DOCTORAATSPROEFSCHRIFT

2006 | Faculteit Toegepaste Economische Wetenschappen

Road Safety, Risk and Exposure in Belgium: an Econometric Approach

Proefschrift voorgelegd tot het behalen van de graad van
Doctor in de Toegepaste Economische Wetenschappen, te verdedigen door:

Filip A.M. VAN DEN BOSSCHE

Promotor: Prof. dr. Geert WETS
Copromotor: Prof. dr. Koen VANHOOF

D/2006/2451/4

universiteit hasselt

The beauty of time…

Geswinde Grijsart die op wackre wiecken staech,
De dunne lucht doorsnijt, en sonder seil te strijcken,
Altijd vaert voor de windt, en ijder nae laet kijcken,
Doodtvijandt van de rust, die woelt bij nacht bij daech;

Onachterhaelbre Tijdt, wiens heten honger graech
Verslockt, verslint, verteert al watter sterck mach lijcken,
En keert, en wendt, en stort Staeten en Coninckrijken;
Voor ijder een te snel, hoe valdij mij soo traech?

Mijn lief, sint ick u mis, verdrijve' ick met mishaeghen
De schoorvoetighe Tijdt, en tob de lange daeghen
Met arbeidt avontwaerts; uw afzijn valt te bang.

En mijn verlangen can den Tijdtgod niet beweghen.
Maer 't schijnt verlangen daer sijn naem af heeft gecreghen,
Dat ick den Tijdt, die ick vercorten wil, verlang.

Pieter Corneliszoon Hooft (1581 - 1647)

Preface

In one of his recent books ("*Er zijn geen economische problemen*", 2001, Davidsfonds, Leuven), the Belgian politician and professor of Economics Mark Eyskens refers to his doctoral dissertation (1962) as a necessary sin of youth that every academic has to commit and for which no exemption or forgiveness exists. This statement is helpful in realising that a doctoral dissertation can only reflect the writer's knowledge at a certain point in time. While this may be frustrating once in a while, the transience of knowledge is at the same time an indication of the scientific progress that is made in the research domain. Especially in road safety research, this is only to be applauded.

Saying that the ending of this doctoral dissertation only brings about feelings of personal and scientific pride would be a sign of selfish and inappropriate pretension. It is thanks to the help and support of many people that this work could be finished, and feelings of gratitude are therefore also in order here.

In the first place, I am indebted to the internal members of the jury, my promoter prof. dr. Geert Wets, my co-promoter prof. dr. Koen Vanhoof and dr. Tom Brijs. They offered me the opportunity to carry out the scientific research described in this doctoral dissertation on a project for the Flemish Policy Research Centre for Traffic Safety that was supported by the Flemish Government.

I am very grateful to the external members of the jury, dr. Jacques Commandeur, dr. Lasse Fridstrøm and prof. dr. Marc Gaudry. The conversations and fruitful discussions with them have been valuable in gaining insights into time series analysis and its applications in road safety research. Also, this text improved considerably from their critical review and constructive comments.

Also, I owe many thanks to my colleagues at IMOB – Hasselt University, for their continuous support and the happy moments we shared together.

Of course, I am indebted to my parents, my family and my friends. They patiently bore the pressures caused by the commitment to bring this doctoral research to a favourable conclusion. Finally, I owe a large debt of gratitude to my wife Veerle, for her patience and understanding, for all the times that my mind was with this work, while it should have been with her.

Filip Van den Bossche
20 June 2006

Table of contents

List of tables

List of figures

Chapter 1 Introduction

1.1 The road safety problem

Thousands of pages have been written on the problem of road safety, but there is not one word, even not one sentence, that can express the pain, the sorrow and the irreparable loss caused by road accidents. Road safety is a worldwide problem, with consequences for public health, social life and economic prosperity of a country. One realizes the proportions of the problem when consulting the statistics published by the World Health Organization (WHO, 2004). The number of people killed in road traffic crashes each year is estimated to be around 1.2 million, and, without increased efforts, this number is expected to rise by 65% between 2000 and 2020. The number of injured persons may be as high as 50 million. That is the population of 5 of the world's largest cities.

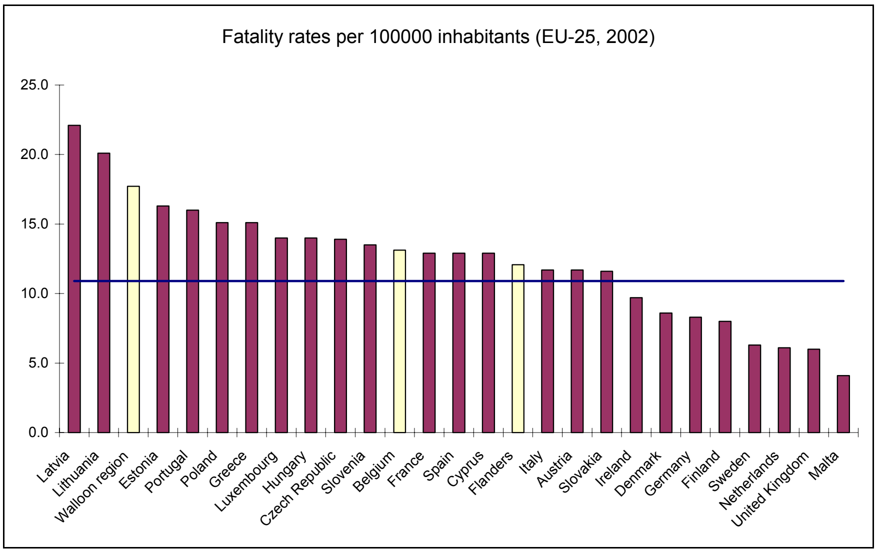

FIGURE 1: Fatality rates for European countries

Road traffic crashes are among the most prominent death causes all over the world. Deaths from road traffic injuries account for 25% of all injury deaths. Especially children and young adults take a prominent position in road crash statistics. Over 50% of the road traffic deaths are young adults (15-44 years old). For children aged 5-14 years and young people aged 15-29 years, road traffic injuries are the second leading cause of death worldwide. Apart from the human suffering caused by road crashes, it has been calculated that the direct economic costs of road crashes amount to $ 518 billion. For European Union countries, the estimated annual costs exceed € 180 billion.

On a European scale, the Belgian road safety performance is below average. Compared to the rest of Europe, the Southern and Eastern EU countries, together with France and Belgium, have relatively high fatality rates (ETSC, 2003). Between 1997 and 2001, the fatality risk in Belgium hardly reduced. In FIGURE 1, fatality rates per 100 000 inhabitants are shown for the 25 European countries (data for 2002, from EU-CARE and Statistics Belgium). The average fatality rate is at 10.9, which is clearly below the figure registered for Belgium. The graph also shows the fatality rates for Flanders and for the Walloon region. While Flanders shows a record comparable to that for Belgium as a whole, the Walloon region has a considerably higher fatality rate. Only Latvia and Lithuania show a higher rate.

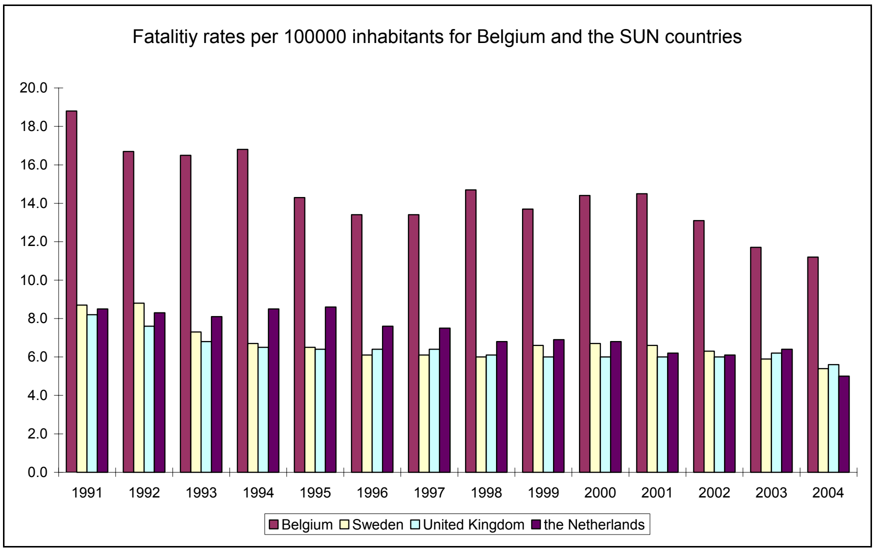

FIGURE 2: Fatality rates for Belgium and the SUN countries

However, road safety is not a country-related issue that stops at the border. From the graph, it is seen that 17 of the 25 countries listed are above average, indicating that Belgium is not an isolated incident. Road safety is an issue that should be dealt with on all policy levels, from the local government to the European institutions and world organisations. This has been understood during the last decennia, and many initiatives are taken by leading organisations.

Road accidents and victims are, in fact, a negative by-product of a transport system that starts from (a demand for) activities and ends in a set of traffic patterns and in a certain use of the available road infrastructure. The specific properties of this process are driven by human choices and by the situational opportunities and restrictions in terms of time and transport mode. In fact, the transport system and the way in which it is used are determined completely by an ongoing interplay between demand and supply. The fact that people travel is not, in itself, problematic. On the contrary, it is an indication of a highly developed economic, social and cultural life of the citizens of a country. People have to travel to go to school, to go to work or to participate in any social or cultural activities.

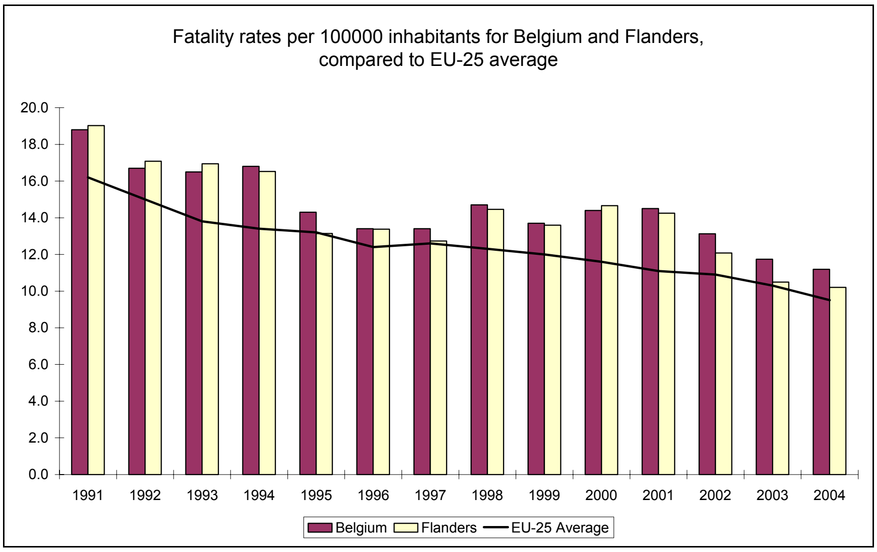

FIGURE 3: Belgium and Flanders compared to EU-25 average

On the other hand, traffic also has negative consequences. As long as people travel along public roads, there have been road accidents and victims. Although the local and federal governments in Belgium and abroad put in a great effort to increase traffic safety, also in the context of their sustainable transport policy, road safety is still an issue on the political agenda. Traffic accidents restrict people in their activities, and they limit other road users in their mobility.

FIGURE 1 also shows that some countries are able to reduce the level of road accidents and victims to a much lower level. The best performing countries in Europe are (apart from Malta) Sweden, United Kingdom and the Netherlands.

They are called the SUN countries, and they are known for their accident records that are far below the European average.

FIGURE 2 compares the Belgian fatality rate to that of the best performing countries over the last 14 years. The fatality rate in Belgium is, fortunately, decreasing, but the same trend is present in the other countries. Reaching the level of road safety of the best performing countries is, indeed, a difficult task.

FIGURE 3 shows the fatality rates per 100 000 inhabitants for Belgium and Flanders, compared to the EU-25 average for the years 1991-2004. The graph shows the overtaking manoeuvre in Flanders, resulting in a fatality rate that has been below the Belgian rate for several years now. The decreasing trend during the last years makes the EU-25 average fatality rate a realistic target for Belgium and Flanders. However, one should realize that, the higher the level of road safety, the harder it will be to obtain an additional road safety improvement. Road safety therefore is not a project; it is an ongoing concern for every government and every citizen worldwide.

1.2 Approaches towards the road safety problem

In the majority of the European countries, the number of injuries and deaths in road traffic has been reduced over the last decennia, in spite of the increasing amount of traffic. According to (OECD, 1997a), one reason for this promising evolution is to be found in the fact that road safety strategies and countermeasures are based on systematic research.

Many research programmes in road safety are developed around certain components of the problem. An example of a systematic approach towards road safety is the triad of E's: Engineering (making safer vehicles and roads), Enforcement (reduce unacceptable risk taking behaviour) and Education (changing attitudes and enhancing road skills). Later on, the triad became a quatrain, by adding Evaluation (assess the quality and results of road safety activities), which is getting more important when scarce resources are to be divided over a set of road safety problems. Using these concepts as a starting point for research ensures a complete overview of all aspects that are related to road safety.

However, the approaches to road safety research can be described on a higher level of abstraction. In (OECD, 1997a, 1997b), four approaches towards road safety research are specified. A first approach is a purely descriptive one, where the main task is to describe the magnitude of the road safety problem. Usually, road safety problems are described in terms of exposure, accident risk and injury consequence. This topic will be further discussed in section 1.3. Instead of focusing exclusively on the number of accidents (consequence), it is necessary to consider the activities that generate the road safety problems (exposure). The accident risk dimension is then the relationship between the level of exposure and the accidents. Due to the random character of the occurrence of accidents, statistical methods are preferred to analyse road safety data.

A second approach is to explain and predict the level of road safety by means of analytical macro models. In (OECD, 1997a), six categories of factors may influence changes in the reported accident occurrence: external factors (like the weather), socio-economic factors (for example, unemployment), transportation (like the infrastructural properties), the data collection system, randomness involved in accident occurrence and countermeasure interventions. The main interest of a researcher usually lies in the assessment of the effectiveness of countermeasures. This can be done by predictive modelling on cross-sectional or time series data, or by effectiveness evaluations in before-after studies. While cross-sectional models start from spatial variation, a time series model considers temporal variation in the data. In other words, a cross-sectional model is concerned with road safety at specific locations, while a time series study looks at trends in road safety over time. For cross-sectional studies, it is obvious to mention the eminent work of Ezra Hauer (Hauer, 1997) and Rune Elvik (Elvik & Vaa, 2004). These authors provided road safety workers with a common framework to analyse the effect of road safety interventions at specific locations (road segments, intersections, etc.). A very popular application nowadays of the cross-sectional approach is black spot analysis, which aims at ranking and treating road sites with high potential for improvement. In the time series approach, the work of Marc Gaudry (Gaudry & Lassarre, 2000) and Siem Oppe (Oppe, 1989, 1991) can be mentioned as guiding examples for the research community. Most of the analytical models, both cross-sectional and temporal, are based on the application of advanced statistical techniques. However, (OECD, 1997a) warns against the blind application of statistics. Every model should be a means to answer the research questions, and the data used in the model should be based on theoretical arguments and professional knowledge.

A third approach includes risk factor models, or analytical models at the individual level. These models start from the triad of road user, vehicle and environment. Indeed, it is commonly assumed that the road safety problem is caused by the continuous interaction between these traffic system components.

As this approach is working from the micro-level, risk factor modelling strongly relates with psychology, sociology, engineering etc., emphasising the multi- and interdisciplinary character of road safety research. Most of the models, however, focus on one specific aspect of road safety. Within the risk factor models, a distinction is made between human factor models and technical models. Human factor models study the impact of individual variables (reaction time, personality, etc.) on road safety, concentrate on the road user task analysis (driving skills, suitability of road users, etc.) or study the attitude and behaviour of road users. Technical models investigate the relationship between road safety and vehicle characteristics (vehicular models), road design (infrastructural models) or traffic characteristics (traffic models).

A fourth approach contains models that are related to the accident consequences. This is, again, a very broad research field, that can roughly be divided in models that consider the consequences of road traffic accidents as a road safety problem from a public health perspective, or as an economic issue. Accident consequence models try to answer the question how the injury consequences of accidents can be reduced. In this context, topics as changes in the road design, vehicle characteristics, active and passive safety, rescue speed, etc. are analysed. Further, the models are concerned with the factors that influence the consequences of accidents (the road user, the vehicle, the road situation, speed, etc.). Also, various schemas have been developed to classify the different injury levels in a unified way. Apart from the application of statistical methods on real road safety data, accident consequence model often try to detect, understand and prevent injuries by means of experiments and simulated collisions.

From this overview, it should be clear that many research efforts and road safety models can hardly be classified in only one of these categories. In a sense, this categorisation is experienced as an artificial framework. Roughly speaking, the models developed in this manuscript are mainly considered as analytical macro levels (which is the second approach), although some models involve road characteristics, road user properties, and so on. The main message brought by this classification is, of course, that every modelling approach towards road safety should be taken from a scientific point of view. Road safety research should not be done off the top of one's head. Whichever approach is followed, the road safety problem should be taken seriously, based on fundamental theory, sound modelling and expert knowledge.

1.3 General concepts: exposure, risk and loss

While the approaches presented in the previous section are all oriented towards a specific part of the road transport system, it is generally accepted that road safety can be described in three principal dimensions: exposure, risk and consequence (OECD, 1997b). In other sources, this triad is called "Exposure / Risk / Severity" (COST 329, 2004) or "Exposure / Frequency / Severity" (Gaudry & Lassarre, 2000). The first dimension describes the magnitude of the activity which results in accidents (the exposure), measured in terms of number of trips, number of vehicle kilometres or trip duration. It accounts for the number of potentially dangerous situations, or the exposure to risk. The second dimension is the probability of an accident or the risk, given a certain level of exposure. The third dimension is the accident consequence. Changes in one of these dimensions will change the entire safety situation. The three dimensions are naturally related to one another in a multiplicative way:

$$Fatalities = (Exposure) \times \left(\frac{Accidents}{Exposure} \right) \times \left(\frac{Fatalities}{Accidents} \right) \qquad (1)$$

In (COST 329, 2004), the same relation between the road safety dimensions is given, but is also applied to the number of accidents as follows:

$$Accidents = (Exposure) \times \left(\frac{Accidents}{Exposure} \right) \qquad (2)$$

By definition, risk is thus the ratio of the expected number of accidents to the level of exposure. However, in many studies, risk is also used to refer to the number of fatalities divided by exposure (Lassarre, 2001; Oppe, 1991). Although there is clearly no agreement on terminology in this respect, the underlying idea of decomposing road safety outcomes is always present.

When the number of fatalities is studied directly, it is clear that the underlying information on exposure, risk and consequences is lost. However, these dimensions can provide useful explanations of the trends in the number of fatalities. Indeed, an increase in the number of fatalities can be caused by a rise in the level of exposure, an increase in the risk level, a change in the accident consequences or a combination of some of these factors. As an example, consider the introduction of airbags in cars. This safety measure will reduce the number of fatalities in accidents, but if road users compensate for the higher safety level by changing their driving habits, the number of accidents might increase. Other measures may decrease the number of accidents, but at the same time increase the severity of accident consequences. These opposite forces remain hidden if risk and accident consequences are not studied explicitly. Also, if a safety measure leads to a re-distribution of traffic over the different road networks, the safety level may be directly influenced by the changes in exposure. The quantification of these impacts provides useful information for road safety policy makers.

1.4 Objectives and organisation of the manuscript

This manuscript contains a number of analyses of road safety, exposure and risk applied to Belgian data. Road safety research is not new in Belgium, and many efforts have been made to support the road safety policy by research results obtained on Belgian data. However, this is not the case for the evolutionary study of road safety. Most studies have been cross-sectional in nature, or analysed only partly the relationship between road safety (e.g. accidents or victims), exposure (e.g. kilometres driven) and a risk component.

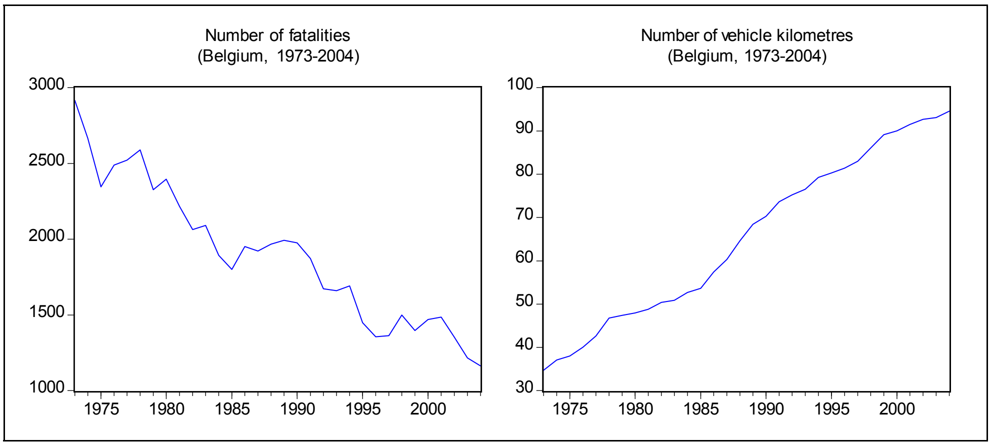

FIGURE 4: Number of fatalities and vehicle kilometres in Belgium

This document takes a time series approach towards exposure and risk, and offers a thorough analysis of Belgian road safety from an evolutionary point of view. That is, all variables used in the subsequent chapters are "time series", which means that they are gathered sequentially and with a regular frequency over time. For example, yearly data on the number of victims is related to the yearly number of kilometres driven, as shown in FIGURE 4. Clearly, the number of fatalities shows a decreasing trend since 1973, while the number of vehicle kilometres has been increasing ever since.

The major objective of this work is to clarify, for the Belgian road safety situation, the relations between the concepts of road safety, exposure and risk in the time domain. After the introduction to the world of road safety research in this chapter, Chapter 2 provides an overview of the field of time series road safety models (Van den Bossche & Wets, 2003a). The chapter gives a motivation for this type of modelling in road safety research, introduces a model typology and reviews the most important results from the relevant literature. In Chapter 3, the nature of exposure measures is described, together with the problems that are often encountered with this type of data. Apart from these introductory chapters, the major contributions of this manuscript are organised as follows.

1. To start, in Chapter 4, an effort has been made to identify and gather possible sources of exposure data for the Belgian road safety situation. An overview is given of the available data for macroscopic road safety research in Belgium, at all levels of aggregation. In particular, time series were gathered to allow the kind of analysis that has been described above. Apart from the description of data sources that were available from official government statistics and transport authorities, a monthly measure of exposure is developed. This variable measures the number of kilometres driven on Belgian roads, based on fuel deliveries, fuel efficiency and the vehicle park. Further, Belgian road safety data (number of crashes per time period, number of fatalities, etc.) and various explanatory variables are introduced in this chapter. The main objective of the research is now to investigate the different ways in which the gathered data can be used in the road safety – exposure – risk framework. That is, given the data that could be found or derived to represent road safety, exposure and road risk, the basic relationship between these components can be established in the subsequent chapters.
2. Chapter 5 discusses possible modelling approaches to analyse the trends in the number of fatalities on a highly aggregated level. By means of yearly data, the relation between road safety, exposure and risk is modelled for the Belgian situation. The objective of this exercise is to apply some basic and more advanced statistical models and to compare their performance in terms of model fit and forecasting accuracy. These models are also used for aggregated road safety forecasting, and are therefore highly relevant for road

safety policy makers. An example of a basic model for aggregated road safety analysis was developed in (Van den Bossche et al., 2005b) and in (Van den Bossche et al., 2005c).

3. In Chapter 6, road safety and exposure are analysed at a lower level of aggregation in time, namely on monthly instead of yearly data. Using monthly data has the advantage of a higher number of observations in time, which allows the development of explanatory road safety models that can be used to associate the trends in road safety with the developments of certain factors over time. Also, monthly observations are less aggregated and are therefore less fraught with aggregation bias than yearly observations. Starting from descriptive ARIMA and state space models for Belgian data, the analysis is gradually extended with explanatory variables. A first model only uses calendar data as covariates, as was done in (Van den Bossche et al., 2006), but this model is then further elaborated to end up with a full explanatory road safety model for Belgium. Previous work on this topic is published in (Hermans et al., 2006a, 2006b; Van den Bossche & Wets, 2003b; Van den Bossche et al., 2004; Van den Bossche et al., 2005a).
4. As it is recognized that road safety programs are oriented towards specific aspects of the road system, a number of disaggregated models is presented in Chapter 7 to assess the relationship between road safety, exposure and risk. These models are again developed on yearly data. Models will be presented for categories of road users by age and gender and by type, for different transport modes and for two-sided accidents.
5. As the reader will be able to conclude from the development of the various chapters, certain types of analysis are not possible with the available exposure data. More specifically, one cannot analyse the trends in road safety for vulnerable road users, simply because their exposure to road risk has not been measured over time. Therefore, some cross-sectional analyses of Flemish road safety and travel survey data are shown in Chapter 8 (Van den Bossche et al., 2005b).

Throughout the various chapters of this manuscript, the reader will get an idea of the possibilities and limitations of the available road safety data sources in Belgium. Actually, since all available time series data sources have been used in this work, it gives an overview of state-of-the-art road safety and risk models that can be applied to the Belgian data, at various levels of aggregation. The global conclusions of the research presented in this manuscript are written down in Chapter 9.

Chapter 2 Time series road safety models

2.1 Introduction

Investigating time series of road safety data is, for various reasons, a rewarding effort. As stated in the introduction, governments are more than ever concerned with the negative consequences of road traffic. The loss of human lives and economic capital due to road accidents forced governments to plan and execute road safety actions (COST 329, 2004). In the early years, until the early seventies, the planned actions were of a reactive nature. The main objective was to stop or slow down the negative developments in road safety. Afterwards, the focus changed to a more strategic reaction. It was commonly accepted that strategic plans for future safety improvements had to be developed, based on the knowledge of the past.

In (OECD, 1997a), the evolution in road safety research is described by means of 4 successive periods or paradigms. First, when the first cars appeared on the roads, they were considered as a replacement for the horse-drawn carriages from that time. Being new in the every day life of citizens, the cars were taken as the main source of unsafety. Second, when more and more cars appeared on the roads, together with more accidents and victims, the need to master the traffic situation, consisting of a number of cars, became clear. It was recognised that it was inadequate to consider cars as the unit of unsafety, and that the complicated traffic situation should be controlled in order to reduce accidents. In a next step, the scope of analysis was broadened from individual traffic situations to the entire traffic system. That is, unsafety is now regarded as an unwanted side effect in a large system. Around this time, the major road safety interventions (safety belts, laws on impaired driving, speed limits, etc.) were introduced to eliminate the risk factors from the road traffic system. To succeed in selling these (sometimes unpopular) new measures, they were combined in programmes that promised a higher level of road safety. Fourth, in the eighties, the scope of road traffic as a system was again experienced as being too narrow. At this time, it was recognised that an increasing mobility was inseparably related to the level of road risk. Realising that the increasing motorized traffic does not only affect road safety, the scope of analysis was enlarged, and traffic safety policy became part of a transport policy. Other modes of transport than motorized traffic were brought to the attention of both road users and policy makers, and the concept of sustainable transport was introduced.

Roughly speaking, the first two periods in this overview can be classified as a reactive approach to road safety, while the strategic issues were introduced little by little starting from the third period. This change in policy is mainly characterised by the introduction of strategic plans and programmes, together with explicit road safety targets. It is clear that, by putting quantitative targets, the noncommittal attitude towards the road safety problem is no longer accepted and underlines the need for ambitious, yet realistic targets. In order to set these targets and to specify or adapt safety programmes accordingly, it is necessary to measure the developments in road safety and to understand the underlying processes of exposure and risk. This is the main motivation for the use of time series analysis techniques in road safety research. A sound analysis of the evolution in time is needed to set future road safety targets and to assess the efforts made to achieve them.

In this chapter, an introduction to the various aspects of time series road safety models is given. First, the main areas of application of the models are described. Next, an attempt is made to define a typology of time series road safety models. In this context, a distinction is made between deterministic versus stochastic models and descriptive versus explanatory models. These topics are discussed in the next two sections. For the explanatory models, an overview is given of the most reported effects of explanatory variables. Last, the special DRAG family of explanatory road safety models is introduced. Given the extensiveness of the field of time series road safety models, and the lack of structure among the various models, the author does not pretend to give a complete and all-embracing overview of the field. This chapter should be read as a conceptual introduction to the subsequent chapters of this work.

2.2 Areas of application

Given the strategic importance of time series analysis, described in broad outline above, four main areas of application can be distinguished (COST 329, 2004). The first is a descriptive analysis of the road safety situation. The second approach extends the descriptive models with explanatory variables in order to understand the developments in road safety. The third application is to predict the level of safety in the future. The latter can combined with both descriptive and explanatory models. A fourth area of application is the (inter)national comparison of the road safety improvements in different geographical regions. The application areas will be shortly introduced here, and more details are provided in the subsequent chapters.

2.2.1 Describing road safety

The first use of time series analysis is to describe the trends in traffic safety by means of descriptive models. Looking at the transport system as a whole, the kind and the number of trips made by road users will determine the level of exposure and the associated risk. Descriptive time series models can be roughly divided in two categories. The first group of models considers the number of accidents or fatalities as the result of the combined action of exposure and risk. When the trend in road safety is studied, it is decomposed in these underlying factors and estimated as the result of their movements. In the second group of descriptive models, the road safety outcome is analysed directly without looking at the underlying patterns in exposure and risk. In general, a descriptive model will be used to assess whether or not the actual trends are in line with the objectives. In the transport system, the number of trips made by road users and initiated by their activities, can be considered as the input for the system. The throughput is then the specific travel pattern that is generated by the combined effect of the movements and choices made by the various road users (characterised by traffic concentration on certain roads, congestion, speed,...). This process, which determines whether the outcome is a success (a safe trip) or a failure (an accident), is highly surrounded by random fluctuations. Therefore, a statistical time series approach to studying road safety evolution might be preferred to a year-to-year comparison of the number of crashes. Moreover, when a certain effect is expected of a measure that was introduced to reduce the number of fatalities or crashes, it might well be that this effect will not be visible after one year. Deriving any value judgement from the comparison of two consecutive years can thus result in a completely wrong assessment of that safety measure. The analysis of a longer time series with the appropriate time series analysis techniques will therefore lead to well-founded conclusions.

2.2.2 Explaining road safety

Instead of only looking at the evolution in time of the road safety indicators, explanatory models also incorporate the trends in the factors that might influence the level of road safety. In (COST 329, 2004), a distinction is made between first order and second order accident factors. The first order factors are the level of traffic, the demography and the level of economic activity. The level and the structure of exposure (or traffic) and the distribution of exposure over the various transport modes might influence the number and the kind of accidents. Also demographic changes can have an effect on exposure and road safety. For example, the sharp rise in the ageing population increases the group of vulnerable road users and undoubtedly influences the use of transport modes. A last first-order factor is the economic activity. Changes in the level or the structure of economic life (e.g. unemployment) may result in a different transport pattern and road safety level.

Apart from these first order factors, there is a whole set of other variables that may influence the road transport system. For example, in the DRAG model (Gaudry, 1984, revised 2002) that will be introduced later in this text, various categories of explanatory variables are considered: fuel prices, motorisation quantity and characteristics, road network properties (transport modes, laws and road safety measures, infrastructure, climate,…), road user properties (age and gender, level of consumption, leisure activities, alcohol consumption,…) and specific variables related to the calendar (number of weekends and working days in a month, leap years,…). It is clear that, because of the large amount of possible variables, modelling these issues is not an easy task. Explanatory models will be discussed in more detail further in this text.

2.2.3 Predictive models

Both descriptive and explanatory models can, in the end, be used for predicting the level of road safety in the future. However, depending on the kind of prediction that is needed, one model will be better suited than another. For example, if an explanatory model is used, then predictions for the road safety outcomes can only be made if future values for the explanatory variables are available. The frequency of the data used in the model will determine the kind of prediction that can be made. According to (COST 329, 2004), models developed on yearly data are less suited for predicting short-term road safety developments or effects of special measures. Typically, the type of forecast and the forecasting horizon will depend on the objective of the researcher and the characteristics of the available data. If a modeller is interested in the recurring seasonal pattern of a series, a too high level of aggregation (like one observation per year) will not be sufficient to make monthly predictions. If on the other hand only highly aggregated predictions are needed, yearly data are useful. Also the data at hand determines the prediction possibilities. Obviously, the properties of the series (seasonal pattern, volatility, trend, etc.) will influence the quality of the predictions. The length of the available series is a determining factor in the choice of the prediction horizon. Especially the ratio of the number of predicted values to the number of observations in the sample cannot be too high. Usually, this issue, related to the frequency of the series, cannot be influenced, as it is determined by the official institutions that gather the data. Typically, official statistics are only available on a yearly basis, and every interpolation of these series to another frequency will probably reduce the data quality.

Sometimes, explanatory models are used to predict future values of the series. At first sight, one would expect that an explanatory model, explicitly taking information on accident causes into account, should lead to better predictions than a purely descriptive model and, as such, be more policy relevant. Although prediction is indeed a possible and interesting application of explanatory models, one can spare oneself the trouble of looking for explanatory factors if the model will only be used for predicting road safety, as the prediction accuracy is often comparable. Also, predicted values for the explanatory variables are needed, which increases the uncertainty in the predictions for the dependent variable.

2.2.4 Road safety (inter)national comparison

On a policy level that goes beyond the country borders, there might be an interest in a comparison of the road safety developments in various countries. Looking at the current initiatives at these levels (COST 329, 2004; WHO, 2004), this application of road safety models is highly appreciated. Also other authors did some efforts in this respect, see for example (Commandeur, 2002; Lassarre, 2001; Oppe, 1991). However, the same philosophy can be followed within a country. There might be good reasons to assume that road safety will not evolve in the same manner in all regions (states, provinces, etc.) of a country.

If differences in trends are observed, then an explanatory model can provide further insights into the similarities and differences in the explanatory factors that can be associated with the given developments. It is clear that this approach can also be very instructive for countries that do not present very good results in terms of road safety. Of course, a comparison between different countries goes with problems that transcend those encountered at a national level. Apart from the difficulties, experienced by all countries, to gather the relevant data, the differences in quality, the seemingly comparable content and the varying availability and quality of the variables are serious issues in this respect. These problems do not, however, detract from the potential benefits offered by such an effort.

