535	-1.215	0.883
α_{21} (Female)	0.317	0.317	-0.303	0.938
α_{22}(Age 17-24)	1.045	0.386	0.288	1.801
α_{23}(Age 25-34)	0.429	0.356	-0.270	1.127
α_{24} (Age +55)	-0.742	0.543	-1.807	0.323
t	0.046	0.009	0.028	0.064
σ	0.760	0.080	0.603	0.916
LL value	-1633.283			
AIC	1657			
Obs	1820			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

The results of adding “Size” variables to Model D2.B1 are reflected in Table 5.26 where the model labels and definitions are the same as those in the case of frontal collisions (see Section 5.3.3.1). The results show that both models have a better goodness of fit than that of Model D2.B1 (measured by log-likelihood through the AIC) as well as statistically significant estimated coefficients for variable “Size”. This confirms that there is an effect of vehicle size beyond that of mass ratio in front to side collisions as well. The negative coefficient of size in these models, which are statistically significant, shows that vehicle size is protective. The goodness of fit of Model D2.B4 is significantly better than that of Model D2.B3. This suggests that, similar to the case in frontal collisions, “vehicle length × vehicle width“ represents the influence of vehicle size on injury probability better than does “vehicle length”.

Table 5.26: Maximum likelihood estimation results: adding the effects of vehicle size

Parameters	Model D2.B3				Model D2.B4			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
α_{10}	-2.358	0.945	-4.211	-0.505	-2.124	0.680	-3.457	-0.791
α_{20}	-1.900	0.944	-3.750	-0.049	-1.673	0.678	-3.003	-0.344
$\alpha_{11} = \alpha_{21}$ (Female)	-0.307	0.086	-0.475	-0.139	-0.309	0.085	-0.476	-0.143
$\alpha_{12} = \alpha_{22}$ (Age 17-24)	0.009	0.110	-0.206	0.225	0.005	0.109	-0.208	0.218
$\alpha_{13} = \alpha_{23}$ (Age 25-34)	-0.104	0.166	-0.430	0.222	-0.128	0.164	-0.450	0.194
$\alpha_{14} = \alpha_{24}$ (Age +55)	0.552	0.109	0.338	0.766	0.541	0.109	0.327	0.755
$\alpha_{15} = \alpha_{25}$ (Size)	-0.388	0.162	-0.706	-0.070	-0.191	0.051	-0.291	-0.092
t	0.040	0.012	0.017	0.063	0.039	0.012	0.016	0.062
σ	0.794	0.139	0.521	1.067	0.775	0.165	0.451	1.100
LL value	-1637.36				-1633.468			
AIC	1655				1651			
Obs	1820				1820			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

Because front to side collisions, unlike frontal collisions, are not symmetric, it was hypothesised that the effects of vehicle size could be different for vehicles 1 and 2 depending on their point of first impact. In particular, it can be argued that for vehicle 2 whose point of first impact is side, “vehicle width” represents the influence of vehicle size better than “vehicle length × vehicle width“. To test these hypotheses, Models D2.B5 and D2.B6 were estimated to include separate coefficients for “Size” for each vehicle. In Model D2.B5 the variable “Size” for each vehicle is represented by “vehicle length × vehicle width“ while in Model D2.B6, “Size” in vehicle 2 is represented by “vehicle width”. The estimation results reflected in Table 5.27 show that although both coefficients for “Size” (α_{15} and α_{25}) are statistically significant, estimating separate coefficients for “Size” for each vehicle does not improve the goodness of fit of the model (the AIC of 1653 and 1654 for Models D2.B5 and D2.B6, respectively, compared to the AIC of 1651 for Model D2.B4). Therefore Model D2.B4, in which “Size” has a similar effect in both vehicles, remains the best model.

The fitted values of parameters in Model D2.B4 were used to estimate driver injury probabilities for different values of explanatory variables using Equations 5.33 and 5.34 where in these equations, C_1 and C_2 are given by Equations 5.43 and 5.44, and u and $g(z)$ are given by Equations 5.37 and 5.38, respectively. As was mentioned earlier, the integration is calculated numerically using Simpson’s rule over the values of z in the interval [0,40] with increments of 0.01.

Table 5.27: Maximum likelihood estimation results: adding separate effects of vehicle size

Parameters	Model D2.B5				Model D2.B6			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
α_{10}	-1.997	0.720	-3.409	-0.586	-2.275	0.740	-3.726	-0.824
α_{20}	-1.862	0.756	-3.344	-0.380	-1.475	1.063	-3.558	0.609
$\alpha_{11} = \alpha_{21}$ (Female)	-0.311	0.085	-0.477	-0.144	-0.323	0.084	-0.489	-0.158
$\alpha_{12} = \alpha_{22}$ (Age 17-24)	0.002	0.109	-0.211	0.216	0.022	0.109	-0.191	0.235
$\alpha_{13} = \alpha_{23}$ (Age 25-34)	-0.126	0.165	-0.449	0.196	-0.123	0.163	-0.442	0.197
$\alpha_{14} = \alpha_{24}$ (Age +55)	0.540	0.109	0.326	0.754	0.541	0.109	0.327	0.755
α_{15} (Size)	-0.211	0.062	-0.333	-0.090	-0.188	0.060	-0.306	-0.070
α_{25} (Size)	-0.169	0.063	-0.293	-0.044	-0.973	0.437	-1.830	-0.117
t	0.039	0.012	0.016	0.061	0.045	0.013	0.020	0.069
σ	0.783	0.164	0.461	1.105	0.721	0.168	0.391	1.051
LL value	-1633.303				-1634.33			
AIC	1653				1654			
Obs	1820				1820			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

ii) Effects of vehicle mass

The estimated injury probabilities for some examples of front to side crashes are shown in Table 5.28 where examples are defined depending on the values of mass ratio and speed limit; these are for crashes with drivers in the reference category (male drivers aged 35-54). Similar to the case of frontal collisions, an average value of "Size" is calculated for the given values of mass using the relationship shown in Figure 5.6; these were used in estimating P_1 and P_2. As specified above, the first point of impact for vehicle 1 and 2 are, respectively, front and side.

Table 5.28: The effect of mass ratio (μ) on driver injury probabilities (P_1 and P_2) in front to side collisions

Crash		μ (m_2/m_1)	Speed limit	P_1	P_2	$R=P_1/P_2$
1	m_1=2000 m_2=1000	0.5	40	0.037	0.206	2.80
2	m_1=1500 m_2=1000	0.67	40	0.061	0.185	0.33
3	m_1=1000 m_2=1000	1.0	40	0.112	0.155	0.72
4	m_1=1000 m_2=1500	1.5	40	0.138	0.088	1.57
5	m_1=1000 m_2=2000	2.0	40	0.156	0.056	0.18
6	m_1=2000 m_2=1000	0.5	60	0.062	0.311	0.20
7	m_1=1500 m_2=1000	0.67	60	0.100	0.280	0.36
8	m_1=1000 m_2=1000	1.0	60	0.179	0.232	0.77
9	m_1=1000 m_2=1500	1.5	60	0.222	0.137	1.62
10	m_1=1000 m_2=2000	2.0	60	0.251	0.087	2.88

As the results show for the crashes between cars of the same mass (crash 3 and 8), the injury probability of the driver of car 2, whose first point of impact is side, is greater than that of car 1, whose first point of impact is front. This is expected as more crumple room is available for the driver of vehicle 1 compared to that for the driver of vehicle 2. Similar to that for frontal collisions and as expected, the results also show that the probability of injury increases with speed limit; this represents the effect of the closing speed of the vehicles involved in the collision on driver injury probability.

iii) Effects of vehicle size

It was shown that vehicle size has a protective effect above the effect of mass ratio in front to side collisions where the effects were similar in vehicles 1 and 2. It was also shown that the best variable representing the effect of vehicle size is the product of vehicle length and vehicle width. The estimated effects of vehicle size, based on Model D2.B4, are shown for a few examples of front to side crashes in Table 5.29; these are for crashes with drivers in the reference category (male drivers aged 35-54). The results show that, for example, increasing "Size" for vehicle 2 from 8 m^2 to 10 m^2 in a front to side collision between cars with the same mass, all other factors being constant, decreases probability of its driver being killed or seriously injured from 20.1% to 16%. In each case, the effect of increasing vehicle size is to offer protection to the driver of that vehicle without affecting the injury risk of the other driver.

Table 5.29: The effect of vehicle size ("Length × Width" in m^2) on driver injury probabilities (P_1 and P_2) in front to side collisions

Crash		μ (m_2/m_1)	Speed limit	P_1	P_2	$R=P_1/P_2$
1	$m_1=1000$ $Size_1=6$ $m_2=1000$ $Size_2=6$	1.0	60	0.193	0.250	0.77
2	$m_1=1000$ $Size_1=6$ $m_2=1000$ $Size_2=8$	1.0	60	0.193	0.201	0.96
3	$m_1=1000$ $Size_1=8$ $m_2=1000$ $Size_2=6$	1.0	60	0.154	0.250	0.62
4	$m_1=1600$ $Size_1=8$ $m_2=1600$ $Size_2=8$	1.0	60	0.154	0.201	0.77
5	$m_1=1600$ $Size_1=8$ $m_2=1600$ $Size_2=10$	1.0	60	0.154	0.160	0.96
6	$m_1=1600$ $Size_1=10$ $m_2=1600$ $Size_2=8$	1.0	60	0.122	0.201	0.61

iv) Effects of driver factors

The fitted coefficients of driver age and driver gender were used to estimate their partial effects on driver injury risks in both vehicles in front to side collisions. The results for some example collisions where they are different in a driver factor are reflected in Table 5.30. In general, the effects are similar to those for frontal collisions; that is, a female driver has a lower risk of injury than a male driver when involved in similar front to side crashes and driver injury probability increases with driver age.

Table 5.30: The effects of driver age and gender on driver injury probabilities (P_1 and P_2) in front to side collisions

Crash		μ (m_2/m_1)	Speed limit	P_1	P_2	$R=P_1/P_2$
1	m_1=1000 Driver$_1$=male aged 35-54 m_2=1000 Driver$_2$=male aged 35-54	1.0	60	0.179	0.232	0.77
2	m_1=1000 Driver$_1$=male aged 35-54 m_2=1000 Driver$_2$=female aged 35-54	1.0	60	0.179	0.195	0.92
3	m_1=1000 Driver$_1$=female aged 35-54 m_2=1000 Driver$_2$=male aged 35-54	1.0	60	0.149	0.232	0.64
4	m_1=1000 Driver$_1$=male aged 35-54 m_2=1000 Driver$_2$=male aged +55	1.0	60	0.179	0.313	0.57
5	m_1=1000 Driver$_1$=male aged +55 m_2=1000 Driver$_2$=male aged 35-54	1.0	60	0.245	0.232	1.05

5.3.3.3. Front to back collisions

i) Maximum likelihood estimation results

In analysing front to back collisions, vehicles 1 and 2 are labelled so that the first point of impact of vehicle 1 is front while the first point of impact of vehicle 2 is back. Similar to the previous analyses, the simplest model form that includes no driver or vehicle effect except mass ratio μ (i.e. $C_1 = \alpha_{10}$, $C_2 = \alpha_{20}$) was estimated for different closing speed distributions to find the distribution form that best describes the closing speed v. The maximum likelihood estimation results for normal distribution of v and log-normal distribution of v are shown in Table 5.31 and 5.32, respectively. Figures 5.9 and 5.10 show how maximum log-likelihood varies by coefficient of variation of closing speed (C) (in the normal distribution) and standard deviation of logarithm of closing speed (σ) (in the log-normal distribution), respectively.

As the figures show, only in Model D2.B0 a maximum point is available for a positive value of σ (0.002). Similar to the case in front to front and front to side collisions, the best model with highest log-likelihood value is Model D2.B0, in which the log-likelihood is better than the next best model. This model fitted the data substantially better than the corresponding one that did not use speed limit in the model for the distribution of closing speed. Therefore, Model D2.B0 was expanded to include variables related to driver and vehicle characteristics.

Table 5.31: Maximum likelihood estimation results: Normal distribution of *v*

Parameters	Model D1.A0				Model D1.B0			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
α_{10}	-4.429	0.289	-4.994	-3.863	-4.507	0.372	-5.237	-3.778
α_{20}	-4.270	0.277	-4.812	-3.727	-4.358	0.365	-5.074	-3.642
m'	4.562	0.503	3.578	5.547	-	-	-	-
s	-	-	-	-	0.094	0.017	0.062	0.126
C	0.000	0.620	0.000	1.215	0.038	1.047	0.000	2.090
LL value	-714.86				-704.771			
AIC	723				713			
Obs	877				877			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

Table 5.32: Maximum likelihood estimation results: Log-normal distribution of *v*

Parameters	Model D2.A0				Model D2.B0			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
$\propto_{10}$	-4.420	0.288	-4.985	-3.855	-4.452	0.298	-5.036	-3.868
$\propto_{20}$	-4.261	0.276	-4.803	-3.719	-4.304	0.288	-4.869	-3.739
θ	4.569	0.503	3.583	5.555	-	-	-	-
t	-	-	-	-	0.093	0.010	0.074	0.112
σ	0.000	0.355	0.000	0.697	0.002	0.333	0.000	0.655
LL value	-715.54				-703.965			
AIC	724				712			
Obs	877				877			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

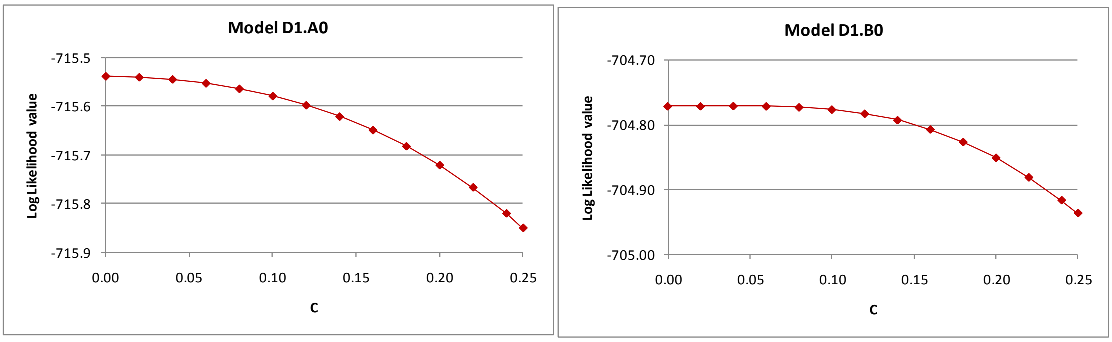


Figure 5.9: Maximised log-likelihood versus coefficient of variation of closing speed (*C*): normal distribution

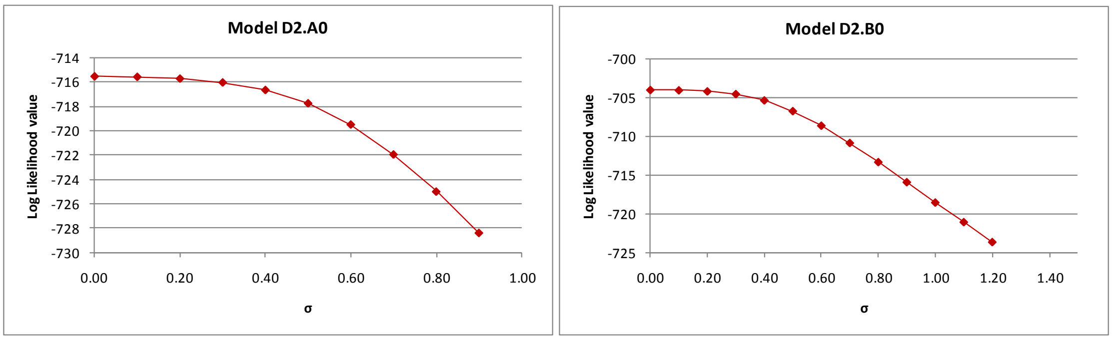


Figure 5.10: Maximised log-likelihood versus standard deviation of $\log_e$ (closing speed) (σ): log-normal distribution

The variables related to driver age and driver gender were added to Model D2.B0 to investigate the effects of these factors. These contribute to C_1 and C_2 in Equations 5.33 and 5.34 as shown by Equation 5.13. A similar modelling approach as that explained for frontal collisions was taken here. The maximum likelihood estimation results for Models D2.B1 and D2.B2 are shown in Tables 5.33 and 5.34, respectively (see Section 5.3.3.1 for the definition of these models).

Comparison of the maximum likelihood estimation results for Models D2.B0, D2.B1, and D2.B2 reflected in Tables 5.32 to 5.34 shows that Model D2.B1 has the best goodness of fit (measured by log-likelihood through the AIC). The interesting point to note about this model is that adding variables related to driver factors increases the estimated value of σ from 0.002 to 0.327 and decreases its standard error from 0.333 to 0.159. Therefore, Model D2.B1 is the best model that represents the effects of drivers' age and gender on driver injury probability in front to back collisions. The sign and significance of the estimated parameters are similar to that of front to side collisions.

Table 5.33: Maximum likelihood estimation results: adding the effects of driver characteristics

Parameters	Model D2.B1			
	Est.	Std. Error	CI-	CI+
α_{10}	-4.512	0.378	-5.253	-3.771
α_{20}	-4.370	0.369	-5.093	-3.647
$\alpha_{11} = \alpha_{21}$ (Female)	-0.262	0.110	-0.478	-0.045
$\alpha_{12} = \alpha_{22}$ (Age 17-24)	0.079	0.151	-0.216	0.375
$\alpha_{13} = \alpha_{23}$ (Age 25-34)	0.057	0.160	-0.258	0.371
$\alpha_{14} = \alpha_{24}$ (Age +55)	0.772	0.157	0.464	1.080
t	0.070	0.018	0.034	0.106
σ	0.327	0.159	0.015	0.639
LL value	-690.667			
AIC	707			
Obs	877			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

Table 5.34: Maximum likelihood estimation results: adding the effects of driver characteristics and driving style

Parameters	Model D2.B2			
	Est.	Std. Error	CI-	CI+
α_{10}	-4.104	0.454	-4.994	-3.215
α_{20}	-3.961	0.448	-4.840	-3.082
α_{11} (Female)	-0.190	0.220	-0.621	0.240
α_{12}(Age 17-24)	0.132	0.273	-0.403	0.667
α_{13}(Age 25-34)	-0.063	0.241	-0.536	0.410
α_{14} (Age +55)	-0.070	0.329	-0.716	0.575
α_{21} (Female)	0.073	0.220	-0.358	0.503
α_{22}(Age 17-24)	0.092	0.272	-0.441	0.625
α_{23}(Age 25-34)	-0.109	0.243	-0.587	0.368
α_{24} (Age +55)	-0.907	0.323	-1.540	-0.273
t	0.090	0.010	0.070	0.110
σ	0.000	0.911	0.000	1.785
LL value	-685.8283			
AIC	710			
Obs	877			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

The results of adding "Size" variables to the model are reflected in Table 5.35 where the model labels and definitions are the same as those in the case of frontal collisions (see Section 5.3.3.1). Although the coefficients for "vehicle length" in Model D2.B3 is statistically significant, adding this variable to Model D2.B1 does not improve the goodness of fit of the model and increases the standard error of the model constants substantially resulting in non-significant constant coefficients (α_{10} and α_{20}). Therefore the best model in the case of front to back collisions is Model D2.B1. Unlike the case of front to front and front to side collisions, the data does not show any effect of vehicle size over that of mass ratio in front to back collisions.

Table 5.35: Maximum likelihood estimation results: adding the effects of vehicle size

Parameters	Model D2.B3				Model D2.B4			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
α_{10}	-0.554	1.054	-2.619	1.511	-3.591	0.790	-5.140	-2.043
α_{20}	-0.378	1.060	-2.456	1.700	-3.441	0.792	-4.992	-1.889
$\alpha_{11} = \alpha_{21}$ (Female)	-0.218	0.112	-0.437	0.000	-0.249	0.112	-0.468	-0.030
$\alpha_{12} = \alpha_{22}$ (Age 17-24)	0.006	0.156	-0.299	0.312	0.067	0.152	-0.232	0.366
$\alpha_{13} = \alpha_{23}$ (Age 25-34)	0.081	0.165	-0.242	0.404	0.069	0.162	-0.249	0.386
$\alpha_{14} = \alpha_{24}$ (Age +55)	0.791	0.159	0.480	1.102	0.785	0.158	0.475	1.095
$\alpha_{15} = \alpha_{25}$ (Size)	-0.795	0.205	-1.196	-0.393	-0.091	0.070	-0.227	0.046
t	0.025	0.014	-0.002	0.051	0.052	0.022	0.010	0.095
σ	0.718	0.242	0.244	1.192	0.441	0.193	0.062	0.820
LL value	-689.65				-689.80			
AIC	708				708			
Obs	877				877			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

The estimated values of parameters in Model D2.B1 were used to predict driver injury probabilities for different values of the explanatory variables using Equations 5.33 and 5.29 where in these equations, C_1 and C_2 are given by Equations 5.43 and 5.44, and u and $g(z)$ are given by Equations 5.37 and 5.38, respectively. As before, the integration is calculated numerically using Simpson's rule over the values of z in the interval [0,40] with increments of 0.01.

ii) Effects of vehicle mass

The estimated injury probabilities for a few examples of front to back crashes are shown in Table 5.36 where examples are defined depending on the values of mass ratio and speed limit; these are for crashes with drivers in the reference category (male drivers aged 35-54). As noted above, the first point of impact for vehicle 1 and 2 are, respectively, front and back.

For the crashes between cars of the same mass (crash 3 and 8), the results show that the injury probability of the driver of car 2, whose first point of impact is back, is slightly greater than that of car 1, whose first point of impact is front. Similar to that for other collision types and as expected, the results also show that the probability of injury increases with speed limit; this represents the effect of the closing speed of the vehicles involved in the collision on driver injury probability.