2.3 Typologies of time series road safety models

In the literature, there is no common framework that can be used to classify the existing time series road safety models. In (Cameron, 1997), reference is made to a schema proposed by Hakkert and McGann (1996). They distinguish between macro-models, meso-models and micro-models, as follows: "*A macro-model describes the development of the road safety situation on a highly aggregated level, generally using national statistics. The models mostly present a description of the trend over time without pretence of giving any explanation for details of change. Meso-models try, still on an aggregated level, to include some explanatory variables. Micro-models attempt to introduce more explanatory variables and generally treat only a small segment of the overall safety picture*". This typology is shown in TABLE 1.

Hakkert and McGann focus on two distinctive factors of road safety models. The first is the level of aggregation. Some models consider highly aggregated data, while other models are oriented towards specific segments of road safety. Aggregated models are useful to monitor the total number of fatalities in road crashes for a country or a well-defined (large) region, relating the number of fatalities to the national level of exposure. Disaggregated models investigate changes in parts of the transport system (like a specific type of road) or look at the safety of subgroups of road users. Note that these models are still aggregated in time, as they consider road safety outcomes for aggregated time units (month, year,...). The second factor is the amount of explanatory power in the model. Macro-models tend to be more descriptive, which means that they focus on separating the trend and the random component in a series of observations. Contrary to descriptive models, explanatory models try to explain the developments in road safety by measuring the effects of a set of explanatory variables. According to Hakkert and McGann, this is typical for meso-models and micro-models.

TABLE 1: Typology of road safety models (Hakkert & McGann, 1996)

	Descriptive	**Explanatory**
Aggregated	MACRO	MESO
Disaggregated		MICRO

Although this typology is an interesting starting point to structure the proliferation of road safety models, it offers only a partial representation of the distinguishing model characteristics. The level of aggregation of a model is one aspect of the kind of application for which the model is built. From the same point of view, it is also interesting to distinguish between direct versus indirect models and long term versus short term models. While direct models are directly applied to final outcomes of road safety in terms of crashes or victims, an indirect model will consider the number of fatalities as the combination of exposure, risk and loss, as was discussed in Chapter 1. If a model is needed to predict road crashes and fatalities without insight in the separate movements of exposure and risk, a direct model is developed. If more details are required on how exposure, risk and losses influence the level of road safety, an indirect model is preferred. Another application-based distinction that can be made is between short term and long term prediction models. Some models are suited to provide forecasts for the next few months, while other models can be used to predict the level of road safety in 2010. Depending on the kind of prediction that is required, a specific modelling approach will be preferred.

The distinction between descriptive and explanatory models has a special meaning in the statistical sense. That is, the statistical formulation of an explanatory model may be quite different from a descriptive model. From a statistical point of view, models may be further categorised as deterministic versus stochastic models and univariate versus multivariate models. In a deterministic model, it is assumed that the observations follow a deterministic trend. For example, including a linear trend variable in a regression equation for the (decreasing) number of fatalities implies that the dependent variable will decrease at a fixed rate. A quadratic trend allows the number of fatalities to decrease first and increase again after some time, or vice versa. If a deterministic trend is not a realistic assumption or leads to autocorrelated residuals, models with a stochastic trend may be used. In this case, it is assumed that the trend itself is affected by random fluctuations. This may be a more natural assumption, although studies have shown that some evolutions in road safety are quite well estimated as exponential or logistic trend curves, as will be shown later in this text. Another statistical distinction between models is based on whether the road safety outcomes are modelled in a univariate or multivariate way. In a univariate model, only one road safety outcome (or dependent variable) is modelled at a time. In a multivariate model, a vector of dependent series is considered, which makes it possible to account for correlations between the series.

It is clear that organising the existing models in separate categories will not be an easy task. In TABLE 2, the various approaches from a statistical point of view are shown, together with the common techniques used to estimate these models. In the literature, examples of most of these techniques applied to road safety problems can be found. Note that this schema is only valid for models based on time series data.

TABLE 2: Typology of road safety models

	Descriptive	**Explanatory**
Deterministic	(Non-) linear regression	(Non-) linear regression G(N)LM SUR models
Stochastic	Polynomial splines Moving averages AR(I)MA models VAR models (G)ARCH models Structural models	AR(I)MAX models VARMAX models Transfer function models Intervention analysis Structural models

In TABLE 2, the classical regression techniques are categorised as deterministic, descriptive or explanatory. When a regression model includes only time as an independent variable (i.e. a deterministic trend), it is considered as a descriptive model. It becomes an explanatory model if more explanatory variables are included. The Generalized (Non-)Linear Models (or G(N)LM) extend the classical regression models to models with error terms that follow a distribution from the exponential family. The Seemingly Unrelated Regression (SUR) model is in fact the multivariate extension of the classical regression model. Polynomial splines and moving average techniques are considered as stochastic models, although no real parametric or distributional assumptions should be made. The Autoregressive (Integrated) Moving Average models (or AR(I)MA models), and their multivariate extensions (VAR models) are based on the famous methodology introduced by Box and Jenkins (1976). When explanatory variables are introduced in this model, they are called respectively AR(I)MAX and VARMAX models. Also the transfer function models and the first models for intervention analysis are based on this framework. The structural models, introduced by Harvey (1989), can be used both as a descriptive and an explanatory approach. Even as a descriptive model, the structural models show the underlying components of a series which is indeed, in itself, very descriptive.

In the past, many authors put in a great effort to structure the field of macro models in road safety research and to provide a framework to highlight the main differences and similarities between the various models. The eminent paper of Hakim et al. (1991), published in Accident Analysis and Prevention (AAP), offers an instructive review of macro models for road accidents. In the same issue of AAP (Haight, 1991), various contributions in this field were published, among them (Broughton, 1991; Fridstrøm & Ingebrigtsen, 1991; Oppe, 1991). Another basic document that offers a historical review of both aggregated and disaggregated models for traffic and safety developments is written by the COST 329 research group of the European Commission (COST 329, 2004). In addition to this, literature reviews can be found in (Christens, 2003) and in (Scuffham, 2001). An overview of a special class of models, based on the DRAG model that was developed by Marc Gaudry (Gaudry, 1984, revised 2002) is given in (Gaudry & Lassarre, 2000) and, in the larger context of macro models, in (Van den Bossche & Wets, 2003a).

2.4 From deterministic to stochastic models

The typology presented above illustrates the high diversity in models and their main distinguishing factors. The division into deterministic and stochastic models is in the first place based on the statistical properties of the underlying model, but at the same time largely corresponds to the historical developments of these models. While the first examples of macroscopic road safety models were deterministic in nature, a shift can be observed towards stochastic models in the course of time. This section presents an overview of the milestones in the development of deterministic and stochastic models.

2.4.1 Deterministic models

The first models that were developed to investigate the evolution in road safety over time were deterministic in nature. The most famous deterministic, univariate, descriptive time series models in road safety were introduced by Oppe (Oppe, 1991) and developed on yearly road safety data for various countries. These are indirect models, which can be used for long-term predictions. These models were extended in various ways. For example, (Cameron, 1997) introduced some (explanatory) intervention variables to explain major changes in the level of road safety.

Regression analysis has been commonly applied to relate the number of accidents or fatalities to the level of exposure and to other explanatory factors. According to (COST 329, 2004), the first application of classical linear regression techniques to describe road safety developments can be found in (Recht, 1965). Several authors argued afterwards that the underlying assumptions of the classical linear regression model are quite easily violated when applied to road accident data. Given the specific nature of road accident data, arguments can be found to advance the proposition that accident counts follow a probability law. Indeed, accidents are random and unpredictable as a unit. Moreover, each single accident is, by definition, unpredictable.

As accident and fatality counts are discrete and non-negative numbers, the Poisson distribution appears suitable. The Poisson regression model belongs to the family of generalized linear models (GLM), in which the residuals are assumed to follow a distribution from the exponential family. The classical regression model with normal errors belongs to the GLM class of models, but also models with Poisson, Binomial or Gamma distributed errors are possible. Very popular is the log-linear model, which has been applied frequently in road safety research (Christens, 2003; Fridstrøm & Ingebrigtsen, 1991; Greibe, 1999; Kulmala, 1995; Michener & Tighe, 1992). However, as stated in (Jovanis & Chang, 1986), when a long time period and a large study section are considered, it may be reasonable to approximate the occurrence of accidents by a normal distribution. For large means, the Poisson distribution converges to the normal. In this case, there is only a very small probability of zero accidents in a time interval, and the large mean together with a relatively small variance will make zero or negative values unlikely. Applications can be found in (Abbas, 2004; Joshua & Garber, 1990; Peltzman, 1975; Zlatoper, 1984). In (Peltzman, 1975), a traffic safety analysis is performed in a multiple regression, using yearly data. Time is included as an explanatory variable. For all variables except time, logarithmic values are included. That is, the model assumes that risk changes exponentially over time, which is the same assumption as in (Oppe, 1991). In (Abbas, 2004), the linear, power, logarithmic, exponential and quadratic polynomial functional form were used to predict the expected number of accidents, injuries and fatalities. In his conclusion, he warns the reader that making forecasts entails extrapolating outside the range of real observations, and that the models can be used for short term forecasts of 1-3 years.

Although the normal distribution should not be a problem when the mean of the dependent variable is large, still, working with log-linear models is in a sense a natural choice (Fridstrøm et al., 1995), as it makes sure that the expected outcome, be it the number of accidents or fatalities, is always a positive number. Moreover, an additive model on the log-scale corresponds to a multiplicative model on the original scale, implying that a road safety outcome is seen as the product of an exposure measure and an indication of the risk.

2.4.2 Stochastic models

While most of the recent deterministic models for time series data explicitly take time into account by including a trend variable, they do not correct for the presence of serially correlated observations. In order to produce unbiased parameter estimates of the residual variance and, consequently, meaningful confidence intervals for the regression coefficients, models that explicitly treat time dependencies are needed. Two main modelling approaches are used in this respect. The first is the ARIMA approach, familiarised by Box & Jenkins (1976). The second approach uses unobserved component models, introduced by Harvey (1989).

In ARIMA models, autocorrelation is explicitly treated by including autoregressive and moving average components. These components act as a filter, which captures all relevant information in the series and results in a white noise process. The modelling procedure requires stationary series, which means that, in essence, trends and seasonal components should be removed before applying the filter. This is mostly done by differencing the series. Once the model is specified, the differencing operation is reversed, resulting in predicted values for the original series. The model can be easily extended with explanatory variables, leading to a model in which regression and time series properties are combined. As in any other research domain, the ARIMA methodology is widely used in road safety research. Atkins (1979) used Box-Jenkins time series analysis and intervention analysis to determine the influence of compulsory car insurance, company strikes and a change in the policies of insurance companies on the number of traffic accidents on freeways in British Columbia. In (Wagenaar, 1984), the relationship between changes in economic conditions and motor vehicle crash involvement, thereby taking into account the influence of exposure, was identified using ARIMA and dynamic regression time series modelling procedures. Scott (1986) analysed monthly series of British road accident data with regression models for two-vehicle accidents and ARIMA models for single vehicle accidents. According to Scott, because of the weak correlation in the residuals, the ARIMA models did not appreciably better represent the series compared to the classical regression. Examples of ARIMA models for injury accidents in Spain can be found in (COST 329, 2004). In (Bergel, 1992), ARIMA models with explanatory variables were presented for France. Raeside & White (2004) constructed pure ARIMA models for monthly time series data related to road traffic and pedestrian casualties and fatalities in Great Britain. In (Lassarre, 1986), ARIMA models are estimated on monthly numbers of accidents and fatalities in order to evaluate the effects of introduced speed limits and compulsory seat belt wearing. Ledolter & Chan (1996) investigated the effect of a change in the speed limit on the rural interstate highway system to 65 miles per hour in Iowa. They developed a log-linear model with deterministic seasonal terms for quarterly accident data and a first-order autoregressive model for the noise term. More recently, Christens (2003) presented results of ARIMA models that were developed for two accident series, related to changes in traffic, the number of young people and speed. Regression models with an ARIMA structure on the error term applied to Belgian data can be found in (Van den Bossche et al., 2004; Van den Bossche et al., 2005a). Also, the DRAG family of models (Gaudry & Lassarre, 2000) is essentially made up of regression equations with an autoregressive structure on the error term, which can, simultaneously, be corrected for heteroskedasticity. As will be explained further, the DRAG models are explanatory models, including a relatively large number of explanatory variables, whose effects on the exposure, the frequency and the severity of accidents are estimated by econometric methods.

Instead of filtering the trend and seasonal component, as is done in ARIMA models, state space time series models consider all distinct components of a series as dynamic processes. Descriptive and explanatory models that are based on the relationship between fatalities, exposure and risk were developed by (Lassarre, 2001) and more recently by (Bijleveld & Commandeur, 2004; Bijleveld, Commandeur, Gould et al., 2005). The latter authors use multivariate state space models to estimate simultaneously the level of exposure and the number of accidents and/or fatalities. The first explanatory structural time series application in road safety research can be found in (Harvey & Durbin, 1986). In their study, monthly data on road casualties in Great Britain are analysed to assess the effects of the seat belt law. Johansson (1996) tested the effect of a lowered speed limit on the number of accidents on Swedish motorways, using extended Poisson and Negative Binomial count data models. He incorporated a large number of explanatory factors in a structural time series model. Scuffham developed structural time series models to investigate the changes in the trends and seasonal patterns of fatal crashes in New Zealand in relation to changes in economic conditions (Scuffham, 2003; Scuffham & Langley, 2002). Scuffham included a relatively large set of explanatory variables, among them the unemployment rate, the percentage young males, the volume of beer consumption and several road safety laws. Similar models were developed by Christens (2003). In a first model, he tested the significance of socio-economic factors on three accident series. In another, a state space intervention analysis was used to evaluate the Danish automatic mobile speed camera experiment. For the Belgian data, a comparison was made between ARIMA and state space models to investigate the frequency and severity of accidents (Hermans et al., 2006a, 2006b). For a given set of explanatory variables, very similar results were obtained.

2.5 Descriptive and explanatory models

In the previous section, an overview of both deterministic and stochastic models was given. Looking at the references, it is clear that the more advanced models are also the most recent ones. One can certainly speak of a kind of evolution from deterministic to stochastic models. When looking at descriptive versus explanatory models, it is difficult to find a similar evolution. Both kinds of models have been developed over the last decades, and the choice between them was inspired by the objectives of the study. Therefore, to this very day, descriptive and explanatory models are developed to increase the insight in road safety evolutions and to provide forecasts. This section provides an overview of the main streams of research in the field of both descriptive and explanatory models, without the ambition of being exhaustive. Some of these models were already mentioned in the previous section to highlight the statistical differences. The discussion here is focussed on the families of models that will be developed further in this text for the Belgian road safety situation.

2.5.1 Descriptive road safety models

The descriptive modelling efforts to predict the number of fatalities that can be found in the literature are usually based on one of two theoretical schemas (COST 329, 2004). One approach is based on economic growth theory, the second on learning theory. Both approaches are similar in the sense that no (or at most a very small set of) explanatory variables are used and that nationwide developments in road safety are modelled. They differ in the sense that the second approach is derived from a priori argumentation from the macroscopic road safety theory. Also, the second approach is much more recent than the first, and therefore is more in line with the current thinking about road safety.

The first approach is based on economic growth theory, and started in 1949 with an equation proposed by Smeed (1949). According to Smeed, the number of fatalities per inhabitant is related to car ownership and population. Using F as the number of fatalities, V as the number of registered vehicles and P as the population level, "Smeed's Law" is formulated as:

$$F = 0.0003\sqrt[3]{VP^2} = 0.0003V^{1/3}P^{2/3} \tag{3}$$

This equation can be seen as a Cobb-Douglass production function, where the number of fatalities is "produced" by the number of vehicles and the population, with constant returns to scale. In this line of reasoning, other studies followed (Peltzman, 1975; Zlatoper, 1989). However, this law has been the target of criticism for many years. First, the time window of analysis changed. Compared to the period of analysis in Smeed's work, the level of exposure significantly increased and significant progress in safety performance has been made over the years. However, this argument was refuted in (Adams, 1987). According to Adams, Smeed's formula is still valid for more recent observations, and it is considered as a useful generalisation of the relationship between death rates and exposure. On the other hand, it is clear that the decrease in fatalities at the beginning of the seventies is contrary to the expectations of Smeed's formula. Haight (1984) explained this decrease in the number of fatalities as the effect of a saturation in the road traffic rather than an increase in the level of road safety. Also, the Smeed model has no indicator of the progress in safety performance in terms of road engineering, vehicle construction, driver training, traffic laws, etc. Lassarre (2001) proposed an extension of the framework, taking safety progresses into account. A second source of criticism is the fact that the formula is not derived from a priori reasoning (Oppe, 1991). The model is the outcome of a completely data-driven approach. As a consequence, it is hard to give an interpretation to the results. This also raises the question of the general applicability of the formula. It has, in the end, been accepted that Smeed's formula cannot be considered as a "law" (Broughton, 1991; Oppe, 1991).

A second approach to predicting the annual number of fatalities is based on learning theory. According to (COST 329, 2004), this approach was introduced by Minter (1987). He modelled the evolution of the number of fatalities by two kinds of learning curves. Later, Oppe (1991) explained the evolution in fatality rates over time as the result of a collective learning process that is determined by the evolution in exposure and risk. His reasoning is based on a theory that assumes that societies learn to control the unsafety of the road transport system in an exponential way. In the Oppe approach, it is assumed that exposure and traffic safety can be described as growth processes in time. Exposure is seen as the result of a production system, derived from economic and social transportation demands. It is assumed, however, that the development of exposure has an upper level, indicating the physical boundaries of the road system. Because of practical reasons, there must be a sort of "maximum capacity" for the traffic system. These assumptions correspond to an S-shaped curve, like the logistic model. For the fatality rates, a (negative) exponential trend is assumed. This may be seen as a collective learning process (COST 329, 2004), caused by the ever-increasing knowledge of the traffic safety problem and the constant improvement of the safety performance of the road transport system. In comparison with some decades ago, cars and roads are better equipped, traffic safety education has improved and legislation and enforcement have increased. Learning of individual road users results from this community learning process. The number of fatalities is then, by definition, equal to the product of the risk and exposure curves.

The "Oppe model" has been applied for many countries, Poland being one of the most recent examples (Oppe, 2001). Also, the classical framework proposed by Oppe has been extended by many authors. In (Cameron, 1997), the risk model is enriched with dummy variables to account for the introduction of major road safety measures. Their "Modified Oppe" model fits the data substantially better than the original one. Broughton (1991) proposed a similar model, in which the introduction of the compulsory seat belt wearing (in 1983) and the impact of drink-drive legislation (in 1968) were tested. Bijleveld and Oppe (1996) propose some functional form extensions, allowing for more flexible relationships between fatalities and risk. In (Commandeur, 2002; Commandeur & Koornstra, 2001), the analysis of aggregated road safety data is based on the same framework of exposure and risk, but instead of the logistic function the asymmetric S-shaped Gompertz curve is used.

All models that are inspired by the Oppe approach have the deterministic curve fitting procedure in common. That is, these models assume underlying functional relationships for fatalities, exposure and risk. Instead of using a deterministic trend, Lassarre (2001) introduced a stochastic trend by means of structural models. With this approach, other patterns of evolution than an exponential or logistic one are captured (COST 329, 2004). The measure of exposure is treated as an explanatory variable in the fatalities equation, and can be tested on its proportionality in relation with the number of fatalities. Although the models developed by Lassarre are more flexible in nature than the Oppe models, it is still based on the assumption that exposure can be considered as a deterministic and faultlessly observed variable. As this is not the case in reality, it makes sense to model the level of exposure as a second dependent variable. In recent work (Bijleveld, Commandeur, Gould et al., 2005; Bijleveld, Commandeur, Koopman et al., 2005), a multivariate framework is used to measure accident risk and exposure simultaneously. Both risk and exposure are assumed to be unobserved, which explicitly recognises the fact that these numbers are never measured without error. This approach is called the multivariate unobserved components time series framework, and models are indicated as LRT (Latent Risk Time series models), as they focus on the developments in a latent risk variable.

2.5.2 Explanatory factors in aggregated road safety models

In (OECD, 1997b), some broad categories of factors influencing traffic accident counts are listed. First, the number of accidents depends on some autonomous factors that cannot be influenced on a short-term and countrywide level (weather and state of technology belong to this category). Second, economic conditions like unemployment and income are part of the general climate in which accidents occur. Although these issues are sometimes subjected to political intervention, they are rarely oriented towards road safety improvement. A third category covers the size and the structure of the transportation sector, which is often closely related to exposure (infrastructure, vehicle park...). Fourth, the accident countermeasures, formalized in laws and regulations, are explicitly brought into being to reduce the risk of road accidents. Fifth, the accident counts also depend on the data collection system. Changes in collection strategies may produce fictitious increases or decreases in accident counts. A last influence is the random variation in accident counts. Since accidents are, by definition, unwanted events, they cannot be fully predicted. Therefore, part of this phenomenon will always be inexplicable.

Although it is intuitively appealing to assume that these factors have an influence on the number of accidents, it would be instructive to get a confirmation of this influence. Given the large number of possible factors, it is not easy to get a clear view on the reasons for the trends in traffic safety. Because of the randomness involved in accident occurrence, the investigation of influential factors should be stochastic in nature. Econometric explanatory models provide a means to test the impact of influential factors. The factors summarized above can be combined in an explanatory model and tested for their (positive or negative) contribution to traffic safety. This makes the models quite appealing to practitioners, who are typically interested in actively increasing the level of traffic safety.

In (OECD, 1997b), the importance of this kind of models in road safety has been extensively described. The wide arsenal of econometric modelling techniques can be very effective in taking into account various influences on aggregate accident figures. This approach is especially useful when many factors are to be tested. Moreover, since accidents are unwanted events, controlled (or "designed") experiments cannot be used. Accidents are, by definition, non-experimental. Because of the random character, a probabilistic view on the accident process is quite natural.

2.5.3 A class of explanatory road safety models: the DRAG family

Explanatory road safety models mainly show similarities and differences concerning the statistical assumptions that are made. From the application point of view, clearly there is no unifying framework that keeps all these models together. It seems that possible close likeness between models is based on pure coincidence. However, one exception that deserves to be mentioned in this respect is the DRAG approach, developed by Marc Gaudry (Gaudry, 1984, revised 2002; Gaudry et al., 2000) and followed by many other road safety researchers. The DRAG framework offers a structured, yet very specific methodology towards the development of explanatory models.

In accordance with the classical decomposition of road safety in terms of exposure, risk and consequences, that was presented in Chapter 1, the DRAG models start from a layered structure in which each of these components can be regressed on a set of relevant explanatory variables. DRAG is a family of models that explain the **D**emand for **R**oad use, **A**ccidents and their **G**ravity. The DRAG model constitutes a very ambitious attempt to explain the development of aggregate exposure, accidents and their severity over time. DRAG models have a well-defined structure, use flexible form regression analysis and are calibrated with monthly time series data defined over a country or region.

In the DRAG models, the number of victims is decomposed into three elements, namely exposure, frequency and severity, which themselves become the objects to be explained. This means that an explanation of the number of victims is effectively derived from the separate explanation of the three terms of the identity, as in:

$$\text{VICTIMS} \leftarrow \begin{bmatrix} \text{Demand for road use (DR)} & \leftarrow & [\text{- -}, X_1] \\ \text{Accident Frequency (A)} & \leftarrow & [\text{DR}, X_2] \\ \text{Accident Gravity (G)} & \leftarrow & [\text{DR}, X_3] \end{bmatrix} \begin{matrix} \text{Exposure} \\ \text{Frequency} \\ \text{Severity} \end{matrix} \quad (4)$$

Here, X_1, X_2, and X_3 are three (possibly different) sets of explanatory variables. This (at least) three-layer recursive structure of explanation, involving road use, accident frequency and severity, is a common feature of all members of the DRAG family (COST 329, 2004). In contrast to many other models, road use is not considered as exogenous, but is explained by a set of variables. Moreover, each dimension may be further split into various sub-categories: type of road use for ***DR***, (gasoline or diesel), category of accident for ***A*** (fatal, injury, etc.) and a measure of severity (mortality, morbidity) for ***G***. Categories of explanatory variables include prices, vehicle availability and characteristics, network characteristics (legal regimes, modal mix, weather, etc.), consumer characteristics and activity levels or trip purposes (employment, shopping, etc.).

A structure as defined in equation 4 makes it possible to search for evidence of risk substitution among exposure, frequency and severity risk dimensions. For instance, if snow is included as an explanatory factor in the three groups of explanatory variables X_1, X_2, and X_3, it might lead to less driving (***DR*** decreases) and, at the reduced exposure level, to more accidents (***A*** increases) that are less severe (***G*** decreases): the net impact on the number of road victims results from the relative strength of these potentially offsetting effects.

Beside the very large number of explanatory variable taken into account, another distinguishing feature of the DRAG model is its use of Box-Cox transformations to relax the linearity assumption usually embedded in a regression model. These transformations are given by:

$$x^{(\psi)} = \begin{cases} \dfrac{x^{\psi} - 1}{\psi} & \text{for } \psi \neq 0 \\ \ln(x) & \text{for } \psi = 0 \end{cases} \quad (5)$$

The parameter ψ is called the Box-Cox parameter. Different values of ψ correspond to different curvatures or functional forms for the relation between the dependent and independent variables. For example, the Box-Cox transformation includes the cubic ($\psi = 3$), quadratic ($\psi = 2$), linear ($\psi = 1$), square root ($\psi = 0.5$), logarithmic ($\psi = 0$), and reciprocal ($\psi = -1$) functional forms as special cases. In general, the Box-Cox regression model takes the following form:

$$y_t^{(\omega)} = \sum_{j=1}^{J} \beta_j x_{j,t}^{(\psi_j)} + u_t \tag{6}$$

All parameters (ω, ψ_j and β_j, j=1,2,..., J) for a given equation are estimated simultaneously. Thus, the data are allowed to determine, apart from the regression coefficients, the optimal functional form (within the Box-Cox family of monotonic functions). A log-linear model can be viewed as a special case of the Box-Cox regression model, in which one sets $\omega = 0$ and all ψ_j pre-specified and (by assumption) known.

The BC-GAUHESEQ algorithm, developed by Liem, Gaudry, Dagenais & Blum (2000) further allows general heteroskedasticity and autocorrelation structures, that are defined on the residual term u_t:

$$\begin{aligned} u_t &= \left[\exp\left(\delta_0 + \sum_{m=1}^{M} \delta_m z_{m,t}^{(\lambda_{z_m})} \right) \right]^{-\frac{1}{2}} v_t \\ v_t &= \sum_{l=1}^{L} \rho_l v_{t-l} + w_t \end{aligned} \tag{7}$$

In this expression, the $z_{m,t}$ are heteroskedasticity factors and v_t represent homoskedastic yet possibly autocorrelated residuals. Therefore, an autoregressive structure can be imposed for v_t, such that w_t is white noise (that is, independent and normally distributed homoskedastic residuals).

For the explanatory variables in the DRAG models, typically elasticities are calculated. The elasticity of an explanatory variable measures the percentage change in the dependent variable caused by a one percent change in the independent variable of interest. The elasticities offer the advantage that they are not measured in any particular unit and therefore have a clear and univocal interpretation. Assuming no heteroskedasticity variables $z_{m,t}$, the elasticity can be written as:

$$\eta_{x_{j,t}} = \frac{\partial y_t}{\partial x_{j,t}} \frac{x_{j,t}}{y_t} = \beta_j \frac{x_{j,t}^{\psi_j}}{y_t^{\omega}} \tag{8}$$

Several DRAG-type models have been developed, all based on the same framework of a multi-layered structure, including several explanatory variables, with Box-Cox transformations for both the dependent and the independent variables and autocorrelation and heteroskedasticity corrections for the residuals. The models developed up to now are (Gaudry & Lassarre, 2000):

1. DRAG (Demande Routière, les Accidents et leur Gravité), covering the state of Québec, Canada (Gaudry, 1984, revised 2002) and further developed as the DRAG-2 model (Gaudry et al., 1995).
2. SNUS (Straßenverkehrs-Nachfrage, Unfälle und ihre Schwere), authored by Gaudry and Blum (Gaudry & Blum, 1993), covering Germany.
3. DRAG-Stockholm (Demand for Road use, Accidents and their Gravity in Stockholm), authored by Tegnér and Loncar-Lucassi (Tegnér & Loncar Lucassi, 1997), covering the Stockholm county of Sweden.
4. TAG (Transports routiers, Accidents et Gravité), authored by Jaeger and Lassarre (Jaeger, 1999; Jaeger & Lassarre, 1997), covering France.
5. TRULS (TRafikk, ULykker og deres Skadegrad), by Fridstrøm (Fridstrøm, 1997, 1999), covering Norway .
6. TRACS-CA (Traffic Risk And Crash Severity - CAlifornia), authored by McCarthy and covering California (McCarthy, 2000).

2.5.4 Review of selected empirical results

The number of models in which the effects of explanatory variables are tested increased over the last 20 years. Reviews of empirical results of these models can be found in (Gaudry & Lassarre, 2000; Hakim et al., 1991; Scuffham, 2001). In this section, some of these results are summarized. The overview is directed towards the variables that will also be tested in the subsequent chapters of this work, and is by no means a complete overview of the existing models. Note that, along with the effects of explanatory variables on accidents and victims, also effects on exposure are reported.

2.5.4.1 Economic factors

Economic factors appear to be related to the number of accidents, the number of injuries and to the distance travelled. As in many social sciences, there is no underlying economic theory that indicates which explanatory variables should be considered in a model for road accidents (Scuffham, 2003). Unemployment rate is often used as a proxy for economic conditions. Other possible economic indices are disposable income, automobile production, gross national product, manufacturing, consumption per capita, retail sales and interest rates (Hakim et al., 1991). According to Scott (1986), economic indicators only have an indirect influence on accident data, via changes of the characteristics of the traffic and road environment. Other authors (Hakim et al., 1991; Scuffham, 2003) state that economic factors can have both an indirect and a direct effect on the number of fatalities and road crashes.

In many models, the effect of unemployment on distance travelled and road safety is analysed, often resulting in a coefficient with a negative sign. As mentioned in (Hakim et al., 1991), Hoxie & Skinner (1985) found that, in periods of recession (with high unemployment rates), the number of fatalities with young drivers is significantly lower, even without a significant reduction in miles driven. This is explained by the fact that young drivers drive less in recessionary times (thereby reducing the number of fatalities), but their impact on total mileage is to be neglected. In the DRAG-2 model (Fournier & Simard, 2000), the ratio of the number of unemployed to the number of people with driver's licences is included. An increase in this ratio results in a decrease in the distance travelled and in the number of accidents. Scuffham (2003) tested lagged effects of unemployment and found that an increase in the unemployment rate lead to a reduction in fatal crashes in the following period, whereas the level of the unemployment rate in the current period was not significant. According to (Blum & Gaudry, 2000), unemployment has a small and highly significant negative effect on road use, but only a moderate effect on casualties, probably because of institutional regulations. In (Tegnér et al., 2000), it is shown that an increase in the number of employed persons results in larger distances travelled because of higher income, higher rate of car ownership and consequently more private consumption and leisure activities. In (Jaeger & Lassarre, 2000), the increase in traffic risk caused by a rise in unemployment is negligible. This result is not in accordance with most of the other models, and may indicate differences in the variables measured and in social protection systems.

A second key economic variable is income. This is related to the level of unemployment, in the sense that a higher unemployment rate may affect the traveller's ability to pay for travel (Hakim et al., 1991). This income elasticity of demand for travel is low for work trips, but is quite high for young drivers and recreational trips (Hoxie et al., 1984; Partyka, 1984; Peltzman, 1975; Wagenaar, 1984). Peltzman (1975) uses a demand-oriented explanation for the negative relationship between economic growth and the number of road accidents. As income rises, the demand for safer cars increases, which leads to fewer accidents. Also, a higher income on the supply side increases the budget for investment in infrastructure and road maintenance. However, in (Blum & Gaudry, 2000), where household income is used as an economic indicator, a rise in the income results in an increased vehicle ownership, which in turn increases road use demand and the number of accidents. In (Johansson, 1996), disposable income hardly contributes to the explanation of the number of casualties. Further, it has no significant effect on the number of fatal and minor injury accidents and a positive significant effect on the number of severe injury accidents. For the vehicle damage accidents, the significance depends on the statistical model structure.

Apart from unemployment and income, some authors use GDP as an economic indicator. Christens (2003) tested the effect of Gross National Product (GNP) and found that an increasing GNP is associated with additional fatal accidents. On the other hand, Scuffham (2003) reported a very strong negative effect (a reduction in risk) of increases in Gross Domestic Product (GDP) on the number of fatal crashes.

2.5.4.2 Gasoline prices

The general idea behind the effect of gasoline prices on road safety, as formulated in (Harvey & Durbin, 1986), is that higher prices may induce drivers to drive more slowly, which is then expected to reduce accident rates. The authors found that a 1% rise in the petrol price leads to a 0.31% reduction in casualties. Hoxie et al. (1984) reported an inverse relationship between gasoline prices and accidents, as well as between gasoline prices and the number of non-work-related trips. Scott (1986) also concluded that petrol prices are strongly (negatively) related to many accident series, except for those where two-wheeled vehicles were involved. Similar results are obtained in (Fournier & Simard, 2000). In (Johansson, 1996), a petrol price index has no effect on the different types of accidents considered. According to Jaeger & Lassarre (2000) and Tegnér et al.

(2000), the number of kilometres driven decreases with a rise in gasoline prices. McCarthy (2000), however, did not find any impact of the real gasoline price on the demand for travel. This is explained by relating the gasoline price to the opportunity cost of travel, which is a generalised cost, including monetary and time costs. In the same model, the number of accidents decreases with a rise in the gasoline price. Note that the effect of changes in gasoline prices on accidents is not direct. The price of gasoline determines its demand, which in turn affects the number of accidents.

2.5.4.3 Young drivers

Young drivers are considered as a high-risk group, having a higher probability of involvement in car accidents with injuries. In (Fournier & Simard, 2000), an increase in the number of young drivers, between 16 and 24 years old, results in a rise in the number of road accidents. Quite often, the topic of young drivers has been related to the effect of the minimum legal drinking age on accidents. Wagenaar (1983) stated that approximately 20% of all alcohol-related crashes involving young drivers can be prevented by removing legal access to alcoholic beverages. Hoxie & Skinner (1987b) showed that raising the minimum drinking age can save a considerable amount of young lives. Scuffham (2003) found that the proportion of young males (aged 15-24 years) is directly related to the number of fatal crashes, with a positive sign.

2.5.4.4 Weather

Many road safety models considered the effect of weather conditions on the number of accidents and fatalities. As the weather is related to the geographic properties of the area of concern, it is expected that the results will vary among the models. Results may also vary according to the time period considered. A climatologic variable studied on a daily level may provide completely different insights than when studied on a monthly level. This is clearly illustrated in (Eisenberg, 2004), where the relationship between precipitation and accidents is investigated on a daily and a monthly basis. On a monthly basis, a negative relationship is found between precipitation and accidents, while on a daily level a strong positive effect is estimated. However, it is still instructive to study weather effects on road safety over time, as part of a larger set of influential variables. The most common weather variables are rainfall and temperature, although snow and frost are also considered in some models.

In (Scott, 1986), higher rainfall and cold temperatures are found to be associated with more accidents. In the original DRAG model (Gaudry, 1984, revised 2002), severe weather conditions reduce the number of deaths and have mixed effects on the number of injuries. Cold weather increases the number of material accidents, and reduces considerably the number of fatal accidents. According to the DRAG-2 model, developed for Quebec (Gaudry et al., 1995), an increase in temperature would result in a moderate increase in distance travelled. Higher temperatures also significantly reduce the number of accidents with property damage and the number of victims injured. On the other hand, the number of fatal accidents and the number of victims killed increase with temperature. Also a moderate decline is noticed in the number of bodily injured accidents and the morbidity and mortality rates. In the SNUS-2.5 model for Germany (Blum & Gaudry, 2000), an increase in temperature implies a strong decrease in material damage accidents and an increase in the number of bodily injury accidents and in morbidity and mortality rates. It also strongly increases the road demand. There are strong effects of sunshine on the frequency and severity of accidents, and the presence of rain has larger proportionate impacts than the amount of rain. In California (McCarthy, 2000), average rainfall significantly affects the frequency and severity of crashes, but the direction of the effect is negative for fatal crashes and mortality and positive for non-fatal crashes and morbidity.

In Norway, injury accidents become less frequent when the ground is covered by snow (Fridstrøm, 2000). The risk reduction is larger the deeper the snow is. This "snowdrift effect" reduces the frequency of single vehicle injury accidents, but increases the risk of head-on collisions. The number of injury accidents goes up during days with snowfall, but at the same time the severity is reduced. This is a "risk compensation effect". The monthly number of days with temperatures dropping below zero has a favourable effect on the number of accidents, especially on the most severe injuries. This effect is much stronger for bicyclists and motorcyclists than for pedestrians, suggesting a reduction in the two-wheeler exposure. Similarly, higher rainfall reduces injury accident counts because of the reduced exposure among unprotected road users.

Results on Belgian data (Van den Bossche et al., 2004) show that precipitation increases the number of light injuries and the corresponding number of accidents, but does not affect the serious injuries or fatalities. On the other hand, the percentage number of rainy days in a month has a similar influence on all dependent variables, resulting in an increase in accidents and victims. Further, a higher number of days with thunderstorm significantly increases the light injury outcomes. More days with frost and less sunny hours in a month decrease all road safety outcomes. The percentage number of days with snow or sunlight was not significant.

2.5.4.5 Legislation

Most researchers are interested in analysing the possible long-term effects of legislative intervention measures while statistically controlling for non-intervention variables hypothesized to be associated with accidents. Sometimes, however, one is interested in the effectiveness of a particular intervention. This is done in partial (before-after or treatment-control) studies. Some of the possible intervention variables are discussed below.

<u>Speed limits and speed variance</u>

Several authors, for example Partyka (1984), showed that reducing the speed limits appears to be related to a reduction in fatalities. Also the severity of injuries is positively related to the allowed speed. Blum & Gaudry (2000) obtained similar results in their model for Germany. According to McCarthy (2000), however, increased speed limits slightly reduce risk exposure. In his model for California, higher speeds have no effects on fatal crashes, but a strong positive impact on the frequency of non-fatal injury crashes. Keeping everything else constant, there were fewer fatalities per fatal crash and fewer injuries per non-fatal injury crash after the increased speed limits law. But if a crash occurs, a higher speed results in more serious injuries. According to Christens (2003), a reduction of the urban speed limit is associated with a decrease in the total number of fatal accidents. In (Scuffham, 2003), it was found that a 1% increase in the open road speed limit was associated with a 1.2% increase in fatal crashes.

Lave (1985) investigated the effect of speed variance rather than speed limit on highway fatality rate. According to Lave, there is a strong statistical relationship between fatality rate and speed variance. However, Levy & Asch (1989) concluded that efforts should be directed to slowing down high-speed drivers rather than speeding up slower drivers. The empirical studies of Fowles & Loeb (1989) and Snyder (1989) reject the importance of speed variance; they report on a positive and statistically significant effect of speed on fatalities. Later, Jaeger & Lassarre (2000) consider speed as a description of the incidence of traffic risk of an individual's behaviour in terms of control over the vehicle. They conclude that speed limits have significant effects, but these effects are due more to a reduction in the dispersion of speeds than to a drop in average speed.

Drinking and driving

Several authors (Hoxie & Skinner, 1985; Loeb, 1987; Zlatoper, 1984) found that the consumption of alcohol is positively related to accidents with fatalities. Blum & Gaudry (2000) state that beer consumption is both a social variable (thereby increasing road demand) and a factor changing the frequency and severity mix of accidents. Increased beer consumption increases the number of fatalities, but decreases the frequency of accidents. Most of those that drink and drive, have only consumed little and compensate to prevent accidents, while those who drink a lot may increase their risk. A similar conclusion for wine consumption is obtained by (Jaeger & Lassarre, 2000). An increase in wine consumption per adult implies, ceteris paribus, an increase in the demand for road use for recreational purposes. Also the number of injuries and accidents increases with higher wine consumption. In his model for Norway, Fridstrøm (2000) found positive relationships between alcohol consumption and accidents of every degree of severity, thereby concluding that the restrictive Norwegian alcohol policy has prevented road accidents and fatalities. Tegnér et al. (2000) found that low consumption levels of alcohol seem to reduce the number of light and severe injuries. At higher levels of consumption, the accident risk augments rapidly. Also McCarthy (2000), investigated the effect of beer, distilled spirits and wine consumption on traffic safety in his TRACS-CA model for California. Only the beer consumption increased the demand for road use. A rise in consumption of distilled spirits per capita leads to more fatal crashes and more fatalities per fatal crash. The wine consumption per capita increases the frequency of non-fatal injury and material only crashes.

Although it is clear that excessive alcohol consumption increases accident occurrence and the risk of being killed or injured, Gaudry (2000) tested the possibility of a J-shaped relation between alcohol consumption and road safety. That is, the risk may at first be reduced by a small consumption of alcohol (because of risk compensation or reduced aggressiveness), but will strongly increase with higher blood alcohol concentration. In this case, the effect of higher alcohol consumption will depend on its distribution among drivers (Gaudry, 1993).

In several studies (Cook & Tauchen, 1984; Loeb, 1987), a negative relationship was found between both the number of accidents and the severity of injuries and the minimum age for purchasing alcohol. Hoxie et al. (1984) used the minimum age for the legal drinking of alcohol as an intervention variable, and found it to be statistically significant in explaining fatalities. Evans (1990) states that eliminating alcohol drinking may significantly reduce traffic fatalities. Also McCarthy (2000) studied the impact of the availability of alcohol. An increase in the number of alcohol licences per month generates an increase in non-fatal injury and materials only crashes.

Mandatory seat belt use laws

Several empirical studies have shown that seatbelt legislation can significantly reduce the number of fatalities and the severity of injuries (Campbell & Campbell, 1986, 1988; Friedland et al., 1987; Harvey & Durbin, 1986; Hoxie & Skinner, 1987a; Rutherford, 1987; Van den Bossche et al., 2004; Van den Bossche et al., 2005a; Williams & Lund, 1988). Friedland et al. (1987) concluded that safety belts provide full protection against fatalities in accidents that occurred at speeds lower than 60 miles per hour. He also found that seatbelt use leads to serious reductions in injuries to drivers and front-seat passengers. Moreover, it reduces the number of brain injuries. Hoxie and Skinner (1987a) have shown that many occupant fatalities can be avoided when seat belt use laws are mandated. Harvey & Durbin (1986) developed an intervention analysis, using structural time series models, for monthly data on road casualties in Great Britain in order to assess the effects of the introduction of the seat belt law on casualty rates. As far as the number of killed and seriously injured road users is concerned, Harvey and Durbin found a reduction of 23% for car drivers and 30% for front seat passengers.

Blum & Gaudry (2000) found that seat belt use reduces both the frequency and the severity of bodily injury accidents, but increases the frequency of material damage accidents. According to the results in (Fridstrøm, 2000), an increase in the number of car drivers not wearing the seat belt will increase the number of car occupant injuries and the number of fatalities. Seat belts seem to be more effective in preventing less severe injuries than in saving lives. Results in (Jaeger & Lassarre, 2000) show that there is a risk compensation effect, since the use of seat belts leads to an increase in speed. This is known as driver behaviour retroaction. In (McCarthy, 2000), seat belt use did not affect the frequency of fatal crashes, but increased the incidence of non-fatal injury and material only crashes. When a fatal crash occurs, more persons are killed, but when an injury crash occurs, fewer individuals suffer an injury.

2.6 Conclusion

This chapter offers an introductory overview of the field of time series road safety models. The application of these models is in line with a strategic view towards road safety management, and is highly useful in relation with the very difficult task of setting realistic road safety targets for the future. Depending on its specific properties, a model can be used to describe, explain, predict or compare road safety developments on an aggregated level. Starting from a simple typology developed by Hakkert & McGann (1996), an overview of the distinctive properties of time series road safety models is given. In particular, models were categorised based on their statistical (deterministic versus stochastic) and application (descriptive versus explanatory) properties.

Although the statistical techniques used in this area of research are quite common, it is clear that the research field of the models described in this chapter is rather unstructured in some respects. First, the choice of a statistical technique is not always straightforward. Many classical regression models were, for some reason, developed on time series data. In the course of time, more and more authors shifted to the models that were specifically developed for time series problems. This aspect, therefore, is clearly related to the evolution in the field.

Second, among the explanatory models, the choice of explanatory variables varies considerably. Whereas some models are specifically oriented towards the analysis of one category of variables (typically economic indicators or law interventions), other models combine a large number of variables representing a variety of categories of influential factors, as is typically the case in the DRAG models. Related to this aspect, it is found that many authors are not clear about the specific contents of the variables used in their models. This, in turn, affects the way in which the results can be interpreted and compared, as was confirmed in the overview of empirical results. This is especially true for the socio-economic variables, and to a smaller extent for the interventions (laws and regulations).

Third, the models differ in the frequency of the observations. Clearly, in a seasonal model (on monthly or quarterly data), different aspects should be considered compared to a model developed on yearly data. Moreover, the effect of many variables is related to the level of aggregation on which it is measured. This is especially true for weather variables, but also the effect of other factors like economic indicators might be concealed in the frequency structure of the data. Therefore, the objective of the model and the frequency of the data are, in a sense, related. Fortunately, it is observed that most of the explanatory models for various countries were developed on the basis of monthly data, while descriptive models typically use yearly data. When predicting road safety is the objective of the model, using explanatory variables can be experienced as a burden, in the sense that this often limits the length of the prediction horizon and that the data needs are much higher.

Examples of attempts to develop models that are based on a unified framework were found in the studies made by (Oppe, 1989) and (Bijleveld, Commandeur, Gould et al., 2005) for the descriptive models and by (Gaudry, 1984, revised 2002) for the explanatory models. These authors put a great effort to structure the time series road safety models. Oppe created a very simple and straightforward conceptual model that can be used to study developments in road risk and exposure. The models developed by Gaudry are, in a sense, based on the same concepts, but are highly demanding in terms of data and model complexity.

It is clear that, due to the complexity of the application, both in terms of data needs and modelling approaches, the field of time series road safety models has not yet attained its full development, and will continue to grow, hopefully in a structured way.

Chapter 3 Exposure in macro models

3.1 Introduction

As a further development of the previous chapter, this section gives attention to the specific role of exposure in macro models. Sometimes, exposure is considered as an endogenous variable, explained by other factors in the model, while in other studies it is seen as an explanatory variable next to socio-demographic variables, economic indicators or weather conditions. It is a driving variable in almost all models, just because of its crucial role in the relation between fatalities and road risk. However, it is, at the same time, one of the most difficult variables to measure and is therefore rarely obtained in a format and aggregation level that is useful for time series modelling. As a result, very different exposure measures are found in macroscopic models.

In (WHO, 2004), exposure is defined as the amount of movement or travel within the system by different users or a given population density. Exposure can thus be seen as the number of potentially dangerous situations on which the number of accidents and victims depend (COST 329, 2004). Because of the complex relation between exposure and the multitude of activities by which it is initiated (like travel for work, education or leisure), it is impossible to measure the level of exposure without error. Theoretically, exposure is often expressed in terms of the number of trips, the distance travelled or the duration of travel. Even these definitions are only a partial representation of the real exposure, because interactions between these terms can result in different transport patterns. In practical studies, proxies like traffic counts, the number of vehicle kilometres, the number of vehicles or fuel consumption are used. The multi-dimensional concept of exposure cannot be measured in all its aspects, and therefore these measures or indicators are commonly accepted as valid representations of the level of exposure.

In this chapter, focus is on the way exposure measures are generally used in macroscopic road safety models. First, a short review of the common ways of getting exposure data is given and some typical problems of exposure data are listed. Next, an overview is given of the exposure measures used in macroscopic time series models, together with the results obtained for the exposure variable.

3.2 Deriving measures of exposure

The literature on exposure measures is as large as that on road safety. This is easily explained by the fact that exposure is also key information in mobility studies that are not explicitly related to road safety, but rather to environment, economics, travel behaviour, and so on. However, the treatment of exposure measures in macroscopic road safety models is less widespread. Apart from a special issue of Accident Analysis and Prevention, published in 1982 (Haight, 1982), most references were written in the nineties (COST 329, 2004; OECD, 1997b). One of the main reasons for the increased importance in the context of road safety research is the attention given to road safety in an international context. The European Government, the OECD and the WHO are increasingly interested in comparisons of road safety among their member states. In order to provide a common basis for comparison, road safety figures are usually scaled using some measure of exposure. Especially when international comparisons are to be made, it is essential that exposure measures are derived along common principles.

Based on the available literature on road safety research, two basic methods for the collection of exposure data can be found (Wolfe, 1982). The first is to obtain data while trips are in progress, as is done with mechanical traffic counters. Also human observations and automatic cameras can be used. Based on these data, average annual daily traffic (AADT) is calculated, sometimes for different types or periods of days (week or weekend, day or night, etc.). Obviously, only vehicle types can be counted, but not the number of occupants. However, the number of vehicle kilometres can be used as a measure of exposure in time series or cross-sectional studies by extrapolating it into time and/or space. It is, indeed, the most appropriate way of data gathering to obtain continuous measurements over time. With the second method, data is gathered after the trips are completed, using in-person interviews, telephone interviews or mail questionnaires. This is usually done in travel habit surveys. People are surveyed about their travel habits (time, route choice, transport mode) for selected time periods. From this information, vehicle kilometres, but also person kilometres, travel time and number of trips can be derived. However, this data is less easily extrapolated for time series studies, as it has typically a registration period of a few days. Also, depending on the type and size of the survey, it is often difficult to relate the distance travelled to the route taken (OECD, 1997b). On the other hand, travel surveys are often the only sources of information available for data on exposure per type of road user, especially for non-motorized transport modes and vulnerable road users. In Chapter 8 of this manuscript, some examples are given on the use of travel survey data in road safety research for Flanders.

In more recent work, and due to the lack of traffic counters on more regional road networks, exposure measures are often derived from data on fuel sales, fuel efficiency and vehicle park (COST 329, 2004). This estimation procedure is possible for any type of motorised transport. Often, these procedures also use the rough estimates of yearly distance travelled from periodic surveys. In some of the DRAG models, a similar framework for the estimation of kilometres travelled is developed (Fridstrøm, 2000; Jaeger & Lassarre, 2000). In (Cardoso, 2005), a method is presented to estimate yearly national traffic volumes using data on vehicle fleet and fuel sales of the studied country and mathematical models fitted to existing data from other countries on fuel consumption, vehicle fleet and traffic volume. The COST 329 report (2004) shows further examples for Germany and France. When producing exposure data from fuel sales, it is assumed that the distance travelled is related to fuel sales, which makes it necessary to correct for the increasing engine efficiency of cars. However, it is also interesting to note that the weight of the cars increased over the same time period, again resulting in higher fuel consumption per kilometre driven. In Chapter 4, a very simple procedure is proposed to obtain monthly kilometres driven for Belgium, derived from fuel deliveries, average fuel economy and vehicle park.

3.3 Problems with measures of exposure

Unfortunately, exposure measures are often the primary source of annoyance for many traffic safety researchers. First, as mentioned by Wolfe (1982), the most easily obtained exposure measures are rarely the most desirable ones. Total population, the number of registered vehicles or the number of licensed drivers are not always good proxies for exposure. Depending on the scope of the analysis, some exposure measures may be more or less relevant. For example, fuel consumption is meaningless as a measure of exposure for a black spot analysis. Second, some measures of exposure are simply not available. In Flanders (Belgium), before the first travel surveys were conducted in the late nineties (Hajnal & Miermans, 1996), there was nothing else than the traffic counts, transformed in a yearly index of distance travelled. Moreover, the exposure data are rarely in a format that can be used in whatever type of analysis. Often historical exposure measures are not available over a long time period. For studies on specific groups of road users, the traffic counts are either not relevant or too aggregated. Third, typically in aggregated studies, it is not easy to find a measure of exposure that matches the traffic accidents studied. Traffic counts are usually available on a regular basis for highways, but not for the whole road network. For local and provincial roads, exposure measures are often based on an extrapolation of a sample of traffic counts on some predefined locations. More aggregated measures of exposure are easier to find, but sometimes too rough to analyse the relationship between accident occurrence and exposure. Another problem is that road safety is generally not the primary objective when gathering travel data. Traffic surveys are usually not designed for the analysis of accident risk but for sustainable transport organisation or travel behaviour studies.

Apart from the practical issues, the quality and reliability of exposure data is often low. This limits the level of detail of road accident analysis (COST 329, 2004). First, since traffic data are often based on surveys, they are only representative under restrictive assumptions. Second, they are mostly not available for all kinds of traffic. In traffic counting systems, the distinction between light and heavy motorised traffic is either not available or not reliable. Third, whereas traffic counts for motorways are mostly known, counts for regional and local roads are less frequently provided. Fourth, exposure data on a sufficiently low level of aggregation are almost never available. Even if we have data for a longer period of time, we are never sure that the data collection method remained unchanged for the considered period.

3.4 Using exposure measures in road safety research

In almost every macro model for road safety, a measure of exposure is used. However, there is a large variety in measures, and even if the measure is named identically, there might be substantial difference in the way it is constructed. Exposure measures vary considerably, and it is usually not possible to judge the quality of these variables. In (Hakim et al., 1991), the authors state that "as aggregate mileage increases, the overall number of accidents with injuries increases as well". According to the authors, this positive relationship is confirmed in other studies. In this section, some results for the exposure variable in various models is gathered.

In the DRAG-2 model for Quebec (Fournier & Simard, 2000), data on fuel sales of gas and diesel, expressed in litres, are associated with energy efficiency of vehicles (litres per 100 kilometres). Also, the efficiency reducing effect of cold winters and changes in types of vehicles on the road are taken into account.

According to the authors, the total distance travelled is the most important variable for explaining changes in the number of accidents, the severity and the number of victims. A 10% increase in distance travelled resulted in a 7.8% increase in bodily-injury accidents and a 7.2% increase in victims injured. However, the effect on morbidity was a 0.8% decrease, although less statistically significant. For fatal accidents and victims killed, an inverted U-shaped relationship was found, indicating that traffic safety can increase with exposure. A plausible assumption is that the severity of accidents will at first increase with exposure, but at a certain point it will decrease, as higher exposure will decrease speed and reduce the severity. This results in an inverted U-shaped relationship between exposure and severity (Gaudry & Lassarre, 2000). For mortality, the relation was also inversely U-shaped, although not statistically significant.

Fridstrøm et al. (1995) used data on gasoline sales as a proxy for exposure. For injury accidents, a coefficient close to one was obtained, while for fatal accidents and fatalities this value was significantly smaller than one, suggesting a less than proportional relationship. These lower values were explained by the hypotheses that average severity of accidents decreases with traffic volume due to lower speeds and that learning effects come into play. In the TRULS-1 model for Norway (Fridstrøm, 2000), measures of exposure for various groups of road users are used. These are derived in a complex econometric system in which, among others, traffic counts, benchmark traffic volumes, fuel sales, weather conditions and vehicle mix are combined. In the model, the injury accident frequency had an elasticity of 0.494 with respect to motor vehicle kilometres. This result is conditional on the assumption that the ratio of vehicle kilometres to road kilometres remains unchanged. Also, a higher heavy vehicle share in the traffic volume leads to a higher frequency of injury accidents. Motorcycle exposure influences the motorcycle accidents, but has only a small effect on overall accident frequency.

In the DRAG-Stockholm-2 Model (Tegnér et al., 2000), exposure (vehicle-kilometres) for gasoline driven passenger cars was based on monthly gasoline sales within the Stockholm County and on fuel efficiency. For the periods where no data were available, estimates were obtained from a multiple regression on total gasoline sales, Gross National and Regional Product and population. Also, fuel efficiency of passenger cars and the fluctuations in efficiency due to temperature differences between winter and summer are taken into account. In the first models, the number of bodily injury accidents was not proportional to exposure. The number of road accidents increased with the number of vehicle-kilometres, but was reduced in congested situations. At low levels of exposure, the number of light and severe injuries and fatalities decreased at first, but in congested situations the proportion of severe injuries and fatalities seemed to increase. Since these results were not logical, a new model was formulated, unfortunately resulting in a non-plausible U-shaped relationship for severe accidents and fatalities. For the severity models, a concave inverted U-shape was obtained. However, the overall performance of these new models was so poor that the authors considered the first models as superior, even with the wrong U-shaped structure for exposure.

In the TAG-1 Model (Jaeger, 1999; Jaeger & Lassarre, 2000) for France, the number of kilometres travelled by all road vehicles was calculated on the basis of petrol and diesel sales. The TAG-1 model is also an example of a rigorous framework in which monthly exposure data are derived. Total mileage has a significant positive impact on both injury and fatal accidents. Similar results are found for the number of fatalities, serious injuries and light injuries. On the other hand, the gravity rates for minor and serious severity were not significantly linked with exposure. Exposure to risk was positively correlated with the number of accidents and deaths.

The TRACS-CA Model for California (McCarthy, 2000) used the total number of vehicle miles travelled on state highways as an index of risk exposure for traffic on all roads in California, based on the assumption that traffic on state highways is highly correlated with the total vehicle miles travelled. Risk exposure was an important and statistically highly significant determinant of highway safety. Fatal and materials only crashes had relatively high elasticities with respect to exposure. Non-fatal injury crashes were less sensitive to risk exposure. Risk exposure increased crash frequencies, as well as mortality and morbidity. Also the fatality rate and the non-fatal injury rate were increased by exposure.

In the SNUS-2.5 model for Germany (Blum & Gaudry, 2000), road demand was expressed by the kilometres driven with gasoline and diesel, based on monthly gasoline and diesel consumption and the consumption rates. The number of accidents clearly depended on exposure, with a positive sign. However, the estimated parameters are less than one in every equation. In the equation for accidents with light and severe material damage, the elasticities are 0.24% and 0.31% respectively, while it is 0.08% for the total of material damage accidents. The elasticity is even lower for accidents with personal damage (0.2%) and 0.46% for total accidents.

Also other time series models in traffic safety include in some way an exposure index. For example, Lassarre (1986) uses a monthly traffic volume index estimated from a census in a predictive model of the severity of accidents (number of deaths). He found a unitary elasticity for traffic volume. In his comparison of ten European countries (Lassarre, 2001), the same author found a less than proportional effect for West Germany (0.895), Italy (0.449) and the Netherlands (0.544), while for France (1.640), Finland (1.959), the UK (1.194) and Sweden (1.173) a more than proportional effect was reported. This is an interesting result, as the best performing (or SUN) countries show quite different elasticities for exposure (although no statistical results on the proportionality were reported).

Other authors, for example Ledolter & Chan (1996), model the accident risk as the ratio of accident counts and exposure level, thereby presuming a priori that the coefficient of exposure is equal to 1. Another approach is presented in (Johansson, 1996). He investigated the effect of a lowered speed limit on the number of accidents with fatalities, injuries or vehicle damage. Because exposure is not known, it is modelled as a latent variable instead of using proxies like gas deliveries or the number of registered vehicles. The latent exposure measure is determined by gasoline prices, disposable income and an error term. The fact that exposure is a latent variable, next to other explanatory factors, implies that the direct and indirect effects of variables are combined in one parameter, making it impossible to separate exposure and other effects. Also, in this model structure, the latent variable is identical to a model in which no exposure measure is included. In (Scuffham, 2003), also unobserved components are used, but the exposure variable is explicitly modelled. Using a general-to-specific approach on current values and lagged differences of the dependent and independent variables, a model was developed for the number of fatal crashes. The current period exposure effect was highly significant with an elasticity higher than 3%, and the negative significant lagged differences of exposure indicate that as distance travelled increases, the risk of a crash increases but at a decreasing rate. With the given model, a long-run multiplier of exposure can be calculated, giving a value of 1.39, which is again a more than proportionate effect. Also in the various (both deterministic and stochastic) models developed by Christens (2003), elasticities higher than 1 are found.

Another solution to the low exposure data availability and quality in many countries is offered by induced exposure and cancellation of exposure techniques (COST 329, 2004). Induced exposure methods are based on the idea that a relative measurement of driver crash risks can be derived from accident counts. With this method, exposure data are estimated by making a number of assumptions about the relationships among the available accident data. For example, one can assume that general driver's characteristics are proportional to the characteristics of innocent drivers in the accident data (Li & Kim, 2000). The advantage of the induced exposure method is that the same level of detail as the casualty data used can be obtained. In Great Britain, for example, the method was applied to get exposure data for large and small motorcycles on rural and urban roads. A disadvantage is that there is a danger of circularity in the method. Accident data are used to estimate exposure, which is subsequently used to analyse the accident and fatalities data. Also, the assumptions made can often hardly be verified.

The cancellation of exposure technique is based on taking the ratio of accident risks for two groups of road users, such that the exposure part cancels out . By taking for example severity ratios (fatalities per injury accident) for various types of accidents, comparison of these ratios will reveal the circumstances that lead to severe accidents. Also, the occurrence of accidents under two types of conditions can be studied, always under the assumption that the level of exposure in the different conditions is comparable. This is, of course, a very strong assumption. On the other hand, the ratios are easily calculated from accident databases and in itself often provide useful information.

Apparently, the effect of exposure and the way it should be constructed is not as clear as one would expect from theory. In general, the exposure elasticities obtained from various macro models for road safety are positive, but as for proportionality, very different results are found. Also, the results of a (more complex) inverse U-shaped relationship between exposure and accidents are not necessarily in accordance with expectations. In some models, exposure even has an illogical or unexpected sign.

3.5 Conclusion

The fact that different results are obtained in different countries or on different data sets is not difficult to accept. Indeed, each country has a different road transport system and therefore the effect of the exposure to risk should not be comparable as such. A more general problem, however, is the fact that the definition of the exposure measure often varies considerably, making it very difficult to compare the results, even in terms of elasticities. Sometimes huge efforts are done to construct a valid exposure measure, while in other models only very simple indicators are used. All models use a variable named “exposure”, but they are often constructed differently. As shown for example in (OECD, 1997b), different exposure measures can lead to diverging results.

In (Scuffham & Langley, 2002), the total number of fatal crashes was standardised by three different proxies for exposure, namely the total population, the number of registered vehicles and the distance travelled. By doing so, a proportional relation between risk and crashes is again assumed. The results show that a different exposure measure may lead to different results in the effects of the other explanatory variables. It is therefore important to indicate what kind of exposure is measured in the model. Obviously, road safety researchers must make shift with the exposure data they got. This is perhaps even more true for time series than for cross-sectional studies. Indeed, in this domain, it is never possible to regain data of past periods.

Clearly, all this has some implications for the use of exposure data in time series studies. First, if a variable is not measured in the past, it cannot be recovered later. Experiments cannot be done in the past. In a black spot analysis, for example, it is still possible, albeit not easy, to take a sample of exposure data at the location of interest. Second, the aggregated approach to road safety makes it difficult to set up a data gathering or sampling exercise. Not all series are available for a long enough analysis period. Some data are available, but the frequency might be wrong or useless. Third, the granularity of the measure can be too high or low for the kind of road safety data that should be analysed, and this is not always easily changed. Fourth, the analysis period is often determined by the availability of exposure data. This may also be a reason for differing results among studies. For example, while the first DRAG model considers data from 1956 onwards, Belgian data on exposure can only be used for the years 1986 and later. Another time horizon or time series length can influence the results. Also, it is not difficult to see that the road transport system, and maybe also the relation between exposure and road safety, changed over time, which may lead to different results for the same country in newer studies.

While it is commonly agreed that exposure data is highly important and interesting in road safety studies, further details are often necessary to bring into light the effect of the exposure measure used and to give a correct interpretation to the results obtained from the model.

Chapter 4 Data for macroscopic road safety models in Belgium

4.1 Introduction

In this chapter, an overview is given of the data that is available in Belgium and in Flanders for the development of macroscopic road safety models. Traffic safety analysis usually starts from two main information sources (OECD, 1997b): crash and casualty data on the one hand and exposure data on the other hand. Although crash data is basic information in road safety analysis, exposure data is considered as key information for every kind of road safety research. Therefore, the first task of a road safety researcher is to list the possible variables related to crashes and exposure. This effort is as important as it is difficult, and the research possibilities are often limited by the data. Indeed, the level of detail that can be achieved in a model depends to a large extent on the data availability and quality.

In macroscopic models, road crashes and exposure are often seen as parts of a larger road system. The level of exposure and the occurrence of crashes cannot be seen as isolated concepts, but should be related to the environment in which they evolve. In this environment, factors like the weather, the introduction of road safety laws or the general economic climate might influence the level of exposure or the number of crashes counted. Therefore, a macroscopic model also contains "explanatory" variables that describe the larger system in which traffic and crashes occur.

As an introduction to the models that will be developed in the subsequent chapters, an overview is given of the available statistics in Belgium that are useful in macroscopic modelling. The sources for data on exposure, road crashes and explanatory factors will be described. Note that this overview is focused on the application in macroscopic time series models. It is therefore by no means an inventory of available data sources in Belgium. A more general description of road safety data sources in Flanders can be found in (Van Hout et al., 2004).

4.2 Data on exposure

As explained in the previous chapters, exposure is a key variable in road safety research. The kind of analysis that can be performed therefore depends to a large extent on the availability of exposure data. In particular, a time series model needs observations of exposure over several years or months, while for a cross-sectional model different observations of exposure at a certain point in time might be more useful. It goes without saying that historical time series data are harder to find, as the data cannot be gathered in an ad-hoc experiment. For the models developed in this study, mainly time series data are needed. Therefore, an overview is given of the different data sources that can be used in this respect. If exposure data is not available for a certain type of analysis, a proxy variable may be used.

4.2.1 Data per type of road

4.2.1.1 Vehicle kilometres per type of road

The first and most important measure of exposure in Belgium is the yearly number of vehicle kilometres, published by the Belgian government (Federal Government Service for Mobility and Transport). This statistic is calculated for motorways, regional roads and local roads. The total distance travelled is based on fuel consumption (De Borger & De Borger, 1987).

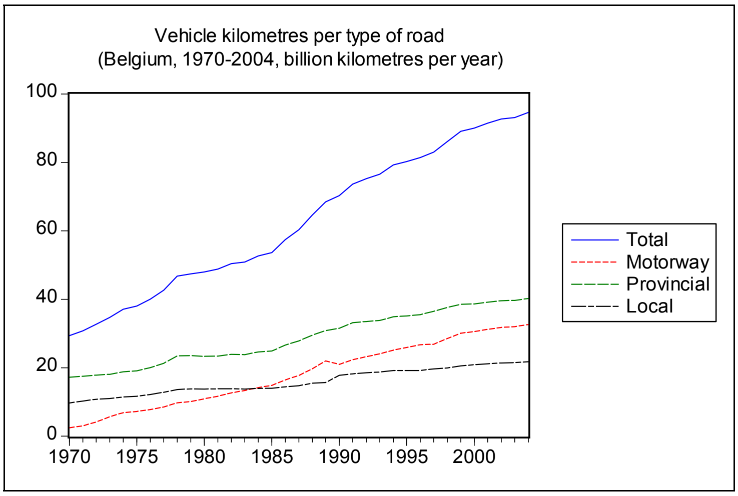

FIGURE 5: Vehicle kilometres by road type

The distribution of traffic over the different road networks is done by the "GcLR" method (Labeeuw, 2005). This method determines the global traffic evolution, based on an estimate of the traffic intensity on each road section and on traffic counts in the (spatial) neighbourhood or in the (temporal) past. Each year, the number of vehicle kilometres per type of (motorized) road user and per type of road are calculated. The number of vehicle kilometres travelled per type of road is shown in FIGURE 5. The distances travelled on the different types of road are all monotonically increasing. However, transport on local roads is quite constant, while motorways and provincial roads show substantial increases.

4.2.1.2 Road network length

Apart from the number of vehicle kilometres for each road type, the length of the different road networks is known (Labeeuw, 2005). Although these series could be used as an indicator for exposure, it is certainly a less accurate measure compared to the number of vehicle kilometres. The road network length does not take into account the traffic density. This is especially important for motorways, where the distance travelled is increasing faster than the length of the motorway network leads to suspect. Moreover, the official time series for road length are put together from different data sources. Among them, various inconsistencies can be found.