Table 5.36: The effect of mass ratio (μ) on drive injury probabilities (P_1 and P_2) in front to back collisions

Crash		μ (m_2/m_1)	Speed limit	P_1	P_2	$R=P_1/P_2$
1	m_1=2000 m_2=1000	0.5	40	0.030	0.098	0.31
2	m_1=1500 m_2=1000	0.67	40	0.037	0.080	0.47
3	m_1=1000 m_2=1000	1.0	40	0.051	0.059	0.88
4	m_1=1000 m_2=1500	1.5	40	0.071	0.043	1.66
5	m_1=1000 m_2=2000	2.0	40	0.087	0.035	2.51
6	m_1=2000 m_2=1000	0.5	60	0.051	0.229	0.23
7	m_1=1500 m_2=1000	0.67	60	0.071	0.183	0.39
8	m_1=1000 m_2=1000	1.0	60	0.111	0.124	0.89
9	m_1=1000 m_2=1500	1.5	60	0.165	0.080	2.06
10	m_1=1000 m_2=2000	2.0	60	0.209	0.059	3.55

The estimated probabilities in Table 5.36 are generally less than the corresponding ones for front to front crashes (Table 5.19). This was expected because of vehicles here travelling in the same direction and the consequent reduction in closing speed. This only applies to the cases where mass of the vehicles are the same.

iii) Effects of driver factors

The estimated coefficients of driver age and driver gender were used to estimate their partial effects on driver injury risks in both vehicles in front to back collisions. The results for some example collisions where they are different in a driver factor are reflected in Table 5.37. In general, the effects are similar to those for other collisions; that is, a female driver has a lower risk of injury than a male driver and driver injury probability increases with driver age.

Table 5.37: The effects of driver age and gender on driver injury probabilities (*P_1* and *P_2*) in front to back collisions

Crash		μ (m_2/m_1)	Speed limit	P_1	P_2	$R=P_1/P_2$
1	m_1=1000 Driver$_1$=male aged 35-54 m_2=1000 Driver$_2$=male aged 35-54	1.0	60	0.111	0.124	0.89
2	m_1=1000 Driver$_1$=male aged 35-54 m_2=1000 Driver$_2$=female aged 35-54	1.0	60	0.111	0.101	1.10
3	m_1=1000 Driver$_1$=female aged 35-54 m_2=1000 Driver$_2$=male aged 35-54	1.0	60	0.089	0.124	0.72
4	m_1=1000 Driver$_1$=male aged 35-54 m_2=1000 Driver$_2$=male aged +55	1.0	60	0.111	0.222	0.50
5	m_1=1000 Driver$_1$=male aged +55 m_2=1000 Driver$_2$=male aged 35-54	1.0	60	0.201	0.124	1.61

5.3.4. Secondary safety performance of makes and models

As was mentioned in Chapters 1 and 2, there are two distinct measures of safety performance for a vehicle that is involved in a two-car crash: "Secondary Safety Performance" which is linked to the injury risk to the occupants of that vehicle, and "Aggressivity Performance" which is linked to the injury risk that the vehicle imposes to the occupants of the colliding vehicle. It was shown in the previous section that vehicle mass significantly contributes to both secondary safety and aggressivity performance of the vehicle (having both protective and aggressive effects in two-car collisions).

The methodology explained in Section 5.3.1 was used to investigate whether there are any specific effects of vehicle make and model on driver injury probability in frontal two-car collisions over and above the effects of mass (as represented by mass ratio). In order to investigate such effects, variables related to the makes and models of the vehicles involved in the collisions should be included in C_1 and C_2 as described by Equations 5.43 and 5.44.

Examination of frontal two-car crash data in which at least one of the drivers is KSI showed that makes and models in this dataset have a sparse distribution; hence, the sample size (crash involvements) for the majority of makes and models is too small to result in reliable estimates. It was also mentioned earlier in this chapter (Section 5.2.2) that in discussing the DfT (2006) method for estimating safety indices as a measure of secondary safety performance of vehicles, Broughton (1996c) recommended that it is more sensible to concentrate on the secondary safety estimates of makes and models based on “all casualties” rather than KSI as it is shown to be highly correlated with the estimates based on KSI casualties and it is more discriminating because of the much larger number of accidents used in the estimation. Therefore for this analysis, the dataset explained in Section 5.3.2 was expanded to include frontal collisions in which at least one of the drivers is slightly injured, seriously injured, or killed; this included 12,730 collisions occurred during 2000-2006. It should be noted that this dataset is dominated by slight injuries.

The previous analyses for different collision types consistently showed that a log-normal distribution for closing speed in which mean is related to speed limit, referred to as distribution D2.B (see Section 5.3.1 and Table 5.13 for details), is the best form of distribution to describe the closing speed. Therefore, the same form of distribution was used here to investigate the specific effects of makes and models. Three models were estimated. In the first step, the simplest model form that includes no driver or vehicle effects except mass ratio μ was estimated; this model is referred to as Model S1. Then the variables related to driver age and driver gender were added to this model; the resulting model is labelled as Model S2. The maximum likelihood estimation results for these two models are shown in Table 5.38. The results show that including the driver age and gender variables substantially improves the goodness of fit of the model (measured by log-likelihood value through the AIC). The sign and significance of these variables are also consistent with the findings in the previous section.

Table 5.38: Maximum likelihood estimation results: Log-normal distribution of v

Parameters	Model S1				Model S2			
	Est.	Std. Error	CI-	CI+	Est.	Std. Error	CI-	CI+
$\alpha_{10} = \alpha_{20}$	-3.708	0.145	-3.992	-3.424	-2.819	0.175	-3.161	-2.476
$\alpha_{11} = \alpha_{21}$ (Female)	-	-	-	-	-0.836	0.038	-0.911	-0.761
$\alpha_{12} = \alpha_{22}$ (Age 17-24)	-	-	-	-	-0.034	0.048	-0.129	0.061
$\alpha_{13} = \alpha_{23}$ (Age 25-34)	-	-	-	-	0.560	0.330	-0.087	1.207
$\alpha_{14} = \alpha_{24}$ (Age +55)	-	-	-	-	0.287	0.051	0.187	0.386
t	0.073	0.006	0.060	0.085	0.072	0.005	0.062	0.082
σ	1.141	0.051	1.041	1.240	1.172	0.049	1.076	1.268
LL value	-13114.12				-12850.36			
AIC	13120				12864			
Obs	12730				12730			

CI- and CI+ show, respectively, lower and upper bound of 95% confidence interval of the estimates

In the next step the fixed effects for makes and models were added to Model S2. The estimation results of this model, labelled as Model S3, are shown in Table 5.39. Make and model categories were defined for the car models with a minimum of 100 records (crash involvements) in the dataset; this is an arbitrary threshold which results in 43 make and model categories accounting for about 85% of all vehicles involved in crashes. The remaining 15% were placed in the category "other". Effect coding method was used to code make and model variables in order to avoid an arbitrary choice of a single vehicle make and model as the reference group and to allow estimation of relative effects of different makes and models (see Chapter 4 Section 4.3.3 for the definitions and details of effect coding method).

Table 5.39: Maximum likelihood estimation results: adding fixed effects of makes and models

Parameters.	Model S3			
	Est.	Std. Error	CI-	CI+
$\alpha_{10} = \alpha_{20}$	-2.453	0.239	-2.922	-1.985
$\alpha_{11} = \alpha_{21}$ (Female)	-0.848	0.039	-0.924	-0.771
$\alpha_{12} = \alpha_{22}$(Age 17-24)	-0.024	0.049	-0.120	0.072
$\alpha_{13} = \alpha_{23}$(Age 25-34)	0.438	0.293	-0.136	1.013
$\alpha_{14} = \alpha_{24}$ (Age +55)	0.285	0.052	0.183	0.386
t	0.069	0.005	0.059	0.080
σ	1.200	0.056	1.090	1.311
FORD FIESTA	-0.005	0.082	-0.165	0.155
VAUXHALL ASTRA	0.164	0.075	0.017	0.310
VAUXHALL CORSA	0.026	0.089	-0.149	0.201
FORD MONDEO	-0.119	0.086	-0.287	0.050
FORD ESCORT	0.033	0.092	-0.147	0.213
FORD FOCUS	-0.099	0.090	-0.276	0.079
VAUXHALL VECTRA	-0.087	0.091	-0.266	0.092
FIAT PUNTO	0.013	0.107	-0.197	0.223
RENAULT CLIO	-0.229	0.110	-0.445	-0.013
CITROEN SAXO	0.190	0.130	-0.065	0.445

Table 5.39: (continued)

PEUGEOT 206	0.084	0.118	-0.148	0.316
PEUGEOT 306	0.227	0.113	0.006	0.448
VOLKSWAGEN GOLF	-0.061	0.110	-0.277	0.155
NISSAN MICRA	0.327	0.136	0.061	0.594
RENAULT MEGANE	0.224	0.119	-0.008	0.457
ROVER 200/400	-0.108	0.124	-0.352	0.136
PEUGEOT 406	-0.061	0.124	-0.303	0.181
PEUGEOT 106	0.280	0.153	-0.021	0.580
HONDA CIVIC	0.284	0.133	0.024	0.544
RENAULT LAGUNA	-0.314	0.138	-0.585	-0.044
VOLKSWAGEN POLO	0.028	0.152	-0.269	0.325
BMW 3 series	-0.109	0.144	-0.391	0.172
FORD K	0.072	0.177	-0.276	0.419
ROVER 25/45	0.030	0.164	-0.292	0.352
VOLKSWAGEN PASSAT	0.267	0.154	-0.035	0.570
FORD KA	0.102	0.187	-0.265	0.469
NISSAN ALMERA	-0.017	0.174	-0.358	0.324
CITROEN XSARA	0.204	0.180	-0.150	0.558
TOYOTA AVENSIS	0.629	0.174	0.287	0.971
AUDI A4	-0.362	0.196	-0.745	0.021
NISSAN PRIMERA	0.060	0.194	-0.320	0.439
HONDA ACCORD	0.351	0.201	-0.042	0.744
PEUGEOT 307	0.039	0.218	-0.388	0.466
LAND ROVER DISCOVERY	-0.468	0.228	-0.914	-0.022
LAND ROVER FREELANDER	-0.334	0.235	-0.795	0.127
CITROEN XANTIA	-0.677	0.246	-1.159	-0.194
MERCEDES C CLASS	-0.237	0.232	-0.692	0.218
FIAT BRAVA	0.079	0.244	-0.400	0.558
SKODA OCTAVIA	0.467	0.227	0.023	0.912
MINI MINI	0.015	0.269	-0.513	0.542
TOYOTA COROLLA	-0.818	0.277	-1.361	-0.275
SEAT IBIZA	-0.116	0.275	-0.655	0.422
VAUXHALL OMEGA	0.123	0.256	-0.379	0.625
LL value	-12800.71			
AIC	12901			
N	12730			

Estimated coefficients that are statistically significant at 5% level are highlighted

Comparison of the estimation results for Model S3 and Model S2 shows that adding fixed effects of makes and models does not improve the goodness of fit of the model significantly: the AIC is increased from 12864 to 12901. Besides, the estimated effect of the majority of makes and models (32 out of 43) are not found to be statistically significant (the statistically significant effects are highlighted in the table). These confirm that there is no justification to include these make and model effects in the model.

In general, the modelling results based on “all casualty” frontal collision data during 2000-2006 suggest that there is no effect of make and model over and above that of mass ratio. Thus we conclude that the net secondary safety performance of different makes and models in two-car crashes is mainly explained by the effect of mass ratio.

5.4. Summary and conclusions

The analysis of driver injury risk in this chapter was divided into two main parts. In the first part, the relationship between vehicle mass and its relative driver injury risk in the 2000-2004 British fleet was investigated and the results were compared with the results of a similar analysis based on 1989-1992 period.

Modelling results showed that mass can explain a high proportion of variation in driver injury risk for cars when the driver injury risk was defined as the proportion of drivers injured when involved in two-car crashes where at least one of the drivers is injured. Based on the modelling results, a 100 kg increase in mass would decrease risk of injury to the driver in a two-car injury crash between 2.6% and 3.2%. This effect was derived from injury crashes in Great Britain from 2000 to 2004. Comparison of these results with results from a 1989-1992 period showed that the effect of mass on driver injury risk in fleet has changed considerably between these periods. This suggests that characteristics of the fleet, and in particular the distribution of mass within the fleet, is an important factor in determining the relationship between mass and secondary safety performance of individual vehicles at each point in time.

There were, therefore, two principal findings of the performed analysis on the effect of mass, which were in agreement with Equations 5.2 to 5.7. The first was that an increase in vehicle mass is associated with an increase in vehicle secondary safety performance as represented by the defined driver injury risk. The second is that an increase in the mass of a specific vehicle could be detrimental to the secondary safety performance of other vehicles within the fleet, all other things being constant. Based on these findings, it was concluded that in order to fully understand the likely changes in crash injury outcome of fleet as a result of a change in mass distribution within the fleet, a detailed analysis of two-car crashes is required to investigate both protective and aggressive effects of mass in crashes. Besides this, it was argued that it would be ideal to represent vehicle secondary safety performance by absolute driver injury risk of vehicles where in a two-car crash the injury risk in one vehicle is independent from the injury risk in the other vehicle in crash.

The analysis was extended to a disaggregate cross-sectional analysis of two-car crashes where a novel methodology was introduced to estimate partial effects of mass on absolute driver injury risk in each of the vehicles in the crash. In the introduced methodology, driver injury probability is described by a logistic function that includes, for each vehicle involved in the crash, the velocity change (defined as a function of mass ratio and closing speed) as well as various driver and vehicle characteristics. Because data on the speed of the vehicles prior to the crash is not available, a distribution for closing speed is assumed the parameters of which are estimated in model estimation process. The methodology uses the conditional joint injury probabilities in two-car crashes as the basis of analysis to solve the issue related to lack of data on crashes where no driver is injured; these conditional joint probabilities are used to form the likelihood function. The parameters describing the driver injury probability in each vehicle are estimated by maximising the likelihood function over the two-car crash dataset.

Three types of crashes were analysed separately: front to front crashes, front to side crashes, and front to back crashes. For all the crash types, it was found that the distribution form that best describes the closing speed is a log-normal distribution in which mean is related to the speed limit of the road. For all collision types, the results confirmed that in a two-vehicle collision, the probability of injury of the driver of vehicle 1 increases with speed limit and with increasing mass ratio ($\mu = m_2/m_1$) while the probability of injury of the driver of vehicle 2 increases with speed limit and with decreasing mass ratio. The results showed that, for example, if two cars with a similar mass (1000 kg) crash into each other in a road where the speed limit is 60 mile/hr, the probability of each driver being killed or seriously injured is about 13.5%. However, if car 2 had a mass twice that of car 1 (2000 kg compared to 1000 kg), the probability that the driver of car 1 (lighter car) is killed or seriously injured would increase to about 19.4% while the probability that the driver of car 2 (heavier car) is killed or seriously injured would decrease to about 4.8%.

Another novel aspect of the analysis based on the introduced methodology in this chapter was separating the effect of vehicle mass from that of vehicle size on absolute driver injury risks of the vehicles involved in a two-car crash, where vehicle size is represented by “vehicle length × vehicle width”. The results confirmed that there is a protective effect of vehicle size above and beyond that of vehicle mass for front to front and front to side crashes; the data did not show any effect of vehicle size in front to back crashes.

The findings on the effects of vehicle age and gender on the probability of driver injury were generally consistent for all collision types. The effects of driver age were in accordance to the prevailing wisdom; that is, a younger driver suffers less injury than an older driver when involved in a comparable crash due to relatively better physical strength and less vulnerability. However, the findings on effects of driver gender were not in accordance with this prevailing wisdom. One might expect female drivers are generally more vulnerable than male drivers when involved in similar crashes. Instead, it was shown consistently for all collision types that a female driver will probably suffer less injury than a male driver when involved in similar crashes. Given the stability and consistency of this effect between different collision types and different levels of injury ("KSI" and "all injuries"), one possible explanation could be given by the type of cars female drivers tend to driver compared to male drivers. For example, they might tend to drive model variants that are newer or have better secondary safety features.

The introduced methodology was also used to investigate whether there are any specific effects of vehicle makes and models on driver injury probability in frontal two-car collisions over and above the effects of mass (as represented by mass ratio). The analysis results based on frontal collisions in which there is at least one driver injury (of any level) during 2000-2006 suggested that there is no statistically significant (5%) effect of make and model over and above that of mass.

The estimated effects of vehicle mass and size in two-car crashes presented in this chapter will be used to investigate likely changes in injury outcome of two-car crashes in the UK fleet of 2007 with a number of variations in different mass distribution scenarios. The results of these analyses will be presented in Chapter 6 of this thesis.

CHAPTER 6. SAFETY AND ENVIRONMENTAL CONSEQUENCES OF CHANGES IN FLEET MASS DISTRIBUTION

The objective of this chapter is to investigate likely safety and environmental consequences of changes in mass distribution within the vehicle fleet by estimating the partial effects of a number hypothetical mass distributions defined relative to a base mass distribution. This chapter is organised as follows. A brief background is given in the first section (6.1). The methodology and the data used for the analysis are explained in the second (6.2) and third (6.3) sections, respectively. The fourth section (6.4) reflects the analysis results. The chapter ends by providing a brief discussion on the results and some concluding remarks (6.5).

6.1. Background

The characteristics of a fleet of vehicles within a country is continuously changing over time as new cars enter the fleet, some become older, and others leave the fleet. Besides, the vehicle usage pattern does not always remain constant over time. These changes influence safety and environmental outcomes of the vehicle fleet in different ways. One of these key changes relate to the mass distribution of vehicles within the fleet. The effect of such a change on overall safety and fuel economy of a vehicle fleet has generated a lot of debates amongst policy makers on whether there is a conflict between the overall goals as a result of the trade-off between fuel economy and safety performance in individual vehicles design within the fleet which is imposed by vehicle mass. This will be discussed in this chapter.

The conclusions of many of previous studies that have investigated this issue are based on aggregate analysis of several observations over a number of years. Regardless of the limitations in the methodologies used, which were discussed in detail in Chapter 2, such analyses suffer from a common important problem: the influence of different contributing factors that change alongside mass over time are not fully controlled. Therefore, the conclusions on the effects of changes in the composition of vehicle fleet do not reflect the isolated influence of changes in vehicles' mass on overall fleet safety and fuel economy. A different approach is used here; the partial effects of a number of hypothetical mass distribution scenarios, where the effect is only the result of a change in the mass distribution of vehicle fleet holding all other factors constant, on overall fleet fuel economy and safety is estimated. Such estimates are unlikely to be achieved exactly in reality as a number of other contributing factors including vehicle ownership and vehicle usage pattern are also likely to change over time as the mass distribution changes; however, they provide the necessary basis to formulate policies related to the vehicle fleet that aim at reducing overall fuel consumption or the number of crash injuries and fatalities where no adverse impact on either side would be acceptable.

6.2. Methodology

To investigate the safety and environmental consequences of different mass distribution scenarios, an incremental approach is introduced that estimates only the relative changes from a base case (for which observed data is available) in overall fuel consumption and crash injuries as a result of a hypothetical change in vehicles' mass in the fleet, holding all other factors constant. An incremental approach is consistent with the methodologies used in Chapters 4 and 5 to estimate the partial effects of mass on fuel consumption and secondary safety performance, the results of which are used as part of the introduced method in this chapter. Besides, such an approach includes all the key characteristics of the base vehicle fleet ,with respect to which the relative changes are estimated.

6.2.1. Fuel consumption of a vehicle fleet

Overall fuel consumption of a vehicle fleet depends on the total distance travelled by vehicles and the fuel consumption rate of different types of vehicles according to the following equation:

$$Y = \sum_{s,d} Y_{sd} = \sum_{s,d} N_s \times \overline{D}_{sd} \times \overline{F}_{sd} \tag{6.1}$$

where,

Y is the overall fuel consumption of vehicle fleet (in volume),

Y_{sd} is the total fuel consumption of vehicles in design segment s driven under driving cycle d (in volume),

N_s is the number of vehicles in design segment s,

$\overline{D}_{sd}$ is the mean distance travelled of vehicles in design segment s when driven under driving cycle d, and

$\bar{F}_{sd}$ is the mean fuel consumption rate of vehicles in design segment *s* when driven under driving cycle *d* (in volume per unit of distance travelled).

The greater the number of fuel consumption categories (defined by design segment *s* and driving cycle *d*) in Equation 6.1, the more precise the estimate of overall fuel consumption. The following explains how this equation is used as the basis to estimate the relative change in overall fuel consumption as a result of a change in fleet characteristics which accordingly changes the determinants of fleet fuel consumption shown in Equation 6.1. Taking one of the fuel consumption categories as the reference (denoted by subscript r), the following parameters are defined for the base vehicle fleet (denoted by superscript B). They reflect the relative differences of fuel consumption components (see Equation 6.1) between different fuel consumption categories and the reference category:

$$n_s^B = \frac{N_s^B}{N_r^B} \tag{6.2}$$

$$d_{sd}^B = \frac{\bar{D}_{sd}^B}{\bar{D}_r^B} \tag{6.3}$$

$$f_{sd}^B = \frac{\bar{F}_{sd}^B}{\bar{F}_r^B} \tag{6.4}$$

where,

n_s^B reflects the relative number of vehicles in design segment *s* in the base fleet,

d_{sd}^B reflects the relative mean distance travelled by vehicles in design segment *s* when driven under driving cycle *d* in the base fleet, and

f_{sd}^B reflects the relative mean fuel consumption rate of vehicles in design segment *s* when driven under driving cycle *d* in the base fleet.

Therefore, the relative total fuel consumed in a fuel consumption category *sd* in the base fleet is written as

$$\frac{Y_{sd}^B}{Y_r^B} = n_s^B \times d_{sd}^B \times f_{sd}^B. \tag{6.5}$$

A change in each of the fuel consumption components shown in Equation 6.1 changes the overall fuel consumption of vehicle fleet. An alternative fleet (denoted by superscript A) is considered where these components differ from those in the base fleet. As discussed before, the objective of this incremental approach is to estimate the relative change in overall fuel consumption with respect to that in the base fleet. The following parameters reflecting the relative changes in fuel consumption components between the base fleet and alternative fleet for each of the fuel consumption categories are defined:

$$\alpha_s = \frac{N_s^A}{N_s^B} \tag{6.6}$$

$$\beta_{sd} = \frac{\bar{D}_{sd}^A}{\bar{D}_{sd}^B} \tag{6.7}$$

$$\gamma_{sd} = \frac{\bar{F}_{sd}^A}{\bar{F}_{sd}^B}\,. \tag{6.8}$$

Using Equations 6.1 to 6.8, the ratio of overall fuel consumption in the alternative fleet (Y^A) to that in the base fleet (Y^B) can be written as

$$\frac{Y^A}{Y^B} = \frac{\sum_{s,d}(\alpha_s \times \beta_{sd} \times \gamma_{sd})\left(N_s^B \times \bar{D}_{sd}^B \times \bar{F}_{sd}^B\right)}{\sum_{s,d} N_s^B \times \bar{D}_{sd}^B \times \bar{F}_{sd}^B} = \frac{\sum_{s,d}(\alpha_s \times \beta_{sd} \times \gamma_{sd})\left(n_s^B \times d_{sd}^B \times f_{sd}^B\right)}{\sum_{s,d} n_s^B \times d_{sd}^B \times f_{sd}^B}\,. \tag{6.9}$$

In this equation, α, β, and γ reflect relative changes in fuel consumption components from the base fleet to the alternative fleet for different fuel consumption categories, and n, d, and f reflect relative differences of these components between different fuel consumption categories within the base fleet. This general equation is used as the basis to estimate the effects of various scenarios on overall fuel consumption of vehicle fleet.