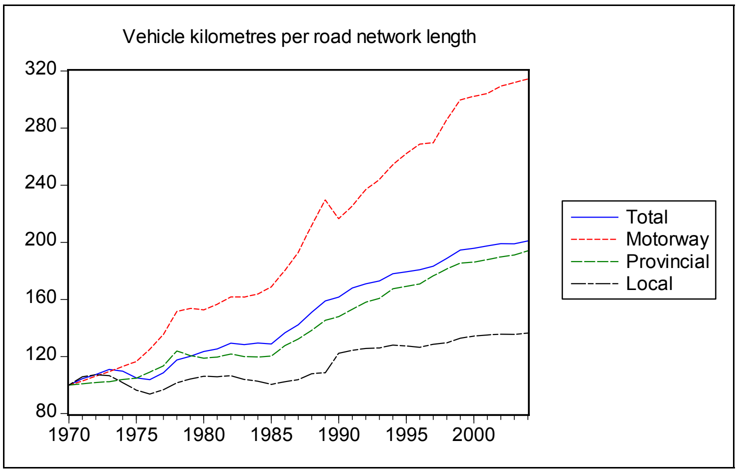

FIGURE 6: Vehicle kilometres per road network length (index)

Also, it is known that the definitions of the road networks changed over time, leading to artificial shifts in the length of the road networks. However, it is instructive to consider the evolution of the number of kilometres driven per km road network. An index of ratio of the vehicle kilometres to the road network length is shown in FIGURE 6. The ratio in 1970 is set equal to 100 and subsequent values are calculated relatively to this first year. Note the steady increase in the index for the motorway network. The density on local roads is increasing much slower.

4.2.1.3 Traffic counts on motorways

Next to the yearly official statistics on the vehicle kilometres per type of road, monthly motorway traffic counts are available (Federal Government Service for Mobility and Transport). Monthly averages have been calculated for the whole territory of Belgium, giving an indication of the traffic flows on motorways. These counts are shown in FIGURE 7.

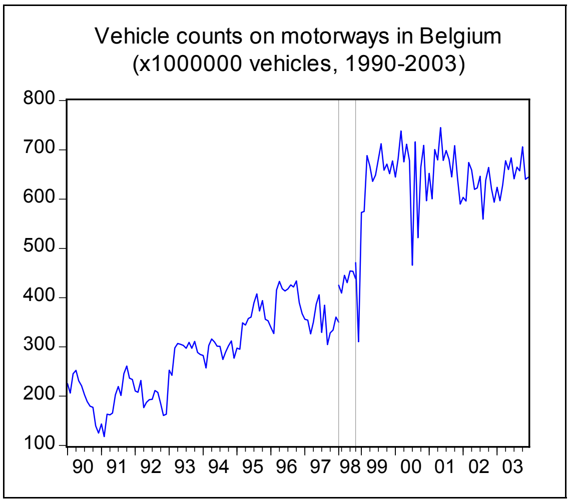

FIGURE 7: Vehicle counts on motorways in Belgium

However, these counts are not available for the road networks, and there is no indication at all of the number of kilometres driven by the counted vehicles. Moreover, the series shows a significant jump at the beginning of the nineties, which has been caused by a change in the count registration system and has nothing to do with increased exposure. The two vertical bars indicate sample breaks. These values are missing in the original data set. It is therefore clear that the series of vehicle counts on highways shows some properties that are not related at all to the developments in exposure. Using these counts as a proxy for exposure is not desirable.

4.2.2 Exposure data per type of road user

4.2.2.1 Vehicle kilometres per type of road user

The distance travelled by motorised road user type is shown in FIGURE 8. Data are available for personal cars, trucks, busses and motorcycles. Cars and trucks are read from the left axis, busses and motorcycles from the right axis.

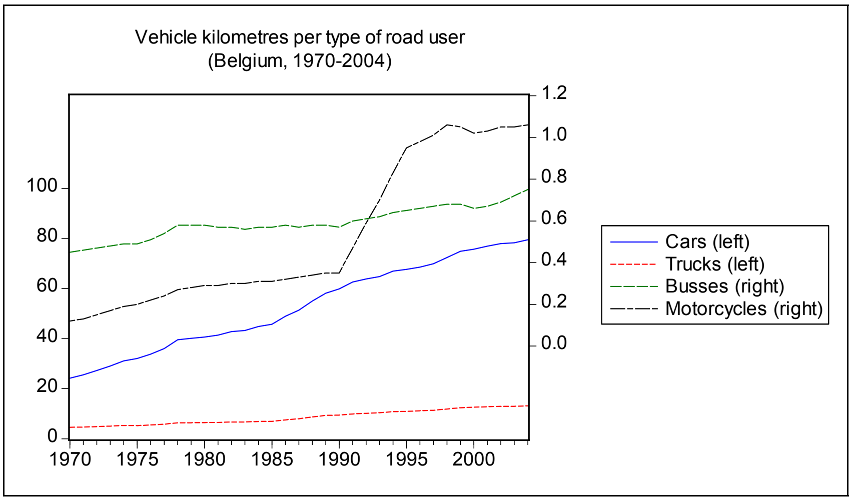

FIGURE 8: Vehicle kilometres by road user type

The same time series are shown in FIGURE 9 on separate graphs. Over the last 30 years, the distance travelled increased considerably for all types of road users. Comparing the values in 1970 with those in 2004, an expansion factor can be calculated for each type. The number of busses and trucks is respectively 1.67 and 2.88 times higher in 2004. Cars travel by far the largest distance from all vehicle types, with an expansion of 3.29. From 1990 onwards, the motorcycles show a steady increase in the distances travelled. Compared to 1970, the distance travelled by motorcycles is 8.83 times higher in 2004.

4.2.2.2 Vehicle park per type of road user

In the context of exposure of different types of road users, also the data on the vehicle park (available from the Federal Government Service for Mobility and Transport) is interesting. The yearly statistics give an overview of the number of registered cars, busses, motorcycles and trucks. The vehicle park data are shown in FIGURE 10. The number of cars, trucks and motorcycles seem to follow the corresponding number of kilometres driven, but this is not the case for the busses. While the distance travelled increased over time, the park reached a maximum in the late seventies, and decreased until the late nineties. The quality of this variable might be doubtful. In general, it is expected that the vehicle park will be a less accurate measure of exposure than the number of vehicle kilometres, because it does not take into account the use of the vehicles.

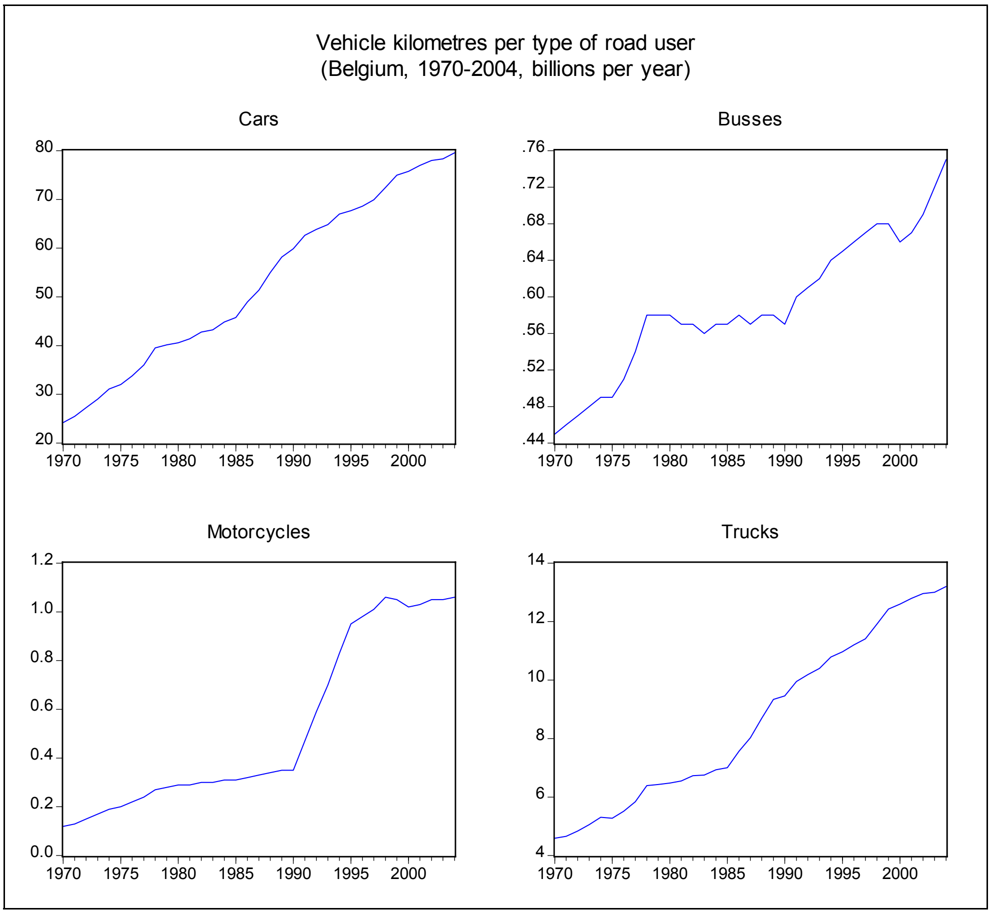

FIGURE 9: Vehicle kilometres by road user type

It can be useful, however, to calculate the ratio of the number of vehicle kilometres to the vehicle park. This is done in FIGURE 11. These graphs show that the number of kilometres driven per vehicle increases over time for the cars and the busses. The distances travelled by trucks and motorcycles are realized by a decreasing number of vehicles since 1990 for trucks and since 1995 for motorcycles. Because of the increasing park for these vehicles, each vehicle travels less on average.

Sometimes, data on the vehicle park on a monthly basis are needed. These statistics are not available, but a proxy for the monthly changes in the vehicle park can be created using the monthly vehicle registrations. If it is assumed that the deletions of vehicles is constant over time, then the seasonality in the vehicle park will resemble that of the registrations (Jaeger, 1999).

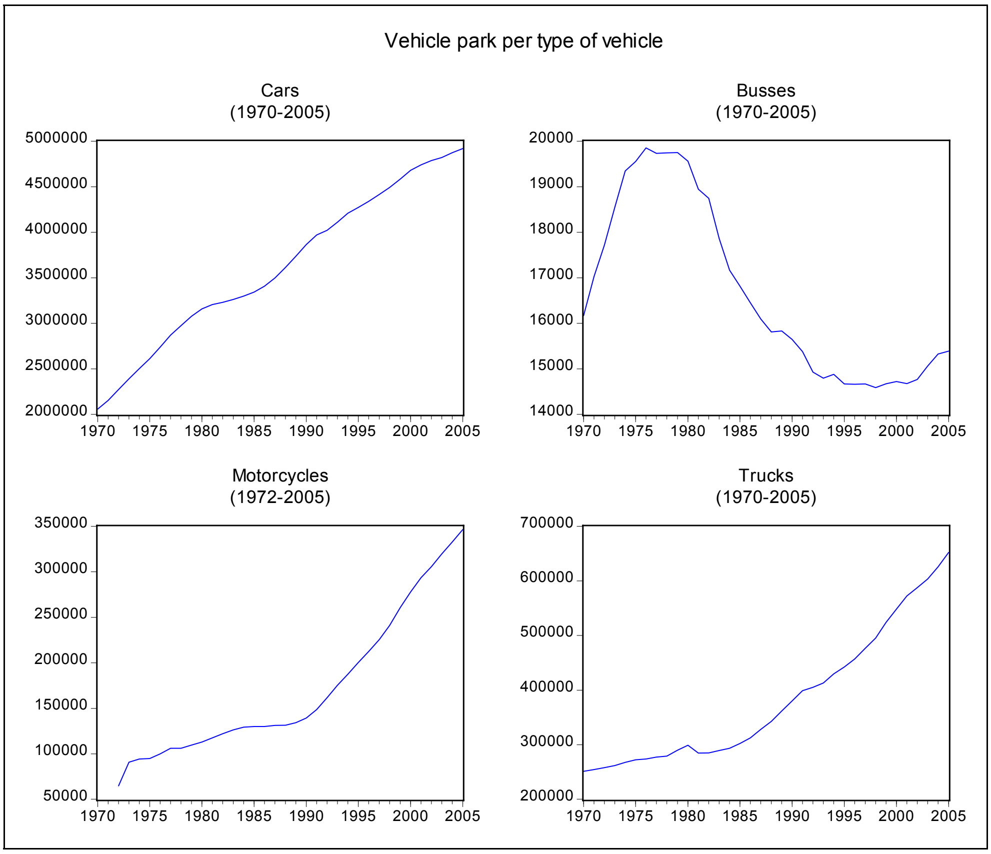

FIGURE 10: Vehicle park per type of vehicle

For the personal cars, the lorries and the motorcycles, the yearly changes in the vehicle park were distributed over the months according to the monthly registrations. The yearly statistics on the vehicle park are created every year on 1 August. It is therefore assumed that the park statistics represent the situation in July every year. The difference between two July values is proportionally distributed over the remaining months according to the monthly percentage of registrations.

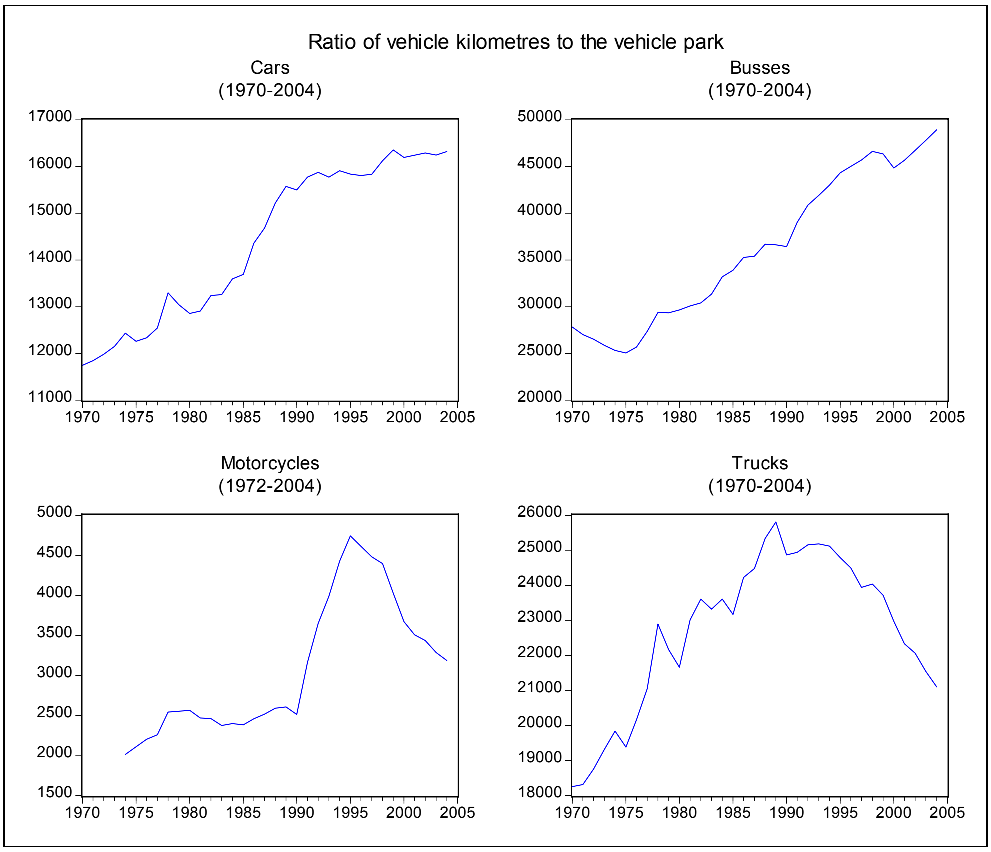

FIGURE 11: Ratio of vehicle kilometres to the vehicle park

In FIGURE 12, the created series are shown on the left graph. The vehicle park of cars is projected on the right axis, while the number of trucks and motorcycles are read from the left axis. All types of vehicles show an increase over time, and the cars and lorries clearly outnumber the park of motorcycles. However, the right graph shows an index of the evolution of the number of cars, lorries and motorcycles. All observations in July 1973 are set equal to 100 and later values are calculated proportional to this starting point. From this graph it is clear that the increase in the number of motorcycles is strikingly larger than for cars and lorries. Having a comparable growth rate until the early nineties, the number of motorcycles completely diverges in comparison with the number of lorries and cars. It is not inconceivable that this evolution in the composition of the vehicle park will have an effect on road safety.

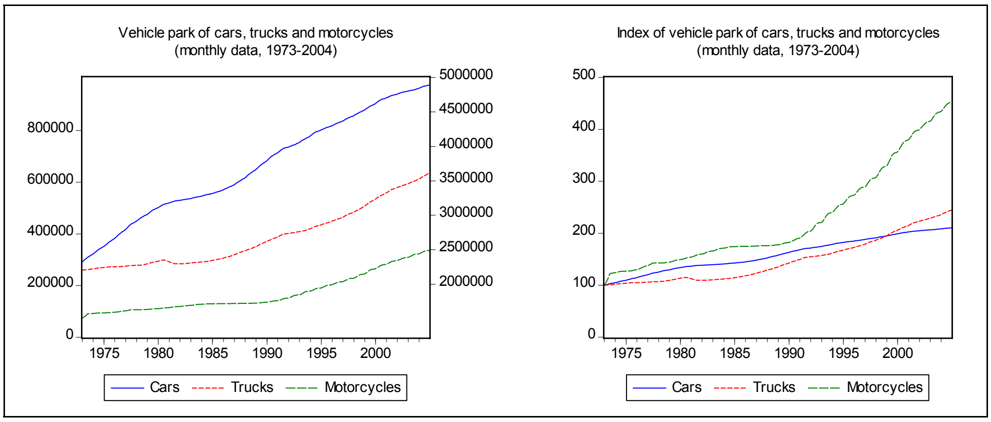

FIGURE 12: Vehicle park of cars, lorries and motorcycles

4.2.3 Additional sources for exposure data

All official statistics presented up to now suffer from three important restrictions. First, they do not report the distance travelled by non-motorized modes of transport (pedestrians, bicyclists, etc.). The statistics are useful for aggregated studies on road safety, but they are limited in the number of road user groups that can be analysed. Especially in relation to road crashes, it is important to have an indication of the exposure of vulnerable road users. Second, no information is given on the special characteristics of the road users, although it is known that road safety and exposure differ substantially over various age and gender categories. Third, with the yearly data, there is no possibility of measuring seasonal (monthly) variation in exposure and road safety. If the level of exposure and the occurrence of crashes show a specific distribution over the months of a year, this would never be seen with yearly data. To deal with these restrictions, some other data sets will be introduced. First, the Flemish travel survey data can be used to obtain exposure on vulnerable road users, albeit not in a time series format. Second, population data offers insight in the age/gender distribution of the road users. A solution to the absence of monthly exposure data is given in the next section.

4.2.3.1 Flemish travel survey data

To measure the exposure of non-motorized transport, and to distinguish between various groups of road users in terms of age and gender, studies are done on travel survey data. In the Flemish travel survey for the year 2000 (Zwerts & Nuyts, 2004), trips of road users (car drivers, car passengers, pedestrians, bike and motorbike riders, public transport users) were registered for the period

January 2000 – January 2001. It is based on a random sample of 2823 households, including 7638 people who were more than 6 years old. In total, 21031 trips were registered. People participating in the survey were given diaries to record the trips made during two specified days in the survey period, but only the data of the first day are retained because of the less accurate registrations on the second day (Zwerts & Nuyts, 2004). Especially short trips, shopping activities and the transport mode were less often reported than on the first day.

Since data was recorded on person level, also the age and gender of the respondents are available. From the travel survey data, an exposure measure (the number of kilometres travelled) is derived for various user groups based on the average daily number of kilometres and available population statistics (population density for each age-gender combination). As the travel survey has been conducted in one year, no time series can be derived from it. For investigations on this level of detail, a cross-sectional setting will be used (see Chapter 8).

4.2.3.2 Population data

A useful data source that allows making a distinction between road user groups in terms of age and gender is the official population database from the Belgian government. The population is a key indicator of the Belgian economy, maintained by the Belgian National Bank (Belgostat). Also from a road safety point of view, it is not illogical to consider the risk according the age and gender of the road users. It is known that young persons are often involved in road crashes. Moreover, looking at the official statistics (D'Hondt, 2002), road crashes are clearly one of the most important death causes for young persons. In 2001, approximately 36% of the fatal road victims was between 10 and 29 years old (CARE, 2004).

The official statistics show the number of inhabitants for each age/gender category. Although it is not a perfect measure of exposure (nothing is said about distances travelled), the population data provides an upper bound of the number of road users. Given the assumption that not the whole population travels (especially for older people), the risk of being killed will be underestimated. However, current changes in the structure of the population might be reflected in the data.

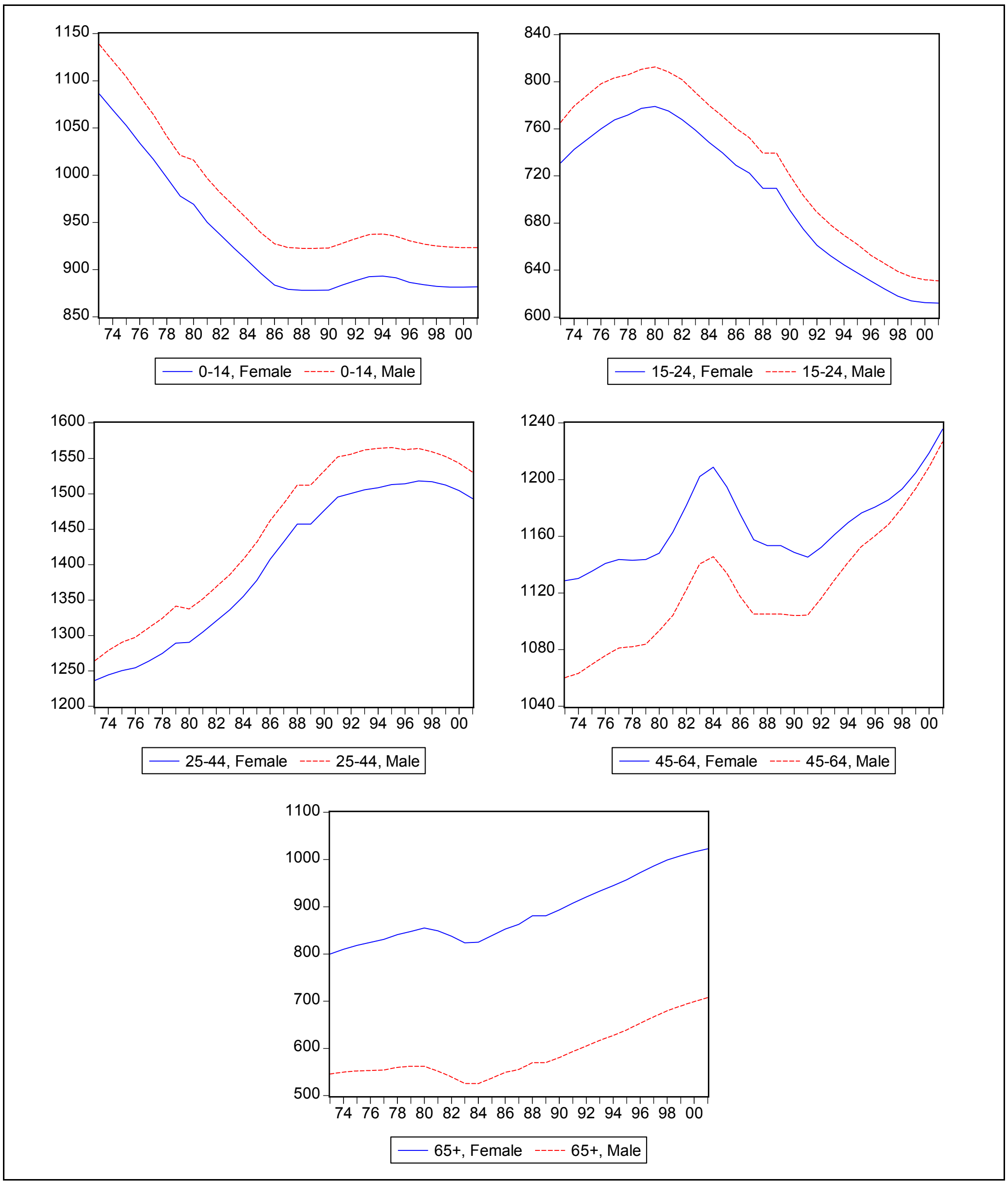

FIGURE 13: Population data for various age/gender groups

As an illustration, the population number for five age categories are given in FIGURE 13: 0-14, 15-24, 25-44, 45-64 and 65 and older. It can be seen that the youngest group of road users has been decreasing for about 25 years, and seems to be stabilising in the late nineties. The group of older people is growing. It is also interesting to note that females are the smallest group among the younger people, and the largest group among the older. It is clear that these changes in population might also be reflected in the level of exposure to road crashes.

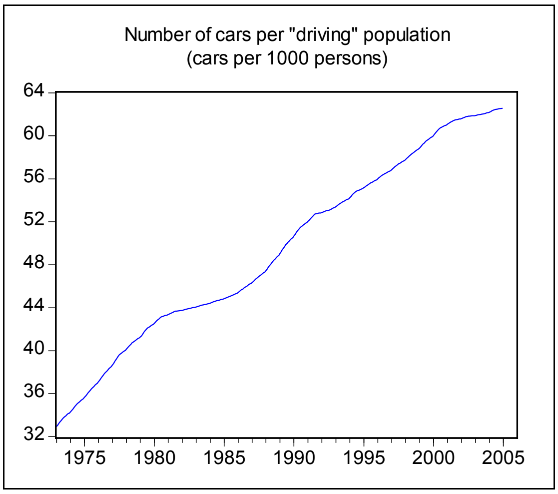

FIGURE 14: Number of cars per "driving" population

If monthly population data are needed, the SAS procedure EXPAND can be used to change the data from yearly to monthly frequency. The procedure fits cubic spline curves to the data to form continuous-time approximations of the series (SAS Institute Inc., 2004a). In combination with the (created) monthly data on the vehicle park, the number of vehicles per (part of the) population can be calculated. For example, it might be interesting to know the number of cars per person in the "driving" population. Given the available population statistics, persons aged 15-74 may be roughly considered as the driving population. The evolution of the ratio of the car park to the driving population is shown in FIGURE 14.

4.2.4 Creation of monthly exposure based on fuel consumption

The overview of available exposure measures given above indicates that not all variables are suited for every kind of road safety analysis. For an analysis on a monthly basis, the traffic counts on motorways are the only available statistics. Unfortunately, they offer a partial representation of traffic on the Belgian motorway network and are erroneous for some periods. Therefore, an effort is needed to develop a monthly exposure measure from another proxy variable that is registered every month, namely the statistics on fuel deliveries for domestic consumption in the transport sector, obtained from the Belgian Ministry of

Economics (Department of Energy). In combination with yearly exposure data and using some common sense assumptions, a measure of exposure on a monthly basis can be derived.

In this section, a method is presented to calculate an exposure measure (number of kilometres driven) for Belgium, based on the monthly fuel deliveries, the average fuel economy of cars and the vehicle park. In short, the procedure is as follows.

1. The monthly fuel sales (in metric tons) per type of fuel are corrected for extreme December values and transformed into litres.
2. Average fuel economy by fuel type is calculated based on a weighting scheme that takes into account the vehicle park.
3. Based on the results of step 1 and step 2, the number of kilometres per fuel type per month is calculated. The sum of these values is an initial estimate of the number of kilometres driven per month. The sum over the months per year is an initial estimate of the number of kilometres driven per year.
4. A correction factor for each year is calculated by comparing the initial estimate of the number of kilometres driven per year and the corresponding official statistic.
5. A final estimate per month is obtained by multiplying the estimated values with the correction factor.

The first three steps of the procedure result in a number of kilometres that is basically derived as the quantity of fuel deliveries in litres divided by an average fuel economy (litres consumed per 100 kilometres driven). As the fuel deliveries are on a monthly basis, the resulting number will be an initial estimate of the number of kilometres driven in each month. However, on a monthly basis, there is no other quantity that can be used to assess the quality of this monthly number. The only reference that one can think of is the officially reported yearly number of kilometres driven. In order to bring the results in line with the official yearly statistics, correction factors are calculated by comparing the sum of the obtained values for each year with the corresponding official value. This is done in the last two steps of the procedure. In the subsequent sections, the procedure is applied to the Belgian data.

It is clear, however, that using fuel deliveries as a basis for exposure calculation has some pitfalls, as mentioned for example in (COST 329, 2004). First, fuel prices may lead to changes in fuel sales, especially at the border of the country.

This may lead to the phenomenon that Belgian people buy their fuel in the neighbouring countries and subsequently travel on Belgian roads. Also, it is known that Belgium is a transit country, meaning that many kilometres on Belgian roads are driven by foreigners with fuel that was bought elsewhere. Both situations can result in an underestimation of the number of kilometres driven, when calculated on the basis of fuel delivery data. Another problem is that there might be a time lag between the fuel delivery and the consumption in a vehicle. That is, fuel can be bought in one month and only effectively consumed in the next month. These transitory effects in space and in time are typical of fuel delivery data, and are not explicitly corrected for in the analysis. However, using fuel deliveries as a proxy is still felt as a logical choice in the context of this study, as there is, at the moment of writing, no other valid alternative.

4.2.4.1 Monthly fuel sales correction

From official statistics (Federal Government, Ministry of Economics, Department of Energy), the monthly consumption of gasoline, diesel and LPG used for transport by all kinds of vehicles is obtained. The data are available for the period January 1986 - December 2004. For gasoline, the series consists of the deliveries of various kinds: leaded or unleaded, high-octane or low-octane, and so on. The time series are presented in FIGURE 15.

Some problems arise with these series. First, the data are in metric tons, instead of litres. The fuel sales have to be transformed, taking into account the average density of the fuel products. Second, some extreme values are observed in the series. Unexpected high or low values in December months are caused by accounting principles. The last month of the year is often used to correct the registration of deliveries of the year. These corrections also cause the negative value for LPG in December 1993.

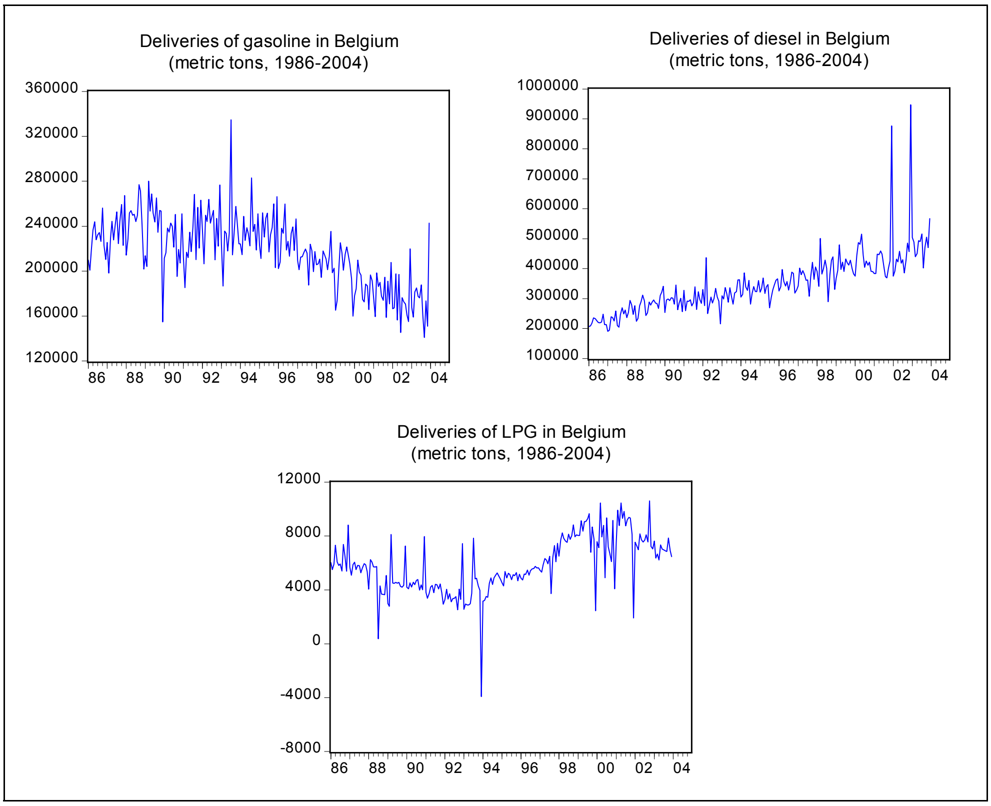

FIGURE 15: Deliveries of gasoline, diesel and LPG in Belgium

To transform the series from metric tons into litres, average densities (at 15°C) of 745 km/m³ for gasoline, 830 kg/m³ for diesel and 510 kg/m³ for LPG are used. These values were communicated by the Department of Energy from the Ministry of Economics. As an example, consider a delivery of 5000 metric tons of Diesel. According to the densities in the table, this results in:

$$5\,000 \text{ metric tons} = 5\,000 \times \frac{1\,000 \text{ kg}}{\left(\dfrac{745 \text{ kg}}{1\,000 \text{ litres}}\right)} = 6\,711\,409 \text{ litres} \tag{9}$$

To correct for the extreme values in the series, December observations in seven years were given missing values (1989 and 2003 for gasoline, 2001 and 2002 for diesel and 1993, 1999 and 2001 for LPG).

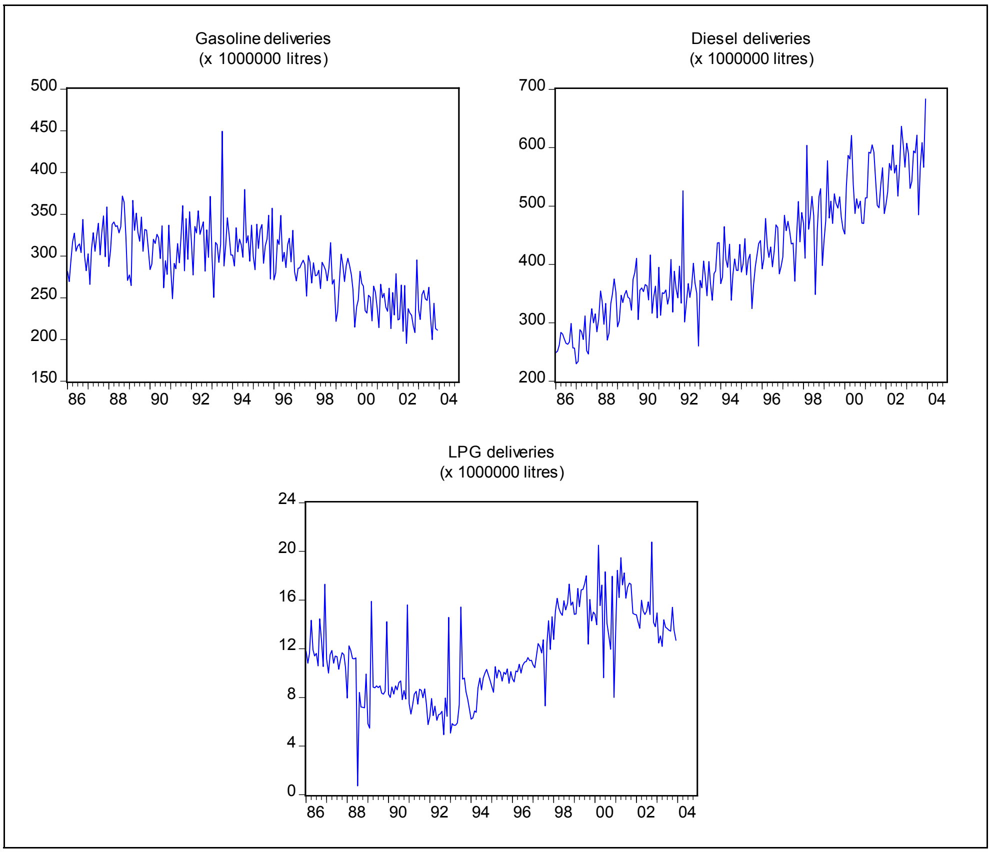

FIGURE 16: Corrected deliveries of gasoline, diesel and LPG in Belgium

The SAS procedure EXPAND was used to interpolate the missing values. The procedure fits cubic spline curves to the non-missing values to form continuous-time approximations of the series (SAS Institute Inc., 2004a), including the missing values. Because the changes made to the seven December months affects the totals of the corresponding year, the correction of the December month is divided over the other months of the year. The graphs for DIESEL, LPG and GASOLINE with the corrected months now have less extreme values and only positive deliveries, as can be seen in FIGURE 16. These series will now be used to estimate the number of kilometres driven.

4.2.4.2 Calculation of average fuel economy

Data on the average fuel economy of various kinds of vehicles are not available in Belgium, at least not in the form of time series. From the official statistics in the Netherlands (Statistics Netherlands), data on the average fuel economy for personal cars per fuel type was obtained. Because there is no similar data set available for Belgian cars, it is assumed that fuel consumption for Dutch and Belgian cars is comparable. The Dutch data set starts in 1987 and ends in 1997, and should therefore be extended by estimation back to 1986 and up to 2004 in order to fit the available data on fuel deliveries. The estimated values are obtained from three nonlinear (exponential) models to estimate the values of the missing years. The model is of the form $y_t = c + \exp(ax_t + b)$, where y_t is the fuel economy to be estimated and x_t is an index for the year (1985 is set equal to 1). The estimated equations are as follows:

$$\begin{aligned} &\textit{Gasoline}: \quad y_t = 8.31 + \exp(-0.39\,x_t + 0.05) \\ &\textit{Diesel}: \quad y_t = 6.29 + \exp(-0.07\,x_t + 0.19) \\ &\textit{LPG}: \quad y_t = 9.90 + \exp(-0.19\,x_t + 1.25) \end{aligned} \tag{10}$$

This estimation procedure results in a yearly average fuel economy per fuel type, as shown in FIGURE 17. The models only give a rough fit to the data (not all parameter estimates are significant), but they are useful to show the trend.

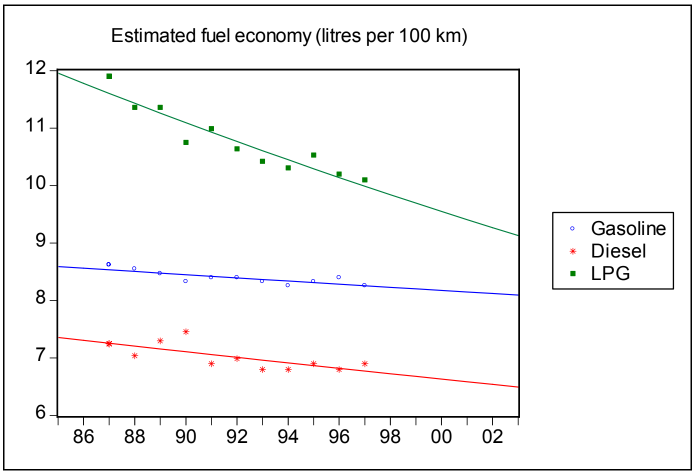

FIGURE 17: Estimated fuel economy curves

The obtained figures, however, are only valid for passenger cars. Data on fuel economy for motorcycles and trucks for the period of analysis are not found. In a Dutch publication (Transport en Logistiek Nederland, 2002), a comparison is made between the average fuel economy of trucks in 1965 (49 litres per 100 km) and 1998 (33 litres per 100 km). Also other sources mention fuel economies of that order. Based on these few figures, we assume that trucks and busses have an average fuel consumption that is four times larger as the fuel consumption of passenger cars. Further, it is assumed that motorcycle fuel consumption is comparable to that of passenger cars.

Starting from the estimated values of the fuel economy of passenger cars and the assumptions made for trucks, busses and motorcycles, an average fuel economy for each type of fuel will be derived. Because the distribution of the fuel consumption over the various kinds of vehicles is not available, the vehicle park is used to calculate an average weighted fuel economy for each type of fuel. Consider, as an example, the vehicle park for the year 2000 (statistics for August 1st, 2000). TABLE 3 contains the distribution of the vehicle park over the three main kinds of fuel (column percentages). It can be seen that personal cars represent the largest part of the vehicle park. Among the diesel vehicles, the trucks account for about 26%.

TABLE 3: Vehicle park per vehicle and fuel type (2000)

Vehicle type	GASOLINE	DIESEL	LPG
Personal car	2 732 352	1 867 351	59 059
	89.03%	73.27%	86.14%
Bus	214	14 347	47
	0.01%	0.56%	0.07%
Truck	68 905	666 796	9 453
	2.25%	26.16%	13.79%
Motorcycle	267 529	34	0
	8.72%	0.00%	0.00%
Total	3 069 000	2 548 528	68 559

The fuel economy of the various vehicle types is subsequently weighted according to the distribution of vehicle types over the various types of fuel in the vehicle park. For gasoline, a fuel consumption for personal cars of 8.31 litres per 100 km is noted, so the weighted average gasoline consumption equals 8.31*(97.75%) + 33.24*(2.25%) = 8.87 litres. For the year 2000, an average weighted gasoline consumption of 8.87 litres per 100 km is therefore taken into account. In the same way, averages of 12.07 litres for diesel and 14.23 litres for LPG are obtained. For the other years, the calculation is similar. Average fuel consumption values for the months between the obtained values are estimated using linear interpolation. Note that interpolation is done between the July months. Since the weighting schema is based on the vehicle park (measured on

1 August each year), it is assumed that the weighted average fuel economy is valid for the month of July each year, and interpolation is done for the other months. The final weighted fuel economy curves are shown in FIGURE 18.

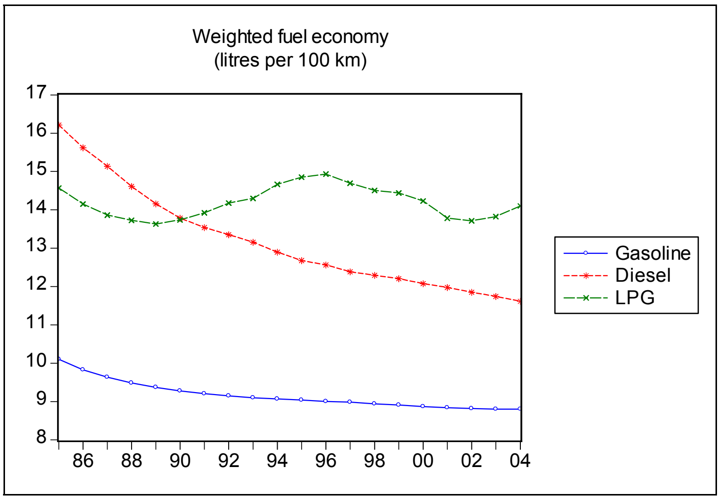

FIGURE 18: Weighted fuel economy

4.2.4.3 Estimation of the number of kilometres

Using the data on fuel deliveries (in litres) and fuel economy (in litres per 100 km), a measure of kilometres driven per month for each fuel type can be calculated by dividing the fuel deliveries by the fuel economy. The sum of these values per month is an initial estimate of the number of kilometres driven. However, to bring the estimated number of kilometres in line with the official yearly statistics (Federal Government Service for Mobility and Transport), a correction factor is calculated. This factor equals the ratio of the reported number of kilometres to the sum of the estimated monthly numbers of the year. In a sense, the correction factor can be seen as a measure of accuracy of the proposed method. It is applied to all monthly figures to make sure that the yearly number matches the official statistic, supposing the latter to be correct.

FIGURE 19 shows the correction factors for each year. These are scaled around zero by subtracting the correction factor from one. A positive correction factor means that the estimated number of kilometres is higher than the official statistic for the given year. For example, the year 1986 shows a correction factor of 2.1%, meaning that the estimated value was higher than the official statistic, and a correction of (1-0.021) = 0.979 was needed to make them correspond. That is, the official statistic for 1986 is 97.9% of the estimated value. It can be seen that in general only small corrections are needed (a maximum of 6% in absolute value), indicating that the values obtained with the proposed procedure are quite close to the official statistics.

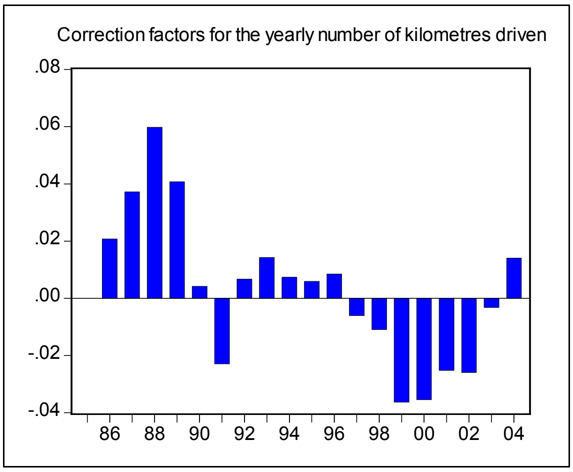

FIGURE 19: Correction factors for the yearly number of kilometres driven

It is interesting to note that there seems to be a certain pattern in the correction factors. Until the early nineties, the estimated values were systematically above the official statistics, while the reverse is true for more recent years. This indicates that the official statistics and the fuel deliveries are related in a specific, yet unknown manner. Although this finding deserves further attention, it will not be investigated here.

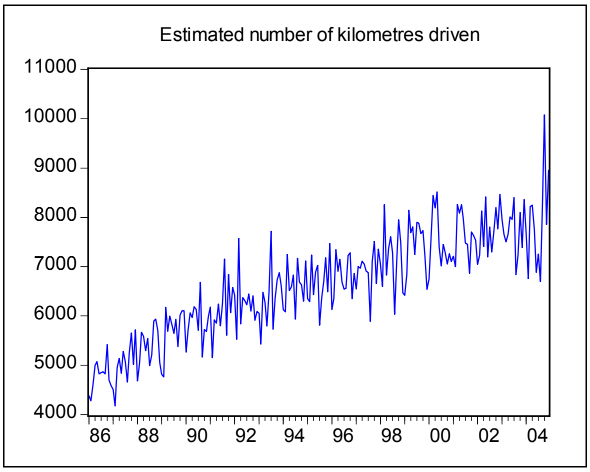

FIGURE 20: Estimated number of kilometres driven

Once the correction factors are applied to the monthly data, we obtain an estimate for the monthly number of kilometres driven in Belgium. The result is shown in FIGURE 20. This series can now be used as an input in road safety models for monthly data. In Chapter 6, models are developed based on monthly data. Although it is technically possible to include a yearly measure of exposure in a model for the monthly number of fatalities, it is desirable to include variables with the same frequency to capture possible seasonal effects in the data. The measure of exposure that was derived here, will therefore be used in these models to explain developments in road safety on a monthly basis.

4.3 Data on road crashes

Statistics on road crashes are always at the basis of road safety studies. If research is expected to enhance the insights in road safety, then a consistent framework for crash reporting is not an option but a necessity. Data on crashes are, in principle, available from three sources (COST 329, 2004): police registrations, hospitals and insurance companies. In Belgium, the main source of information is the registration done by police officers when a crash occurs. The data sets gathered by hospitals and insurance companies are not easily accessible. They could, however, provide an added value to the existing data sources, as the police registrations suffer some difficulties. The quality of the data gathered by hospitals and insurance companies is usually high, because these data are important for the correct settlement of medical or juridical dossiers. A disadvantage of these sources is the partial representation of the problem, as the data is generally not available for every crash. Also, the data sources are mostly not accessible for research purposes, because of the presence of personal information. The use of these data is considered as an intrusion on the victim's privacy. Moreover, the data sets contain crucial information to support the future business of these companies, and are therefore kept private. In this section, the available Belgian crash data, the registration process of these data and the known problems of erroneous registration and under-reporting are discussed.

4.3.1 Definitions and registration procedure

In Belgium, *traffic accidents* are defined as accidents that occurred on a public road, which are reported to the police and which lead to casualties (OECD, 1998). Although not explicitly stated, this definition is more or less in line with the definition of a traffic accident in the Geneva Convention: "an accident which occurred or originated on a way or street open to public traffic; which resulted in one or more persons being killed or injured and in which at least one moving vehicle was involved" (COST 329, 2004). From 1973 onwards, material damage only crashes are no longer included in the official Belgian statistics. An *injury accident* is an accident occurring on a public road and involving at least one moving vehicle, recorded by the police and leading to one or more persons being injured (fatal or non-fatal). To define a *fatal injury*, Belgium follows the standard definition given by UN/ECE (Economic Commission for Europe, 2005): any person who was killed outright or who died within 30 days as a result of the accident. For a *serious injury*, a hospitalisation of at least 24 hours is required. A *slightly injured* is a person injured in an accident who is not fatally or seriously injured.

The majority of road crashes is settled by mutual agreement between the parties involved. In this case, there is no police intervention. This is especially true for crashes with material damage only, but one can assume that also part of the injury crashes is never reported. If a police officer intervenes, an official report of the crash is made. This report is an objective description, which serves as a basis for further legal and administrative treatment of the crash.

For injury crashes, an analysis form ("Analyseformulier voor verkeersongevallen met doden of gewonden", abbreviated as VOF) is completed, based on the information in the official report. This form is, in ideal circumstances, filled in at the location of the crash. Especially the description of the manoeuvres and the location is preferably recorded immediately. After the crash, the analysis form is send to the National Institute for Statistics (NIS). The NIS is responsible for the further treatment and publication of the crash data.

4.3.2 Data problems

Although the police forces are supposed to assess the scene of the crash, one has to come to the conclusion that the corresponding statistics are not complete. It is generally known that not all traffic accidents that should be reported to the police are also actually reported. This phenomenon, called under-registration, is a general problem in most countries, and its magnitude is hard to determine. In the past, the partial registration of road crashes has not always been considered as a major problem, as the registration was not entirely done for road safety research (OECD, 1997b). In (COST 329, 2004), several causes for under-reporting are given. First, the involved road users may not be aware of the fact that the accident should be reported to the police. This may be the case when the injury is so minor that it does not require any further attention. Second, it may be that the injuries are not apparent at the time of the crash, but only appear some time later. Consequently, the victim is not regarded as "lightly injured", and no reporting is done. Third, it is possible that people forget reporting the accident because they are rescuing victims, calling medical assistance or providing any other kind of help at the scene of the accident. Often, everyone expects others to report the accident. Fourth, the persons involved in the accident might have their own reasons for not reporting it to the police. This may be because of fear of prosecution, or an expected increase in the future insurance payments when no-claims bonuses are missed. Fifth, the police files do not contain all accidents. Depending on the kind of crash, the police will or will not intercede at the location of the crash. This is especially true for damage only crashes, where normally no police intervention is required. Also police work load issues may lead to erroneous or incomplete reporting. Sometimes, there may even be a kind of selectivity in the reporting, because of a special interest in particular crashes or in non-reporting.

If a registration of a crash is done, also another form of under-reporting might be present. With the current analysis form, some important variables are not registered at all. Questions concerning the responsibility for the crash or the efficiency of the emergency services are not asked. Of course it is not always possible to determine the essential (causal) variables at the place and time of a crash, but a well-considered analysis form should at least give an indication of all major aspects known from previous research. As it is practically impossible to register the crash circumstances in all details, it is necessary to make a profound selection of the registered variables, in agreement with the police officers who do the work. This should make it possible to balance the feasibility and the desirability of the registration of certain information in the crash data base.

The under-reporting issue can lead to several problems in the road safety statistics (COST 329, 2004). First, the modal split influences the average reporting rate. If the level of reporting is very low for bicycle crashes, then the completeness of the registration will depend on the distance travelled by bicyclists and the corresponding number of crashes.

Second, the reporting rate can vary by the type of road users involved. In a comparison of police and hospital statistics for Norway, Lereim (1984) found differences in registration of about 50% for accidents with motor vehicles, and 30% for accidents without motorised road users. The level of reporting was highest for four-wheel motorized road users and pedestrians, and very low for bicycle-only accidents. In (De Mol & Boets, 2003), an overview is given of comparison between hospital data and police records for the SUN countries (Sweden, United Kingdom and the Netherlands) and the United States. The objective of the research was to obtain more complete and representative accident statistics by linking and comparing hospital data with police records. The research showed that not all countries analysed in the report are working in the same manner. In the Netherlands only a periodic coupling of the data sources is done, whereas in Great Britain the coupling is limited to specific regions. In the United States, efforts are made to create a data warehouse structure, while Sweden automated the coupling of medical reports and accident data.

Third, the level of reporting may vary with the severity of accidents. There is a clear relationship between the gravity of a crash and the percentage of crashes reported. In a British study (James, 1991), crash data from police reports and hospitals were compared. It turned out that crashes with fatalities were almost completely reported, whereas registration was between 30% and 90% for crashes with serious injuries and between 20% and 80% for crashes with light injuries. It was also found that the reporting rate depends on the road user group involved. The rate is highest for vehicle occupants and pedestrians, and lowest for bicyclists. In (Alsop & Langley, 2001), under-reporting of seriously injured motor vehicle traffic crash victims in New Zealand was investigated with respect to crash, injury, demographic and geographic factors, by comparing police records with hospital data. They found that in 1995, less then two thirds of all hospitalised victims were recorded by the police. Also, the reporting rates varied significantly by age, injury severity, length of stay in hospital, month of crash, number of vehicles involved, whether or not a collision occurred, and geographic region. In many other papers, hospital and police data are compared (Harris, 1990; Lopez et al., 2000; Rosman & Knuiman, 1994). From these studies, it is clear that the more serious the injury, the higher the probability of being registered.

Fourth, for the registered crashes, one is not always certain that the reported data effectively describe the crash. Some characteristics may be registered incorrectly, which makes it hard to recover the rights and the wrongs of the crash. Examples are the exact location of the crash or the fact whether or not seat belts were used. Other questions on the analysis form leave so much room for subjective interpretation that two police officers might come up with a completely different answer for the same question. This may lead to biased results concerning the gravity of the crash, often based on the experience of the police officer. Some issues are so vaguely described or so complex to evaluate that a correct answer is almost impossible. Also the length of the analysis form and the number of questions may lead to incorrect or incomplete crash registration.

These shortcomings in the registration of crash data show that the process might be improved in various ways. A main issue is of course that registration should be done shortly after and at the location of the crash. This may exclude important registration errors or incomplete forms. A related issue is the once-only registration of the data. If crash data is recorded partly at the accident, partly in the office, in the hospital or at the insurance company, the probability of errors and inconsistencies increases. Further, data sources at hospitals and insurance companies should be made accessible for scientific research. Given the state-of-the-art possibilities of database management, it should be possible to extract tables that do not suffer from privacy issues. Researchers are not interested in the private information on road victims (private address, number plate, etc.), but try to find patterns in the data based on general characteristics. It does not seem advisable to shield these data any longer from the research community merely for privacy reasons. Lastly, given the multidisciplinary character of road safety, it is desirable to have the possibility to link different data sources. In particular, it would be interesting to connect crash information with weather data, infrastructural properties or vehicle engineering. This allows the investigation of road safety in all its aspects.

4.3.3 Improving data sources

The Belgian governments, both at the local and federal level, recognised the problems related to the crash data. In the past 5 years, many initiatives were started to improve the quality of data gathering and processing. First, the information gathering process is more and more supported by information systems. The registration of road crashes is made part of the ISLP system (Integrated System for the Local Police). This is an integrated system that allows the single registration of crashes in one location, minimizing the probability of inconsistencies in the data. The system is gradually being introduced in the various police zones of the country. Also the police officers are more and more equipped with electronic devices that facilitate the registration of crashes, which will considerably enhance the quality of the crash data in the future. Also the National Institute of Statistics puts in great efforts to register the road safety data as correctly as possible and to speed up the publication of the statistics.

Second, in the research project "Innovative Spatial Analysis Techniques for Traffic Safety" (Steenberghen et al., 2003), various universities worked together on the localisation of road crashes by means of GIS-based applications. Starting from the raw crash data from the National Institute of Statistics (NIS), the regional road administrations in Flanders and the Walloon region check the location attributes of crashes, and correct them if necessary. The localisation of crashes starts from either information on the street name and house number (address matching) or on the road number and hectometre number (dynamic segmentation). The results from this project can be used to optimize the investments in road infrastructure, directed by prioritization of the treatment of hazardous roads and intersections.

Third, the State-General for Road Safety (Staten-Generaal voor de Verkeersveiligheid, 2001) pointed out the necessity of having reliable and complete sources of crash data. Subsequently, the research project "Exploitation of Road Safety Data" was funded by the Federal Research Policy (Kinet et al., 2004). In this project, the existing registration systems were analysed and compared with European initiatives on standardising road crash information. Next, the needs of the involved parties (police, government, road maintenance authorities, researchers, etc.) were studied and translated into an IT format that can be used as a basis for the development of new information systems for crash data registration.

Fourth, initiatives on the European level demand attention for the data problems. In the eighties, the "International Road Traffic and Accident Database" (IRTAD) was set up as a tool for international comparisons and national road safety development assessment. To date, the IRTAD database is an example of an international collection of crash data, with detailed, up-to-date and consistent time series data on road safety. It is currently managed by the joint OECD/ECMT Transport Research Committee. Another European initiative is CARE (**C**ommunity database on **A**ccidents on the **R**oads in **E**urope). This database started in the nineties and contains information on road crashes resulting in death or injury. Compared to other databases, CARE offers a high level of disaggregation, with data on individual accidents. The purpose of the CARE system is to provide a powerful tool which would make it possible to identify and quantify road safety problems throughout the European roads, evaluate the efficiency of road safety measures, determine the relevance of Community actions and facilitate the exchange of experience in this field (CARE, 2004). The data of the CARE database has extensively been used in various road safety studies, and is currently at the basis of the activities of the SafetyNet consortium . This is an integrated project that serves as a "Road Safety Observatory", as it was expressed by the European Commission in its White Paper on road safety (European Commission, 2001). The objective of SafetyNet is the creation of a co-ordinated set of data resources that will support the Commission in its road safety policy. These European initiatives clearly indicate the increasing international awareness of the importance of road safety monitoring. More and more, national statistics will be used for international comparison among (European) countries. It is imperative to work on a coordinated policy with regard to standardisation, definition and dissemination of crash data.

4.3.4 Motivation for the use of official statistics

In the subsequent chapters, data on road crashes will be used in a diversity of formats. The data are, however, always derived from the official database on road safety, published by the Belgian National Institute of Statistics. There are some reasons for working with these data. The most obvious one is that this is the only available database on road crashes. It is very probable that hospital data and information from insurance companies are available as well, but at the time of writing, these data sets are not publicly accessible and not linked with the official crash database. Also, the official database, gathered from the police registrations, is the only one that, at least theoretically, should be complete. Every database from a hospital or an insurance company will only cover a subset of crashes. A second reason is to be found in the type of analysis chosen for our research. In (almost) all studies, time series data is used, either on yearly or monthly data. To obtain a data set that is large enough for statistical treatment, data on a considerable number of years should be available. This is the case for the official statistics, and it is questionable that the same level of detail would be available in hospital or insurance data. Third, as indicated in section 4.3.2 above, the quality of registration increases with the gravity of the crash. Data on fatalities and, to a lesser degree, crashes with seriously injured road users are assumed to be registered quite correctly. Therefore, focus will be on these groups of crashes and fatalities (although some models for slightly injured victims can be developed as well). Moreover, in the context of time series, it can be assumed that the level of under-reporting might be consistent over time. This has been investigated by Maas and Harris (1984) for the Netherlands. The authors compared the incomplete police data on road victims with hospital data. They found that the extent of under-reporting was constant in the late seventies, and that, in spite of some differences, a similar general data structure appeared in both databases. The authors concluded that the police data is reliable for time series studies, even for various transport mode / age group combinations. Although we are aware of the drawbacks and the limitations of the official statistics, they will be used throughout this document.

4.3.5 Description of the available data

Statistics on road safety in Belgium have been regularly published since the early fifties. However, some major changes have been made in the reporting practices throughout the years (BIVV, 2002). Until 1972, the number of fatalities was derived from the statistics on death causes, based on the number of death announcements. From 1973 onwards, these statistics are based on the number of victims reported by the office of the public prosecutor to the National Institute of Statistics. In 1971, the definition of a seriously injured victim was changed to the one that is still used today. From 1973 onwards, a new registration system was brought into use. In 1990 another change occurred in the registration system, but no changes were made to the definitions. To make sure that enough years of data can be used in the analyses, without major changes in definitions or registration procedures, it is decided to use the data on road crashes from 1973 onwards.

At the time of writing, these statistics are available up to and including the year 2004. Aggregated statistics on a yearly or monthly basis are available for the number of victims killed, seriously injured or lightly injured, and the corresponding number of crashes. For each year, detailed statistics are available on the specific circumstances of the crash. In particular, it is possible for these years to analyse subgroups of crashes or road users.

In the introductory chapters, it was emphasised that the models developed here are macroscopic in nature. Therefore, road crashes will not be investigated in all detail. Also, it was mentioned that, whenever possible, road safety data will be analysed in combination with a measure of exposure. The road safety data that can be used and the level of detail of the analysis will therefore be determined by the availability of exposure data. On a yearly basis, exposure data are available in the form of the number of vehicle kilometres in total, per type of road and per type of road user. Corresponding sets of crash data include the total number of crashes and victims, the number of crashes and victims per type of road (motorways, regional roads and local roads), the number of crashes and victims per type of road user (personal cars, motorcycles, buses and trucks) and the number of crashes and victims per gender and age category.

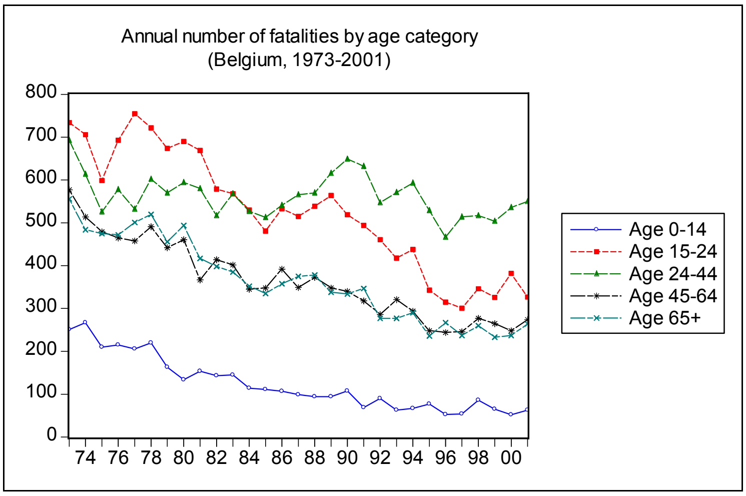

FIGURE 21: Annual number of fatalities by age category

As an example, consider the evolution in time of the number of fatalities in 5 different age groups: 0-14, 15-24, 25-44, 45-64 and 65+. The data are extracted from the official Belgian road safety statistics, and the curves are shown in FIGURE 21. Young persons (15-44 years old) clearly have the largest share in road fatalities for the last 30 years. Older and younger persons have a lower number of fatalities. This presentation of road fatalities shows the added value of analyses that make a distinction according to the properties of the road users. Some examples of this kind of analysis will be shown in Chapter 7.

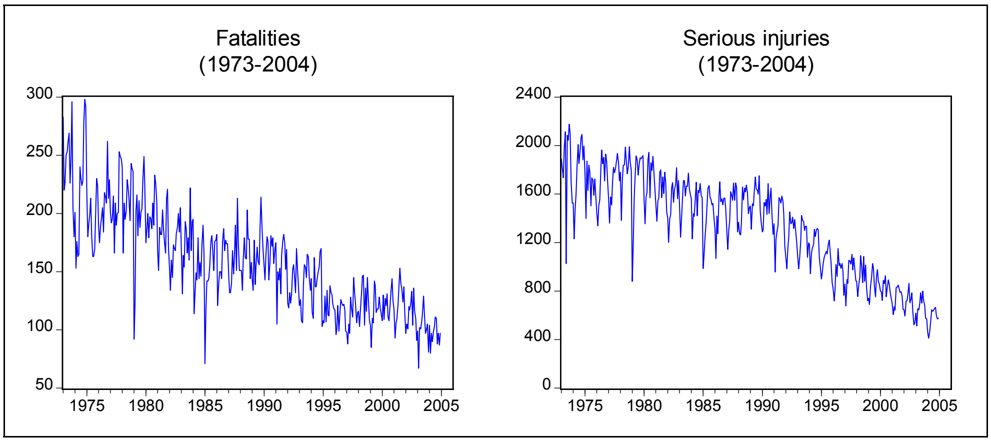

FIGURE 22: Number of fatalities and persons seriously injured in Belgium

On a monthly basis, there currently is no valid statistic available that can be used as an exposure measure. However, as shown in section 4.2.4, a monthly indicator of the number of kilometres driven can be derived for the whole of Belgium from monthly fuel sales, vehicle park and average fuel consumption. As this is an aggregated exposure measure, it can be used in the analysis of aggregated road safety, namely the total monthly number of crashes and victims in Belgium. The monthly time series for the number of persons killed or seriously injured are shown in FIGURE 22.

4.4 Explanatory variables

When explanatory models for road safety are developed, a set of covariates is needed that can be used to explain (part of) the variation in the number of road crashes or victims. The use of explanatory variables is, however, not always possible, for several reasons. If models are developed using yearly data, the number of observations is usually small, which seriously restricts the number of explanatory variables that can be used. If monthly data are used, the number of variables that can be included is generally larger. However, monthly time series on explanatory variables are often not available. When developing macro models, the researcher typically has to decide on a trade-off between model frequency and the number of variables in the model. In this section, an overview is given of several classes of explanatory variables: economic variables, weather conditions, laws and calendar variables. These variables are selected from various official sources, based on their quality, availability and accessibility. Note that in some models, exposure will be treated as an explanatory variable as well. The exposure measures that are available were discussed in section 4.2.

4.4.1 Economic variables

4.4.1.1 Fuel price

A first economic variable that will be considered is the average fuel price per kilometre. Data on fuel prices are available for many years (Ministry of Economic Affairs, Department of Energy). The fuel price used in the analyses is an average price calculated from the prices of the different kinds of fuel (medium, normal, normal unleaded, super, super 95, super 98, diesel, LPG) and includes all taxes. As prices are registered in nominal values, they partly reflect variations due to inflation. To remove this effect, real prices are calculated by correcting the prices for the changes in the consumer price index (with base year 2004). The real prices are then divided by the number of kilometres driven to obtain an approximation of the fuel price per kilometre. The time series of the average real fuel prices per kilometre is shown in FIGURE 23.

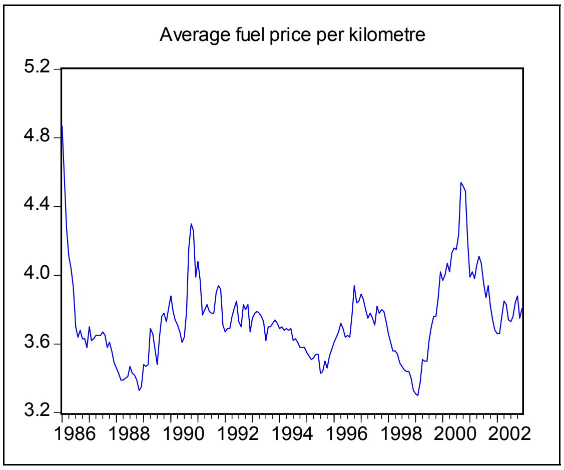

FIGURE 23: Average fuel price

4.4.1.2 Unemployment

A second economic indicator is the degree of unemployment (data from Eurostat, obtained from the Belgian National Bank, Belgostat). This variable is shown in FIGURE 24 for the years 1983-2005.

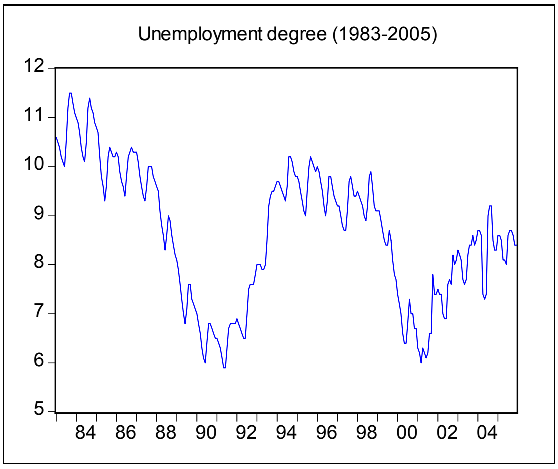

FIGURE 24: Degree of unemployment in Belgium

The degree of unemployment expresses the number of unemployed persons as a percentage of the active population. The unemployment variable is used as an economic indicator as well as a measure of social development in a country. It might have an influence on both the level of exposure and the level of road safety. The graph shows that unemployment is oscillating in periods of about ten years, together with smaller monthly (seasonal) fluctuations.