While the parameters reflecting the relative number of vehicles (α and n) and mean distance travelled by vehicles (β and d) can be directly calculated from the vehicle registration and vehicle usage data, respectively, the parameters related to the relative mean fuel consumption rate of vehicles (γ and f) are more difficult to measure. A method that is based on the estimated fuel consumption models (explained in Chapter 4) is now introduced to estimate these parameters.

Consider different car types within any fuel consumption category where car types *i* includes all the car models with similar design features and hence the same range of fuel consumption rates. Equation 6.4 can be extended to the following:

$$f_{sd}^B = \frac{\bar{F}_{sd}^B}{\bar{F}_r^B} = \frac{\sum_i {p_s^B}_i \times {F_{sd}^B}_i}{\sum_j {p_r^B}_i \times {F_r^B}_i} = \frac{\sum_i {p_s^B}_i \times {f_{sd}^B}_i}{\sum_j {p_r^B}_i \times {f_r^B}_i} \tag{6.10}$$

where,

${p_s^B}_i$ is the proportion of cars of type *i* in the design segment *s* in the base fleet,

$F^B_{sd_i}$ is the fuel consumption rate of car type *i* in the fuel consumption category *sd* in the base fleet,

$p^B_{r\ i}$ is the proportion of cars of type *i* in the reference fuel consumption category in the base fleet,

$F^B_{r\ i}$ is the fuel consumption rate of car type *i* in the reference fuel consumption category in the base fleet,

$f^B_{sd_i}$ is the relative fuel consumption rate of car type *i* in the fuel consumption category *sd* in the base fleet when compared to the fuel consumption rate of a reference car type within the reference fuel consumption category, and

$f^B_{r\ i}$ is the relative fuel consumption rate of car type *i* in the reference fuel consumption category in the base fleet when compared to the fuel consumption rate of a reference car type within the reference fuel consumption category.

Similarly, Equation 6.8 can be extended to the following:

$$\gamma_{sd} = \frac{\bar{F}^A_{sd}}{\bar{F}^B_{sd}} = \frac{\sum_i p^A_{s\ i} \times F^A_{sd_i}}{\sum_i p^B_{s\ i} \times F^B_{sd_i}} = \frac{\sum_i p^A_{s\ i} \times f^A_{sd_i}}{\sum_i p^B_{s\ i} \times f^B_{sd_i}} \tag{6.11}$$

where $f^A_{sd_i}$ is the relative fuel consumption rate of car type *i* in the fuel consumption category *sd* in the alternative fleet when compared to the fuel consumption rate of a reference car type within the reference category in the base fleet.

To calculate the values of relative fuel consumption rates in Equations 6.10 and 6.11 ($f^B_{sd_i}$, $f^B_{r\ i}$, and $f^A_{sd_i}$), the estimated fuel consumption models explained in Chapter 4 are used. It was shown in Chapter 4 that for the fuel consumption category *sd*, fuel consumption rate is estimated based on the following general equation:

$$\bar{F}^B_{sd} = exp(a_0 + \sum aX) \tag{6.12}$$

where X represents a set of design variables (i.e. mass, engine size, year of manufacture, Euro emission standard). This equation was estimated separately for 8 defined fuel consumption categories using 2 estimated statistical models (one for each of urban and extra-urban driving cycles) and 4 interaction terms which were included in each statistical model (see Table 4.8 in Chapter 4).

The relative fuel consumption rate of car type i in category sd in the base fleet ($f^{B}_{sd_i}$) is calculated using the following equation:

$$f^{B}_{sd_i} = exp[\Delta a_0 + \sum(\Delta a)(\Delta X)] \tag{6.13}$$

where Δa represents the difference in the estimated coefficients of different design variables between car type i and the reference car type within the reference fuel consumption category (petrol cars with manual transmission), and ΔX represents the difference in the values of different design variables between car type i and a defined reference car type within the reference category. Similarly, the value of $f^{A}_{sd_i}$ in Equation 6.11 is calculated as

$$f^{A}_{sd_i} = exp[\Delta a_0 + \sum \Delta a(\Delta X)] \tag{6.14}$$

where ΔX represents the difference in the value of design variables within design segment s between the base and alternative scenario (e.g. change in vehicle mass).

Having calculated relative fuel consumption rates of different car types (f_{sd_i}) within different fuel consumption categories and for each of base and alternative fleets, the values of f^{B}_{sd} and γ_{sd} can be calculated using Equations 6.10 and 6.11 and be used in Equation 6.9 to estimate the relative change in the overall fuel consumption of the base fleet as a result of a change in vehicles design (e.g. vehicle mass distribution) within the fleet.

6.2.2. Safety of a vehicle fleet

According to the evidence from the literature as discussed in Chapter 2 and the findings in Chapter 5 on the relationship between vehicle mass and secondary safety, which are in agreement with Equation 5.1, the safety effect of a change in vehicles' mass in fleet mainly relates to the resulting changes in the overall injury outcome of two-vehicle crashes where risk of injury to the occupants of each vehicle depends on the relative mass of the involved vehicles (see Equation 5.1). It was also discussed in Chapter 2 that there is no strong evidence suggesting a direct effect of vehicle mass on the risk of crash involvement of the vehicles. Therefore, the effect of a change in the vehicle mass distribution within the fleet on the total number of driver casualties in two-car crashes is investigated under the assumption that the likelihood of vehicles being involved in crashes is not influenced by the changes in vehicles' mass. The estimated effect on the injury outcome of two-car crashes would largely represent the overall safety outcome of the change in the mass distribution of fleet (Buzeman et al., 2008).

The total number of driver casualties in two-car crashes (to a defined injury level) depends on the injury risk to the drivers who are involved in different types of crashes and the total number of these crashes according to the following equation:

$$I = \sum_k I_k = \sum_k N_k \left(P_{1_k} + P_{2_k}\right) \tag{6.15}$$

where,

I is the total number of driver casualties,

I_k is the total number of driver casualties in crash category *k*,

N_k is the total number of crashes in crash category *k*, and

P_{1_k} and P_{2_k} are, respectively, absolute driver injury risks in vehicles 1 and 2 (as defined in Chapter 5) in crash category *k*.

In Equation 6.15, crash categories are defined according to the driver, road, and crash characteristics that contribute to the injury risk.

As discussed earlier in this chapter, the objective of the analysis is to estimate the partial effect of a change in vehicle design, particularly mass distribution, within the fleet on total number of driver casualties. In order to estimate such an effect, all other factors that contribute to the number of driver injuries are kept constant. As mentioned earlier, an incremental approach is used that estimates the relative change in the total number of driver casualties with respect to the base case as a result of a change in the design of the vehicles. Since the risk of crash involvement of the vehicles is assumed to remain constant between the base and alternative case, the total number of crashes in each category (N_k) remains the same. However, the severity of crashes between the two cases could be different as a result of the changes in vehicles' mass. The ratio of total driver casualties in the alternative case (I^A) to that in the base case (I^B) can be written as the following:

$$\frac{I^A}{I^B} = \frac{\sum_k a_k \times I_k^B}{\sum_k I_k^B} \tag{6.16}$$

where a_k reflects the relative number of driver casualties between the alternative (A) and base (B) case in crash category k; it is calculated according to the following equation:

$$a_k = \frac{I_k^A}{I_k^B} = \frac{P_{1k}^{\ A} + P_{2k}^{\ A}}{P_{1k}^{\ B} + P_{2k}^{\ B}} \tag{6.17}$$

where, $P_{1k}^{\ B}$ and $P_{1k}^{\ A}$ are the absolute driver injury risks in vehicle 1 in two-car crash category k in the base and alternative case, respectively, and $P_{2k}^{\ B}$ and $P_{2k}^{\ A}$ are the absolute driver injury risks in vehicle 2 in two-car crash category k in the base and alternative case, respectively. The values of absolute risk in two-car crashes (P) for each vehicle and for each scenario are estimated based on the values of mass ratio and other driver and vehicle factors using the modelling results presented in Chapter 5 Section 5.3.3. It should be noted that in deriving Equation 6.17, the number of crashes in crash category k (N_k) is assumed to be the same between the base and the alternative case (where the only difference is fleet mass distribution).

Having estimated a for all the defined categories of two-car crashes, the change in the total number of driver casualties from the base case to the alternative case is estimated using Equation 6.16. In this equation, the total driver casualties in the base case in each crash category (I_k^B) is obtained from the base case crash data.

6.3. Base vehicle fleet data

The following introduces the base vehicle registration data, outlines distribution of vehicles in fleet by different design factors, and explains how this data was used to estimate the base mass distribution.

6.3.1. Vehicle registration data

As it was discussed in Chapter 3, a dataset of vehicle registration in Great Britain in the last quarter of 2007 was developed that included cross-sectional data on various design aspects of registered makes and models in fleet. According to the data, a total of 30,536,224 cars were registered in British fleet in the last quarter of 2007. This was chosen as the base fleet to investigate safety and environmental consequence of a number of hypothetical mass distributions.

Tables 6.1 to 6.4 reflect distribution of registered cars by various design factors. Distribution of cars by year of registration reflected in Table 6.1 suggests that in 2007, about 30% of registered cars were newer than 3 years old while about 52% of them were older than 5 years (when the registration date is used as a proxy for manufacture date). Table 6.2 shows the number of registered cars by engine size band, which is a contributing factor to the fuel consumption rate of the car. According to the data, the engine size band of 1750 cc to 2000 cc is the most popular engine size band when about 80% of registered cars had an engine size between 1000 cc and 2000 cc. Only about 15% of registered cars had an engine size of over 2000 cc which is normally associated with higher fuel consumption rates. Table 6.3 shows that the Hatchback body type is the most popular body type in Britain accounting for about 58% of registered cars followed by the Estate and Saloon body types. Multi-Purpose Vehicles (MPV) accounted for only about 6% of registered cars in 2007 British fleet. Table 6.4 shows proportion of registered cars by fuel type and transmission type. According to the results presented in Chapter 4, these are the two design variables that significantly influence fuel consumption rate of vehicles depending on their mass and engine size. The data shows that about 76% of cars in 2007 consumed petrol and manual transmission was more popular than automatic transmission when it accounted for about 76% of all registered cars. In particular, manual petrol cars were the most popular category of cars where they formed about 58% of all registered cars in 2007.

Table 6.1: Vehicle registration in the base fleet (2007) by year of registration

Year of registration	Number of registered cars	Percent of registered cars	Cumulative percent of registered cars
<1994	2,520,882	8.3	8.3
1994	844,231	2.8	11.1
1995	1,089,955	3.6	14.7
1996	1,371,805	4.5	19.2
1997	1,682,548	5.5	24.7
1998	1,907,589	6.2	30.9
1999	2,003,468	6.6	37.5
2000	2,131,714	7.0	44.5
2001	2,400,563	7.9	52.4
2002	2,537,508	8.3	60.7
2003	2,525,737	8.3	69.0
2004	2,499,335	8.2	77.2
2005	2,369,733	7.8	85.0
2006	2,267,575	7.4	92.4
2007	2,383,581	7.8	100.0
Total	30,536,224	100.0	100.0

Table 6.2: Vehicle registration in the base fleet (2007) by engine size band

Engine size band (cc)	Number of registered cars	Percent of all registered cars
500 – 749	60,588	0.2
750 – 999	1,270,952	4.2
1000 – 1249	3,544,446	11.6
1250 – 1499	5,805,042	19.0
1500 – 1749	5,376,847	17.6
1750 – 1999	9,948,478	32.6
2000 – 2249	895,058	2.9
2250 – 2499	1,458,523	4.8
2500 – 2749	406,054	1.3
2750 - 2999	869,169	2.8
3000 - 3999	604,579	2.0
4000 - 4999	199,099	0.7
5000 - 5999	52,041	0.2
6000 - 6999	22,306	0.1
Total	30,536,224	100.0

Table 6.3: Vehicle registration in the base fleet (2007) by body type

Body type	Number of registered cars	Percent of all registered cars
Cabriolet	930,269	3.0
Coupe	781,678	2.6
Estate	4,073,759	13.3
Hatchback	17,824,674	58.4
MPV	1,888,814	6.2
Saloon	3,774,767	12.4
Other	1,262,203	4.1
Total	30,536,224	100.0

Table 6.4: Proportion of vehicle registration in the base fleet (2007) by fuel type and transmission type

Fuel type	Transmission type			Total
	Manual	Automatic	Other	
Petrol	58.3%	10.3%	7.3%	75.9%
Diesel	17.9%	2.5%	3.4%	23.8%
Other	0.1%	0.1%	0.1%	0.3%
Total	76.3%	12.9%	10.8%	100.0%

6.3.2. Base mass distribution

Although the vehicle registration data included information on many design features of vehicles, it did not include data on vehicle mass and size. Section 3.3.2.3 in Chapter 3 explained how mass and dimension data were assigned to different registered makes and models in the vehicle registration data to make a sample dataset of registered makes and models that included about 73% of all registered makes and models. This sample was then used as the basis to estimate the base vehicle mass distribution.

To examine how well the developed sample dataset represented the full vehicle registration data, distribution of registered vehicles by various design features were compared between the sample and full data. The results reflected in Figures 6.1 to 6.4 show a close match between the two in terms of the proportion of registered cars by engine size, body type, fuel type, and transmission type. This suggests that the sample dataset of registered cars is reasonably representative of the full registration data; therefore it was used to estimate the mass distribution of cars in the base fleet.

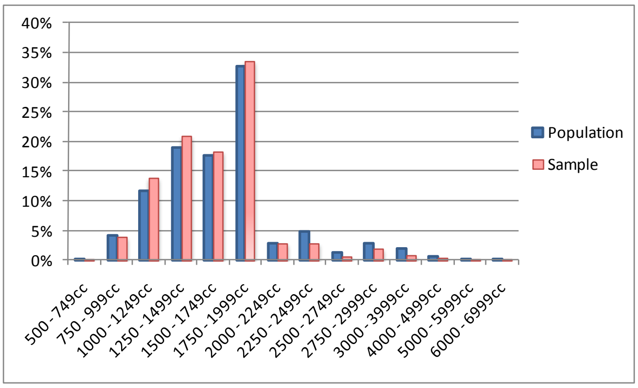


Figure 6.1: Distribution of vehicles by engine size band between the sample and full data

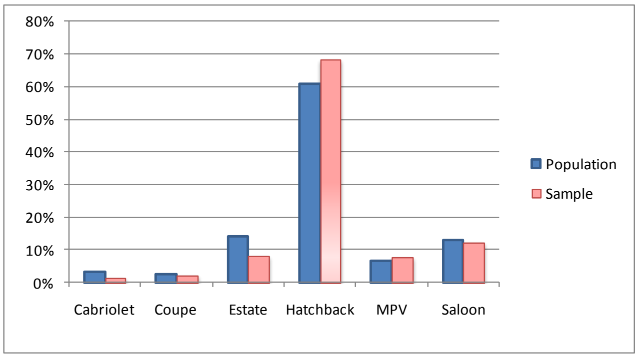


Figure 6.2: Distribution of vehicles by body type between the sample and full data

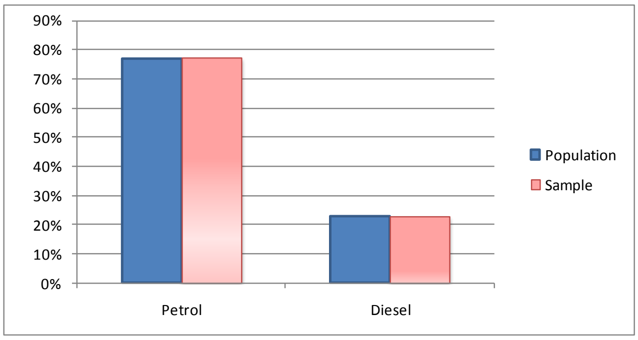


Figure 6.3: Distribution of vehicles by fuel type between the sample and full data

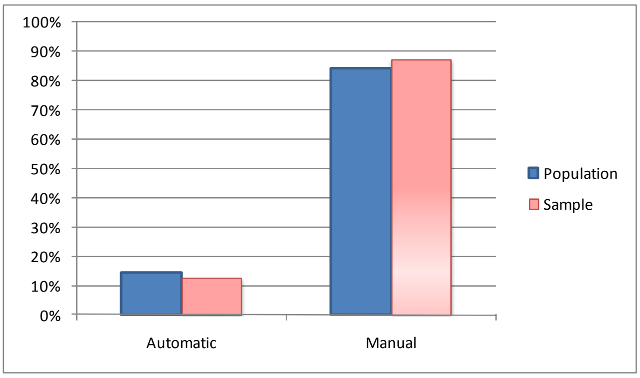


Figure 6.4: Distribution of vehicles by transmission type between the sample and full data

Table 6.5 and Figure 6.5 show the mass distribution of the base fleet when registered cars are grouped into mass categories with intervals of 100 kg. Examination of mass data revealed that choosing a mass range of 50 kg to represent mass distribution results in a jagged histogram with statistical fluctuations due to paucity of samples in each mass category while choosing a range greater than 100 kg (e.g. 200 kg) results in a relatively flat histogram imposing the risk of not reflecting the underlying distribution properly. Therefore, the choice of 100 kg intervals to represent mass distribution seemed to be an appropriate choice.

Table 6.5: Distribution of registered cars by mass in the sample registration data

Mass range	Frequency	Percent	Cumulative Percent
700-800	146,411	0.7	0.7
800-900	2,230,349	10.0	10.6
900-1000	2,892,825	12.9	23.6
1000-1100	3,216,791	14.4	38.0
1100-1200	3,641,655	16.3	54.3
1200-1300	3,249,731	14.5	68.8
1300-1400	3,225,685	14.4	83.3
1400-1500	1,591,594	7.1	90.4
1500-1600	1,061,488	4.8	95.1
1600-1700	451,724	2.0	97.2
1700-1800	305,555	1.4	98.5
1800-1900	132,107	0.6	99.1
1900-2000	60,286	0.3	99.4
2000-2100	21,785	0.1	99.5
2100-2200	36,965	0.2	99.7
2200-2300	27,023	0.1	99.8
2300-2400	4,154	0.0	99.8
2400-2500	21,754	0.1	99.9
2500-2600	9,490	0.0	99.9
2600-2700	11,616	0.1	100.0
2700-2800	1,831	0.0	100.0
Total	22,340,819	100.0	100.0

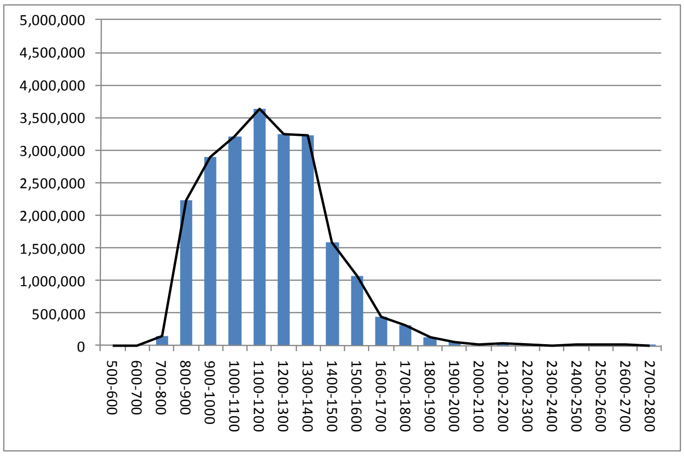


Figure 6.5: Vehicle mass distribution in the sample vehicle data

The category with the highest proportion of registered cars is the mass range of 1100 kg to 1200 kg (16.3% of registered cars). About 60% of cars have a mass ranging between 1000 kg and 1400 kg while about 90% of cars are lighter than 1500 kg. The average mass in the base fleet is about 1190 kg. This mass distribution, referred to as the base mass distribution, was used as the reference to examine the safety and environmental effects of a number of hypothetical mass distributions that were defined relative to the base mass distribution. These are explained in detail in the next section.

6.4. Scenario testing

The methodology explained in Section 6.2 was used to estimate partial effects of some alternative mass distributions defined relative to the base distribution (explained in Section 6.3.2). The relative changes in overall fuel consumption and total number of driver casualties from the base fleet with the base mass distribution were estimated. According to the methodology explained in Section 6.2, the base data required for the analysis should include cross-sectional vehicle registration data (which was explained in the previous section), vehicle distance travelled data, and two-car crash data. Ideally, all the base data should belong to the same time period. However, due to lack of such a match in the available data, there is a difference of one year between some parts of the base data. While the vehicle registration data belongs to the last quarter of 2007, the distance travelled data and two-car crash data belong to 2006. The following defines the alternative mass distribution scenarios and explains how the base data was used to estimate the partial effects of these scenarios.

6.4.1. Mass distribution scenarios

Three hypothetical alternative fleet downsizing scenarios are defined according to their mass distribution. They are characterized relative to the base mass distribution shown in Figure 6.5. In practice, reductions to vehicles' mass in fleet can be made in different ways; either by different cars being replaced by smaller and lighter cars in the fleet or by changes made in car design towards using lighter materials while vehicle size is maintained.

As was discussed in Chapter 4, the estimated fuel consumption models can only be used to estimate the within make and model effects of mass. This limits the range of mass distribution scenarios whose effects on fuel consumption of vehicle fleet can be quantified in this study. Only the fuel consumption effects of those scenarios can be quantified where the change in mass distribution is the result of a change in the distribution of model variants within makes and models while the distribution of makes and models within the fleet is kept constant.

For each scenario, two cases are examined. In the first case, vehicle mass is changed but vehicle size is maintained (the relationship between vehicle size and mass in the fleet is changed) while in the second case, vehicle size is also changed accordingly with vehicle mass (the relationship between vehicle size and mass in the fleet is maintained). If μ^B and σ^B denote respectively the mean and standard deviation of the base mass distribution, then the following hypothetical scenarios are defined.

1. Uniform fleet downsizing (S1)

This is a scenario that is generally in favour of fleet fuel consumption and emission reduction policies; however, its influence on overall safety has been subject to conflicting and inconsistent arguments as discussed in Chapter 2, Section 2.4. In this study, this scenario is defined according to an alternative mass distribution characterised by parameter ω when it is compared to the base mass distribution according to the following rule:

$$\forall m: \quad m^A = \omega m^B \therefore \ (\mu^A = \omega \mu^B \ , \quad \sigma^A = \omega \sigma^B) \tag{6.18}$$

where m^B and m^A are individual vehicles' mass in the base and alternative fleet, respectively, and ω is a parameter ranging between 0 and 1 that reflects the proportional reduction in average mass in fleet. Figure 6.6 shows the resulting mass distribution in this scenario for two example values of ω (0.8 and 0.9) when it is compared to the base mass distribution.

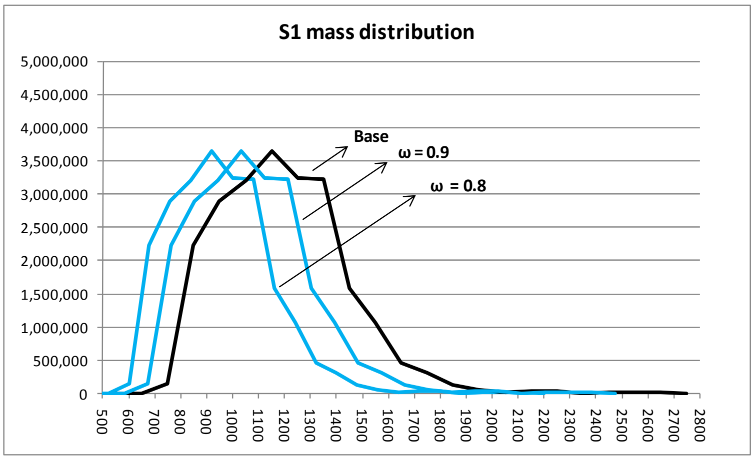


Figure 6.6: Fleet mass distribution in S1 scenario (ω=0.8 and ω=0.9)

As mentioned earlier, two cases are examined separately for scenario S1 according to the relationship between vehicle mass and size in the fleet:

$S1_a$: Vehicle size is maintained.

$S1_b$: Vehicle size is changed.

In the $S1_a$ scenario, it is assumed that the relationship between vehicle mass and size is different from that in the base fleet and, as a result, cars in the base fleet are replaced by lighter cars of the same size. On the other hand, $S1_b$ scenario assumes that the relationship between vehicle mass and size is the same as that in the base fleet and, as a result, cars in the base fleet are replaced by lighter cars which are also smaller in size. It should be noted that $S1_a$ and $S1_b$ scenarios are only expected to have different effects on driver casualties as it was found in Chapter 5 that vehicle size has a significant effect on injury risk; whilst, vehicle size was not found to have a significant effect on vehicle fuel consumption in Chapter 4.

2. Symmetric reduction in fleet diversity (S2)

Reduction in fleet diversity is generally regarded as a policy in favour of fleet safety; however, there are inconsistencies in the methodologies used to quantify its effects as discussed in Chapter 2. Besides, its detailed effect on overall fuel consumption has not been investigated. For the symmetric diversity reduction scenario, the following mass distribution is defined which is characterised by parameter θ relative to the base mass distribution:

$$\forall m: \quad m^A = \mu^B + \theta(m^B - \mu^B) \therefore (\mu^A = \mu^B \ , \quad \sigma^A = \theta\sigma^B) \tag{6.19}$$

where θ is a parameter ranging between 0 and 1 that reflects the proportional reduction in variance of mass in fleet. The resulting mass distributions for two example values of 0.6 and 0.8 for parameter θ are compared with the base mass distribution in Figure 6.7. Similar to the previous scenario, the following two cases are examined separately for this scenario:

$S2_a$: Vehicle size is maintained.

$S2_b$: Vehicle size is changed.

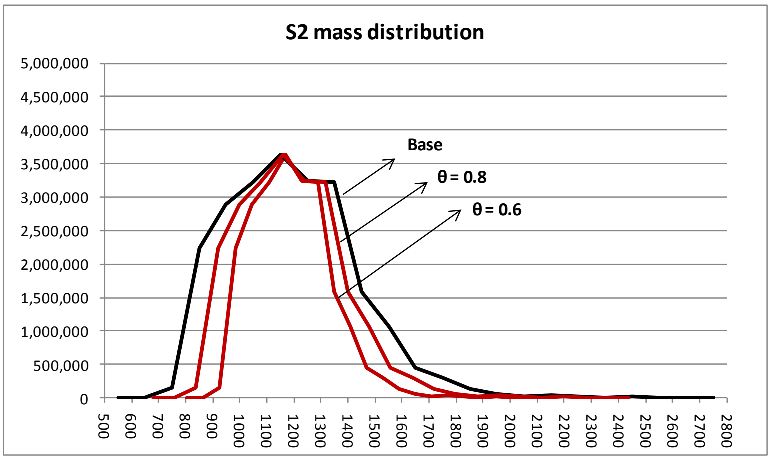


Figure 6.7: Fleet mass distribution in S2 scenario (θ=0.8 and θ =0.9)

3. Asymmetric reduction in fleet diversity (S3)

This scenario is similar to scenario S2 except that the reduction in fleet diversity is not uniform. Based on whether individual vehicles' mass are greater or less than average mass in fleet, their mass is reduced according to the following rules:

$$\forall m: \quad \begin{array}{rl} m^A = m^B & if\ m^B \leq \mu^B \\ m^A = \mu^B + \theta(m^B - \mu^B) & if\ m^B > \mu^B \end{array} \tag{6.20}$$

where θ is a parameter ranging between 0 and 1. The S3 mass distributions for two example values of 0.6 and 0.8 for parameter θ are compared with the base mass distribution in Figure 6.8. Similarly, the following two cases are examined separately:

$S3_a$: Vehicle size is maintained.

$S3_b$: Vehicle size is changed.

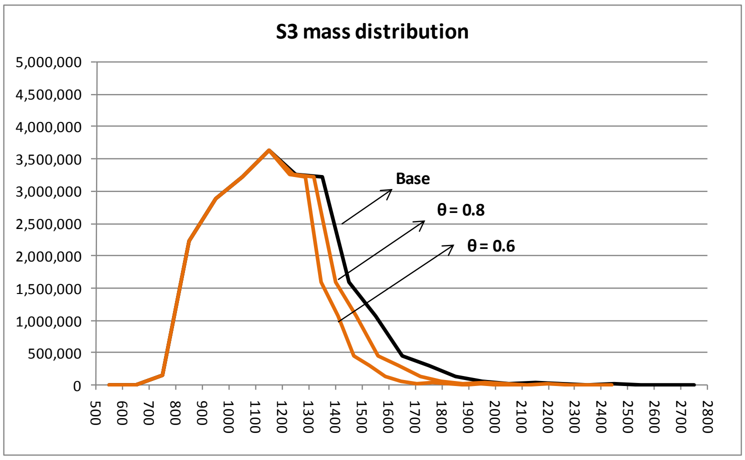


Figure 6.8: Fleet mass distribution in S3 scenario (θ =0.8 and θ =0.9)

6.4.2. Likely effects of defined scenarios

For each scenario, four hypothetical distributions were examined based on some example values for parameters ω and θ where their partial effects on fleet fuel consumption and total driver casualties are investigated. Table 6.6 shows the examined distributions together with mean and standard deviation of mass for each distribution. The example values of parameters ω and θ were chosen in a way to introduce a practical range of changes in mass for each scenario based on the observed relationship between vehicle mass and size in the base fleet. For example, as the base fleet data suggests, a uniform reduction of more than 20% in vehicles' mass while vehicle size is maintained would introduce an engineering challenge in vehicle design in many cases. While in all the alternative scenarios, average and standard deviation of mass is less than that in the base fleet, the lowest average and standard deviation belong to uniform downsizing (S1) and symmetric reduction in diversity (S2) scenarios, respectively. The following sections explain the estimated partial effects of these alternative mass distributions on overall fuel consumption of the vehicle fleet and the total number of driver casualties in two-car crashes.

Table 6.6: Characteristics of mass distribution scenarios

Mass Distribution	Parameters	Mean	Standard Deviation
Base	$\omega = \theta = 1$	1192	242
S1	$\omega = 0.80$	954	194
	$\omega = 0.85$	1014	206
	$\omega = 0.90$	1073	218
	$\omega = 0.95$	1133	230
S2	$\theta = 0.60$	1192	145
	$\theta = 0.70$	1192	170
	$\theta = 0.80$	1192	194
	$\theta = 0.90$	1192	218
S3	$\theta = 0.60$	1154	187
	$\theta = 0.70$	1164	200
	$\theta = 0.80$	1173	214
	$\theta = 0.90$	1183	228

6.4.2.1. Fleet fuel consumption

The methodology explained in Section 6.2.1 was used to estimate the partial effects of defined mass distribution scenarios, where the effect is the result of a change in mass distribution holding all other factors constant, on overall fuel consumption of the 2007 British passenger car fleet. It should be remembered that, as discussed in Section 6.4.1, the estimated changes in fuel consumption are based on the assumption that the change in fleet mass distribution is the result of a change in the distribution of model variants within makes and models while the distribution of makes and models within the fleet is constant (e.g. a change in vehicle design by manufacturers to reduce mass of their model variants). This is unlikely to be the case in reality and any change in fleet mass distribution is likely to be the result of a change in the distribution of makes and models as well. It was discussed in Chapter 4 that the estimated fixed effects of makes and models on fuel consumption are correlated with the effects of mass between makes and models (see Tables 4.9 and 4.10 and Figure 4.12). Therefore, any estimated saving in fuel consumption in this study as a consequence of a reduction in vehicles' mass within the fleet could be an underestimation of what is expected if the distribution of makes and models does not remain constant.

Based on the fuel consumption modelling results presented in Chapter 4, eight fuel consumption categories were defined as shown in Table 6.7. The first category (manual petrol cars driven under urban driving cycle) was chosen as the reference fuel consumption category.

Table 6.7: Defined fuel consumption categories

Design segment (s)	Category number	
	Driving cycle (d)	
	Urban	Extra-urban
Manual Petrol cars	1(r)	5
Automatic Petrol cars	2	6
Manual Diesel cars	3	7
Automatic Diesel cars	4	8

Equation 6.9 was used to estimate the relative change in overall fuel consumption of the vehicle fleet as a result of a change in the base mass distribution, holding all other factors constant. The following explains how different parameters in this equation were estimated.

- *Estimating n_s^B and d_{sd}^B (Equations 6.2 and 6.3):*

In order to estimate n_s^B and d_{sd}^B (defined in Equations 6.2 and 6.3 respectively), data on the total number of registered cars and the mean distance travelled of cars for the base year by fuel consumption category is required. The number of registered cars by fuel consumption category was obtained from the base vehicle fleet data explained in Section 6.3.1. To estimate the relative mean distance travelled by cars in different categories, the vehicle data from the Great Britain National Travel Survey (NTS), which is a continuous survey of households with field work being implemented every month of the year, was used. The NTS vehicle data, which includes about 9000 records per year, has information on annual distance travelled by cars as well as information on different vehicle design variables including fuel type; however, the data does not include transmission type of the vehicles. Due to the lack of annual distance travelled data by transmission type of the vehicles, it was assumed that for a given fuel type, mean distance travelled by manual cars is not different from that by automatic cars. The NTS data also provides information on the proportion of annual urban and extra-urban driving by fuel type. These estimates are available separately for petrol and diesel cars from Transport Statistics Great Britain (TSGB, 2007). These data were used to estimate the mean annual distance travelled by cars in 2006 (which was the latest available year of data at time of the study) by fuel consumption category.

The number of registered cars and estimates of annual mean distance travelled by fuel consumption category were used to estimate n_s^B and d_{sd}^B using Equations 6.2 and 6.3. The results are shown in Table 6.8. As the results show, manual petrol cars are the most popular category of cars in the fleet while the annual mean distance travelled by diesel car is about 1.6 times greater than that by petrol car.

Table 6.8: Base fleet vehicle registration and usage data by fuel consumption category

Design segment	Registration		Mean distance travelled			
			Driving cycle			
			Urban		Extra-urban	
	N_s^B	n_s^B	$\overline{D}_{sd}^B$ (mile)	d_{sd}^B	$\overline{D}_{sd}^B$ (mile)	d_{sd}^B
Manual Petrol	17,802,808	1.000	3119	1.000	4706	1.509
Automatic Petrol	3,153,493	0.177	3119	1.000	4706	1.509
Manual Diesel	5,467,815	0.307	4960	1.590	7483	2.399
Automatic Diesel	755,775	0.042	4960	1.590	7483	2.399

- *Estimating f_{sd}^B (Equation 6.4):*

The relative mean fuel consumption rate of cars in each fuel consumption category in the base fleet, f_{sd}^B, was calculated using Equation 6.10. In this equation, different car types (denoted by subscript i) within each fuel category are defined on the basis of having a similar design, and hence, a similar average fuel consumption rate. These car types were defined based on the design variables included in the estimated fuel consumption models in Chapter 4 (i.e. engine size, mass, year of manufacture, Euro emission standard). The number of defined car types within each design segment is shown in Table 6.9.

Table 6.9: The number of defined car types within each design segment

Design segment (s)	Total number of car types (i)
Manual Petrol	451
Automatic Petrol	380
Manual Diesel	410
Automatic Diesel	306

The estimated fuel consumption models (see Tables 4.9 and 4.10 in Chapter 4) together with Equations 6.10 and 6.13 were used to estimate the values of f_{sd}^B for different fuel consumption categories when, by definition, $f_r^B = 1$. These values, shown in Table 6.10, reflect the relative mean fuel consumption rate of different fuel consumption categories in the base fleet when the manual petrol category is taken as the reference category. As the results show, manual diesel cars driven under the urban driving cycle have the highest relative mean fuel consumption rate in the base fleet.

Table 6.10: Relative fuel consumption rates in the base fleet (f_{sd}^{B})

Design segment	Driving cycle	
	Urban	Extra-urban
Manual Petrol	1.000	0.617
Automatic Petrol	1.428	0.530
Manual Diesel	0.960	0.546
Automatic Diesel	1.186	0.760

- *Estimating α_s and β_{sd} (Equations 6.6 and 6.7):*

The total number of vehicles and mean distance travelled by vehicles in each fuel consumption category are kept constant between the base and alternative scenarios; therefore, the values of α_s and β_{sd} in Equation 6.9 are, by definition (Equations 6.6 and 6.7), equal to 1.0.

- *Estimating γ_{sd} (Equation 6.8):*

In order to estimate the ratio of annual mean fuel consumption rate of cars in the alternative fleet to that in the base fleet for each of the fuel consumption categories, γ_{sd}, a similar approach to that used in estimating f_{sd}^{B} was used. For each of the alternative fleets, cars registered in each design segment were grouped into types i according to their design. The values of γ_{sd} were then estimated separately for each mass distribution scenario using the estimated coefficients of the fuel consumption models explained in Chapter 4 together with Equations 6.11 and 6.14. The results are shown in Table 6.11.

Table 6.11: Ratio of mean fuel consumption rate of alternative to base by fuel consumption category (γ_{sd})

Fuel consumption category	Mass distribution S1				Mass distribution S2				Mass distribution S3			
	ω				θ				θ			
	0.80	0.85	0.90	0.95	0.60	0.70	0.80	0.90	0.60	0.70	0.80	0.90
Manual Petrol / urban cycle	0.980	0.985	0.990	0.995	1.002	1.001	1.001	1.000	0.998	0.998	0.999	0.999
Automatic Petrol / urban cycle	0.948	0.961	0.974	0.987	0.984	0.988	0.992	0.996	0.981	0.986	0.990	0.995
Manual Diesel / urban cycle	0.949	0.961	0.974	0.987	0.985	0.989	0.993	0.996	0.984	0.988	0.992	0.996
Automatic Diesel / urban cycle	0.907	0.929	0.952	0.976	0.939	0.954	0.969	0.984	0.939	0.954	0.969	0.984
Manual Petrol / extra-urban cycle	0.971	0.978	0.985	0.993	1.003	1.002	1.002	1.001	0.997	0.998	0.999	0.999
Automatic Petrol / extra-urban cycle	0.946	0.959	0.972	0.986	0.984	0.988	0.992	0.996	0.981	0.986	0.990	0.995
Manual Diesel / extra-urban cycle	0.936	0.951	0.967	0.983	0.982	0.986	0.991	0.995	0.980	0.985	0.990	0.995
Automatic Diesel / extra-urban cycle	0.897	0.921	0.947	0.973	0.933	0.949	0.966	0.983	0.980	0.985	0.990	0.995

It should be noted that since the fuel consumption data in Chapter 4 did not suggest any significant effect of vehicle size on fuel consumption, the two cases of maintaining and changing vehicle size within each mass distribution scenario (see Section 6.4.1) result in identical effects on fleet fuel consumption. The results show that in the uniform downsizing scenario (S1), the mean fuel consumption rate in all the fuel categories decreases (ratio of less than 1) as a result of a reduction in mass of all the vehicles in fleet. This is also the case in S3 scenario (asymmetric reduction in diversity). Similarly, this is because in this scenario, mass either remains constant or decreases depending on its value relative to the average mass in fleet (see Equation 6.23). On the other hand, the results show an increase (ratio of greater than 1) in mean fuel consumption rate from the base to alternative for some fuel categories in S2 scenario (symmetric reduction in diversity). As the Equation 6.22 and Figure 6.7 show, in this scenario mass of some of the lighter cars in fleet is increased. A considerable proportion of cars in the first design segment (manual petrol cars) are relatively lighter and smaller compared to those in the other design segments. This explains the estimated slight increase in the mean fuel consumption rate of cars in this design segment.

Overall effects (Equation 6.9):

Table 6.12 shows the partial effects of the defined mass distribution scenarios on the overall fuel consumption of base fleet as relative changes and percent changes estimated using Equation 6.9. The greatest savings in fuel are, as expected, related to the uniform fleet downsizing scenario (S1) where the overall mass reduction is the highest (up to 4% reduction for a 20% uniform reduction in mass). Care should be taken in interpreting these results. These are not the expected reductions in the overall fuel consumption of a future fleet that has a mass distribution as that of the defined mass distributions. This is because other contributing factors including vehicle ownership and usage pattern are also likely to change over the transition period. These estimates are in fact the expected reductions in fuel consumption if the mass distribution of the base fleet would be replaced by the defined alternative mass distributions when all other factors remained constant (all the cars in the base fleet were replaced by cars with a different mass). For example, if all the cars were about 90% of the mass of the base fleet; the annual fuel consumption of cars would be expected to be about 2% lower than that in the base fleet.

Table 6.12: Estimated effects of mass distribution scenarios on overall fuel consumption of base fleet

Mass distribution	Parameters	Y^A/Y^B	Percent change
S1	ω = 0.80	0.960	-4.00
	ω = 0.85	0.970	-3.01
	ω = 0.90	0.980	-2.02
	ω = 0.95	0.990	-1.02
S2	θ = 0.60	0.992	-0.78
	θ = 0.70	0.994	-0.59
	θ = 0.80	0.996	-0.40
	θ = 0.90	0.998	-0.20
S3	θ = 0.60	0.990	-1.02
	θ = 0.70	0.992	-0.77
	θ = 0.80	0.995	-0.51
	θ = 0.90	0.997	-0.26

As was discussed earlier, these estimated reductions in fleet fuel consumption are based on the assumption that the distribution of makes and models within the fleet is not changed. Therefore they only reflect the effects of changes in mass distribution within makes and models. For example, the average reduction in fuel consumption as a result of a change in vehicle design by manufacturers to reduce mass of their model variants (e.g. use of lighter materials in design). In reality, a fleet downsizing scenario is likely to be accompanied by changes in the distribution of makes and models within the fleet as well (e.g. an increase in the proportion of makes and models that are typically lighter and smaller). Therefore, the estimated changes in fuel consumption shown in Table 6.12 are likely to underestimate the savings that would be achieved in reality for the defined scenarios.

6.4.2.2. Driver casualties in two-car crashes

The methodology explained in Section 6.2.2 was used to estimate the partial effects of defined mass distribution scenarios on the total number of driver casualties in two-car crashes in Great Britain in 2006. Equation 6.16 was used to estimate the ratio of total number of driver casualties in the alternative case to that in the base case (I^A/I^B) where the only difference between the base and alternative case is the mass distribution of the vehicle fleet. It is assumed that risk of crash involvement, and hence the total number of crashes, is the same for the base and alternative case. Crash categories in Equation 6.16 (denoted by subscript k) were defined according to the main factors influencing risk of injury to the drivers in crashes as detailed in Chapter 5. These include speed limit of the road, direction of impact, drivers' age, and drivers' gender. Consistent with the injury risk models presented in Chapter 5, the level of driver injury considered in the analysis is KSI (Killed or Seriously Injured). Table 6.13 reflects the distribution of KSI drivers by speed limit and type of impact in Great Britain in 2006. As the table shows, the greatest number of driver casualties belong to frontal crashes and amongst these types of crashes, those on roads with a speed limit of 60 mile/hr are greatest. It should be remembered that the number and distribution of crash involvements by various factors are the same for the base and alternative case; however, the injury outcome of these crashes are changed as a result of a change in the mass of the colliding vehicles.

Table 6.13: Distribution of KSI drivers by speed limit and type of impact

Type of impact	Speed limit (mile/hr)				Total
	20 or 30	40 or 50	60	70	
Front to Front	484	233	934	29	1680
Front to Back	189	76	93	94	452
Front to Side	671	215	591	140	1617
Total	1344	524	1618	263	3749

According to Equation 6.16, in order to estimate the ratio of overall driver casualties in the alternative case to that in the base case (I^A/I^B), an estimate of this ratio is required separately for each crash category (denoted by a_k in Equation 6.16). a_k was estimated for each crash category using Equation 6.17. In using this equation, the relative changes in the absolute driver injury risks in vehicles 1 and 2 from the base case to the alternative case are required for all the crash categories. These were estimated using the modelling results reflected in Chapter 5, Sections 5.3.3.1, 5.3.3.2, and 5.3.3.3.