4.4.2 Weather conditions

The climatologic conditions in a country will undeniably have an impact on both the level of exposure and the level of road safety. Depending on the weather, the traffic can have a different composition. A sunny weekend may result in more recreational traffic than a normal weekend. Frost and snow may restrain people from travelling. These decisions of road users may in turn affect the level of road safety. But there is also a direct effect of the weather conditions on road safety. Due to slippery roads, rainy days may be more risky than sunny days, and road users may adapt their speed or driving behaviour accordingly.

To test for climatologic conditions in the models, four weather variables are considered, all measured on a monthly basis: the average temperature (in degrees Fahrenheit), the number of days with frost, the number of days with precipitation and the number of days with snow. The variables are gathered by the Belgian Royal Meteorological Institute and published by the National Institute of Statistics. The first three variables are recorded in the climatologic centre in Ukkel (in the centre of Belgium), while the number of days with snow is measured for the whole of Belgium. The four time series of climatologic variables are presented in FIGURE 25. Note that the climatologic variables are highly seasonal and might therefore be related to the seasonality in road safety data.

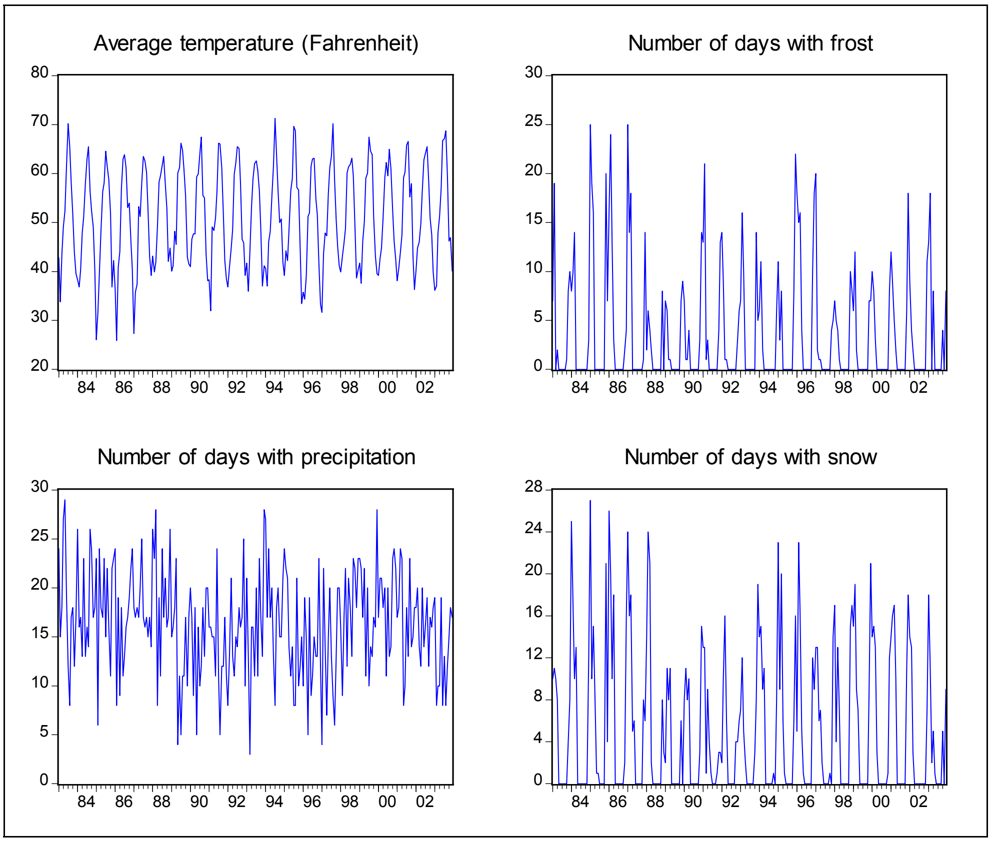

FIGURE 25: Climatologic variables for Belgium

4.4.3 Laws and regulations

Five dummy variables are used in the models to assess the impact of laws and regulations that came into force at a certain date. It is not always possible to isolate the effect of a single measure, because sometimes more regulations become effective at the same moment in time. Nevertheless it makes sense to test whether policy enforcement at a certain point in time is indeed effective. The regulations considered are summarized below. The main focus of the laws included in the models is on seat belt wearing, speed and impaired driving. For each of the dummy law variables, the value is equal to zero before the law, and switches to one as from its introduction.

– From June 1975 onwards, the use of seat belts is mandatory. In the beginning, the law stated that car drivers and passengers should use the seat belt if they are available. This means that the effect of this law has only become gradually visible.

- In January 1991, some regulations were introduced to improve the position of vulnerable road users in traffic. Among the list of regulations, one can find that (1) bicyclists of age less than 9 are allowed to cycle on pavements or raised shoulders, (2) vehicles should be extremely careful in the neighbourhood of crossings for pedestrians or bicyclists, (3) bicyclists are allowed to drive side by side (except in some specified circumstances) and (4) if a crossing for bicyclists is available, bicyclists should use it, taking account of possibly approaching cars. Finally, apart from the measures for vulnerable road users, rear seat passengers in cars are required to wear a seat belt.
- One year later, in January 1992, some regulations were introduced on the load of vehicles, cycling tourists and speed. A very important law is the speed limit of 50 km per hour in urban areas, and other specific speed limits for different user groups, like motorcycles and trucks. Also, 90 km per hour became the new speed limit at road sections with at least 2x2 lanes without a raised shoulder or any other separation of the opposing lanes.
- Starting from December 1994, the 0.05% maximum alcohol level was imposed and higher fines were written out for a 0.08% or higher alcohol level.
- In January 1998, the first speed cameras were installed at intersections.

4.4.4 Calendar variables

Calendar data offers the undeniable advantage of availability of the data. Some of the road safety models discussed in the subsequent chapters include one or more calendar variables.

4.4.4.1 Trend

In some models, a time variable can be used to reflect a linear trend. This variable simply equals the year or the time point of the observation.

4.4.4.2 Seasonal

Seasonal dummy variables can be used to handle seasonality in monthly data series. The seasonal pattern is represented by the variables JAN_t, FEB_t, …, NOV_t, where each of these equals 1 in the given month, and zero otherwise. As, in general, *k* groups can be distinguished with *k*-1 dummies, no variable is included for the month December. Thus, for monthly data, 11 dummy variables are needed. The coefficients reflect the average difference in the dependent variable between the given month and the omitted month. Note that by the use of seasonal dummy variables, it is implicitly assumed that the seasonal component is identical in each year.

4.4.4.3 Leap year

The seasonal dummy variables can take care of the length of the months. However, the effect of a leap year is not taken into account, as it does not occur every year. Therefore, an extra variable can be introduced, taking a value of one in February of the leap year, and zero otherwise.

4.4.4.4 Trading day pattern

The number of road crashes and the corresponding number of victims may vary according to the day of the week. To correctly forecast the number of road crashes and victims, the number of Mondays, Tuesdays, etc. in each month is taken into account. Trading day effects reflect variations in monthly time series due to the changing composition of months with respect to the number of times each day of the week occurs (Bell & Martin, 2004). In each month, there are four weeks plus usually one, two or three more days. Each weekday occurs at least 4 times in a month, but some days will occur 5 times. The composition of the calendar will affect the data for the month. If, for example, a shop is only open on weekdays, then sales will be higher if some weekdays occur five times in a month. Especially monthly time series that are totals of daily activities (like the records of road crashes), are often influenced by the weekday composition of the month. Details on the construction of trading day models are given in (Findley et al., 1998), based on (Young, 1965), (Cleveland & Grupe, 1983) and (Bell & Hillmer, 1983). Trading day patterns can be included in different forms, as is shown in (Soukup & Findley, 2000). In this manuscript, a parsimonious form of the trading day pattern is used, as provided by (Gómez & Maravall, 1996). This form captures the trading day effect in one variable. Since we are primarily interested in the difference between weekdays and weekends, we propose a trading day variable TD_t that is defined as follows :

$$TD_t = \sum_{j=Mon}^{Fri} D_{jt} - \frac{5}{2} \sum_{j=Sat}^{Sun} D_{jt} , \tag{11}$$

where D_{jt} indicates the number of times the j-th day occurs in month t. This formula forces the weights of the different days of the week to sum up to one. It also requires that the weights for all weekdays and all weekend days are the same. For example, if the parameter for the trading day variable TD_t equals -0.005, then each weekday is given a negative weight of -0.005, while each weekend day is weighted as (-2.5)*(-0.005)=0.0125, which indicates that months with more weekend days may be more dangerous than months with more weekdays.

4.4.4.5 "Special day" measure

Some periods of the year are characterized by more or other traffic than other periods. The traditional holiday periods often cause days of holiday rush, especially on the highways to the tourist locations in Belgium and in the neighbouring countries. Also holidays related to Christmas and Easter and the public holidays cause very specific traffic patterns. It is to be expected that periods with this "special" traffic have a different road safety profile than other periods.

TABLE 4: Contents of the "Special day" measure

Celebration days		**Fixed Christian holy days**	
New Year	1 Jan	Assumption	15 Aug
Labour day	5 May	All Saints' Day	1 Nov
Flemish holiday	11 Jul	All Souls' Day	2 Nov
National holiday	21 Jul	Evening of Christmas Eve	24 Dec
Walloon holiday	27 Sep	Christmas Day	25 Dec
Cease-fire	11 Nov	Second Christmas Day	26 Dec
New Year's Eve	31 Dec		
Holiday periods		**Summer holidays**	
Autumn half-term	WE 1: Fri, Sat	First school day	
	WE 2: Sat, Sun	Last school day	
Christmas vacation	WE 1: Fri, Sat	First day of summer holiday	
	WE 2: Fri, Sat, Sun	Last day of summer holiday	
	WE 3: Sat, Sun	First summer holiday WE (Fri, Sat, Sun)	
Spring half-term	WE 1: Fri, Sat	Last summer holiday WE (Sat, Sun)	
	WE 2: Sat, Sun	Last July weekend (Fri, Sat, Sun)	
Easter holidays	WE 1: Fri, Sat	First August weekend (Fri, Sat, Sun)	
	WE 2: Fri, Sat, Sun		
	WE 3: Sat, Sun		
	Easter Monday		
Ascension holidays	Wed, Thu, Fri, Sun		
Whitsun(tide) holidays	Fri, Sat, Mon		

To account for these differences, a variable is developed based on the "density indicator" from the Flemish Automobile Association (VAB). The measure classifies each day of the month as either a "normal" or a "special" day, and then sums up the number of special days over the month. These days are selected on the basis of the experience of the VAB road experts and the spread of public holidays, religious feasts and school holidays over the months. In TABLE 4, the rules on which this variable is based are shown. The variable takes into account the celebration days, Christian holidays, holiday periods and summer holidays. Information on holidays is taken from the official holiday tables published by the Ministries of Education. Especially for holiday periods, the special days can change over the months. For example, Easter holiday is moving every year between March and April. The corresponding number of special days therefore also changes with the months and the years. As monthly data are analysed, these shifts over the months may affect road safety.

The variable is constructed on a monthly basis for the years 1973-2008. In a model with monthly road safety data, this variable can indicate the effect of an additional special day in a month. The moving Easter period may lead to more (or different) traffic in April of year *T* compared to the same month in year *T*+1. Because these effects change from one year to another, they cannot be captured by a deterministic seasonal pattern, and therefore should be accounted for by an additional variable. This is the objective of the "special day" variable.

4.5 Conclusion

In this chapter, an overview is given of the data that are available in Belgium for the development of macroscopic road safety models. This overview is by no means a complete list of variables. It is at best an attempt to gather useful information of various data sources. Every kind of research has specific data needs, and this overview is entirely focused on data for macroscopic modelling.

From the list of variables considered in this section, it is clear that road safety is a multidisciplinary issue. Not only information on road safety is needed, but also data sources on population, legal issues, weather, etc. might provide useful information for road safety research on an aggregated level. After all, road crashes do not happen in a cocoon that is screened off from the environment. On the contrary, road safety is an essential part in the combined action of road demand and road supply, and should be analysed as such.

This line of reasoning is the underlying motive for the sometimes discouraging quest for data. In Belgium, there is no tradition in the longitudinal analysis of road safety. It must be said that many data sources do exist, but they all suffer from two main drawbacks. First, the data is not always easily accessed. Especially for older series, it is not clear where they can be found and who is responsible. For more recent data, the government is working to make up lost ground. Many electronic sources are set up, which increases data quality and certainly improves the accessibility. Second, if the data are available, they are mostly not gathered with the purpose of road safety research. Often the frequency is not adapted, or the units are difficult to interpret in the context of road safety. This is one of the reasons for creating a monthly measure for exposure based on available fuel deliveries. The work of combining data is therefore as hard as the gathering itself.

The data discussed here will be used in various analyses in the next chapters. The choice of variables in a specific model is determined by the scale and the objectives of the model. Often, these factors go hand in hand. For example, a model developed on yearly data will usually be less explanatory than a model on monthly data, simply because the frequency of observations is larger and thus more explanatory variables can be included. This finding does not give a verdict on the quality of any of these models. For example, it might be more meaningful to test for certain aggregated effects on yearly data than on disaggregated monthly data. The choice for a certain model therefore depends on the objectives and the availability of data. The results of the models should then be interpreted and used with these restrictions in mind.

Chapter 5 Aggregated models for yearly exposure and risk

5.1 Introduction

One of the main objectives in road safety is to describe and monitor the number of fatalities in road crashes by looking at trends over the years. Many models have been developed to explain the long-term evolution in the number of fatalities (COST 329, 2004). Aggregated models, on a countrywide level, allow gaining a better insight in the evolution of fatalities over time.

The annual number of fatalities is still an important indicator of the level of road safety in a country. Policy makers use these statistics to indicate past trends and to create rough indicators of the future evolution. In many countries, the government formulated clear long-term quantitative objectives in terms of the number of fatalities. In Belgium and in Flanders, as well as on the European level, the goal is to half the number of fatalities by 2010. The Belgian States General for road safety (Staten-Generaal voor de Verkeersveiligheid, 2001) recommended the objective of halving the number of fatalities by 2010, compared to the average number of fatalities for the years 1998 – 2000. This boils down to a maximum of 750 fatalities in 2010. In Flanders, the government expects making up arrears compared to the best performing countries in Europe by 50%, taking into account the own ambitions of these leading countries in terms of road safety. Broadly speaking, this implies a reduction in the number of fatalities by 50% compared to 1999, or a maximum of 375 fatalities in 2010 (Ministerie van de Vlaamse Gemeenschap, 2001). In (Van Schagen, 2000), an overview is given of the quantitative targets formulated in some OECD countries. These targets are in terms of annual fatalities, and should therefore be supported by models on the same level of aggregation.

The choice for annual data offers the advantage of having the data readily available from the official government statistics. Data on road crashes and on the annual level of exposure are published regularly. In comparison with a more disaggregated analysis (on monthly data for example), the effort needed to obtain the data is much smaller. However, the number of annual observations is usually small. For Belgium, annual data that can be compared over time is available from 1965 onwards, for both the number of fatalities and the number of vehicle kilometres. A low number of observations implies that models should be kept simple, in the sense that number of estimated parameters should be reasonable compared to the number of observations in the data set. Also, introducing too many variables will lead to unstable predictions because of multicollinearity problems. Therefore, models on annual data are often purely descriptive, without the ambition of explaining the observed developments in the model. Usually the models are assessed in terms of goodness-of-fit and forecasting accuracy, but no explanatory power is expected.

In this chapter, some deterministic and stochastic models are introduced. The proposed models are aggregated, both in time (yearly data) and in space (Belgium as a whole). In the introductory chapters, these models have been extensively described in a historical perspective.

5.2 Models to predict the annual number of fatalities

In this section, some of the deterministic and stochastic models described in Chapter 2 will be developed for the Belgian accident and exposure data. The first model is based on the pioneering work of Oppe (1989; 1991), and describes the trends in the number of fatalities by a decomposition in risk and exposure. Based on some extensions proposed in the literature, the Oppe model is then adapted to allow for more general curves for risk and exposure. Various functional forms are considered and interventions are added for the major road safety measures (seatbelt use, speeding and impaired driving). Also, some stochastic extensions of the model are shown. Structural models based on (Lassarre, 2001) are presented and the LRT framework, developed in (Bijleveld, Commandeur, Gould et al., 2005), is used to build models in which both fatalities and exposure are treated as unobserved components.

Using the final models, predictions for the Belgian fatalities up to 2010 will subsequently be derived and compared with the quantitative targets formulated by the government. The objective of the study is to estimate a model that is acceptable both in terms of statistical fit and prediction accuracy.

5.2.1 A starting point for Belgium: the Oppe approach

In a first attempt to model the relation between fatalities, exposure and risk, the approach proposed by Oppe is followed. As explained before, the Oppe model assumes that a learning process is behind the evolution in road safety. That is, society learns to control the undesirable consequences of the traffic system in an exponential way (Cameron, 1997). More formally, the number of fatalities (F_t) can be seen as the product of the level of exposure (V_t) and the level of risk (R_t), or $F_t = V_t \times R_t$. The following notation can be used to model an exponential curve for the risk and a logistic curve for exposure:

$$R_t = \exp(\alpha t + \beta) + \varepsilon_{R,t} \tag{12}$$

$$V_t = \frac{V_m}{1 + \exp(-(at + b))} + \varepsilon_{V,t} \tag{13}$$

In these equations, α, β, V_m, a and b are parameters to be estimated, while $\varepsilon_{R,t}$ and $\varepsilon_{V,t}$ are the residual terms. The years are indicated by t. In the risk function, α indicates the learning rate of the country. In the exposure function, the parameter a is the growth rate over time, with V_m being the upper bound. This model will be applied to the Belgian road safety situation, using official statistics on fatalities and exposure. The fatalities are the yearly number of persons killed in road traffic. The measure of exposure on a yearly basis is the number of vehicle kilometres, as explained in Chapter 4.

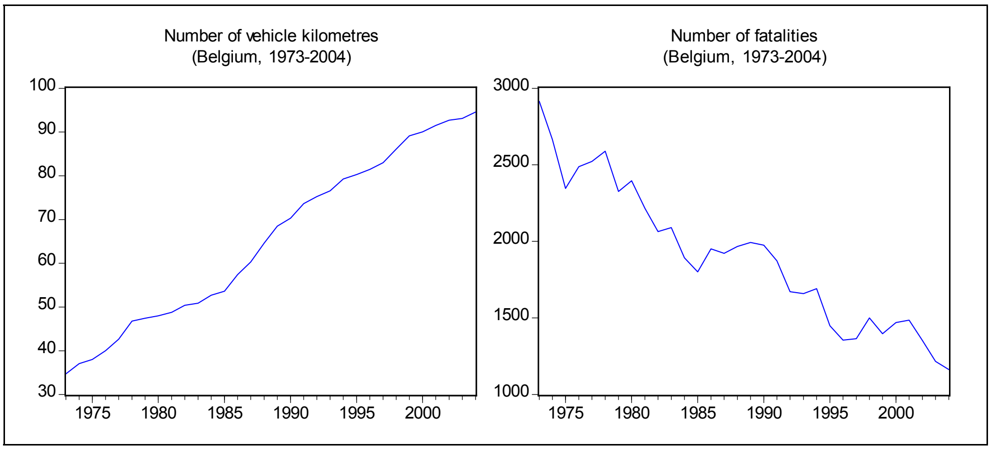

FIGURE 26: Number of fatalities and number of vehicle kilometres

The observed number of fatalities and the number of vehicle kilometres are shown in FIGURE 26. For each variable, three observations at the end of the series are left out of the analysis to test the quality of the forecasts afterwards. Estimation is done in EViews 5.1 (QMS, 2004). In TABLE 5, the parameter estimates are shown.

A comparison between V_m and the latest available value for the amount of traffic gives an indication of the increase in traffic that is still to be expected. In 2004, the number of vehicle kilometres is at 94.56. Compared to the value of 145.66, a growth in traffic of more than 50% before saturation is expected. The parameter *a* is the growth rate of exposure over time, and is around 6%. The larger *a* is, the faster traffic volume grows over the years. In the risk function, α shows the (negative) growth rate per year in risk, which is also around 6%.

TABLE 5: Parameter estimates for the Oppe model

	Coefficient	Std. Error	t-Statistic	Prob.
α	-0.0629	0.0017	-37.1714	0.0000
β	128.4149	3.3494	38.3393	0.0000
V_m	145.6568	15.9784	9.1159	0.0000
a	0.0614	0.0057	10.7932	0.0000
b	-122.2090	11.0994	-11.0104	0.0000

The estimates for the yearly number of fatalities are now derived by multiplying exposure and risk. The results are shown in FIGURE 27. The dots are the observed values, the full lines are the estimated outcomes from the models. The dotted lines represent the 95% confidence intervals. Note that the confidence interval for the number of fatalities is calculated as the product of the confidence intervals for exposure and risk, and is therefore only approximately correct. The curves show the decisive factors in the evolution of traffic safety. The level of the number of fatalities depends on the increase in exposure and the decrease in risk. It is clear that the number of fatalities will decrease only when the decrease in risk is larger than the increase in exposure. However, the better the traffic safety situation, the more efforts it will take to obtain a decrease in risk.

5.2.2 Extending the Oppe approach

In general, the models seem to fit the data fairly well. It is known that the Oppe model is an interesting starting point for the analysis of risk and exposure on a macro level. However, there is still room for improvement. We see that the model for fatalities shows some deviances from the observed values for certain years. Moreover, some of the assumptions that are at the basis of the model might not be valid in practice (COST 329, 2004), as will now be investigated.

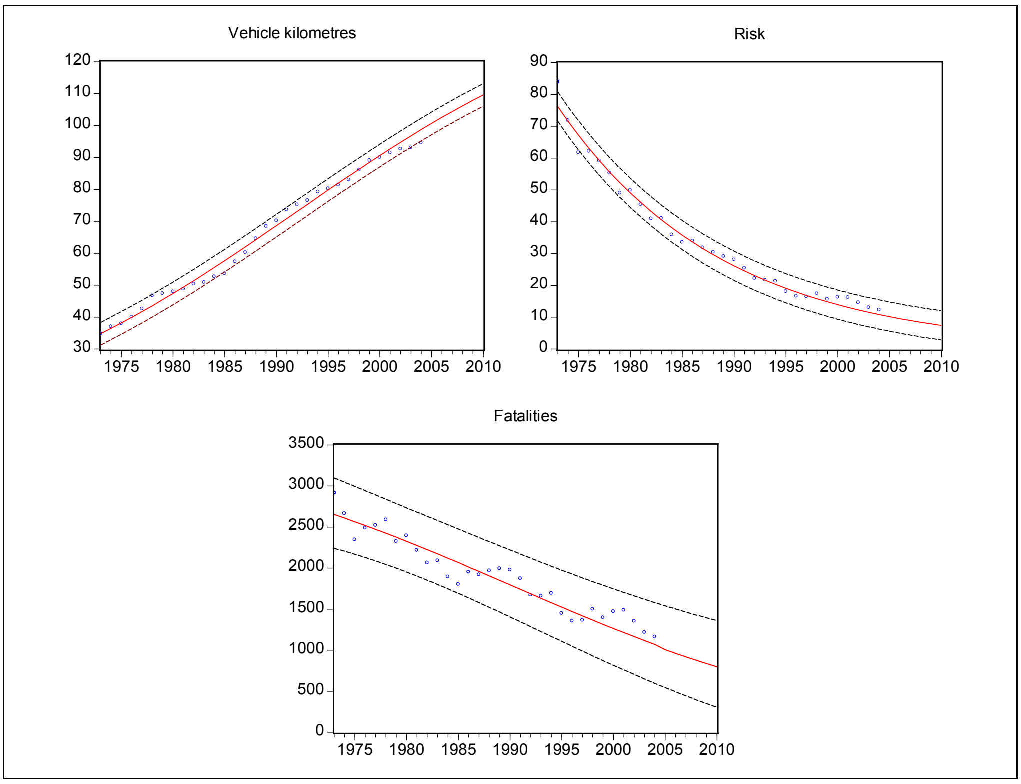

FIGURE 27: Results from the Oppe model

5.2.2.1 Introducing serial correlation terms

First, the curves presented in FIGURE 27 show that the observed values for exposure are systematically overestimated in some periods and underestimated in others. This indicates a problem of autocorrelation in the models, which affects the estimated confidence intervals and is not treated in the Oppe model. The same issue was mentioned in (Broughton, 1991), where the Cochrane and Orcutt method (Cochrane & Orcutt, 1949) was used to eliminate autocorrelation. Another approach is to include autoregressive terms in the models, which are estimated simultaneously with the other parameters in the model. For example, consider the model $y_t = f(x_t, \beta) + u_t$, where f is a general (nonlinear) function of the covariates x_t and the parameters β. If it is assumed that the residuals u_t follow an autoregressive process of order 1, then the model is extended such that $u_t = \phi\, u_{t-1} + \varepsilon_t$ can be estimated simultaneously with the parameters β. To test for serial correlation in the residuals, the Box-Ljung Q-statistic is used at different lags. This statistic is given by:

$$Q_K = n(n+2)\sum_{k=1}^{K}\frac{r_k^2}{n-k} \tag{14}$$

Here, r_k is the lag-*k* autocorrelation coefficient, *n* is the number of observations and *K* is the maximum lag being considered. Under the null hypothesis of no serial correlation, this test statistic is asymptotically distributed as $\chi^2(K\text{-}M)$, where *M* equals the number of parameters estimated in the model. Large autocorrelation coefficients lead to a high Q-statistic. A high value therefore indicates significant autocorrelation and thus rejection of the null hypothesis. More details on serial correlation will be given in section 6.2.1.3 on ARIMA model diagnostics.

5.2.2.2 The Richards curve for exposure

Second, there are some functional form issues that should be mentioned. The logistic exposure curve reflects an upper bound in traffic growth, but is restricted in the sense that it assumes a perfect symmetric behaviour. This assumption may be relaxed by testing a more general class of S-shaped curves. In (Commandeur, 2002; Commandeur & Koornstra, 2001), the Gompertz curve is used for exposure instead of the logistic curve. The Gompertz as well as the logistic curve can be seen as special cases of the Richards curve, which is a very general class of S-shaped curves. In its most general form, this curve can be expressed as (Pereira & Pernías-Cerrillo, 2005):

$$V_t = V_m\left[1+\mu\exp\left(-\gamma(1+\mu)(t-\tau)\right)\right]^{-1/\mu} \tag{15}$$

In this function, V_m, μ, γ and τ are parameters to be estimated. V_m is the upper limit (or saturation level) of the Richards curve. The time period at which the curve has an inflection point is given by τ, and γ is the relative growth rate at time $t = \tau$. The shape parameter μ allows for an asymmetric curve. The logistic curve is symmetric, that is the inflection point occurs at $V_t = V_m/2$. As it does not allow any asymmetry, the shape parameter of the logistic curve is equal to one, resulting in:

$$V_t = V_m\left[1+\exp\left(-2\gamma(t-\tau)\right)\right]^{-1} \tag{16}$$

The Gompertz curve can be asymmetric, with a value at the inflection point equal to $V_t = V_m/\exp(1)$. It is a special Richards curve in which the shape parameter is approaching zero (in the limit), resulting in:

$$V_t = V_m \exp\left(-\exp\left(-\gamma\left(t-\tau\right)\right)\right) \tag{17}$$

As an example, consider the curves in FIGURE 28. The saturation level is $V_m = 100$ and the inflection point is $\tau = 1983$. The relative growth rate equals $\gamma = 0.25$, and for the Richards curve μ is equal to 0.5.

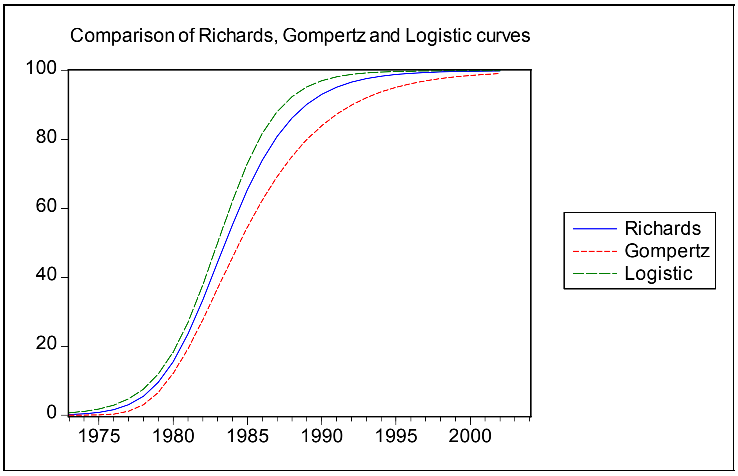

FIGURE 28: Richards, Gompertz and Logistic growth curves

As an extension to the Oppe model, the number of vehicle kilometres will be modelled using a Richards curve. The parameter estimates are shown in TABLE 6. The Richards curve estimates a maximum exposure level of 96, which is considerably smaller than the prediction from the logistic curve. The inflection point in the curve is around the end of 1990, with a relative growth rate of 3.5% at that time. Note that two autoregressive terms are included to make the residual series white noise. That is, the residuals are estimated simultaneously with the other parameters as $u_t = \phi_1 u_{t\text{-}1} + \phi_2 u_{t\text{-}5} + \varepsilon_t$, where ε_t is assumed to be a white noise process.

The logistic and Richards exposure curves can now be evaluated in terms of model fit and forecasting accuracy. The model fit is assessed by the Akaike Information Criterion or AIC (Akaike, 1973), calculated as (QMS, 2004; Quantitative Micro Software, 2004):

$$AIC = -2\left(l/n\right) + 2\left(k/n\right) \quad (18)$$

In this formula, *n* is the number of observations used in the model, *l* is the value of the log-likelihood function and *k* is the number of parameters estimated. A lower AIC value implies a better fit of the model.

TABLE 6: Parameter estimates for the Richards exposure curve

	Coefficient	Std. Error	t-Statistic	Prob.
V_m	95.93135	6.990884	13.72235	0.0000
μ	6.804093	4.114639	1.653631	0.1118
γ	0.035229	0.001404	25.09038	0.0000
τ	1990.819	1.327394	1499.796	0.0000
ϕ_1	0.656798	0.117889	5.571303	0.0000
ϕ_2	-0.347229	0.111233	-3.121642	0.0048

The forecasting accuracy is given by the Mean Absolute Percentage Error (MAPE), calculated as:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|PE_t\right| = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - F_t}{Y_t}\right| \quad (19)$$

Here, *PE* is the Percentage Error, based on the difference between the observed value Y_t of the series to be modelled and the corresponding estimated value F_t. The lower the MAPE, the better are the forecasts of the model.

TABLE 7: Model fit and prediction accuracy for exposure models

	AIC	MAPE	Q_1	Q_6
Logistic exposure	4.12	3.56%	19.60 (0.000)	50.37 (0.000)
Richards exposure	2.93	0.63%	-	5.251 (0.262)

To test both model fit and forecasting accuracy, the data set was split in two. The first part, observations from 1973 up to 2001 were used in the modelling stage. The second part, from 2002 up to 2004, was only used to evaluate the forecast accuracy. In TABLE 7, the AIC and MAPE values are given for both exposure models, together with the Q-statistics for serial correlation at order 1 and 6 (also other orders were tested). The Richards model yields a better fit and according to the MAPE values, has a much better prediction accuracy. The Q-statistics indicate the presence of serial correlation for the logistic exposure, but not for the Richards exposure.

5.2.2.3 Adding a constant to the risk curve

Another assumption that is often questioned is the fact that the exponential curve for the risk, R_t, implies that risk will continue to decrease to zero (COST 329, 2004). This assumption can be relaxed by testing a constant term in the risk function. On the Belgian data, the added constant was significant, slightly improving the fit of the model. The Q-statistics indicate the absence of serial correlation in both risk models. Also, the prediction accuracy on the most recent years of the model with a constant was significantly higher, as can be seen in TABLE 8. Although the exponential decrease in risk over time is in itself a rather strict assumption, the risk curve in FIGURE 26 shows that an S-shaped curve would not be appropriate for the Belgian data. The more general class of (logistic) risk functions that has been suggested in the literature (COST 329, 2004) will therefore not be discussed here.

TABLE 8: Model fit and prediction accuracy for risk models

	AIC	MAPE	Q_1	Q_6
Exponential risk	4.60	13.10%	0.804	4.532
			(0.370)	(0.605)
Exponential risk + constant	4.43	6.06%	0.571	6.260
			(0.450)	(0.395)

5.2.2.4 The modified model

As a comparison with the Oppe model, FIGURE 29 shows the new estimates for exposure, risk and fatalities, together with the observed values and the 95% confidence intervals. The remark concerning the accuracy of the confidence interval for the number of fatalities, made for the Oppe model, also applies here. The estimated values in the figures are one-step ahead forecasts of the dependent variable. For the out-of-sample periods (2002 - 2010), the ARMA errors are assumed to be zero to allow one-step ahead forecasts for these periods.

The graph for exposure illustrates that the possibility of having an asymmetric curve especially improves the fit at the end of the series, as the Richards curve clearly flattens off in line with the observed values.