Given the fact that the full 2006 two-car crash data did not include mass and size of the vehicles, a sample of these crashes for which mass and size data was available (see Section 3.3.2.2 in Chapter 3) was used to estimate a_k in Equation 6.17. This sample was shown to be a good representative of the full data when crashes between 2000-2006 were compared between sample and full data by different crash, road, and driver types (see Figure 3.9 in Chapter 3). To confirm that this is also true when KSI crashes in 2006 are considered only, the proportions of driver KSI were compared by type of impact and speed limit of the road between sample and full data in 2006. The results, which generally show a good agreement between the two datasets, are shown in Figures 6.9 and 6.10.

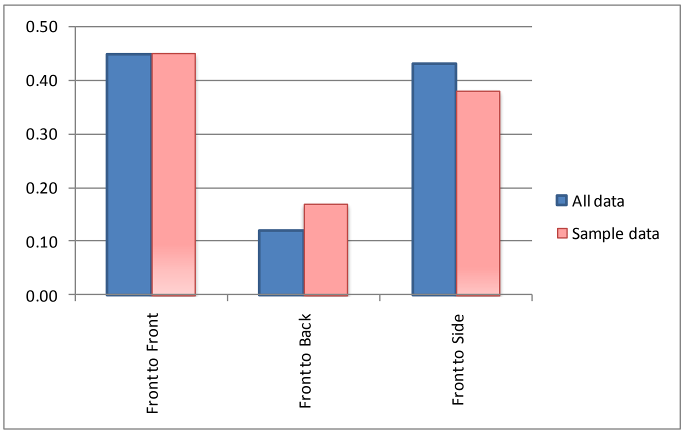


Figure 6.9: Distribution of driver KSI by type of impact in the sample and full data

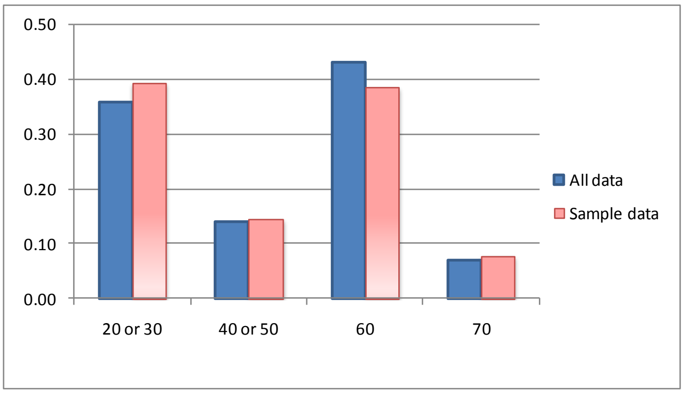


Figure 6.10: Distribution of driver KSI by speed limit in the sample and full data

The changes in vehicles' mass and size according to the defined scenarios in Section 6.4 were applied to the vehicles in different categories of two-car crashes in the 2006 sample dataset. In changing vehicle size, which was as defined in Chapter 5 (Length × Width), an average increase in size as a result of the change in mass was applied based on the estimated relationship between mass and size shown in Figure 5.6 in Chapter 5.

Having estimated a_k for each crash category and separately for each alternative scenario, the overall expected change in the number of driver casualties in two-car crashes from the base fleet (I^A/I^B) was estimated using Equation 6.16. The results in terms of ratios, percent changes, and net values are shown in Table 6.14. These reflect the number of driver casualties in two-car crashes that would have been expected if the cars involved in these crashes had had a different mass as shown by the alternative mass distributions.

Table 6.14: Estimated effects of mass distribution scenarios on the total number of driver casualties in two-car crashes in the base year (2006)

Mass distribution	Parameters	Size maintained?	I^A/I^B	Percent change	Net change in KSI
S1	ω = 0.80	Yes (S1$_a$)	1.000	0.00	0
		No (S1$_b$)	1.094	9.38	352
	ω = 0.85	Yes (S1$_a$)	1.000	0.00	0
		No (S1$_b$)	1.068	6.84	256
	ω = 0.90	Yes (S1$_a$)	1.000	0.00	0
		No (S1$_b$)	1.044	4.37	164
	ω = 0.95	Yes (S1$_a$)	1.000	0.00	0
		No (S1$_b$)	1.020	1.97	74
S2	θ = 0.60	Yes (S2$_a$)	0.989	-1.12	-42
		No (S2$_b$)	0.983	-1.71	-64
	θ = 0.70	Yes (S2$_a$)	0.992	-0.81	-30
		No (S2$_b$)	0.986	-1.43	-54
	θ = 0.80	Yes (S2$_a$)	0.994	-0.59	-22
		No (S2$_b$)	0.989	-1.11	-42
	θ = 0.90	Yes (S2$_a$)	0.997	-0.30	-11
		No (S2$_b$)	0.992	-0.75	-28
S3	θ = 0.60	Yes (S3$_a$)	0.995	-0.48	-18
		No (S3$_b$)	1.005	0.52	19
	θ = 0.70	Yes (S3$_a$)	0.996	-0.36	-13
		No (S3$_b$)	1.003	0.28	11
	θ = 0.80	Yes (S3$_a$)	0.998	-0.24	-9
		No (S3$_b$)	1.001	0.06	2
	θ = 0.90	Yes (S3$_a$)	0.999	-0.12	-4
		No (S3$_b$)	0.998	-0.15	-6

For the first scenario (uniform fleet downsizing), the results suggest that reducing mass of all vehicles in fleet proportionally, keeping their size constant, does not lead to an increase in the total number of driver casualties. This is consistent with the fundamental relationship between velocity change and relative mass of the vehicles in two-car crashes (Equation 5.1 of Chapter 5). When the mass of all the vehicles is reduced proportionally, the first term in this equation (mass proportion) remains constant resulting in no change in the velocity change of vehicles in a collision. However, in the case where the size of the vehicles is reduced alongside mass, this could lead to an increase in the total number of driver casualties as a result of an increase in the risk of injury in crashes because of the reduction in vehicle size.

The results show that a reduction in the diversity of the fleet where vehicle mass is maintained (S2$_a$ and S3$_a$ scenarios), which results in a decrease in the variance of mass in the fleet, is a desirable policy leading to a reduction in the total number of driver casualties in two-car crashes. Reduction in diversity generally tends to decrease the mass ratio of vehicles involved in two-car crashes; therefore, the total number of driver casualties also tends to reduce.

When vehicle size is changed alongside vehicle mass, the estimated effects of the uniform downsizing scenario $S1_b$ shows an increase in the total number of driver casualties. This effect is explained by the result found in Chapter 5 on the protective effects of vehicle size. The estimated effects of changing vehicle size in the $S2_b$ and $S3_b$ scenarios (reduction in fleet diversity) are different. The results for the $S2_b$ scenario show that changing vehicle size according to the change in vehicle mass results in even less number of driver casualties. This suggests that the benefit gained from increasing the size of the smaller and lighter cars in the 2006 vehicle fleet outweighs the disbenefit gained from decreasing the size of the larger and heavier cars. On the other hand, in the asymmetric reduction in diversity ($S3_b$) where, in contrast with $S2_b$ scenario, mass of lighter and smaller cars is kept constant, decreasing size of heavier cars alongside their mass results in an increase in the number of driver casualties (for $\theta \leq 0.8$). This suggests that the disbenefit gained by decreasing the size of the heavier cars in fleet outweighs the benefit gained by decreasing the variance of mass within the fleet in this scenario.

6.5. Conclusions and discussion

Table 6.15 summarises the findings on the partial effects of different mass distribution scenarios on overall fuel consumption and total number of driver casualties. These are the estimated effects as a result of a change in fleet mass distribution, holding all other factors constant.

As it was expected, the most favourable scenario regarding the fleet fuel economy is shown to be the uniform downsizing scenario (S1). Depending on the scale of reduction in mass (represented by parameter ω), considerable reduction in the overall fuel consumption and hence, carbon emissions can be gained. This will be accompanied by no increase in the number of casualties as a result of the change in vehicles' mass if the size of vehicles is maintained (i.e. a change in vehicle design towards using lighter materials). The results on the safety effects of the uniform mass reduction are in contrast to those by Buzeman et al. (1998) who found an increase in the total number of fatalities as well as to those by Broughton (1999) who found a decrease in the total number of injuries and fatalities. As discussed in Chapter 2, besides the shortcomings in the methodologies used, the role of vehicle size is not properly addressed in these studies.

Table 6.15: Estimated effects of mass distribution scenarios on the overall fuel consumption and total number of driver casualties

Mass distribution	Parameters	Size maintained?	Fleet fuel consumption[1]		Driver casualties	
			Y^A/Y^B	Percent change	I^A/I^B	Percent change
S1	ω = 0.80	Yes ($S1_a$)	0.960	-4.00	1.000	0.0
		No ($S1_b$)	0.970	-3.01	1.046	4.6
	ω = 0.85	Yes ($S1_a$)	0.980	-2.02	1.000	0.0
		No ($S1_b$)	0.990	-1.02	1.034	3.4
	ω = 0.90	Yes ($S1_a$)	0.992	-0.78	1.000	0.0
		No ($S1_b$)	0.994	-0.59	1.022	2.2
	ω = 0.95	Yes ($S1_a$)	0.996	-0.40	1.000	0.0
		No ($S1_b$)	0.998	-0.20	1.011	1.1
S2	θ = 0.60	Yes ($S2_a$)	0.990	-1.02	0.948	-5.2
		No ($S2_b$)	0.992	-0.77	0.936	-6.4
	θ = 0.70	Yes ($S2_a$)	0.995	-0.51	0.961	-3.9
		No ($S2_b$)	0.997	-0.26	0.951	-4.9
	θ = 0.80	Yes ($S2_a$)	0.960	-4.00	0.973	-2.7
		No ($S2_b$)	0.970	-3.01	0.967	-3.3
	θ = 0.90	Yes ($S2_a$)	0.980	-2.02	0.987	-1.3
		No ($S2_b$)	0.990	-1.02	0.982	-1.8
S3	θ = 0.60	Yes ($S3_a$)	0.992	-0.78	0.977	-2.3
		No ($S3_b$)	0.994	-0.59	0.978	-2.2
	θ = 0.70	Yes ($S3_a$)	0.996	-0.40	0.983	-1.7
		No ($S3_b$)	0.998	-0.20	0.983	-1.7
	θ = 0.80	Yes ($S3_a$)	0.990	-1.02	0.989	-1.1
		No ($S3_b$)	0.992	-0.77	0.989	-1.1
	θ = 0.90	Yes ($S3_a$)	0.995	-0.51	0.994	-0.6
		No ($S3_b$)	0.997	-0.26	0.994	-0.6

[1] The estimated changes in fleet fuel consumption are based on the assumption that the distribution of makes and models within the fleet is not changed.

On the other hand, reduction in the fleet diversity (S2 and S3) was generally shown to be the most desirable scenario in terms of safety leading to a reduction in the total number of driver casualties (except scenario $S2_b$ where in an asymmetric reduction in fleet diversity, the size of the heavier cars in fleet is decreased alongside their mass). A small decrease in the overall fuel consumption is also achievable in these scenarios depending on the characteristics of the vehicle fleet. A similar safety effect of a reduction in variance of mass in fleet had been suggested by some other studies (e.g. Buzeman et al., 1998; Ross and Wenzel, 2001); however, the magnitude of the estimated effects are different due the different methodologies used and the different vehicle fleets examined.

As was discussed in Chapter 4, the estimated fuel consumption models can only be used to estimate the effects of mass within make and model. As a result of this limitation, the estimated changes in fuel consumption in this study are based on the assumption that the change in fleet mass distribution is the result of a change in the distribution of model variants within makes and models while the distribution of makes and models within the fleet is constant (e.g. only a change in vehicle design by manufacturers to reduce mass of their model variants). It was also discussed that in reality, any change in the mass distribution of vehicle fleet to improve fleet fuel economy is probably achieved by a change in the distribution of makes and models, as well as a change in the mass distribution of model variants within the makes and models; such a change could result in greater reductions in overall fuel consumption than those estimated here.

The estimated outcome of the introduced scenarios in this study was a partial effect where other contributing factors are assumed to remain constant. This is unlikely to be the case in reality. Vehicle ownership and vehicle usage pattern, which change alongside the mass distribution of fleet in the course of time, are important factors that could also influence the fleet fuel economy and safety outcomes. According to Equation 6.1, an increase in each of the total number of vehicles and average distance travelled by vehicles in different fuel consumption categories increases the overall fuel consumption of fleet and vice versa. A change in the distribution of vehicles by different fuel consumption categories also has an effect on the fleet fuel consumption as shown by Equation 6.9. On the other hand, an increase in the overall distance travelled by cars means an increase in the exposure to the risk of vehicle crash involvement, hence results in an increase in the total number of crashes (N_k in Equation 6.15). Further research is required to investigate the effects of vehicle ownership and vehicle usage patterns on fuel consumption and safety in fleet; which must be taken into consideration when policies aiming to increase fleet fuel economy or improve safety through changes in fleet composition are to be formulated.

It was discussed earlier in this chapter that a change in mass distribution within the fleet, holding all other factors constant, is likely to have a significant effect on the injury outcome of two-car crashes only. Mass of a vehicle in a single-vehicle crash with a given speed of impact does not influence the velocity change experienced by the occupants in the crash (Buzeman et al., 1998; Van Auken and Zellner, 2005). However, some evidence from the literature (see Chapter 2, Section 2.3.2) and findings on the protective effect of vehicle size in Chapter 5 suggest that in a single-vehicle crash, an increase in vehicle size increases the safety performance of the vehicle. Therefore, the overall safety outcome of a change in the size of vehicles in fleet might be different from those estimated based on the injury outcome of two-car crashes (e.g. a higher safety benefit is expected through the effects on the injury outcome of single-vehicle crashes of an increase in the size of the vehicles in the fleet). Further analysis is required to investigate the effects of changes in vehicles' size on the injury outcome of single-vehicle crashes.

The results generally showed that an informed change in the mass distribution of vehicles within the fleet not only imposes no trade-off between the fuel economy and safety goals, but also could lead to a desirable outcome in both aspects. However, the effects of some other factors which contribute to both aspects, and tend to change over time when mass distribution changes, should be carefully considered.

CHAPTER 7. CONCLUSIONS AND DISCUSSION

The first section of this chapter (7.1) summarises the main findings and contributions of the study. The policy implications are discussed in the second section (7.2). The third and fourth sections (7.3 and 7.4) outline the study limitations and recommend some future research, respectively.

7.1. Research findings and contribution

A thorough review of the key studies relevant to the issue of trade-off between fuel consumption and secondary safety performance in vehicle design imposed by vehicle mass, which was discussed in Chapter 2, revealed the following main shortcomings in the existing literature:

- The partial effects of vehicle mass, where the effects are isolated from those of engine size and other factors, on fuel consumption rate in different driving cycles are not clear
- There are major methodological issues associated with the most well-known estimated relationship between relative injury risk and mass ratio in two-car crashes as introduced by Evans and Frick (1993)
- The effect of a change in the relative mass of the vehicles in a two-car crash on the absolute injury risk to the drivers of each vehicle, all other factors being constant, is not clear
- The isolated effects of vehicle mass and size on secondary safety performance of vehicles are not fully understood due to the high correlation between the two
- The issue of lack of information on non-injury crashes in the two-car crash data, which prevents a direct estimation of absolute injury risk from the data, has remained a challenge in estimating the effects of different factors on risk of injury in crashes

A number of studies (e.g. Buzeman et al., 1998; Broughton, 1999; Ross and Wenzel, 2001) have used the estimated relationships between mass and each of fuel consumption and secondary safety performance to address the issue of interaction between fuel economy and safety outcomes in fleet and have resulted in different, and sometimes, conflicting conclusions. The knowledge gaps in the literature as mentioned above limit the creditability of these research findings.

Several other studies (e.g. Crandall and Graham, 1989; Kahane, 2003; Noland, 2004; Noland, 2005; Ahmad and Greene, 2005) that have addressed the issue of conflict between fuel economy and safety in the vehicle fleet caused by mass are empirical studies that have used aggregate time-series data. In such studies, the characteristics of the vehicle fleet and the time period to which the data belongs could influence the results. This partly contributes to the inconsistencies in the results of these studies. Besides, since the effect of vehicle mass is not controlled for, it is not clear to what extent the changes in fuel consumption are related to the changes in vehicle mass rather than other contributing factors that tend to change over time.

Five objectives were defined (Chapter 1) in order to address the issue of interaction between environmental and safety performance in vehicle design within the vehicle fleet. These objectives fully address the main knowledge gaps summarised earlier. The defined objectives were as follows:

1. To estimate the effects of vehicle design (particularly mass) on fuel consumption.
2. To estimate the protective and aggressive effects of vehicle mass in two-vehicle crashes separately.
3. To separate the effects of vehicle mass and size on secondary safety performance.
4. To examine whether there are specific design effects of different vehicle makes and models on their secondary safety performance beyond the effect of their mass.
5. To investigate partial safety and environmental consequences of differences in vehicle mass distribution in fleet.

Objective 1 was addressed in Chapter 4 where the effects of vehicle design features, including mass, on fuel consumption were estimated. The partial effects of mass and engine size on fuel consumption in both urban and extra-urban driving cycles were found to be significantly different for different combinations of fuel and transmission types. It was found that a 100 kg increase in mass of model variants within a make and model could increase fuel consumption by 0.9% to 3.1% depending on fuel type, transmission type, and the driving cycle. The greatest partial effect of mass was found for automatic diesel cars when a 100 kg increase in mass would increase typical urban and extra-urban fuel consumption by 2.9% and 3.1%, respectively.

Chapter 5 addressed objectives 2 to 4 through a detailed analysis of two-car crashes. Such an analysis was concluded necessary when the comparison of the relationship between vehicle mass and its relative driver injury risk in fleet in two different time periods showed a significant change (see Figure 5.3) confirming separate protective and aggressive effects of vehicle mass in two-car crashes. A novel methodology was introduced to estimate partial effects of mass on absolute driver injury risk in each of the vehicles in the crash, which had remained unclear in the vehicle safety literature. It was found that, for example, if two cars with a similar mass (1000 kg) crash into each other in a road where the speed limit is 60 mile/hr, the probability of each driver being killed or seriously injured is about 13.5%. However, if car 2 had a mass twice that of car 1 (2000 kg compared to 1000 kg), the probability that the driver of car 1 (lighter car) is killed or seriously injured would increase to about 19.4% while the probability that the driver of car 2 (heavier car) is killed or seriously injured would decrease to about 4.8%.

Another novel aspect of the analysis based on the introduced methodology was separating the effect of vehicle mass from that of vehicle size on absolute driver injury risks of the vehicles involved in a two-car crash, where vehicle size is represented by “vehicle length × vehicle width”. The results confirmed that there is a protective effect of vehicle size above and beyond that of vehicle mass for front to front and front to side crashes; the data did not show any effect of vehicle size in front to back crashes.

The introduced methodology was also used to investigate whether there are any specific effects of vehicle makes and models on driver injury probability in frontal two-car collisions over and above the effects of mass (as represented by mass ratio). The analysis results based on frontal collisions in which there is at least one driver injury (of any level) during 2000-2006 suggested that there is no statistically significant (at 5% level) effect of make and model over and above that of mass.

Chapter 6 addressed the last objective where the partial effects of different mass distribution scenarios on the overall fuel consumption and total number of driver casualties were estimated using an incremental approach that estimated the relative changes compared to a reference mass distribution. It was found that a 20% uniform reduction in mass of all model variants within makes and models in the 2007 British fleet, all other factors including the distribution of makes and models being constant, would reduce the overall fuel consumption by about 4%. This could be accompanied by no increase in the total number of casualties if the size of vehicles is maintained. On the other hand, it was estimated that a 40% reduction in the standard deviation of mass and size in fleet would result in about 6.5% reduction in the total number of killed or seriously injured drivers as well as about 1% reduction in the overall fuel consumption.

These results generally show that the relationship between fuel economy and safety performance in vehicle design within the fleet depends on the characteristics of the vehicle fleet, and in particular, mass distribution within the fleet. It was shown that an informed change in the mass distribution not only imposes no trade-off between the fuel economy and safety goals, but also could lead to a desirable outcome in both aspects, for example, through maintaining mass of the lighter cars within the fleet while decreasing mass of the heavier cars maintaining their size.

The following contributions of this research are specified:

- Isolated estimated effects of mass and engine size on fuel consumption rate for different combinations of fuel and transmission types and for each of urban and extra-urban driving conditions.
- Introduction of a novel methodology that provides a solution to the issue of lack of information on non-injury two-car crashes in national accident data, which has often led to focusing on relative measures of injury risk that are not independent of risk in the colliding cars[1].
- Introduction of independent incremental methods to investigate relative changes in the overall fuel consumption and driver injuries as a result of changes in the characteristics of vehicle fleet

7.2. Policy implications

A uniform fleet downsizing scenario was found to be the most favourable scenario regarding the fleet fuel economy due to a reduction in vehicles' mass within the fleet;

---

[1] See Section 2.3.1.2 for details

however, in order to avoid any adverse safety impact as a result of the vehicle size reduction, the downsizing should only focus on vehicle mass maintaining vehicle size. On the other hand, a reduction in fleet diversity by decreasing variance of mass within the fleet was found to be a favourable scenario regarding safety. As it was shown, a decrease in the overall fuel consumption is also achievable in this scenario depending on the characteristics of the vehicle fleet. According to the findings, increasing the size of lighter cars within the fleet while maintaining their mass on one hand, and decreasing the mass of heavier cars within the fleet while maintaining their size on the other hand is the most desirable scenario in favour of both safety and environmental goals.

There are different ways in which the mass distribution of vehicle fleet could change. For example, a change could be achieved through a development in the design of the new cars by using various mass reduction technologies, which results in a reduction in mass while maintaining vehicle dimensions. A shift in the drivers’ choice towards using a different type of vehicle is another way of changing fleet composition and, in particular, mass distribution. The following discusses possible policy options to change vehicle mass distribution. It is important to note that what follows are only examples and recommendations as formulating new policies is beyond the scope of this study and requires further research, which will be discussed later in this chapter.