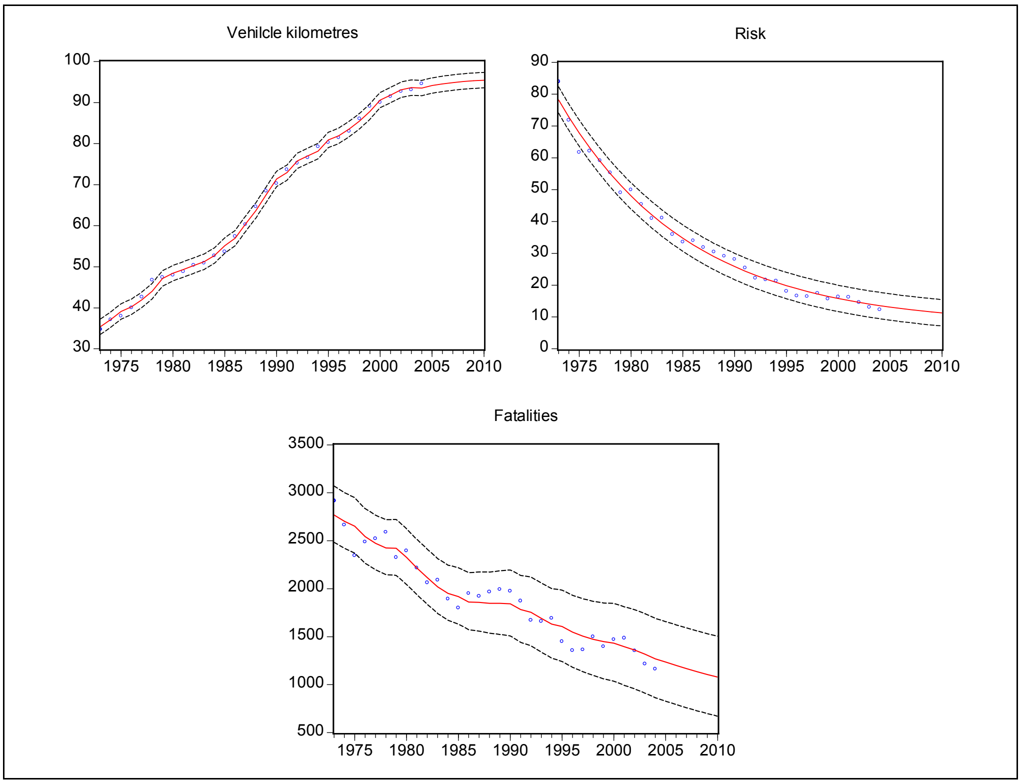

FIGURE 29: Modified Oppe model for exposure, risk and fatalities

The Richards curve is, at first sight, more realistic for short term predictions of exposure than the logistic curve. The smaller growth in exposure results in a more conservative prediction of the fatalities. For 2010, the original Oppe model predicts 797 fatalities, while the modified model ends up with 1079 fatalities. According to the Oppe model, the objective of at most 750 fatalities in 2010 (which is a 50% reduction compared to the average number of fatalities for the years 1998, 1999 and 2000) will more or less be achieved, but clearly this is too optimistic, as the modified model is felt to be more appropriate. More efforts are needed to increase road safety by 2010. However, the most recent road safety statistics for 2003 and 2004, which were not used to develop the model, are hopeful and indicate an even stronger decrease in the number of fatalities.

5.2.3 Alternative risk models

Although the modified model presented in the previous section is an improvement over the original Oppe approach for the considered data, there is still room for further improvement. In the literature, a new generation of risk models has been presented (COST 329, 2004). Apart from the improvements presented above, another set of topics will now be considered.

First, the Oppe model starts from the fundamental relationship $F_t = R_t \times V_t$. This inherently assumes that risk and traffic growth are proportional to one another. To test the validity of this assumption on the Belgian data, the relation is extended with a power for V_t. The more general relationship between exposure, risk and fatalities can then be written as:

$$F_t = V_t^{\eta} R_t \tag{20}$$

If $\eta \neq 1$, then the risk function depends on the level of exposure. This is easily seen when equation (20) is written as $F_t / V_t = V_t^{(\eta-1)} R_t$. A value of η larger than 1 implies stronger increases in fatalities per increase in exposure compared to the proportional relationship (COST 329, 2004).

In this section, some of these extended models will be presented. The logarithm of the fatalities will be modelled directly, using the observed exposure and a family of exponential curves for the risk. Logistic models were also tested, but were felt to be unnatural as the observed risk is not S-shaped for the period under consideration. This was confirmed by the tests; none of the logistic models could outperform the exponential models and illogical parameters estimates were obtained. They will therefore not be retained in the overview of models.

The estimation is done in two steps, as was proposed in (COST 329, 2004). First, an ordinary least squares model is fitted. Then, taking the results of this estimation as starting values, a weighted least squares estimation is done, using the number of fatalities as the weight. The quality of the models is again evaluated by comparing the AIC values. To test for serial correlation in the residuals, the Box-Ljung Q-statistic, introduced in Equation 14, is used. The quality of the predictions is evaluated by looking at the MAPE. Models are estimated on the years 1973-2001, and predictive tests are done on 2002-2004.

The family of risk models that will be considered can, in its most general form, be written as:

$$\log F_t = (1-\eta)\log V_t + \log\left(c + e^{\alpha_1 + \alpha_2 t + \alpha_3 LAW75 + \alpha_4 LAW92 + \alpha_5 LAW95}\right) + \varepsilon_t \qquad (21)$$

The parameter η is used to test the assumption that risk and traffic growth are proportional, while *c* is a constant added to allow for a nonzero minimal risk level and ε_t is the residual term. The variable *t* is a measure of time (in years), and LAW75, LAW92 and LAW95 are 0/1 level dummy variables testing respectively the introduction of laws on seatbelt use, speed and alcohol consumption. In TABLE 9, the most simple model is referred to as EXP. This is the classical exponential model, without a constant (c = 0), without covariates ($\alpha_3 = \alpha_4 = \alpha_5 = 0$) and no extra parameter for exposure (η = 0).

TABLE 9: Diagnostics for alternative models

	Parameter constraints	**AIC**	**MAPE**	**2010**
EXP	$c = \alpha_3 = \alpha_4 = \alpha_5 = \eta = 0$	-2.70	8.38	758
EXP-C	$\alpha_3 = \alpha_4 = \alpha_5 = \eta = 0$	-2.93	5.11	1032
EXP-V	$c = \alpha_3 = \alpha_4 = \alpha_5 = 0$	-2.70	3.28	877
EXP-C-V	$\alpha_3 = \alpha_4 = \alpha_5 = 0$	-2.87	4.97	1024
EXP-VAR	$c = \eta = 0$	-2.91	4.98	807
EXP-C-VAR	$\eta = 0$	-3.51	16.87	1382
EXP-V-VAR	$c = 0$	-2.84	4.35	826
EXP-C-V-VAR	-	-3.47	16.13	1351

From the AIC values in TABLE 9, it can be seen that the models have a similar fit. The best fitting model is EXP-C-VAR, including a constant term and explanatory variables for the main interventions. In terms of forecasting accuracy, the more parsimonious model EXP-V is ranked first. Note that the best fitting model is also the one that gives the worst predictions, which gives a clear indication of the danger of over-fitting in these models. Also, the most simple model, with the exponential risk function as in the Oppe approach, gives a lower fit and only a moderate forecasting accuracy. On the other hand, the most complex models, especially those with a constant term in the risk function, perform quite well in model fitting, but are clearly outperformed by the other models in terms of forecasting accuracy. As for the correlation in the residuals, the Q-statistics indicated for all models that the hypothesis of zero correlation could not be rejected at different lags (tested up to lag 12). The absence of correlation might be explained by the fact that exposure, which is highly autocorrelated, is now treated as an explanatory variable.

It is clear from these results that policy makers should be cautious when analysing the results of the different models. This can also be seen in the predictions of the number of fatalities in 2010 derived from the models. The predicted number ranges from 758 (model EXP) to 1382 (model EXP-C-VAR). The only prediction that comes in the neighbourhood of the 2010 objective (750 fatalities) is the EXP model. As this model was certainly not the most appropriate in terms of model fit and forecasting accuracy, it is questionable whether the target is indeed feasible. The best predictive model (EXP-V) shows a value of 877, which seems to be a more realistic figure. It is also interesting to note that the models with unrealistically high predictions for 2010 are those with a constant term for risk. Apparently, the recent developments in road safety are in disagreement with the assumption that road risk should level off at a nonzero value. Although the constant significantly improves the model fit, its value will be biased by the most recent observations. This effect is clearly present in the Belgian data. Because of the significant drop in the number of fatalities in the last years (2003-2004), all models with a constant predict very badly.

The last four models in the table include explanatory variables for the introduction of important road safety laws. The estimated parameter values for the different models are shown in TABLE 10, with the corresponding *p*-values between brackets.

TABLE 10: Parameter estimates for road safety laws

	Seatbelt	**Speed**	**Alcohol**
EXP-VAR	-0.1169 *(0.0026)*	-0.0533 *(0.2497)*	0.0228 *(0.6260)*
EXP-C-VAR	-0.0717 *(0.0495)*	-0.2704 *(0.0262)*	-0.4958 *(0.1462)*
EXP-V-VAR	-0.1144 *(0.0058)*	-0.0530 *(0.2618)*	0.0176 *(0.7459)*
EXP-C-V-VAR	-0.0772 *(0.0332)*	-0.2499 *(0.0213)*	-0.3995 *(0.1471)*

In all models, the seatbelt law has a significant influence on the number of fatalities. The speed law is significant in the models with a constant, but not in the others. The introduction of the law on alcohol was not significant in any of the models. The statistical test cannot reject the hypothesis that the alcohol law is without effect at the aggregate level. This indicates that either the law is, indeed, ineffective, or the test applied did not have sufficient power to detect any effect, even if this effect would be present. Note that, due to the nonlinear functions in the models, the parameter estimates cannot be directly interpreted in the models with a constant. In the two other models, the parameter estimates can be seen as a percentage effect. According to the models EXP-VAR and EXP-V-VAR, the introduction of the seatbelt law reduced the number of fatalities by about 1-exp(-0.11)=0.1042 or 10.42%.

TABLE 11: Parameter estimates for exposure (η)

	Coefficient	Std. Error	t-Statistic	Prob.
EXP-V	0.3963	0.2849	1.3913	0.1759
EXP-C-V	-0.0892	0.3238	-0.2754	0.7852
EXP-V-VAR	0.0669	0.3297	0.2028	0.8411
EXP-C-V-VAR	-0.1958	0.2430	-0.8060	0.4289

Another aspect that is tested in some of the models is the proportionality assumption of fatalities and exposure. That is, the constraint that the number of fatalities is proportional to the level of exposure is relaxed by estimating a power (1-η) for the level of exposure. TABLE 11 shows the estimated values for η in the four relevant models. Because the model formulation includes a term (1-η), the test statistic shown in the table is to test the hypothesis that $\eta = 1$. According to these models, the hypothesis that exposure is proportionally related to the number of fatalities could not be rejected. In a similar model, testing for $\eta = 0$, the null hypothesis was rejected in all models.

5.2.4 Stochastic trend models

Instead of imposing S-shaped curves for exposure and an exponential or more general logistic curve for risk, Lassarre (2001) used a local linear trend approach. This stochastic trend model belongs to the family of structural models introduced by Harvey (1989), and can be written as a set of equations in which unobserved components are used for the time-varying level μ_t and slope v_t:

$$\begin{cases} \log(F_t) = \mu_t + (1-\eta)\log(V_t) + \sum_i \alpha_i w_{i,t} + \varepsilon_t \\ \mu_t = \mu_{t-1} + \nu_{t-1} + \xi_t \\ \nu_t = \nu_{t-1} + \zeta_t \end{cases} \tag{22}$$

In this structural model, ε_t, ξ_t and ζ_t are white noise processes with standard deviations equal to σ_ε, σ_ξ and σ_ζ. To test the non-unitary elasticity of exposure, the parameter η is included. The parameters α_i measure the effect of the interventions (i.e. in this model the laws on seatbelt use, speed and alcohol). Note that this formulation is still linear, but the parameters for the level and the slope are unobserved components that are now allowed to change over time. The estimation of the variances will indicate whether these components are indeed time-varying for the given dataset. A more detailed introduction to structural models will be given in Chapter 6, section 6.2.2.

Four models are considered in this framework. In the first two models, no parameter is estimated for exposure. This parameter is added in the third and fourth model. The second and the fourth model include the intervention variables for seatbelt, speed and alcohol laws. The fourth model is the most complete one, as it includes both parameters for the intervention variables and a parameter for exposure. The first model, STOCH, can be written as follows:

$$\begin{cases} \log(F_t) = \mu_t + \log(V_t) + \varepsilon_t \\ \mu_t = \mu_{t-1} + \nu_{t-1} + \xi_t \\ \nu_t = -0.0584 \end{cases} \tag{23}$$
$$\sigma_\varepsilon^2 = 0.000579 \quad \sigma_\xi^2 = 0.002719$$

The variance of the slope component did not improve the model and was set equal to zero. The trend declines at a rate of 5.6% per year. The second model, STOCH-VAR, includes parameters for the intervention variables. In this case, the level is fixed, but the slope is allowed to change over time. That is, according to this model, the rate of decline in the trend is changing over time. The model formulation is given in Equation 24.

$$
\begin{cases}
\log\left(F_t\right) = \mu_t + \log\left(V_t\right) \\
\quad - \underset{(0.0417)}{0.0893}\text{LAW75} - \underset{(0.0930)}{0.1216}\text{LAW92} - \underset{(0.0438)}{0.1483}\text{LAW95} + \varepsilon_t \\
\mu_t = \mu_{t-1} + \nu_{t-1} \\
\nu_t = \nu_{t-1} + \zeta_t
\end{cases} \tag{24}
$$

$$
\sigma_\varepsilon^2 = 0.001152 \quad \sigma_\zeta^2 = 0.000046
$$

As an illustration, FIGURE 30 shows the slope for the STOCH-VAR model. Starting from a value of -6%, the slope becomes less negative over time, which is in line with the exponential model.

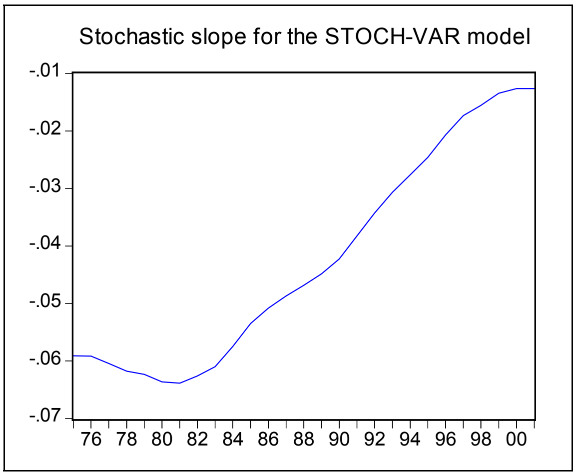

FIGURE 30: Stochastic slope for the STOCH-VAR model

The third model, STOCH-V, includes a proportionality parameter for the level of exposure. Again, the level is treated deterministically and the slope can vary over time:

$$
\begin{cases}
\log\left(F_t\right) = \mu_t + \left(1 + \underset{(0.4810)}{0.5472}\right)\log\left(V_t\right) + \varepsilon_t \\
\mu_t = \mu_{t-1} + \nu_{t-1} \\
\nu_t = \nu_{t-1} + \zeta_t
\end{cases} \tag{25}
$$

$$
\sigma_\varepsilon^2 = 0.001600 \quad \sigma_\zeta^2 = 0.000240
$$

The last model, STOCH-V-VAR, includes both the parameter for exposure and the road safety intervention variables. The best fitting model has a structure that is comparable with the previous one. It is given in Equation 26.

$$\begin{cases} \log(F_t) = \mu_t + \left(1 + \underset{(0.3943)}{0.1081}\right)\log(V_t) \\ \quad - \underset{(0.0417)}{0.0907}\text{LAW75} - \underset{(0.0934)}{0.1192}\text{LAW92} - \underset{(0.0453)}{0.1445}\text{LAW95} + \varepsilon_t \\ \mu_t = \mu_{t-1} + \nu_{t-1} \\ \nu_t = \nu_{t-1} + \zeta_t \end{cases} \tag{26}$$
$$\sigma_\varepsilon^2 = 0.001141 \quad \sigma_\zeta^2 = 0.000047$$

For the stochastic models, TABLE 12 gives an overview of the Q-statistics for serial correlation, the AIC and MAPE values, as well as the predictions for 2010, based on the Richards prediction for exposure. As the Log-Likelihood is calculated differently compared to the deterministic models, the AIC values are not directly comparable. They can be used, however, to compare the stochastic models with one another. The MAPE and the predictions for 2010 can be compared with the deterministic models.

TABLE 12: Diagnostics for alternative stochastic models

		AIC	Q_1	Q_6	MAPE	2010
STOCH	$\alpha_i = \eta = 0$	-1.21	0.170 (0.680)	7.449 (0.281)	3.51	812
STOCH-VAR	$\eta = 0$	-1.56	0.121 (0.727)	3.430 (0.765)	10.05	886
STOCH-V	$\alpha_i = 0$	-1.18	0.016 (0.899)	8.060 (0.234)	8.01	762
STOCH-V-VAR	-	-1.49	0.238 (0.625)	3.544 (0.738)	9.79	876

The best fitting stochastic model is STOCH-VAR. However, as in the case of the deterministic models, the best fitting model is not the one with the most predictive power. The simple stochastic model STOCH results in the lowest MAPE. The range of forecasts produced by the stochastic models is much smaller than for the deterministic models. The predictions are around the same values as for the most sensible deterministic models.

In the models STOCH-VAR and STOCH-V-VAR, the intervention variables were tested. The results are shown in TABLE 13. For the seatbelt law, the results are similar to those obtained in the deterministic models. The introduction of the law resulted in a 9% decrease in fatalities. For speed, the intervention variables are not significant, as was the case in the deterministic models without a constant. Unlike the previous models, the alcohol law shows a significant decrease in fatalities.

TABLE 13: Parameter estimates for road safety laws (stochastic models)

	Seatbelt	**Speed**	**Alcohol**
STOCH-VAR	-0.0893 *(0.0324)*	-0.1216 *(0.1911)*	-0.1483 *(0.0007)*
STOCH-V-VAR	-0.0907 *(0.0297)*	-0.1192 *(0.2019)*	-0.1445 *(0.0014)*

For the proportionality assumption of exposure and fatalities, TABLE 14 shows the results in the stochastic models. The negative sign indicates that a 1% increase in exposure might result in a more than 1% increase in fatalities. However, again the parameter estimates are not significant.

TABLE 14: Parameter estimates for exposure (stochastic models)

	Coefficient	**Std. Error**	**t-Statistic**	**Prob.**
STOCH-V	-0.547240	0.480998	-1.137719	0.2552
STOCH-V-VAR	-0.108085	0.394287	-0.274127	0.7840

An interesting result that can be derived from the structural models is the rate of progress in safety, calculated as the ratio of the derivative of the form of the trend and the elasticity of exposure. According to Lassarre (2001), this is the rate that nullifies the effect on safety of an increase in exposure. This concept is different from the rate of development of risk at constant exposure, which is measured by the slope. Looking for example at the STOCH-V-VAR model, the slope at the end of the analysis period (2001) equals -0.015. The parameter for exposure was 1.1081, yielding a rate of progress of 1.36% which is reasonably low. Given the evolution in the slope component, going from -6% in 1975 to -1.5% in 2001, it is clear that the rate of progress is much lower now than it was in the past. This confirms the hypothesis that much more progress could be gained at the time of major road safety interventions. This was especially the case for the introduction of the seatbelt law in Belgium. Actually, the same philosophy is behind an exponential risk function. The higher the level of road safety, the more difficult it will be to improve it.

5.2.5 Stochastic latent risk models

In the models proposed by Lassarre and discussed in the previous section, exposure is treated as an explanatory variable that is measured without error. Multivariate state space models offer the possibility of treating both fatalities and exposure as latent processes, while the relation between them is still explicitly modelled. The model that will be presented below is a special case of state space methods (Durbin & Koopman, 2001; Harvey, 1989). More specifically, it is a bivariate local linear trend model, and was introduced in road safety research in (Bijleveld & Commandeur, 2004). Using matrix algebra, a multivariate linear state space model can, in its most general form, be written as:

$$\begin{aligned} y_t &= Z_t \alpha_t + \varepsilon_t & \varepsilon_t &\sim NID(0, H_t) \\ \alpha_{t+1} &= T_t \alpha_t + R_t \eta_t & \eta_t &\sim NID(0, Q_t) \\ (t &= 1, 2, \ldots, n) \end{aligned} \tag{27}$$

The first equation is called the observation or measurement equation. Here, y_t contains the observed time series at time point t, and ε_t is a vector with observation errors. These errors have a zero mean and variances and covariances that form the matrix H_t. The second equation is the state equation, in which the state vector α_t is updated. The matrix T_t is the transition matrix and η_t are the state errors, with variances and covariances gathered in the matrix Q_t. Errors in both equations are assumed to be *NID* or "Normally and Independently Distributed".

In the case of fatalities F_t and exposure V_t, this model can be developed as follows:

$$
\begin{aligned}
&\log(V_t) = \mu_t^{(1)} + \varepsilon_t^{(1)} \\
&\mu_t^{(1)} = \mu_{t-1}^{(1)} + \nu_{t-1}^{(1)} + \xi_t^{(1)} \\
&\nu_t^{(1)} = \nu_{t-1}^{(1)} + \zeta_t^{(1)} \\
&\log(F_t) = \mu_t^{(1)} + \mu_t^{(2)} + \varepsilon_t^{(2)} \\
&\mu_t^{(2)} = \mu_{t-1}^{(2)} + \nu_{t-1}^{(2)} + \xi_t^{(2)} \\
&\nu_t^{(2)} = \nu_{t-1}^{(2)} + \zeta_t^{(2)}
\end{aligned}
\tag{28}
$$

The unobserved components $\mu_t^{(1)}$ and $\mu_t^{(2)}$ represent the trends for exposure and risk. As the two observation equations are written in logs, the equation for $\log(F_t)$ can easily be considered as the sum of the log-trend of exposure and the log-trend of risk. Clearly, taking exponentials gives a multiplicative form in which fatalities can be seen as the product of exposure and risk. This model formulation therefore is completely in line with the conceptual framework that is at the basis of the Oppe models and the extensions presented above.

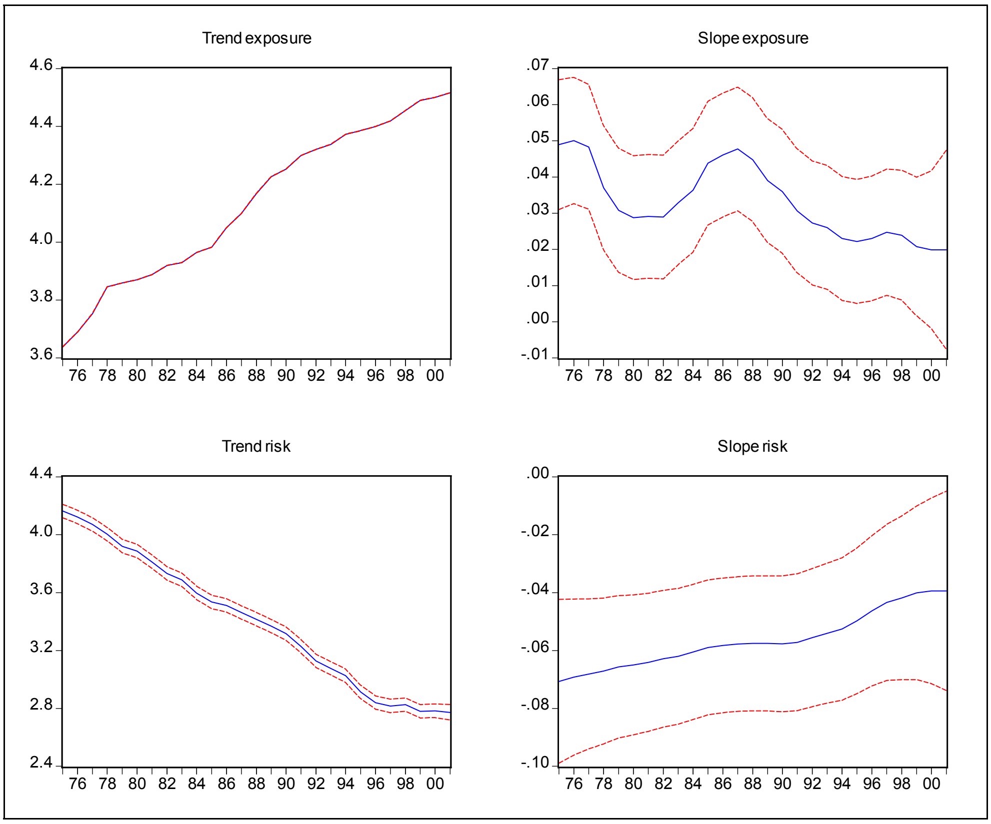

FIGURE 31: Estimated trend and slope for exposure and risk

The multivariate approach of this problem in a state space setting has some nice properties. First, as this approach is a logical extension of the Lassarre (2001) models presented in the previous section, it takes over all advantages of the structural framework presented there. In particular, the variables modelled in the system are allowed to have changing trends and slopes over time. Second, the measure of exposure is treated here as an endogenous variable, with an associated error. Exposure is treated stochastically, as a separate equation in the system. It is therefore recognised that the number of vehicle kilometres is only a measure of exposure that, like any other measure, might be only partially correct. Third, the structure of unobserved components as introduced in Equation 28 gives an estimate of the unobserved risk, without introducing it as a deterministic variable. Indeed, $\mu_t^{(2)}$ is a latent component that takes up the role of the risk. Instead of treating risk as a deterministic component, it is calculated as a nice by-product of the procedure.

The model structure in Equation 28 has been applied to the Belgian data. FIGURE 31 shows the estimated trend and slope of the log-exposure and the log-risk, together with their 95% confidence intervals. It is no surprise that the trend in exposure is upwards, but the slope gives some interesting insights that may not be visible at first sight. In the early eighties, the slope decreases, but goes up again to reach its original level in the early nineties. In 2001, the yearly increase in exposure is estimated to be about 2%.

Although the trend in the risk is clearly decreasing, the slope indicates that the rate of decline is getting smaller over time. It can also be seen from the slope that the developments in risk came to a halt in the years 1988-1990. Also in 2000 and 2001 the risk seems to stagnate.

The estimates for exposure and risk are shown in FIGURE 32, together with predictions up to 2010 for both variables. While estimated exposure results from the development of one latent component, the fatalities are composed of a latent variable that takes exposure into account and another one that represents the risk. Taking exponentials, the estimation of the number of fatalities is a textbook example of how risk and exposure can determine the level of the number of fatalities.

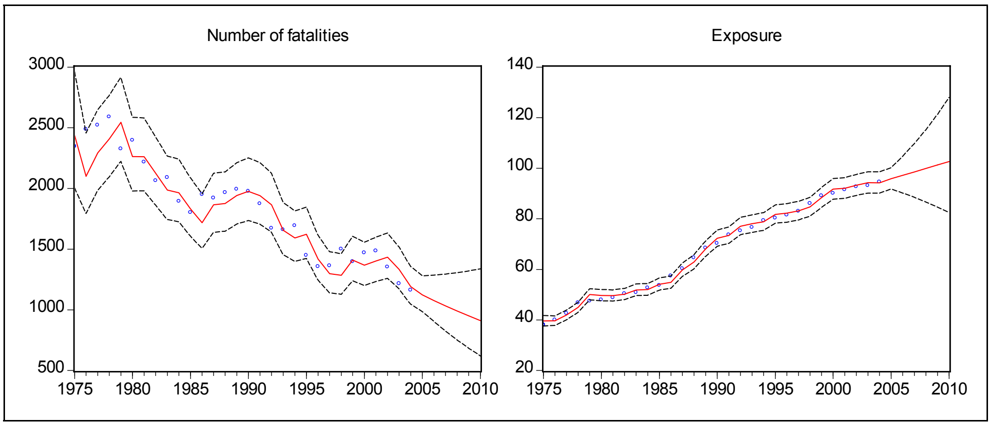

FIGURE 32: Estimated exposure and number of fatalities

The Q-statistics (up to order 12) for this model, calculated for the standardised one-step-ahead prediction residuals for the equations for exposure and fatalities, indicate that the white noise hypothesis cannot be rejected. The model predicts an exposure value of 103 billion vehicle kilometres in 2010, which is only slightly higher than the outcome of the Richards curve. Fatalities are decreasing up to a level of 911 in 2010. Compared to the other models, this is a reasonable prediction. The MAPE value for the predictions on 2002-2004 equals 6.04%, a reasonable value which is in line with the other models. The confidence interval for the out-of-sample forecasts indicates the high uncertainty that goes together with the predictions. The prediction horizon should therefore be not too long.

Although the latent risk model does not really outperform the more classical approaches, it has the undeniable advantage of showing the underlying components (trend and slope) of the series. Also, the fact that no deterministic functional form is imposed feels quite natural and results in at least equally sensible outcomes.

5.3 Conclusion

In this chapter, various models have been developed to determine the relation between the number of fatalities and exposure on yearly data for Belgium. All models are based on the classical assumption that fatalities can be seen as the result of a certain level of exposure and risk. These are the basic concepts behind the Oppe model, which was presented as an introduction. In subsequent models, the various restrictions in the Oppe model were relaxed and tested.

Starting from a range of deterministic models, the idea of stochastic models was introduced and compared with the more traditional models.

This chapter shows that for the Belgian data, models other than the classical Oppe approach may lead to better results. The Richards curve for exposure significantly improved the model fit and the prediction accuracy. Inclusion of a constant term resulted in an improvement of the exponential model for the risk. On the other hand, as stated in (COST 329, 2004), the simple exponential model has the advantage of a small number of parameters, which makes the model easy to solve and less dependent on local deviations. Also, the model generally gives a reasonable approximation of the road safety situation and is simple to interpret. For the Belgian situation, the Oppe model underestimates the trend in the number of fatalities compared with more complex models.

The addition of a constant term in the deterministic models was meant to prevent a possible non-zero risk level in the future. It is interesting to note that the constant is only significant in the models where no parameter for exposure is included. The constant will influence the risk when it flattens off (that is, in the long term), and this happens at the same time that the exposure reaches its upper bound. When no parameter is estimated for exposure, the constant takes up this effect. In all models, the inclusion of a constant significantly improved the fit (compared to the same model without a constant). However, the apparent effect of a constant may be biased by the most recent observations in the sample (COST 329, 2004). This may influence the quality of the predictions when sudden changes in the patterns occur. Indeed, given the significantly smaller number of fatalities for the years 2003-2004, the projections for 2010 obtained from these models are unrealistically high.

The proportionality of the level of exposure and the number of fatalities was tested by comparing models with and without a parameter for the exposure measure. In general, the elasticity of the number of fatalities to the level of exposure is low. The inclusion of this parameter has no impact on the results and it can be assumed to be equal to one. Also in (Lassarre, 2001), generally very low elasticities were found for various countries, although it is not clear whether they are significantly different from 1.

From the models with intervention variables, it was concluded that only the seatbelt law has a clear and unambiguous effect on the number of fatalities. On average, the introduction of this law was associated with a decrease of 9% in the number of fatalities. For the other laws, the effect is less clear. Speed is only significant in the models with a constant. These are also the models where exposure showed an elasticity to the number of fatalities greater than one (although not significant). This indicates a possible relation between the form of the exposure curve and speed limits. For example, the inflection point that was estimated for the Richards exposure curve is not far away from the introduction of the speed law. This does not, however, prove a causal relationship between the two curves. Moreover, it is not unthinkable that, due to the diverse nature of the speed law (50 km per hour in urban areas, speed limits for motorcycles and trucks, 90 km per hour on provincial roads, etc.), it is difficult to measure the effect on a highly aggregated level. A similar conclusion may be valid for the laws on alcohol. This law is only significant in the stochastic models. This may indicate a strong influence of the deterministic functional form on the significance of the variables. However, whereas seatbelt use and speed are directly related to the behaviour of all motorized road users, the law on alcohol controls is of a different kind. The effect of an alcohol law might be strongly related to the frequency and the spreading of tests for alcohol, more than it is the case for seatbelt use and speed, which are obviously related to aspects of the car and the road infrastructure. As the alcohol tests are known to be spread in time, the alcohol law may be a more complex intervention to test than seatbelt use or speed.

Another aspect that has been treated in this chapter is the difference between deterministic and stochastic models. Whereas a deterministic model assumes a fixed trend, the stochastic models allow for a more flexible evolution over time. In general, the deterministic models produce acceptable results, both in terms of model fit and forecasting accuracy. However, it became clear that the specifics of the functional form can have a strong influence on the outcome of a model. For example, it is undesirable that adding a constant to the exponential risk specification inflates the predictions or change the significance of the explanatory variables. Moreover, the stochastic trend models are more natural in the sense that they do not presuppose a certain functional form as the cornerstone of evolutions in road safety and exposure. Also, the stochastic outcome is much more enriched than the deterministic one, as it provides information on the dynamics of the curves (the level and the slope), which are allowed to change over time. In addition, the bivariate extension of the Lassarre models, proposed by (Bijleveld & Commandeur, 2004), are perfectly in line with the basic idea of the multiplicative relation between fatalities, risk and exposure. Although these models do not assume any functional relation in advance, the outcomes are in line with common expectations.

All models presented in this chapter were also used to make a test prediction for three years and an out-of-sample prediction up to the year 2010. The quality of the predictions fluctuate between the models. Especially the predictions made by the deterministic models are quite sensitive to the functional form of the underlying model. Although the addition of a constant results in an improved fit for the models, it diminishes the quality of predictions. Also, the addition of intervention variables did not generally improve the quality. In this context, the stochastic form is interesting. All dynamics in the data are explicitly modelled, without assuming a functional relation beforehand. As a consequence, the predictions are not affected by the functional assumptions.

To conclude, the presented models are useful to make long term predictions in road safety, but due to the different characteristics of the models it is necessary to consider a series of models as a kind of sensitivity analysis. In combination with expert knowledge, it is possible to come up with reasonable road safety predictions.

Chapter 6 Aggregated models for monthly exposure and risk

6.1 Introduction

In this chapter, various models are presented for the description, explanation and prediction of monthly road safety outcomes in Belgium. Contrary to the models in the previous chapter, all subsequent models in this chapter are built on monthly time series data, obtained from the official statistics and described in Chapter 4.