The relationship between vehicle size and mass reflected in Chapter 5, Figure 5.6 shows a considerable variation in mass for a given size. This suggests that there is the potential to decrease mass of many vehicles in fleet whilst maintaining their dimensions. As Ross and Wenzel (2001) discussed, there are a number of mass-reduction techniques (e.g. use of lightweight materials in design, use of lighter high-efficiency propulsion systems) which could be used by manufacturers to reduce the kerb mass of their new car models. Therefore, specific policies could be formulated to encourage design of lighter vehicles in the larger vehicle classes by manufacturers.

An effective policy to promote informed changes in new vehicle design could be through fuel consumption or CO_2 emission regulations that are a function of vehicle size or mass. For example, China has set fuel consumption limits for 16 different passenger car classes according to the vehicle curb mass with all the vehicles falling within a class being subject to the uniform fuel consumption limit of that class (Wang et al., 2010). Although this policy has resulted in a reduction in fuel consumption rate of Chinese cars since its implementation in 2004, the effect of such a policy on mass reduction within the fleet is unclear. This is due to the possibility that manufacturers prefer to increase the kerb mass of the cars falling in the upper ranges of a given mass class in order to move to a higher class, and hence, being obliged to meet a less stringent fuel consumption limit. An alternative system in which the fuel consumption limits are a function of kerb mass as a continuous variable rather than a categorical one seems to be more effective in promoting mass reduction in new vehicle design.

However, a fuel consumption or CO_2 emission regulation policy, which is a function of vehicle mass aiming at reducing mass of the new cars, would not ensure that reductions in mass are not accompanied by reductions in size. Alternatively, such regulations could be designed to be a function of vehicle size. Green (2009) discussed a new vehicle taxing policy which is a function of both fuel consumption rate and vehicle size. He argued that such a policy removes the incentive to buy a smaller car, which tends to be less safe for its occupants. A similar approach could be implemented to introduce fuel consumption or CO_2 emission regulations that are a function of vehicle size (whether as a continuous or categorical variable). This encourages manufacturers to increase the size of their new cars in order to fall in a higher class to meet less stringent limit, as well as to decrease their mass in order to decrease their fuel consumption and CO_2 emission within a given size class. It is important to note that Green (op.cit.) recommended vehicle footprint, defined as the product of track width and wheelbase, as a representative of vehicle size; however, the introduced measure of vehicle size in this study (the product of vehicle length and vehicle width), which was found to better represent the safety effects of vehicle size, is recommended to be used to represent the vehicle size in formulating these policies.

It is also recommended that the findings on the effects of different design features on fuel consumption be considered in setting any fuel consumption or CO_2 emission regulations for the new vehicles (e.g. separate limits for petrol and diesel cars as well as manual and automatic cars within a given mass class).

Other fleet downsizing policies could target consumer car purchase behaviour in different ways. Road taxing policies are a common way of influencing drivers' choice of cars, and hence, the composition of vehicle fleet. For example, the Vehicle Excise Duty (VED) in the UK, which is based on certain CO_2 emission bands for the vehicles registered after March 2001 (for cars registered before this date, VED is based on engine size), is designed to provide an economic incentive to drive cleaner cars. However, this does not seem to be a powerful tool to encourage buying lighter cars partly due to a high variation of vehicle mass within the defined CO_2 emission or engine size bands. Introduction of a mass-based road taxing system to the current CO_2-based system that adds extra taxes to the heavier cars for imposing a higher risk of injury to the other car users in the fleet could be more effective in altering consumers' choice towards driving lighter cars through economic perspectives as well as keeping the environmental benefits of a CO_2-based taxing system as the initial goal of the current system. To be more effective, this could be accompanied by increasing public awareness of the improvement in the secondary safety performance of lighter cars as a result of a reduction in the proportion of heavier cars in fleet. Other alternative policy options include specific heavier car taxing policies, new car pricing policies, and increase in the level of awareness.

Economic incentives are necessary but not sufficient to stimulate behavioural change (Lane and Potter, 2007). Therefore, a combination of policies targeting both new vehicle design through various mass-based regulations and consumer car purchase process through various economic incentives could be a more effective way in achieving an informed change in the mass distribution of vehicle fleet. As pointed out in Chapter 6, such a change takes several years to complete. An increased scrappage rate scheme could also be introduced to minimize the transition period.

It was discussed in Chapter 6 that vehicle usage pattern plays an important role in the overall fuel economy and safety outcomes of the vehicle fleet. Certain policies could also affect the average distance travelled by different types of cars in fleet. Research is required to understand the detailed relationships between vehicle usage pattern and each of safety and environmental goals in fleet. Besides, making policies to change the observed pattern in the usage of cars requires detailed understanding of the effects of contributing factors. While there appears to be a combination of effects, it is difficult to entirely separate the effects of driver type and vehicle type on vehicle usage. This is due to the uncertainty in the extent to which drivers who choose their car with specific design features do so because of their usage patterns or subsequently change their usage as a consequence of their choice of car. Research is required to understand these effects fully.

7.3. Limitations of the research

The estimated effects of mass on fuel consumption were based on official fuel consumption data measured under controlled conditions in the laboratory on a chassis dynamometer. Despite the several advantages of using this type of data to estimate the partial effects of mass as discussed in Chapter 4, Section 4.3, the fuel consumption measurements were limited to just two driving cycles representing typical urban and rural driving conditions. As discussed in Section 4.5, no significant change is expected in the partial effects of mass in practical urban and extra-urban driving conditions that are similar to the European cycles, from those estimated here; however, more research is needed to investigate changes in these mass effects in different driving cycles involving more accelerations/decelerations and factors such as wind, hills and corners. It is also important to note that the fuel consumption rate of vehicles are expected to be different in some real driving cycles as influenced by various driver, road, and environmental factors; therefore, the estimated fuel consumption models presented in Chapter 4 should not be used to predict fuel consumption rate of vehicles driven under practical conditions; however, they can be used to investigate the partial effects of vehicle design features on fuel consumption rate in typical urban and extra-urban driving conditions.

The dataset used to analyse the injury risk distribution in two-car crashes was based on a limited sample of full two-car crashes. It was discussed that a large number of records were eliminated from the final dataset due to lack of design information, in particular mass and dimensions, for the vehicles involved in the two-car crashes. Whilst it was shown that the final sample dataset reasonably represented the full data suggesting no sampling bias, a larger sample size would lead to narrower estimated confidence intervals for the estimates.

One of the disadvantages of STATS19 data is lack of information on the restraint use of the injured occupants. There was initially a variable defined for seat belt usage in STATS19; this was removed after 1993 due to the concerns about the quality of reporting (DfT, 2006b). Seat belt usage could substantially influence the injury outcome of a crash. The majority of the injured drivers during 2002-2006 in Great Britain are expected to have used seat belt; therefore, it is unlikely that the estimated effects are significantly affected as a result of lack of information on this variable.

The estimated partial effects of the defined mass distribution scenarios in Chapter 6 are based on a number of assumptions. As was discussed in Chapter 4, the estimated fuel consumption models in this study can only be used to estimate the effects of mass within makes and models. As a result of this limitation, the estimated changes in fuel consumption in this study are based on the assumption that the change in fleet mass distribution is the result of a change in the distribution of model variants within makes and models while the distribution of makes and models within the fleet is constant (e.g. only a change in vehicle design by manufacturers to reduce mass of their model variants).

In the lack of separate data on annual distance travelled by manual and automatic cars, it was assumed in Chapter 6 that for a given fuel type, the mean distance travelled by manual cars is similar to that by automatic cars. It was also assumed that the proportion of urban and rural driving is the same for petrol and diesel cars as well as for manual and automatic cars. The other limitation associated with the scenario testing process was regarding the base fleet data. Due to lack of a complete match between different available data used for the base year, there was a difference of one year between some parts of the data. While the vehicle registration data belonged to the last quarter of 2007, the distance travelled data and two-car crash data belonged to 2006. The estimated change in the number of driver casualties is based on a sample of two-car crashes in 2006 for which mass and dimension data was available, rather than the full data; however, the estimated proportional changes are applied to the total observed driver casualties to estimate the net differences.

Having acknowledged the main limitations of the study, it is important to note that they are unlikely to have had any considerable influence on the main findings of this research.

7.4. Recommendations for further work

The estimated relationships between vehicle mass and each of fuel consumption and secondary safety performance were used to investigate the expected outcomes if the vehicle fleet had a different mass distribution, all other factors being constant. Whilst this addressed the issue of interaction between safety and fuel economy objectives in vehicle design within the fleet imposed by mass, which was the aim of this research, it also provided the basis to formulate policies resulting in an informed change in the composition of vehicle fleet in favour of both fuel economy and safety goals. However, issues were raised which are required to be addressed through further research before any policy can come into effect.

As discussed in Chapter 6, vehicle ownership and vehicle usage pattern are of the most important aspects of vehicle fleet that change in time, both of which can substantially influence fuel economy and safety outcomes. An informed change in the overall distance travelled by different types of vehicles in fleet could potentially result in further safety and environmental benefits. More research is needed to investigate these. Besides, understanding the detailed relationships between driver type, vehicle type, vehicle ownership, and usage pattern are the key to successful design and implementation of the relevant policies. Therefore, further research is required to understand these relationships as well.

It was assumed in Chapter 6 that a change in vehicle mass alone is unlikely to influence the likelihood of crash involvement of the vehicle; however, it was discussed that a change in vehicle size alongside mass might influence this likelihood (Van Auken and Zellner, 2005). More research is required to investigate such effects in detail. It was also claimed based on the findings documented in the literature that the main effect of a change in mass distribution is on the injury outcome of two-car crashes. Given the possibility of a change in vehicle size alongside mass and the findings on the protective effects of vehicle size in Chapter 5, there is a need to investigate the effects of any change in the fleet composition on the outcome of other types of crashes, especially single-vehicle crashes where vehicle size could potentially have an effect.

Finally, any informed change in the composition of a vehicle fleet such as fleet downsizing or fleet diversity reduction takes several years to complete. During this transitional phase, the proportion of old cars declines as the proportion of new downsized cars increases with a pace which depends on the scrappage rate of the vehicles. The safety and environmental outcomes of fleet during this transitional phase depends on the pattern of changes in the mass and size of the vehicles. There are particular concerns on the safety effects of the transitional phase. This is due to the fact that, as pointed by Broughton (1999), the transition may be non-linear for two-car crashes and, as a result, some old heavier cars be involved in crashes with some new downsized cars which are lighter than those they have been replaced. Therefore, research is required to study the likely outcomes of the transitional phase and to find the best approach resulting in the most desirable outcomes.

REFERENCES

Ahmad, S., Greene, D.L., 2005. The effect of fuel economy on automobile safety: a reexamination. *Transportation Research Record,* 1941, 1-7.

Biggs, D.C., Akcelik, R., 1987. Estimating effect of vehicle characteristics on fuel consumption. *Journal of Transportation Engineering*, 113 (1), 101-106.

Blows, S, Ivers, R.Q., Woodward, M., Connor, J., Ameratunga, S., Norton, R., 2003. Vehicle year and the risk of car crash injury. *Injury Prevention*, 9 (4), 353-356.

Broughton, J., 1996a. The British index for comparing the accident records of car models. *Accident Analysis and Prevention,* 28 (1), 101-109.

Broughton, J., 1996b. Further aspects of the British index of secondary car safety. *Accident Analysis and Prevention,* 28 (6), 791-798.

Broughton, J., 1996c. The theoretical basis for comparing the accident record of car models. *Accident Analysis and Prevention,* 28 (1), 89-99.

Broughton, J., 1999. *The likely effects of downsizing on casualties in car accidents*. Transport Research Laboratory, Report No. 418, TRL Limited, Crowthorne, UK.

Broughton, J., 2003. The benefits of improved car secondary safety. *Accident Analysis and Prevention,* 35 (4), 527-535.

Broughton, J., 2007. *Casualty rates by type of car*. Transport Research Laboratory, Report No. PPR 203, TRL Limited, Crowthorne, UK.

Burgess, S.C., Choi, J., 2003. A parametric study of the energy demands of car transportation: a case study of two competing commuter routes in the UK. *Transportation Research Part D,* 8 (1), 21-36.

Buzeman, D.G., Viano, D.C., Lovsund, P., 1998. Car occupant safety in frontal crashes: A parameter study of vehicle mass, impact speed, and inherent vehicle protection. *Accident Analysis and Prevention,* 30 (6), 713-722.

European road accidents database (CARE), 2009. *Road safety evolution in the EU*. http://ec.europa.eu/transport/road_safety/observatory/statistics/reports_graphics_en.htm, Accessed July 2010.

Claffey, P.J., 1976. *Passenger car fuel conservation*. US Federal Highway Administration, Report No. FHWA-PL-77009, Washington, DC.

Crandall, R.W., Graham, J.D., 1989. The effect of fuel economy standards on automobile safety. *Journal of Law and Economics,* 32, 97-118

De Vlieger, I., 1997. On board emission and fuel consumption measurement campaign on petrol-driven passenger cars, *Atmospheric Environment*, 31(22), 3753-3761

Department for Transport (DfT), 1993. *Cars: make and model: Injury accidents and casualty rates: Great Britain: 1991*. Transport Statistics Bulletin, HMSO, London.

Department for Transport (DfT), 2004. *Instructions for the completion of road accident reports (STATS20)*. Transport Statistics Bulletin, HMSO, London.

Department for Transport (DfT), 2006a. *Cars: make and model: The risk of driver injury in Great Britain: 2000-2004*. Transport Statistics Bulletin, HMSO, London.

Department for Transport (DfT), 2006b. *Review of road accident statistics*. National Statistics Quality Review Series, Report No. 45, Transport Statistics Bulletin, HMSO, London.

Department for Transport (DfT), 2007. *Transport statistics for Great Britain: 2007 edition*. Transport Statistics Bulletin, HMSO, London.

Department of Transport (DfT), 2009. *Reported road casualties Great Britain: 2008*. Transport Statistics Bulletin, HMSO, London.

Department of Transport (DfT), 2010. *Transport carbon reduction delivery plan*. Transport Statistics Bulletin, HMSO, London.

Department for Transport (DfT), 2010. About the Department's executive agencies and trading funds, http://www.dft.gov.uk/about/howthedftworks/aboutthedftexecu-tiveagencies, Accessed December 2010.

Dunn, G., 1989. *Design and analysis of reliability studies*. London: Edward Arnold, Page 111.

Ensfield, C., Bachman, L., Erb, A., Bynum, C., 2006. Evaluating real-world fuel economy on heavy duty vehicles using a portable emissions measurement system. *SAE Technical Paper*, doi:10.4271/2006-01-3543.

Ernst, E., Bruhning, E., Glaeser, K.P., Schmid, M., 1991. Safety in small and large passenger cars: the compatibility problem in head-on collisions. *In: Proceedings of the International IRCOBI Conference on the Biomechanics of Impacts*, 1991, Berlin.

Ernvall, T., Pirtala, P., Hantula, L., 1992. *The effect of car make and model on accidents*. University of Oulu, Road and Transport Laboratory, Report No. SF-90 100 OULU, Finland.

Euro NCAP, 2010. http://www.euroncap.com/faq.aspx, Accessed July 2010.

European Commission (EC), 1980. European Commission Directive 80/1268/EEC of 16 December 1980, the approximation of the laws of the Member States relating to fuel consumption of motor vehicles. *Official Journal of the European Communities*, L 375, 36-45.

European Commission (EC), 1993. European Commission Directive 93/116/EC of 17 December 1993, adapting to technical progress Council Directive 80/1268/EEC relating to the fuel consumption of motor vehicles. *Official Journal of the European Communities*, L 329, 39-53.

European Commission (EC), 1999. European Commission Directive 99/100/EC of 15 December 1999, adapting to technical progress Council Directive 80/1268/EEC relating to the fuel consumption of motor vehicles. *Official Journal of the European Communities*, L 334, 36-40.

European Commission (EC), 2004. European Commission Directive 2004/3/EC of 11 February 2004, amending Council Directives 70/156/EEC and 80/1268/EEC as regards the measurement of carbon dioxide emissions and fuel consumption of N1 vehicles. *Official Journal of the European Union*, L 49, 36-41.

European Commission (EC), 2007. *Results of the review of the community strategy to reduce CO_2 emissions from passenger cars and light-commercial vehicles*. COM(2007)19, Brussels.

Evans, L., 1985. Fatality risk for belted drivers versus car mass. *Accident Analysis and Prevention*, 17 (3), 251-271.

Evans, L., Wasielewski, P., 1987. Serious or fatal driver injury rate versus car mass in head-on crashes between cars of similar mass, *Accident Analysis and Prevention*, 19 (2), 119-131.

Evans, L., Frick, M.C., 1992. Car size or car mass: which has greater influence on fatality risk? *American Journal of Public Health*, 82 (8), 1005-1112.

Evans, L., Frick, M.C., 1993. Mass ratio and relative driver fatality risk in two-vehicle crashes. *Accident Analysis and Prevention*, 25 (2), 213-224.

Evans, L., 1994. Driver injury and fatality risk in two-car crashes versus mass ratio inferred using Newtonian mechanics. *Accident Analysis and Prevention*, 26 (5), 609-616.

Evans, L., Frick, M.C., 1994. Car mass and fatality risk: has the relationship changed? *American Journal of Public Health*, 84 (1), 33-36.

Evans, L., 2001. Causal influence of car mass and size on driver fatality risk. *American Journal of Public Health*, 91 (7), 1076-1081.

Evans, L., 2004. *Traffic safety*. Michigan: Bloomfield Hills. Chapter 4.

Fontaras, G., Samaras, Z., 2010. On the way to 130 g CO2/km – Estimating the future characteristics of the average European passenger car. *Energy Policy*, 38, 1826-1833.

Frampton, R., Welsh, R., Thomas, P., 2002. Belted driver protection in frontal impact: What has been achieved and where do future priorities lie? *Annual Proceedings Association for the Advancement of Automotive Medicine*, 46, 93-109

Gadd, W., 1966. Use of a weighted impulse criterion for estimating injury hazard. *In: Proceedings of 10th STAPP Car Crash Conference*, SAE paper 660973.

Gillespie, T.D., 1992. *Fundamentals of vehicle dynamics*. Warrendale, PA: Society of automotive engineers.

Greene, W.H., 1993. *Econometric analysis*. London: Macmillan.

Greene, D.L., 2009. Feebates, footprints and highway safety. *Transportation Research Part D*, 14 (6), 375-384.

Grime G., Jones I.S., 1970. Car collisions - the movement of cars and their occupants in accidents. *In: Proceedings (1969–1970) of the Institution of Mechanical Engineers*, Automobile Division, 184, Part 2A, No. 5.

Grime, G., Hutchinson, T.P., 1979. Vehicle mass and driver injury. *Ergonomics,* 22 (1), 93-104.

Grime, G., Hutchinson, T.P., 1982. The influence of vehicle weight on the risk of injury to drivers. *In: Proceedings of Ninth International Technical Conference on Experimental Safety Vehicles*, National Highway Traffic Safety Administration, Washington, DC., 726–741.

Gustafsson, H., Hagg, A., Krafft, M., Kullgren, A., Malmstedt, B., Nygren, A., Tingvall, C., 1989. *Folksam car model safety rating 1989-90*. Folksam, Stockholm.

Gyenes, L., Mitchell, C.B., 1994. *The effect of vehicle-road interaction on fuel consumption*. Vehicle-Road Interaction, ASTM STP 1225, American Society for Testing and Materials, Philadelphia, 225-239.

Heisler, H., 1989. *Advanced vehicle technology*. Second edition, Oxford: Hodder Butterworth-Heinemann.

Hosmer, D.W., Lemeshow, S., 2000. *Applied logistic regression.* Second edition, New York: John Wiley and Sons.

Highway Agency, 2007. *Design manual for roads and bridges*. Vol. 11, Section 3, Part 1: air quality, http://www.standardsforhighways.co.uk/dmrb/vol11/section3.htm, Accessed June 2010.

Hutchinson, T.P., Anderson, R.G., 2009. Vehicle mass as a determinant of fuel consumption and secondary safety performance: A comment. Transportation Research Part D, 15(2), 123-125.

Joksch, H.C., 1993. Velocity change and fatality risk in a crash-a rule of thumb. *Accident Analysis and Prevention*, 2 (5), 103-104.

Joksch, H.C., 1998. *Fatality risks in collisions between cars and light trucks*. Report No. DOT-HS-808-802, National Highway Traffic Safety Administration, Washington, DC.

Jones, P.A., Jorgensen, S.H., 2003. The use of multilevel models for the road accident outcomes. *Accident Analysis and Prevention*, 35 (1), 59-69.

Kahane, C.J., 1997. *Relationships between vehicle size and fatality risk in model year 1985–93 passenger cars and light trucks*. Report No. DOT-HS-808-570, National Highway Traffic Safety Administration, Washington, DC.

Kahane, C.J., 2003. *Vehicle weight, fatality risk and crash compatibility of model year 1991–99 passenger cars and light trucks*. Report No. DOT-HS-809-662, National Highway Traffic Safety Administration, Washington, DC.

Kim, D., Washington, S., 2006. The significance of endogeneity problems in crash models: an examination of left-turn lanes in intersection crash models. *Accident Analysis and Prevention*, 38 (6), 1094–1100.

Kirby, H.R., Hutton, B., McQuaid, R.W., Raeside, R., Zhang, X., 2000. Modelling the effects of transport policy levers on fuel efficiency and national fuel consumption. *Transportation Research Part D*, 5 (4), 265-282.

Koppel, S., Charlton, J., Fildes, B., Fitzharris, M., 2008. How important is vehicle safety in the new vehicle purchase process? *Accident Analysis and Prevention*, 40 (3), 994-1004.

Kumara, S.S.P., Chin, H.C., 2005. Application of Poisson underreporting model to examine crash frequencies at signalized three-legged intersections. *Transportation Research Record*, 1908, 46–50.

Kwon, T.H., 2006. The determinants of the changes in car fuel efficiency in Great Britain (1978-2000). *Energy Policy*, 34, 2405-2412.

Lane, B., Potter, S., 2007. The adaption of cleaner vehicles in the UK: exploring the consumer attitude – action gap. *Journal of Cleaner Production*, 15, 1085-1092.

Les, M., Fildes, B., 2001. *Developing vehicle aggressivity rating system*. SAE Technical Paper, 2001-01-3166.

Leung, D.C., Williams, D.J., 2000. Modeling of motor vehicle fuel consumption and emissions using a power-based model. *Environmental Monitoring and Assessment*, 65, 21-29.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives . *Transportation Research Part A*, 44 (5), 291-305.

Maddala, G.S., 2001. *Introduction to econometrics*. London: John Wiley & Sons.

McCullagh, P., Nelder, J.A., 1983. *Generalized linear models*. London: Chapman and Hall.

Miaou, S.P., Bligh, R.P., Lord, D., 2005. Developing median barrier installation guidelines: a benefit/cost analysis using Texas data. *Transportation Research Record*, 1904, 3–19.

Moin, P. 2001. *Fundamentals of engineering numerical analysis*. Cambridge University Press, Cambridge.

Mountain, L., Maher, M.J., Fawaz, B., 1998. The influence of trend on estimates of accidents at junctions. *Accident Analysis and Prevention*, 30 (5), 641–649.

National Research Council (NRC), 2002. *Effectiveness and impact of Corporate Average Fuel Economy (CAFE) standards*. National Academy Press, Washington, DC.

Newstead, S.V., Cameron, M.H., Le, C.M., 2000. *Vehicle crashworthiness and aggressivity ratings and crashworthiness by year of vehicle manufacture: Victoria and NSW crashes during 1987-98, Queensland crashes during 1991-98*. Monash University Accident Research Centre, Report No. 171.

Noland, R.B., 2004. Motor vehicle fuel efficiency and traffic fatalities. *The Energy Journal*, 25 (4), 1-22.

Noland, R.B., 2005. Fuel economy and traffic fatalities: multivariate analysis of international data. *Energy Policy*, 33 (17), 2183-2190.

North, R.J., Noland, R.B., Washington Y.O., Polak, J.W., 2006. Modelling of particulate matter mass emissions from a light-duty diesel vehicle, *Transportation Research Part D*, 11 (5), 344-357.

Pelkmans, L., Debal, P., 2006. Comparison of on-road emissions with emissions measured on chassis dynamometer test cycles. *Transportation Research Part D*, 11 (4), 233-241.

Post, K., Tomlin, J., Pitt, D., Carruthers, N., Maunder, A., Kent, J.H., Bilger, R.W., 1981. *Fuel economy and emissions*. Charles Kolling Research Laboratory, University of Sydney, TN ER-36.

Redsell, M., Lucas, G.G., Ashford, N.J., 1993. Factors affecting car fuel consumption. *Journal of Automotive Engineering*, 207 (D1), 1–22.

Richards, D., Cuerden, R., 2009. *The relationship between speed and car driver injury severity*. Road Safety Web Publication 9, Department for Transport, London, UK.

Ross, M., 1994. Automobile fuel consumption and emissions: effects of vehicle and driving characteristics. *Annual Review of the Energy and the Environment*, 19, 75-112.

Ross, M., 1997. Fuel efficiency and the physics of automobiles. *Contemporary Physics*, 38 (6), 381-394.

Ross, M., Wenzel, T., 2001. *Losing weight to save lives: A review of the role of automobile weight and size in traffic fatalities.* Report No. ACEEE-T013, American Council for an Energy-Efficient Economy.

Ross, M., Patel, D., Wenzel, T., 2006. Vehicle design and the physics of traffic safety. *Physics Today*, 59 (1), 49-54.

Schipper, L., 2008. *Automobile fuel economy and CO_2 emissions in industrialized countries: troubling trends through 2005/6.* EMBARQ, the World Resources Institute Centre for Sustainable Transport, Washington, DC.

Scottish Executive (SE), 2005. *Procedures for submitting road accident data to the Scottish Executive (STATS21 – Scottish Edition).* http://www.scotland.gov-.uk/Topics/Statistics/Browse/TransportTravel/MethodsStats21-2005, Accessed December 2010.

Sorrell, S., 1992. Fuel efficiency in the UK vehicle stock. *Energy Policy*, 20, 766-780.

Stone, R., 1989. *Automobile fuel economy.* London: Macmillan Education.

Taylor, M. P., Young, T.M., 1996. Fuel consumption and emission models for traffic engineering and transport planning applications. *In: Proceeding of Roads 96 Conference, Australian Road Research Board*, 18 (6), Publication No. UCD-ITS-RP-96-18.

Tenny, N., Lam, M., 1985. Estimating fuel consumption from engine size. *ASCE Journal of Transportation Engineering*, 111 (4), 339-357.

Tolouei, R., 2007. *Vehicle mass as a determinant of fuel consumption and safety performance*. CTS Working Paper 01-2007, ISSN 1747-6232, University College London, London.

Tolouei, R., 2008. Modelling the effects of vehicle design on fuel consumption, *Presented at the 40th University Transport Studies Group Conference*, Southampton, UK, 3-5 January 2008.

Tolouei, R., Titheridge, H., 2009. Vehicle mass as a determinant of fuel consumption and secondary safety performance. *Transportation Research Part D*, 14 (6), 385-399.

Tolouei, R., 2010. Secondary safety and aggressivity of makes and models: partial effects of mass and other design features. *Presented at the 42nd University Transport Studies Group Conference*, Plymouth, UK, 3-5 January 2010.

Tolouei, R., 2010. Comparing disaggregate and aggregate analysis of crash data to find the relationship between fatality risk ratio and mass ratio in two-car crashes. *Presented* *at the Fifth IMA International Conference on Mathematics in Transport*, London, 12-14 April.

Tolouei, R., Titheridge, H., Heydecker, B., 2009. Estimating secondary safety and aggressivity performance of makes and models in fleet using a novel methodology based on absolute driver injury risk. *Transportation Planning and Technology*, in review.

Toy, L.E., Hammitt, K.J., 2003. Safety impacts of SUVs, vans, and pickup trucks in two-vehicle crashes. *Risk Analysis*, 23 (4), 641-650.

Van Auken, R.M., Zellner, J.W., 2005. *Supplemental results on the independent effects of curb weight, wheelbase, and track of fatality risk in 1985-1998 model year passenger cars and 1985-1997 model year LTVs*. Report No. DRI-TR-05-01, Dynamic Research, Inc.

Van den Brink, R., Van Wee, B., 2001. Why has car-fleet specific fuel consumption not shown any decrease since 1990? Quantitative analysis of Dutch passenger car-fleet specific fuel consumption. *Transportation Research Part D*, 6 (2), 75-93.

Vehicle Certification Agency (VCA), 2007a., http://www.vcacarfueldata.org.uk-/downloads, Accessed May 2007.

Vehicle Certification Agency (VCA), 2007b., *Fuel consumption testing scheme*. http://www.vcacarfueldata.org.uk/information/fuelconsumptiontestingscheme.asp, Accessed May 2007.

Vehicle Certification Agency (VCA), 2010. *Exhaust emission testing*. http://www.vca-carfueldata.org.uk/information/emissions-testing.asp, Accessed July 2010.

Wang, W.G., Lyons, D.W., Clark, N.N., Luo G.D., 1999. Energy consumption analysis of heavy-duty vehicles for transient emissions evaluation on chassis dynamometer. *Proceedings of Institute of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 213(3), 205-214.

Wang, Z., Jin, Y., Wang, M., Wu, W., 2010. New fuel consumption standards for Chinese passenger vehicles and their effects on reduction of oil use and CO_2 emissions of the Chinese passenger vehicle fleet. *Energy Policy*, 38, 5242-5250.

Ward, H., Robertson, S., Lester, T., Pedler, A., 2002. *Reporting of road traffic accidents in London: matching police force STATS19 with hospital accident and emergency department data*. Transport Research Laboratory, Report No. PR/T/069/2002, TRL Limited, Crowthorne, UK.

Ward, H., Lyons, R.A., Thoreau, R., 2006. *Under-reporting of road casualties: Phase 1*. Road Safety Research Report No. 69. London: Department for Transport.

Watson, H.C.,1982. *Calibration and application of two fuel consumption models*. Fuel consumption modelling for urban traffic management, Australian Road Research Board, Research Report No. 124.

Wenzel, T.P., Ross, M., 2005. The effects of vehicle model and driver behavior on risk. *Accident Analysis and Prevention*, 37 (3), 479-494.

Wonnacott T.H., Wonnacott, R.J., 1990. *Introductory statistics*. Fifth edition, New York: John Wiley.

Wood, D.P., 1997. Safety and the cars size effect: a fundamental explanation. *Accident Analysis and Prevention*, 29 (2), 139-151.

Wood, D.P., Veyrat, N., Simms, C., Glynn, C., 2007. Limits for survivability in frontal collisions: Theory and real-life data combined. *Accident Analysis and Prevention*, 39 (4), 679-687.

Zachariadis, T., 2008. The effect of improved safety on fuel economy of European cars. *Transportation Research Part D*, 13 (2), 133–139.

Zervas, E., 2006. CO_2 benefit from the increasing percentage of diesel passenger cars. Case of Ireland. *Energy Policy*, 34, 2848-2857.

Appendix 1: Variants of Ford Fiesta in VCA fuel consumption dataset

Table A1.1: Variants of Ford Fiesta

ID	Make	Model	Trim	Year	Fuel	Transmission	Engine Size (cc)	Urban fuel cons. (l/100km)	Extra-urban fuel cons. (l/100km)	EU standard	Mass (kg)
1	FORD	Fiesta 2006 Model Year	1.6 Duratorq TDCi	6	Diesel	M5	1560	5.2	3.9	4	1157
2	FORD	Fiesta; October 2005 On	1.6 Duratorq TDCi	7	Diesel	M5	1560	5.2	3.9	4	1157
3	FORD	Fiesta Pre-2004½ Model	1.4 Duratorq TDCi CL (14 inch tyre)	5	Diesel	M5	1399	5.3	3.7	3	1139
4	FORD	Fiesta Pre-2004½ Model	1.4 Duratorq TDCi CL (14 inch tyre)	5	Diesel	ASM	1399	5.3	3.7	3	1139
5	FORD	Fiesta Pre-2004½ Model	1.4 TDCi CL 3 & 5 Door (14 inch tyre)	4	Diesel	M5	1399	5.3	3.7	3	1145
6	FORD	Fiesta Pre-2004½ Model	1.4 TDCi CL 3 & 5 Door (14 inch tyre)	4	Diesel	ASM	1399	5.3	3.7	3	1145
7	FORD	Fiesta 2004½ Model Year	1.4 Duratorq TDCi	5	Diesel	M5	1399	5.4	3.8	3	1139
8	FORD	Fiesta 2004½ Model Year	1.4 Duratorq TDCi	5	Diesel	ASM	1399	5.4	3.8	3	1139
9	FORD	Fiesta Pre-2006 Model Y	1.4 Duratorq TDCi	6	Diesel	M5	1399	5.4	3.8	3	1139
10	FORD	Fiesta Pre-2006 Model Y	1.4 Duratorq TDCi	6	Diesel	ASM	1399	5.4	3.8	3	1139
11	FORD	Fiesta Pre-2004½ Model	1.4 Duratorq TDCi LX; Zetec; Ghia (14 inch tyre)	5	Diesel	M5	1399	5.4	3.8	3	1156
12	FORD	Fiesta Pre-2004½ Model	1.4 Duratorq TDCi LX; Zetec; Ghia (14 inch tyre)	5	Diesel	ASM	1399	5.4	3.8	3	1156
13	FORD	Fiesta 2004½ Model Year	1.4 TDCi 3/5 Door Saloon	4	Diesel	M5	1399	5.4	3.8	3	1145
14	FORD	Fiesta 2004½ Model Year	1.4 TDCi 3/5 Door Saloon	4	Diesel	ASM	1399	5.4	3.8	3	1145
15	FORD	Fiesta Pre-2004½ Model	1.4 TDCi LX; Zetec; Ghia 3 & 5 Door (14 inch tyre)	4	Diesel	M5	1399	5.4	3.8	3	1145
16	FORD	Fiesta Pre-2004½ Model	1.4 TDCi LX; Zetec; Ghia 3 & 5 Door (14 inch tyre)	4	Diesel	ASM	1399	5.4	3.8	3	1145
17	FORD	Fiesta 2004½ Model Year	1.6 Duratorq TDCi	5	Diesel	M5	1560	5.4	4.1	3	1157
18	FORD	Fiesta 2004½ Model Year	1.6 Duratorq TDCi	5	Diesel	M5	1560	5.4	4.1	4	1157
19	FORD	Fiesta Pre-2006 Model Y	1.6 Duratorq TDCi	6	Diesel	M5	1560	5.4	4.1	3	1157
20	FORD	Fiesta Pre-2006 Model Y	1.6 Duratorq TDCi	6	Diesel	M5	1560	5.4	4.1	4	1157
21	FORD	Fiesta	1.6 Duratorq TDCi	7	Diesel	M5	1560	5.4	4.1	4	1157
22	FORD	Fiesta Pre-2004½ Model	1.4 Duratorq TDCi CL (15/16 inch tyre)	5	Diesel	M5	1399	5.5	3.9	3	1139
23	FORD	Fiesta Pre-2004½ Model	1.4 Duratorq TDCi CL (15/16 inch tyre)	5	Diesel	ASM	1399	5.5	3.9	3	1139
24	FORD	Fiesta Pre-2004½ Model	1.4 TDCi CL 3 & 5 Door (15/16 inch tyre)	4	Diesel	M5	1399	5.5	3.9	3	1145
25	FORD	Fiesta Pre-2004½ Model	1.4 TDCi CL 3 & 5 Door (15/16 inch tyre)	4	Diesel	ASM	1399	5.5	3.9	3	1145
26	FORD	Fiesta Pre-2004½ Model	1.4 Duratorq TDCi LX; Zetec; Ghia (15/16 inch tyre)	5	Diesel	M5	1399	5.6	4	3	1156
27	FORD	Fiesta Pre-2004½ Model	1.4 Duratorq TDCi LX; Zetec; Ghia (15/16 inch tyre)	5	Diesel	ASM	1399	5.6	4	3	1156
28	FORD	Fiesta Pre-2004½ Model	1.4 TDCi LX; Zetec; Ghia 3 & 5 Door (15/16 inch tyre)	4	Diesel	M5	1399	5.6	4	3	1145
29	FORD	Fiesta Pre-2004½ Model	1.4 TDCi LX; Zetec; Ghia 3 & 5 Door (15/16 inch tyre)	4	Diesel	ASM	1399	5.6	4	3	1145
30	FORD	Fiesta 2006 Model Year	1.4 Duratorq TDCi	6	Diesel	M5	1399	5.8	3.8	4	1139
31	FORD	Fiesta; October 2005 On	1.4 Duratorq TDCi	7	Diesel	M5	1399	5.8	3.8	4	1139
32	FORD	Fiesta; October 2005 On	1.4 Duratorq TDCi	7	Diesel	ASM	1399	5.9	3.7	4	1139
33	FORD	Fiesta	1.8 TDdi E-Diesel	1	Diesel	M5	1753	5.9	3.7	3	977

34	FORD	Fiesta	1.8 TDdi E-Diesel	2	Diesel	M5	1753	5.9	3.7	3	977
35	FORD	Fiesta	1.8 Turbo E-Diesel	1	Diesel	M5	1753	5.9	3.7	3	977
36	FORD	Fiesta	1.8 TDdi	1	Diesel	M5	1753	6.5	4.2	3	977
37	FORD	Fiesta	1.8 TDdi	2	Diesel	M5	1753	6.5	4.2	3	977
38	FORD	Fiesta	1.8 TDdi + A/C	1	Diesel	M5	1753	7	4.3	3	977
39	FORD	Fiesta	1.8 TDdi + A/C	2	Diesel	M5	1753	7	4.3	3	977
40	FORD	Fiesta	1.8 Turbo Diesel	0	Diesel	M5	1753	7	4.3	3	977
41	FORD	Fiesta	1.8 Turbo Diesel	1	Diesel	M5	1753	7	4.3	3	977
42	FORD	Fiesta; October 2005 On	1.25 Duratec	7	Petrol	M5	1242	7.8	4.7	4	1103
43	FORD	Fiesta 2006 Model Year	1.25 Duratec	6	Petrol	M5	1242	8.2	4.7	4	1103
44	FORD	Fiesta 2006 Model Year	1.4 Duratec	6	Petrol	ASM	1388	8.2	4.9	4	1108
45	FORD	Fiesta; October 2005 On	1.4 Duratec	7	Petrol	ASM	1388	8.2	4.9	4	1108
46	FORD	Fiesta 2006 Model Year	1.4 Duratec	6	Petrol	M5	1388	8.3	5	4	1108
47	FORD	Fiesta Pre-2006 Model Y	1.4 Duratec	6	Petrol	ASM	1388	8.3	5.3	4	1108
48	FORD	Fiesta	1.4 Duratec	7	Petrol	ASM	1388	8.3	5.3	4	1108
49	FORD	Fiesta; October 2005 On	1.4 Duratec	7	Petrol	M5	1388	8.3	5	4	1108
50	FORD	Fiesta 2004½ Model Year	1.4i 16V 3/5 Door Saloon	4	Petrol	ASM	1388	8.3	5.3	4	1105
51	FORD	Fiesta Pre-2004½ Model	1.4i 3 & 5 Door (14 inch tyre)	4	Petrol	ASM	1388	8.3	5.2	4	1102
52	FORD	Fiesta 2004½ Model Year	1.4i Duratec 16V	5	Petrol	ASM	1388	8.3	5.3	4	1108
53	FORD	Fiesta Pre-2004½ Model	1.4i Duratec 16V (14 inch tyre)	5	Petrol	ASM	1388	8.3	5.2	4	1108
54	FORD	Fiesta Pre-2006 Model Y	1.3 Duratec	6	Petrol	M5	1299	8.4	5	3	1107
55	FORD	Fiesta Pre-2004½ Model	1.3i 3 & 5 Door (14 inch tyre)	4	Petrol	M5	1299	8.4	5	3	1107
56	FORD	Fiesta 2004½ Model Year	1.3i 8V 3/5 Door Saloon	4	Petrol	M5	1299	8.4	5	3	1107
57	FORD	Fiesta 2004½ Model Year	1.3i Duratec 8V	5	Petrol	M5	1299	8.4	5	3	1107
58	FORD	Fiesta Pre-2004½ Model	1.3i Duratec 8V (14 inch tyre)	5	Petrol	M5	1299	8.4	5	3	1107
59	FORD	Fiesta Pre-2004½ Model	1.4i 3 & 5 Door (15/16 inch tyre)	4	Petrol	ASM	1388	8.5	5.6	4	1102
60	FORD	Fiesta Pre-2004½ Model	1.4i Duratec 16V (15/16 inch tyre)	5	Petrol	ASM	1388	8.5	5.6	4	1108
61	FORD	Fiesta Pre-2006 Model Y	1.25 Duratec	6	Petrol	M5	1242	8.6	4.9	3	1103
62	FORD	Fiesta 2004½ Model Year	1.25i 16V 3/5 Door Saloon	4	Petrol	M5	1242	8.6	4.9	3	912
63	FORD	Fiesta Pre-2004½ Model	1.25i 3 & 5 Door (14 inch tyre)	4	Petrol	M5	1242	8.6	4.9	3	912
64	FORD	Fiesta 2004½ Model Year	1.25i Duratec 16V	5	Petrol	M5	1242	8.6	4.9	3	1103
65	FORD	Fiesta Pre-2004½ Model	1.25i Duratec 16V (14 inch tyre)	5	Petrol	M5	1242	8.6	4.9	3	1103
66	FORD	Fiesta Pre-2004½ Model	1.3i 3 & 5 Door (15/16 inch tyre)	4	Petrol	M5	1299	8.6	5.3	3	1107
67	FORD	Fiesta Pre-2004½ Model	1.3i Duratec 8V (15/16 inch tyre)	5	Petrol	M5	1299	8.6	5.3	3	1107
68	FORD	Fiesta Pre-2004½ Model	1.4i 3 & 5 Door (14 inch tyre)	4	Petrol	M5	1388	8.6	5.1	3	1102
69	FORD	Fiesta Pre-2004½ Model	1.4i Duratec 16V (14 inch tyre)	5	Petrol	M5	1388	8.6	5.1	3	1108
70	FORD	Fiesta	1.25i 16V (13 inch tyre)	0	Petrol	M5	1242	8.7	5.8	2	912
71	FORD	Fiesta	1.25i 16V (13 inch tyre)	0	Petrol	M5	1242	8.7	5.8	4	912
72	FORD	Fiesta	1.25i 16V (13 inch tyre) - 01 January 2001 to 30 June	1	Petrol	M5	1242	8.7	5.8	2	912
73	FORD	Fiesta	1.25i 16V (13 inch tyre) - 1 July 2001 to 30 April 2002	1	Petrol	M5	1242	8.7	5.8	2	912
74	FORD	Fiesta	1.25i 16V (13 inch tyre) - 01 January 2001 to 30 June	1	Petrol	M5	1242	8.7	5.8	4	912
75	FORD	Fiesta	1.25i 16V (13 inch tyre) - 1 July 2001 to 30 April 2002	1	Petrol	M5	1242	8.7	5.8	4	912

76	FORD	Fiesta	1.3 EFi (13 inch tyre)	0	Petrol	M5	1299	8.7	5.9	3	1107
77	FORD	Fiesta	1.3 EFi (13 inch tyre) - 01 January 2001 to 30 June 2001	1	Petrol	M5	1299	8.7	5.9	3	1107
78	FORD	Fiesta	1.3 EFi (13 inch tyre) - 1 July 2001 to 30 April 2002	1	Petrol	M5	1299	8.7	5.9	3	1107
79	FORD	Fiesta	1.3 EFi (13 inch tyre)	2	Petrol	M5	1299	8.7	5.9	3	1107
80	FORD	Fiesta	1.3 EFi (non-PAS)	0	Petrol	M5	1299	8.7	5.9	2	1107
81	FORD	Fiesta	1.3 EFi (non-PAS) - 01 January 2001 to 30 June 2001	1	Petrol	M5	1299	8.7	5.9	2	1107
82	FORD	Fiesta	1.3 EFi (non-PAS) - 01 July 2001 to 30 April 2002	1	Petrol	M5	1299	8.7	5.9	2	1107
83	FORD	Fiesta Pre-2006 Model Y	1.4 Duratec	6	Petrol	M5	1388	8.7	5.2	3	1108
84	FORD	Fiesta Pre-2006 Model Y	1.4 Duratec	6	Petrol	M5	1388	8.7	5.2	4	1108
85	FORD	Fiesta	1.4 Duratec	7	Petrol	M5	1388	8.7	5.2	4	1108
86	FORD	Fiesta 2004½ Model Year	1.4i 16V 3/5 Door Saloon	4	Petrol	M5	1388	8.7	5.2	3	1105
87	FORD	Fiesta 2004½ Model Year	1.4i Duratec 16V	5	Petrol	M5	1388	8.7	5.2	3	1108
88	FORD	Fiesta 2004½ Model Year	1.4i Duratec 16V	5	Petrol	M5	1388	8.7	5.2	4	1108
89	FORD	Fiesta 2004½ Model Year	1.6i 16V 3/5 Door Saloon	4	Petrol	M5	1596	8.7	5.2	3	1108
90	FORD	Fiesta	1.25i 16V (14 inch tyre)	0	Petrol	M5	1242	8.8	5.7	2	912
91	FORD	Fiesta	1.25i 16V (14 inch tyre)	0	Petrol	M5	1242	8.8	5.7	4	912
92	FORD	Fiesta	1.25i 16V (14 inch tyre) - 01 January 2001 to 30 June	1	Petrol	M5	1242	8.8	5.7	2	912
93	FORD	Fiesta	1.25i 16V (14 inch tyre) - 1 July 2001 to 30 April 2002	1	Petrol	M5	1242	8.8	5.7	2	912
94	FORD	Fiesta	1.25i 16V (14 inch tyre) - 01 January 2001 to 30 June	1	Petrol	M5	1242	8.8	5.7	4	912
95	FORD	Fiesta	1.25i 16V (14 inch tyre) - 1 July 2001 to 30 April 2002	1	Petrol	M5	1242	8.8	5.7	4	912
96	FORD	Fiesta	1.25i 16V (14 inch tyre)	2	Petrol	M5	1242	8.8	5.7	4	912
97	FORD	Fiesta Pre-2004½ Model	1.25i 3 & 5 Door (15/16 inch tyre)	4	Petrol	M5	1242	8.8	5.2	3	912
98	FORD	Fiesta Pre-2004½ Model	1.25i Duratec 16V (15/16 inch tyre)	5	Petrol	M5	1242	8.8	5.2	3	1103
99	FORD	Fiesta Pre-2004½ Model	1.4i 3 & 5 Door (15/16 inch tyre)	4	Petrol	M5	1388	8.8	5.4	3	1102
100	FORD	Fiesta Pre-2004½ Model	1.4i Duratec 16V (15/16 inch tyre)	5	Petrol	M5	1388	8.8	5.4	3	1108
101	FORD	Fiesta 2006 Model Year	1.6 Duratec (4.06 FDR)	6	Petrol	M5	1596	8.8	5.2	4	1130
102	FORD	Fiesta; October 2005 On	1.6 Duratec (4.06 FDR)	7	Petrol	M5	1596	8.8	5.2	4	1130
103	FORD	Fiesta 2006 Model Year	1.6 Duratec (4.25 FDR)	6	Petrol	M5	1596	8.8	5.1	4	1130
104	FORD	Fiesta; October 2005 On	1.6 Duratec (4.25 FDR)	7	Petrol	M5	1596	8.8	5.1	4	1130
105	FORD	Fiesta	1.3 EFi (13 inch tyre)	0	Petrol	M5	1299	9	6	2	1107
106	FORD	Fiesta	1.3 EFi (13 inch tyre) - 01 January 2001 to 30 June 2001	1	Petrol	M5	1299	9	6	2	1107
107	FORD	Fiesta	1.3 EFi (13 inch tyre) - 1 July 2001 to 30 April 2002	1	Petrol	M5	1299	9	6	2	1107
108	FORD	Fiesta Pre-2006 Model Y	1.6 Duratec	6	Petrol	M5	1596	9.1	5.2	3	1130
109	FORD	Fiesta Pre-2006 Model Y	1.6 Duratec (4.25 FDR)	6	Petrol	M5	1596	9.1	5.2	4	1130
110	FORD	Fiesta	1.6 Duratec (4.25 FDR)	7	Petrol	M5	1596	9.1	5.2	4	1130
111	FORD	Fiesta 2004½ Model Year	1.6i 16V 3/5 Door Saloon (4.25 FDR)	4	Petrol	M5	1596	9.1	5.2	4	1108
112	FORD	Fiesta Pre-2004½ Model	1.6i 3 & 5 Door (14 inch tyre)	4	Petrol	M5	1596	9.1	5.2	3	1108
113	FORD	Fiesta 2004½ Model Year	1.6i Duratec 16V	5	Petrol	M5	1596	9.1	5.2	3	1130
114	FORD	Fiesta Pre-2004½ Model	1.6i Duratec 16V (14 inch tyre)	5	Petrol	M5	1596	9.1	5.2	3	1130
115	FORD	Fiesta 2004½ Model Year	1.6i Duratec 16V (4.25 FDR)	5	Petrol	M5	1596	9.1	5.2	4	1130
116	FORD	Fiesta	1.3 EFi (14 inch tyre)	0	Petrol	M5	1299	9.2	6.2	2	1107
117	FORD	Fiesta	1.3 EFi (14 inch tyre) - 01 January 2001 to 30 June	1	Petrol	M5	1299	9.2	6.2	2	1107

118	FORD	Fiesta	1.3 EFi (14 inch tyre) - 1 July 2001 to 30 April 2002	1	Petrol	M5	1299	9.2	6.2	2	1107
119	FORD	Fiesta	1.3 EFi (14 inch tyre) - 01 January 2001 to 30 June	1	Petrol	M5	1299	9.2	6.2	3	1107
120	FORD	Fiesta	1.3 EFi (14 inch tyre)	2	Petrol	M5	1299	9.2	6.2	3	1107
121	FORD	Fiesta	1.3 EFi (14inch tyre)	0	Petrol	M5	1299	9.2	6.2	3	1107
122	FORD	Fiesta	1.3 EFi (14 inch tyre) - 1 July 2001 to 30 April 2002	1	Petrol	M5	1299	9.2	6.2	3	1107
123	FORD	Fiesta Pre-2006 Model Y	1.6 Duratec (Zetec S)	6	Petrol	M5	1596	9.2	5.4	3	1128
124	FORD	Fiesta Pre-2006 Model Y	1.6 Duratec (Zetec S)	6	Petrol	M5	1596	9.2	5.4	4	1128
125	FORD	Fiesta	1.6 Duratec (Zetec S)	7	Petrol	M5	1596	9.2	5.4	4	1128
126	FORD	Fiesta 2004½ Model Year	1.6i Duratec (Zetec S)	5	Petrol	M5	1596	9.2	5.4	4	1130
127	FORD	Fiesta 2004½ Model Year	1.6i Duratec 16V (Zetec S)	5	Petrol	M5	1596	9.2	5.4	3	1108
128	FORD	Fiesta	1.25i 16V (13 inch tyre)	0	Petrol	A	1242	9.3	5.8	2	912
129	FORD	Fiesta	1.25i 16V (13 inch tyre)	0	Petrol	A	1242	9.3	5.8	3	912
130	FORD	Fiesta	1.25i 16V (13 inch tyre) - 01 January 2001 to 30 June	1	Petrol	A	1242	9.3	5.8	2	912
131	FORD	Fiesta	1.25i 16V (13 inch tyre) - 1 July 2001 to 30 April 2002	1	Petrol	A	1242	9.3	5.8	2	912
132	FORD	Fiesta	1.4i 16V	0	Petrol	M5	1388	9.3	6	2	1105
133	FORD	Fiesta	1.4i 16V - 01 January 2001 to 30 June 2001	1	Petrol	M5	1388	9.3	6	2	1105
134	FORD	Fiesta	1.4i 16V - 1 July 2001 to 30 April 2002	1	Petrol	M5	1388	9.3	6	2	1105
135	FORD	Fiesta	1.25i 16V (14 inch tyre)	0	Petrol	A	1242	9.4	6	2	912
136	FORD	Fiesta	1.25i 16V (14 inch tyre)	0	Petrol	A	1242	9.4	6	3	912
137	FORD	Fiesta	1.25i 16V (14 inch tyre) - 01 January 2001 to 30 June	1	Petrol	A	1242	9.4	6	2	912
138	FORD	Fiesta	1.25i 16V (14 inch tyre) - 1 July 2001 to 30 April 2002	1	Petrol	A	1242	9.4	6	2	912
139	FORD	Fiesta Pre-2004½ Model	1.6i 3 & 5 Door (15/16 inch tyre)	4	Petrol	M5	1596	9.4	5.4	3	1108
140	FORD	Fiesta Pre-2004½ Model	1.6i Duratec 16V (15/16 inch tyre)	5	Petrol	M5	1596	9.4	5.4	3	1130
141	FORD	Fiesta Pre-2006 Model Y	1.6 Duratec (4.06 FDR)	6	Petrol	M5	1596	9.5	5.7	4	1130
142	FORD	Fiesta	1.6 Duratec (4.06 FDR)	7	Petrol	M5	1596	9.5	5.7	4	1130
143	FORD	Fiesta	1.6i 16V	0	Petrol	M5	1596	9.5	6	2	1108
144	FORD	Fiesta	1.6i 16V	0	Petrol	M5	1596	9.5	6	4	1108
145	FORD	Fiesta	1.6i 16V - 01 January 2001 to 30 June 2001	1	Petrol	M5	1596	9.5	6	2	1108
146	FORD	Fiesta	1.6i 16V - 1 July 2001 to 30 April 2002	1	Petrol	M5	1596	9.5	6	2	1108
147	FORD	Fiesta	1.6i 16V - 01 January 2001 to 30 June 2001	1	Petrol	M5	1596	9.5	6	4	1108
148	FORD	Fiesta	1.6i 16V - 1 July 2001 to 30 April 2002	1	Petrol	M5	1596	9.5	6	4	1108
149	FORD	Fiesta	1.6i 16V	2	Petrol	M5	1596	9.5	6	4	1108
150	FORD	Fiesta 2004½ Model Year	1.6i 16V 3/5 Door Saloon (4.06 FDR)	4	Petrol	M5	1596	9.5	5.7	4	1108
151	FORD	Fiesta 2004½ Model Year	1.6i Duratec 16V (4.06 FDR)	5	Petrol	M5	1596	9.5	5.7	4	1130
152	FORD	Fiesta 2006 Model Year	1.6 Duratec	6	Petrol	A4	1596	10.2	5.8	4	1130
153	FORD	Fiesta; October 2005 On	1.6 Duratec	7	Petrol	A4	1596	10.2	5.8	4	1130
154	FORD	Fiesta Pre-2006 Model Y	1.6 Duratec	6	Petrol	A4	1596	10.4	5.9	4	1130
155	FORD	Fiesta	1.6 Duratec	7	Petrol	A4	1596	10.4	5.9	4	1130
156	FORD	Fiesta 2004½ Model Year	1.6i 16V 3/5 Door Saloon	4	Petrol	A4	1596	10.4	5.9	4	1108
157	FORD	Fiesta 2004½ Model Year	1.6i Duratec 16V	5	Petrol	A4	1596	10.4	5.9	4	1130
158	FORD	Fiesta 2006 Model Year	2.0 Duratec	6	Petrol	M5	1999	10.4	5.7	4	1165
159	FORD	Fiesta Pre-2006 Model Y	2.0 Duratec	6	Petrol	M5	1999	10.4	5.7	4	1165

160	FORD	Fiesta	2.0 Duratec	7	Petrol	M5	1999	10.4	5.7	4	1165
161	FORD	Fiesta; October 2005 On	2.0 Duratec	7	Petrol	M5	1999	10.4	5.7	4	1165
162	FORD	Fiesta 2004½ Model Year	2.0i Duratec 16V	5	Petrol	M5	1999	10.4	5.7	4	1165

Appendix 2: Estimated fuel consumption models

Table A2.1: Model estimation results (dependent variable: urban fuel consumption)

Variable	Model 4.1			Model 4.2			Model 4.3			Model 4.4		
	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat
Constant	2.357	0.008	308.392	-2.398	0.029	-81.305	1.519	0.008	198.806	1.477	0.019	77.414
Diesel	-0.342	0.005	-71.403	-0.358	0.002	-147.277	-0.418	0.003	-147.079	-0.371	0.004	-87.591
Automatic transmission	0.231	0.005	49.948	0.100	0.002	40.286	0.105	0.003	36.984	0.198	0.004	48.025
Other transmission	0.087	0.009	10.103	0.048	0.004	10.919	0.040	0.005	8.111	0.094	0.008	12.421
Euro II	0.072	0.010	7.191	0.055	0.005	10.840	0.060	0.006	10.442	0.024	0.009	2.769
Euro III	0.073	0.005	13.598	0.021	0.003	7.768	0.021	0.003	6.814	0.043	0.005	9.160
Time	0.002	0.001	1.857	-0.007	0.001	-9.906	-0.011	0.001	-13.825	-0.006	0.001	-5.107
Ln (Engine size)	-	-	-	0.637	0.004	162.247	-	-	-	-	-	-
Mass (kg)	-	-	-	-	-	-	0.00069	0.00001	133.834	-	-	-
Frontal area (m^2)	-	-	-	-	-	-	-	-	-	0.363	0.007	49.348
Models statistics												
Observations	9523			9523			9523			9523		
Null deviance	761			761			761			761		
Residual deviance	388			103			131			305		
Log L value	-20845			-14497			-15654			-19702		
AIC	41705			29013			31325			39422		

Table A2.1: Model estimation results (dependent variable: urban fuel consumption)

Variable	Model 4.5			Model 4.6			Model 4.7			Model 4.8		
	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat
Constant	-1.169	0.035	-33.106	-2.364	0.028	-83.698	1.641	0.013	130.242	-1.190	0.039	-30.780
Diesel	-0.388	0.002	-174.761	-0.366	0.002	-156.537	-0.418	0.003	-148.422	-0.388	0.002	-174.579
Automatic transmission	0.087	0.002	39.262	0.098	0.002	41.286	0.103	0.003	36.383	0.087	0.002	39.289
Other transmission	0.039	0.004	10.248	0.053	0.004	12.622	0.036	0.005	7.250	0.040	0.004	10.330
Euro II	0.056	0.004	12.536	0.042	0.005	8.554	0.068	0.006	11.826	0.055	0.005	12.290
Euro III	0.015	0.002	6.355	0.015	0.003	5.700	0.024	0.003	7.617	0.015	0.002	6.244
Time	-0.010	0.001	-16.214	-0.009	0.001	-13.683	-0.010	0.001	-12.904	-0.010	0.001	-16.273
Ln (Engine size)	0.420	0.005	77.086	0.592	0.004	146.178	-	-	-	0.422	0.006	75.544
Mass (kg)	0.00032	0.00001	51.109	-	-	-	0.00074	0.00001	113.149	0.00031	0.00001	41.216
Frontal area (m^2)	-	-	-	0.124	0.004	28.249	-0.075	0.006	-12.226	0.007	0.005	1.378
Models statistics												
Observations	9523			9523			9523			9523		
Null deviance	761			761			761			761		
Residual deviance	80			95			129			80		
Log L value	-13313			-14105			-15576			-13313		
AIC	26646			28230			31172			26648		

Table A2.2: Model estimation results (dependent variable: extra-urban fuel consumption)

Variable	Model 4.1			Model 4.2			Model 4.3			Model 4.4		
	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat
Constant	1.795	0.007	269.721	-1.808	0.032	-57.305	1.096	0.006	173.310	0.815	0.014	57.029
Diesel	-0.240	0.004	-57.517	-0.255	0.003	-98.250	-0.309	0.002	-131.447	-0.275	0.003	-86.711
Automatic transmission	0.167	0.004	41.292	0.065	0.003	24.516	0.059	0.002	24.885	0.128	0.003	41.505
Other transmission	0.069	0.007	9.291	0.042	0.005	8.995	0.034	0.004	8.215	0.078	0.006	13.820
Euro II	0.095	0.009	10.900	0.079	0.005	14.569	0.081	0.005	16.924	0.040	0.007	6.152
Euro III	0.068	0.005	14.530	0.027	0.003	9.384	0.023	0.003	8.917	0.033	0.004	9.327
Time	0.003	0.001	2.747	-0.004	0.001	-5.077	-0.008	0.001	-12.352	-0.006	0.001	-7.220
Ln (Engine size)	-	-	-	0.483	0.004	115.100	-	-	-	-	-	-
Mass (kg)	-	-	-	-	-	-	0.00058	0.00001	135.863	-	-	-
Frontal area (m^2)	-	-	-	-	-	-	-	-	-	0.405	0.006	73.589
Models statistics												
Observations	9523			9523			9523			9523		
Null deviance	469			469			469			469		
Residual deviance	276			111			89			171		
Log L value	-14028			-9662			-8609			-11724		
AIC	28072			19342			17236			23466		

Table A2.2: Model estimation results (dependent variable: extra-urban fuel consumption)

Variable	Model 4.5			Model 4.6			Model 4.7			Model 4.8		
	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat	Coef.	Std. Error	t-stat
Constant	-0.265	0.035	-7.576	-1.718	0.026	-66.228	0.939	0.010	91.186	-0.738	0.036	-20.443
Diesel	-0.294	0.002	-133.758	-0.273	0.002	-127.155	-0.309	0.002	-134.314	-0.291	0.002	-140.459
Automatic transmission	0.049	0.002	22.367	0.062	0.002	28.154	0.062	0.002	26.768	0.052	0.002	25.239
Other transmission	0.033	0.004	8.717	0.052	0.004	13.625	0.040	0.004	9.816	0.042	0.004	11.671
Euro II	0.079	0.004	17.838	0.051	0.004	11.513	0.071	0.005	14.994	0.063	0.004	14.957
Euro III	0.020	0.002	8.343	0.014	0.002	5.951	0.020	0.003	7.846	0.015	0.002	6.459
Time	-0.007	0.001	-12.478	-0.008	0.001	-13.535	-0.009	0.001	-14.151	-0.009	0.001	-15.656
Ln (Engine size)	0.213	0.005	39.404	0.392	0.004	105.260	-	-	-	0.250	0.005	47.920
Mass (kg)	0.00039	0.00001	63.647	-	-	-	0.00051	0.00001	96.132	0.00026	0.00001	36.766
Frontal area (m^2)	-	-	-	0.245	0.004	60.867	0.097	0.005	19.309	0.008	0.005	1.876
Models statistics												
Observations	9523			9523			9523			9523		
Null deviance	469			469			469			469		
Residual deviance	76			78			85			76		
Log L value	-7858			-8005			-8415			-7857		
AIC	15735			16030			16850			15736		

Appendix 3: Estimated fuel consumption models for selected makes and models

Table A3.1: Summary of model estimation results for selected makes and models (urban driving cycle)

Make and model		All makes and models	BMW 300	BMW 500	Vauxhall Vectra	Vauxhall Astra	Vauxhall Corsa	Mercedes E Class	Mercedes C Class	Ford Mondeo	Ford Fiesta	Ford Galaxy	Audi A6	Audi A3	Renault Laguna
Sample size		9737	490	267	394	340	259	388	245	324	137	91	298	133	218
Estimated coefficients and model statistics															
Model C1	Log (Engine)	0.69	0.71	0.57	0.76	0.66	0.56	0.44	0.48	0.86	0.44	0.62	0.48	0.65	0.66
	Log (Engine) × Diesel	0.20	-0.27	-0.29	-0.35	1.04	NS	NS	NS	-1.04	NS	-0.33	0.36	1.70	0.71
	Log (Engine) × Automatic	-0.19	-0.20	-0.24	NS	-0.31	-0.11	-0.17	-0.19	-0.29	0.29	-0.18	-0.13	-0.43	-0.17
	LL (AIC)	-14224 (28471)	-648 (1315)	-344 (707)	-460 (942)	-249 (546)	-113 (251)	-396 (814)	-170 (360)	-289 (597)	-15 (55)	-16 (51)	-424 (870)	-78 (177)	-137 (294)
Model C2	Mass	0.00073	0.00108	0.00129	0.00021	0.00048	0.00051	0.00044	0.00060	0.00121	0.00036	0.00079	0.00122	0.00082	0.00157
	Mass × Diesel	-0.00004	-0.00039	-0.00040	NS	0.00059	NS	NS	NS	-0.00122	-0.00129	-0.00097	NS	NS	NS
	Mass × Automatic	-0.00009	NS	-0.00024	NS	NS	NS	NS	NS	0.00076	NS	0.00457	-0.00045	-0.00067	NS
	LL (AIC)	-15925 (31874)	-789 (1597)	-407 (834)	-615 (1255)	-474 (967)	-345 (713)	-548 (1118)	-334 (687)	-443 (907)	-50 (124)	-23 (65)	-477 (975)	-118 (258)	-213 (447)
Model C3	Log (Engine)	0.56	0.62	0.47	0.77	0.64	0.55	0.43	0.50	0.85	0.40	1.01	0.45	0.57	0.61
	Log (Engine) × Diesel	NS	-0.23	-0.21	-0.39	1.05	NS	-0.07	-0.18	-0.96	-0.55	-0.76	0.25	NS	0.92
	Log (Engine) × Automatic	-0.26	-0.23	-0.26	NS	-0.31	NS	-0.14	-0.17	-0.42	0.84	0.21	-0.23	-0.30	-0.21
	Mass	0.00022	0.00032	0.00042	NS	0.00010	NS	NS	NS	NS	NS	-0.00464	0.00040	NS	NS
	Mass × Diesel	0.00014	NS	NS	NS	NS	NS	NS	0.00046	NS	-0.00096	0.00456	NS	NS	NS
	Mass × Automatic	0.00009	NS	0.00018	NS	NS	NS	-0.00004	NS	0.00083	-0.00070	NS	NS	NS	NS
	LL (AIC)	-12900 (25832)	-619 (1264)	-314 (654)	-458 (945)	-260 (544)	-112 (256)	-372 (774)	-150 (326)	-249 (524)	7 ()	11 ()	-413 (856)	-71 (172)	-134 (294)

NS: Not statistically significant at 5% level.
Model C1 includes fuel type, transmission type, EU standard, year, and *engine size (with interactions)* as explanatory variables.
Model C2 includes fuel type, transmission type, EU standard, year, and *mass (with interactions)* as explanatory variables.
Model C3 includes fuel type, transmission type, EU standard, year, *engine size, and mass (with interactions)* as explanatory variable