In this chapter, both descriptive and explanatory models are developed. Descriptive models can be used to describe the trends and seasonal fluctuations in the monthly road safety data and to predict future road safety outcomes. Seasonality is an aspect that typically occurs in monthly (or quarterly) data, providing insight into the within-year road safety fluctuations. When exposure data and explanatory variables are available, more elaborate (explanatory) models can be considered. First, it is possible to analyse road safety in terms of exposure and risk. As has been shown in Chapter 5 on yearly data, splitting up the road safety outcomes in terms of exposure and risk can enhance the insights in the underlying processes that lead to a certain safety level. The same is true for an analysis on monthly data. Second, explanatory variables allow measuring the effects of certain macroscopic variables on the level of road safety. If effects of explanatory variables are also measured for the level of exposure, one can distinguish between direct and indirect effects of certain variables on road safety outcomes (Gaudry et al., 2000). If the effect of a variable, such as snowfall, is evaluated in a model for the number of fatalities, a direct effect is measured. But it is not unthinkable that snowfall will also affect the level of exposure, which, in turn, will influence road safety. That is, snowfall will influence the number of fatalities indirectly by its impact on the level of exposure.

It is clear that, depending on the model that is developed, the data needs may be quite high. For a descriptive model of road safety, it is sufficient to have a database with accidents and victims, which is, at a high level of aggregation, usually not too problematic. However, if explanatory models are developed or indirect effects are to be calculated, data issues crop up. Especially with monthly data, it is very hard to find the necessary data. Also, exposure data are not always available in a ready-to-use format. From Chapter 3 it is clear that a diversity of exposure measures is used in macroscopic road safety models. The monthly measure of exposure, which was developed for the Belgian situation in Chapter 4, will be used here in various models.

Once a model is constructed, be it descriptive or explanatory, it can be used for forecasting purposes. It is important to stress that models need not to be complex if the final objective is forecasting. Even the relatively simple descriptive models can be useful in predicting road safety. Moreover, an explanatory model suffers from the drawback that future values are needed if forecasts are to be made.

6.2 Descriptive road safety models

In this section, a variety of models is presented that can be used to describe the evolution in road safety. Descriptive models aim at unravelling the properties of a specific time series. They are useful for monitoring purposes as well as for predicting road safety in the near future. On the other hand, they do not include any explanatory factors related to road safety. Although the lack of explanatory power may be felt as a limitation of descriptive models, it offers at the same time the undeniable advantage of low data needs. No other variables than the series itself are needed to perform the analysis. Especially in the area of road safety, where data is often difficult to find, this is a interesting property.

The descriptive models in this section will be developed on monthly road safety data for Belgium. When road safety data are available on a monthly basis, some specific properties of the series are noted. First, as with yearly data, an increasing or decreasing trend can be present. Second, contrary to the higher frequency data, strong seasonal fluctuations are present. The classical descriptive models try to separate the trend and the seasonal in the data from the irregular pattern. If the trend, seasonal and irregular part of a series are denoted respectively as T_t, S_t and I_t, then a multiplicative decomposition is written as:

$$Y_t = T_t \times S_t \times I_t \tag{29}$$

Alternatively, an additive decomposition is written as:

$$Y_t = T_t + S_t + I_t \tag{30}$$

For positive components, a multiplicative model can be transformed into an additive model by the logarithmic transformation. A decomposition of a time series in the underlying patterns enhances the insight in the level of the series and the magnitude of the (recurring) fluctuations around it. Classical decomposition models can be useful for curve fitting purposes, but they are rarely used for forecasting. Well-known approaches involve calculation moving averages and polynomial splines. Moving averages of odd order k calculates averages of a observation and (k-1)/2 points on either side of it (Makridakis et al., 1998). The average is "moving", because for each next observation the oldest observation is dropped and a new one is included. The higher the order of the moving average, the smoother the result will be. Moving averages are therefore useful to calculate a trend of a series. Polynomial splines use a sequence of polynomials of a low degree (quadratic or cubic) to describe changes in the curvature of the data.

In more advanced models, the random character of the data is taken into account. In this case, a time series Y_t is seen as a random process in the sense that it is not possible to theoretically determine a value for Y_t, but instead a probability distribution is introduced to describe the likelihood of an observed value. Classical regression models with time as an independent variable include a residual term to indicate the random variation in the data, but do not allow the time to influence the series stochastically. That is, the effect of time is the same for each period. A regression model can be used to describe the dependence between a road safety variable and time. It is useful to get an idea of the overall trend in the data, but is dangerous when the sample is extended into the future. One is never sure that the chosen functional form remains valid for out-of-sample observations. For example, when a linear regression is fitted through the (declining) number of fatalities, one can end up with negative numbers in the future.

A modelling approach that has properties of both moving average methods and classical regression is local regression smoothing. Instead of taking averages of a subgroup of points, partial straight lines are fitted and joined to estimate the underlying trend in the data. Analogous to the order of a moving average, a smoothing parameter is used to allow for more or less curvature in the data.

The most famous stochastic time series models are the ARIMA models, introduced by Box and Jenkins. The merit of Box and Jenkins is that they provide a rigorous methodology to time series modelling through identification, estimation and diagnostic checking. ARIMA models consider a time series as a combination of autoregressive and moving average components. The trend and the seasonal component are treated as "noise" in the series, and should be filtered away before models are fit. In fact, the trend and the seasonal component render the series non stationary, and as the ARIMA methodology requires stationary series, they should be removed before estimation.

The separation of a trend and a seasonal from the irregular part of a time series was also at the basis of many seasonal adjustment programs. It is appropriate here to mention the Census Bureau methods, which were developed by the U.S. Bureau of the Census. Throughout the different versions of the methods (X-11, X-11-ARIMA and X-12-ARIMA), the time series decomposition methodology remained the same. In essence, the decomposition involves the application of weighted moving averages. However, the more recent versions combine these techniques with ARIMA models to prevent loss of data at the beginning of the series and extend the series with forecasts to improve the quality of the decomposition. Contrary to the ARIMA methodology, the decomposition techniques show the components of the series as outputs. Especially the seasonal component is used to seasonally adjust the series. On the other hand, the decomposition techniques are less suited for forecasting purposes. Decomposition is a tool for understanding a time series rather than for forecasting.

In the past, the ARIMA methodology has been frequently applied for describing and forecasting many time series. For many years, it was the most outstanding method in education, business and research. However, ARIMA methods suffer from some drawbacks in terms of model specification and interpretation. As will be shown in the subsequent sections, ARIMA models require stationary data, which is a rather stringent assumption. As a consequence, information on the trend and the seasonal component is filtered away. Also, interpretation of the estimated parameters is not an easy task, and missing values or changes in the frequency of the series are not allowed.

The state space methodology deals with most of the drawbacks of ARIMA models. This approach, that originated in the field of control engineering, is extensively described in (Harvey, 1989) and in (Durbin & Koopman, 2001). In the state space approach, the trend, the seasonal component and the irregular are considered as essential parts of the series. They are modelled as independent parts of the series. State space models are very flexible and general. A very large range of models, including the ARIMA models, can be represented as state space models.

From this introduction, it should be clear that the range of possible descriptive models is large. In this section, focus is on the stochastic models. ARIMA and State Space methods will be applied to the Belgian road safety data. A theoretical background is given for both modelling approaches. Next, the models are applied to the series of the number of fatalities, the number of serious injuries and the corresponding number of accidents. All models are developed on data from 1973 up to 2002. The last two years, 2003 and 2004 are used to assess the prediction accuracy. These models will provide a first indication of the underlying processes in road safety.

6.2.1 ARIMA Methodology

The framework of ARIMA processes is based on the work of Box and Jenkins (1976). It is a class of time series models in which time series are modelled in terms of their own past behaviour. This means that the current values of a variable are related to past values (Verbeek, 2000). As prediction is often one of the objectives of a time series analysis, information on the past values of a variable will be used for forecasting the future.

6.2.1.1 ARIMA model components

Consider a time series of *n* observations of some variable, denoted as Y_1, Y_2,…, Y_n. These observations are seen as realizations of random variables with a given joint distribution. In the context of time series, this probability model is called a stochastic process. It is the objective of a time series analysis to find a model which captures the essential characteristics of the stochastic process.

A simple way to model dependence of observations over time would be to say that Y_t is equal to a constant term plus a random variable and one or more (*q*) lagged values of this variable (Verbeek, 2000). This is called a Moving Average process of order *q*, denoted as MA(*q*):

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}\ , \quad \varepsilon_t \sim IID\left(0, \sigma^2\right) \qquad (31)$$

In this expression, $IID(0,\sigma^2)$ denotes independent drawings from an identical distribution with zero mean and variance equal to σ^2. The unobserved random variables ε_t therefore refer to a homoskedastic process with no autocorrelation. Using a backshift operator *B* on Y_t, defined as $B^i Y_t = Y_{t-i}$ (i=1,2,…), this process can be written as:

$$Y_t = \mu + \left(1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q\right)\varepsilon_t\ , \quad \varepsilon_t \sim IID\left(0, \sigma^2\right) \qquad (32)$$

In this expression, the lag polynomial containing the moving average components can be written more compactly as $\Theta(B)$, leading to the following formulation:

$$Y_t = \mu + \Theta(B)\varepsilon_t\ , \quad \varepsilon_t \sim IID\left(0, \sigma^2\right) \qquad (33)$$

Likewise, Y_t can be expressed as a regression in terms of a constant plus one or more (p) weighted previous values of Y_t and an unpredictable component. This process is called an autoregressive relation of order p, denoted as AR(p):

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t\ , \quad \varepsilon_t \sim IID\left(0, \sigma^2\right) \qquad (34)$$

Again using the backshift operator B, this AR(p) process can be written as:

$$Y_t\left(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p\right) = \mu + \varepsilon_t\ , \quad \varepsilon_t \sim IID\left(0, \sigma^2\right) \qquad (35)$$

Introducing a second lag polynomial $\Phi(B)$, now containing the autoregressive terms, the expression can be written as:

$$\Phi(B)Y_t = \mu + \varepsilon_t\ , \quad \varepsilon_t \sim IID\left(0, \sigma^2\right) \qquad (36)$$

In a more general setting, it is possible to include both autoregressive and moving average terms in one equation, leading to an ARMA(p, q) model:

$$\Phi(B)Y_t = \mu + \Theta(B)\varepsilon_t\ , \quad \varepsilon_t \sim IID\left(0, \sigma^2\right) \qquad (37)$$

6.2.1.2 The stationarity assumption

An ARMA model cannot, however, be applied in all circumstances. It is required that the process be stationary. For practical purposes, it is sufficient to have *weak* stationarity, i.e. that the means, variances and covariances of the process are constant through time. It is therefore required that:

$$\begin{aligned} &E(Y_t) = E(Y_{t+s}) && \forall s,t \\ &Var(Y_t) = Var(Y_{t+s}) && \forall s,t \\ &Cov(Y_t, Y_{t+k}) = Cov(Y_{t+s}, Y_{t+k+s}) && \forall k,s,t \end{aligned} \qquad (38)$$

In other words, the series is in equilibrium around a finite mean, and the variance around the mean remains constant over time. Furthermore, the covariances of the series do not depend on the period, but only on the distance in time between the two observations. In general, an ARMA(p,q) model $\Phi(B)Y_t = \mu + \Theta(B)\varepsilon_t$ is stationary if and only if the solutions of the characteristic equation $\Phi(B)$=0 are larger than 1 in absolute value. If a series is non-stationary because the variance is not constant, it often helps to log-transform the data. To obtain a series that is stationary in the mean, the series is differenced. That is, instead of working with the original series, successive changes in the series are modelled. When an ARMA model is built on differenced data, it is called an ARIMA model, where "*I*" indicates the differencing. For example, if first order differences are taken, then the series is called integrated of order 1.

When the series consists of quarterly or monthly observations, it is possible that the departure from stationarity stems from the differences between the same quarter or month in successive years, rather than from differences between successive observations. In these cases, seasonal differencing of order 4 or 12 may be necessary. A combination of first order and seasonal differencing results in a general ARIMA process, that is denoted as ARIMA(p,d,q)$(P,D,Q)_s$. In this formulation, p and q are, respectively, the orders of the AR and MA terms, while d indicates the order of regular differencing. The seasonal AR and MA terms are denoted as P and Q, while D is the order of seasonal differencing and s is the length of the seasonal. For monthly data, s = 12.

To decide whether differencing is necessary for a given time series, stationarity tests can be performed. The most famous stationarity test is the Augmented Dickey-Fuller (ADF) unit root test (Dickey & Fuller, 1979). This test adopts the null hypothesis that the series is not stationary against the alternative that it is stationary. Usually the test can take different forms, depending on whether stationarity is tested around a mean or around a trend. Generally, the ADF regression equation can be written as:

$$\Delta Y_t = \alpha Y_{t-1} + x_t^{'}\delta + \beta_1 \Delta Y_{t-1} + \beta_2 \Delta Y_{t-2} + \ldots + \beta_p \Delta Y_{t-p} + \varepsilon_t \qquad (39)$$

In this expression, $\Delta Y_t = Y_t - Y_{t-1}$, x_t is a vector containing exogenous variables which may consist of a constant or a constant and a trend (QMS, 2004), depending on the form of the non-stationarity. The parameters to be estimated are α, δ, β_1, …, β_p. The number of lagged difference terms is an issue, as sufficient lags should be included to remove serial correlation in the residuals, but statistical software packages like Eviews 5.1 (QMS, 2004) provide an automatic lag length selection. The null and alternative hypotheses to test for a unit root are:

$$\begin{aligned} H_0 &: \alpha = 0 \\ H_a &: \alpha < 0 \end{aligned} \tag{40}$$

The test statistic is the conventional *t*-ratio, which in this case does not follow a classical Student's *t*-distribution. Recently, MacKinnon (1996) calculated critical values for the test. When the null hypothesis cannot be rejected, the process is assumed to be non stationary, and differencing is performed.

6.2.1.3 ARIMA model diagnostics

According to the ideas of Box and Jenkins, a time series model is constructed in three separate steps: identification, estimation and diagnostic checking (Stewart, 1991). Identification is the process by which a theoretical model is selected that corresponds to the characteristics of the observed time series. This is mainly done by investigating the autocorrelation function (ACF) and the partial autocorrelation function (PACF). It also involves the stationarity check and, if necessary, the differencing of the series. After identification, the unknown parameters of the selected model are estimated. During diagnostic checking, some ex-post verification of the model is performed. Typically, the residuals are checked against the basic assumptions of no correlation and homoskedasticity. If the final model is acceptable, it can be used for forecasting purposes.

The most important tool in assessing the specific properties of a time series is the autocorrelation function (ACF), denoted as ρ_k. This function describes the correlation between Y_t and its lag Y_{t-k} as a function of *k* (Verbeek, 2000), and can be defined as:

$$\rho_k = \frac{Cov\left(Y_t, Y_{t-k}\right)}{Var\left(Y_t\right)} \tag{41}$$

It shows how values of the series are correlated with its past values. The ACF is a measure of how much interdependency exists between neighbouring observations (not necessarily next to one another) in a time series. If a variable is completely random, the autocorrelations should be zero for all time lags. If the variable is not completely random, one or more of the autocorrelations will be significantly different from zero.

The theoretical autocorrelation function ρ_k has some interesting properties for specific AR and MA processes. In general, ρ_k equals zero for an MA process of order $q < k$. That is, after q lags, the ACF is zero for an MA(q) model. For an AR(1) process, the autocorrelation coefficients exponentially decline to zero. For higher order AR processes, the ACF is more complex.

The sample autocorrelation function, denoted r_k, gives the estimated autocorrelation coefficients as a function of k. It is estimated as:

$$r_k = \frac{\frac{1}{n-k}\sum_{t=k+1}^{n}\left(Y_t - \bar{Y}\right)\left(Y_t - \bar{Y}_{t-k}\right)}{\frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \bar{Y}\right)_t^2} \tag{42}$$

To test whether a particular r_k value is significantly different from zero, the sampling distribution of the autocorrelations can be used. For a random series, the distribution of the autocorrelation coefficients is approximately normal with a zero mean and a standard error equal to $n^{-1/2}$. Therefore, if a series is white noise, about 95% of all sample autocorrelation coefficients are expected to be smaller in absolute value than $1.96*n^{-1/2}$. If this is not the case, the series is probably not white noise.

Rather than testing one autocorrelation at a time based on the distribution of the autocorrelation coefficient, one can also use the Ljung-Box Q-statistic, introduced in Equation 14 (section 5.2.2), to test a set of autocorrelation coefficients at different lags (Makridakis et al., 1998). The Ljung-Box Q-statistic is used to assess whether or not the residuals are purely random,. Under the null hypothesis of white noise residuals, all autocorrelations r_k up to a given value K should be equal to zero. The statistic is computed as:

$$Q_K = n\left(n+2\right)\sum_{k=1}^{K}\frac{r_k^2}{n-k} \tag{43}$$

where r_k is the lag-k autocorrelation coefficient as defined before, n is the number of observations, and K is the maximum lag being considered. If for an ARMA(p,q) model the null hypothesis is true, then Q_K is chi-square distributed with K-(p+q) degrees of freedom.

Apart from the autocorrelation function, a second useful graph is the Partial Autocorrelation Function (PACF). Partial autocorrelations are used to measure the degree of association between Y_t and Y_{t-k}, when the effects of other (intervening) time lags are removed. The partial autocorrelation coefficient of order *k* for a time series Y_t (t = 1, 2, ..., n) can be obtained by regressing Y_t on its own past values Y_{t-1}, ..., Y_{t-k}. The partial autocorrelation for lag *k* is the estimated coefficient for Y_{t-k} in this autoregressive equation. As with the ACF, all partial autocorrelations should be close to zero for uncorrelated data. If a series is white noise, the estimated partial autocorrelations are approximately independent and normally distributed with a standard error equal to $n^{-1/2}$. Therefore, the same critical values of $1.96n^{-1/2}$ can be used with the PACF to assess whether the series is white noise. For an autoregressive process, the values of the PACF give important information about the order of the process. If the partial autocorrelation show a cut-off after lag *p*, then *p* can be chosen to be the order of the autoregressive process of a time series.

The assumption $\varepsilon_t \sim IID(0,\sigma^2)$ also involves homoskedastic residuals. This is tested by the ARCH LM test (Engle, 1982), which is a Lagrange Multiplier test for autoregressive conditional heteroskedasticity in the residuals. This test is based on the observation that for many time series in practice the magnitude of residuals is related to the magnitude of recent residuals (QMS, 2004). The ARCH LM test statistic is computed from a regression that relates the current squared residuals to the squared residuals of previous periods. In general, to test the null hypothesis that there is no ARCH up to order *q* in the residuals, a regression of the squared residuals is run on a constant and lagged squared residuals up to order *q*. The test statistic is then computed as the number of observations times the R^2 from the test regression and follows, under general conditions, a $\chi^2(q)$ distribution (QMS, 2004). The null hypothesis of the test states that the residuals are homoskedastic.

6.2.2 State space methodology

6.2.2.1 Introduction

In the previous chapter, the state space methodology has been applied to model the level of exposure and the risk in a multivariate framework. In this section, the state space methodology will be described in a more elaborate way, as an introduction to the models that will be developed on the monthly data. Clearly, as in the ARIMA models, the use of monthly data implies some complexity (like for example seasonality) that is not present in models on yearly data. The details on the theory can be found in (Harvey, 1989) and in the more recent work of Durbin & Koopman (2001).

Like ARIMA models, state space models can be considered as dedicated time series models. That is, they are specifically developed to take the time dependencies into account. An important difference with the ARIMA framework is that the underlying components are explicitly modelled instead of filtered away. State space models are typically handled with the Kalman filter, a method of signal processing which provides optimal estimates of the current state of a dynamic system. The univariate basic structural state space models are in fact special cases of a more general class of models, which can be written in compact matrix notation. An observation Y_t of a series at time t is written in state space notation as (Chatfield, 2004):

$$Y_t = h_t^T \theta_t + n_t \tag{44}$$

The state vector θ_t is usually unobservable and contains the state variables as components. The state variables are typically model parameters, like regression coefficients, or parameters describing the state of a system (such as the level or the seasonal). The column vector h_t is known, and n_t is the observation error. It is further assumed that the state θ_t depends on the previous state θ_{t-1}, and that the changes of θ_t through time follow the equation:

$$\theta_t = G_t \theta_{t-1} + w_t \tag{45}$$

Here, G_t is assumed known, and w_t denotes a vector of white noise deviations. Both equations together form the state space model. The first equation is called the observation equation, while the second is the state equation. The errors are generally assumed to be serially uncorrelated and normally distributed.

After formulating the model in terms of its components, the main objective usually consists of estimating the signal, represented by θ_t. The Kalman filter can be used to estimate this unobserved vector. The "Kalman recursion equations" enable the calculation of the one-step forecast errors and the likelihood (Makridakis et al., 1998). This is usually done in two stages, as described for example in (Chatfield, 2004). In the prediction stage, θ_t is forecasted from the data up to time period (t-1). When subsequently the new observation at time t has been observed, the estimator for θ_t can be modified to take account of this extra information. The prediction error of the forecast of X_t is used to update the estimate of θ_t. This is the updating stage of the Kalman filter. The advantage of the recursive character of the Kalman filter is that every new estimate is based on the previous estimate and the latest observation, while at the same time the whole past of the series is taken into account.

6.2.2.2 Structural time series model

A structural time series model is one in which the trend, the seasonal and the error terms are modelled explicitly (Durbin & Koopman, 2001). The Basic Structural Model (BSM) was introduced by Harvey (1989). It consists of a stochastic level, a stochastic slope, a trigonometric seasonal and an irregular. In an extended version, also explanatory variables can be incorporated The models that will be considered further in this section can, in its most general form, be written as follows:

$$\begin{aligned}
Y_t &= \mu_t + \gamma_t + \sum_{j=1}^{J} \beta_{j,t} X_{j,t} + \varepsilon_t & \varepsilon_t &\sim NID\left(0, \sigma_\varepsilon^2\right) \\
\mu_{t+1} &= \mu_t + \nu_t + \xi_t & \xi_t &\sim NID\left(0, \sigma_\xi^2\right) \\
\nu_{t+1} &= \nu_t + \zeta_t & \zeta_t &\sim NID\left(0, \sigma_\zeta^2\right) \\
\gamma_{t+1} &= -\sum_{j=1}^{s-1} \gamma_{t+1-j} + \omega_t & \omega_t &\sim NID\left(0, \sigma_\omega^2\right) \\
\beta_{j,t+1} &= \beta_{j,t} & j &= 1,\ldots,J
\end{aligned} \qquad (46)$$

In this expression, μ_t denotes the level and ν_t is the slope term. The slope determines, as in classical regression, the rate of change in the state, but is now allowed to change over time. The combination of the level and the slope component determines the trend in the model. The term γ_t is the seasonal component, which is made stochastic by adding an error term ω_t to the equation. The effect of the explanatory variables $X_{j,t}$ is measured by the parameters $\beta_{j,t}$. In the equation above, it is assumed that the regression parameters do not change over time, although the framework is easily extended to do so. The error terms in the model are all assumed to be normally and independently distributed with zero mean and a specific standard deviation, denoted σ_ε for the observation equation and σ_ξ, σ_ζ and σ_ω respectively for the level, the slope and the seasonal equations. These standard deviations can be used to assess the assumption that a component is indeed stochastic. If the estimation procedure shows that the standard deviation of a specific component is not significantly different from zero, it can be decided to model it as a fixed component. In such a way, it is clear that deterministic models, like the classical regression models, are special cases of the stochastic state space model. Moreover, note that many of the statistical tests that are frequently used in ARIMA modelling (like the Box-Ljung Q-statistic), are equally useful in a state space setting.

6.2.2.3 State space versus ARIMA models

In this short introduction to state space models, it is not possible to give an overview of the methodology in all details. However, it might be worthwhile to point out here the major differences and advantages of the approach compared to the classical ARIMA methodology. This overview of arguments is based on (Durbin & Koopman, 2001). First, the state space approach is based on a structural analysis of the problem. The components like a trend, seasonal, slope, etc. are modelled separately, and the final model formulation is based on the judgement of the modeller. In any case, the structure of the underlying process remains clear. Second, state space models are flexible. Changes in the structure of the series over time are easily accounted for. Third, state space models are very general. As mentioned above, the include a wide range of existing models. It can be shown that many types of time series models, like the classical regression models and even ARIMA models, can be put into a state space form (Chatfield, 2004). Fourth, the generality of the models makes them very useful in analysing multivariate series, treating missing values and adding explanatory variables. Fifth, state space models are more transparent. The possibility to graph the underlying components for inspection may reveal more information on the observed and analysed data.

On the other hand, one may argue that both methods also have some similarities. Indeed, because of the fact that ARIMA models have a corresponding state space representation, and can also be extended with explanatory variables, it might be assumed that in practice the results of both approaches are (fortunately) comparable. Also, when prediction is the major objective, ARIMA models will still be very useful. However, the ARIMA models are put at a disadvantage when it comes to transparency and flexibility. No information on the components can be derived (as they are filtered away), and the requirement of stationarity, as well as the assumption of difference-stationarity, are obvious weaknesses of the approach.

6.2.3 An ARIMA study on Belgian road accident data

Using the ARIMA methodology described above, descriptive models were developed for the (log of the) following road safety indicators: the number of persons killed, the number of persons seriously injured, the number of accidents with persons killed and the number of accidents with persons seriously injured.

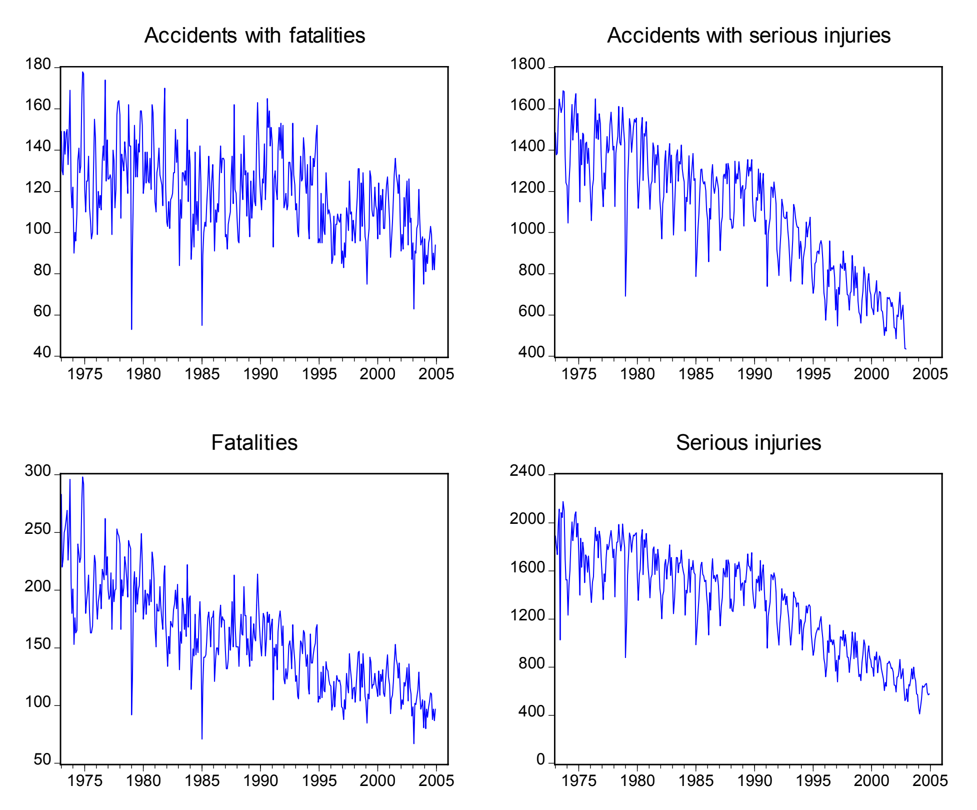

FIGURE 33:Data on fatal and serious injury accidents

The original series are shown in FIGURE 33. The series show a seasonal pattern and a decreasing trend, although the latter is not apparent for the accidents with persons killed. The models will be developed on monthly data for the years 1973-2002. The years 2003 and 2004 will be used to evaluate the forecast accuracy.

For the number of accidents with persons killed and the number of fatalities, the analysis resulted in the following model:

$$\Delta^{(1)}\Delta^{(12)}Y_t = \left(1-\theta_1 B-\theta_2 B^2\right)\left(1-\theta_3 B^{12}\right)\varepsilon_t \qquad (47)$$

Here, $\Delta^{(1)}$ and $\Delta^{(12)}$ indicate respectively first and twelfth order differencing. In general form, this model can be written as ARIMA(0, 1, 2)(0, 1, 1)$_{12}$. For the number of persons seriously injured and the corresponding number of accidents, an order 3 instead of order 2 moving average was needed:

$$\Delta^{(1)}\Delta^{(12)}Y_t = \left(1-\theta_1 B - \theta_2 B^3\right)\left(1-\theta_3 B^{12}\right)\varepsilon_t \tag{48}$$

This is an ARIMA(0, 1, 3)(0, 1, 1)$_{12}$, where the parameter for the second lag is assumed to be zero.

The model diagnostics are shown in TABLE 15. To obtain stationary series, first order and seasonal (order 12) differences were taken. The differenced series did not have a unit root, as was confirmed by the ADF tests. The series were then analysed by looking at the ACF and PACF to assess whether the residuals were white noise. According to the patterns found in the residuals, the appropriate autoregressive and/or moving average terms were added. The statistics Q_6 and Q_{12} are the Ljung-Box Q-statistics for respectively the first 6 and 12 residual autocorrelations. For none of the models the null hypothesis of independent residuals can be rejected. The residuals were also tested for heteroskedasticity by the ARCH LM test. For all final models, the hypothesis of homoskedastic residuals could not be rejected.

The AIC value was defined in Equation 18. The BIC or Schwarz's Bayesian Information Criterion, is an alternative to the AIC, which tends to penalize the model more for additional estimated parameters. It is defined as:

$$BIC = -2\left(l/n\right) + k\left(\log(n)/n\right) \tag{49}$$

where, as before, n is the number of observations used in the model, l is the value of the log likelihood function and k is the number of estimated parameters.

The MAPE was defined in Equation 19, and is used to assess the forecasting accuracy of the models. An alternative measure for forecasting accuracy is the Theil Inequality Coefficient or Theil IC (Theil, 1966). It is defined as:

$$Theil\ IC = \frac{\sqrt{\sum_{t=n+1}^{n+h}\left(Y_{t,pred} - Y_t\right)^2 / h}}{\sqrt{\sum_{t=n+1}^{n+h} Y_{t,pred}^2 / h} + \sqrt{\sum_{t=n+1}^{n+h} Y_t^2 / h}} \tag{50}$$

In this expression, $Y_{t,pred}$ is the predicted value of Y_t and h is the length of the forecasting period. The Theil IC always has a value between 0 and 1, where 0 indicates a perfect fit. An additional property of the Theil IC is that it can be decomposed in a bias proportion, a variance proportion and a covariance proportion. The bias and variance proportion respectively indicate how the mean and variation of the forecast are deviating from the mean and variation of the actual series. The covariance proportion is the remaining percentage that is due to the unsystematic forecasting error (QMS, 2004). It is clear that the covariance proportion, shown in TABLE 15 between brackets, should be as large as possible. Note that this statistic can only be calculated if observed values are available for the forecasting period. The values for the Theil IC are reasonably low for all models, while the covariance proportions are around 80%. The models should therefore be able to produce reasonable forecasts for the safety outcomes.

TABLE 15: Diagnostic checking for ARIMA models

	NACCKIL	**NPERKIL**	**NACCSI**	**NPERSI**
Q_6	3.078 (0.380)	2.128 (0.546)	1.774 (0.621)	2.666 (0.445)
Q_{12}	7.500 (0.585)	8.681 (0.467)	10.214 (0.333)	11.565 (0.239)
ADF	-4.400 (0.000)	-4.707 (0.000)	-5.235 (0.000)	-4.526 (0.000)
ARCH-LM	0.238 (0.625)	0.003 (0.954)	0.000 (0.984)	0.000 (0.989)
AIC	-1.283	-1.335	-2.327	-2.215
BIC	-1.249	-1.301	-2.293	-2.182
MAPE	9.803	9.686	6.065	6.429
Theil IC	0.058 (0.828)	0.057 (0.835)	0.035 (0.805)	0.038 (0.836)

The parameter estimates are shown in TABLE 16. Note the high similarity between the number of accidents and the corresponding number of victims, both for the fatalities and serious injuries. One of the drawbacks of the ARIMA approach is that it is not always easy to give a clear interpretation of the results. A very general interpretation of ARIMA models is of course that each observation of the series can be seen as a weighted average of past observations. For MA models a more precise interpretation can be given. Each forecast from a MA model is an exponentially weighted moving average of a portion of the available data (Pankratz, 1991). For example, in a seasonal MA process ARIMA(0, 0, 0)(0, 1, 1)$_{12}$, an estimate for a variable x_t (say January) gives a certain weight to the same month one year ago (x_{t-12}), a smaller weight to x_{t-24}, and so on. In the case of an ARIMA(0, 1, 2)(0, 1, 1)$_{12}$, these weights for x_t are combined with one period weights (x_{t-1}, x_{t-2},…) and two period weights (x_{t-2}, x_{t-4},…). Each observation is therefore influenced by the value in the two previous periods and by the value in the same month of the previous year.

TABLE 16: Parameter estimates for ARIMA models

	Coefficient	Std. Error	t-Statistic	Prob.
Number of fatal accidents (NACCKIL)				
θ_1	-0.741704	0.051384	-14.43455	0.0000
θ_2	-0.122469	0.048219	-2.539880	0.0115
θ_3	-0.924447	0.017855	-51.77614	0.0000
Number of persons killed (NPERKIL)				
θ_1	-0.708261	0.049386	-14.34123	0.0000
θ_2	-0.151179	0.042474	-3.559301	0.0004
θ_3	-0.932571	0.015218	-61.28232	0.0000
Number of serious injury accidents (NACCSI)				
θ_1	-0.674511	0.043994	-15.33182	0.0000
θ_2	-0.155409	0.043488	-3.573624	0.0004
θ_3	-0.919993	0.017013	-54.07662	0.0000
Number of serious injuries (NPERSI)				
θ_1	-0.668746	0.043937	-15.22058	0.0000
θ_2	-0.143713	0.043780	-3.282623	0.0011
θ_3	-0.944164	0.007760	-121.6736	0.0000

FIGURE 34 shows the predictions for the four dependent variables. The ARIMA predictions follow the seasonal pattern and, apart from the random fluctuations, they seem to predict the data quite well. However, the graphs give the impression of overestimating the number of accidents and victims for the years 2003 and 2004, especially for the number of fatalities. This may be explained by the significantly lower number of accidents in these years.

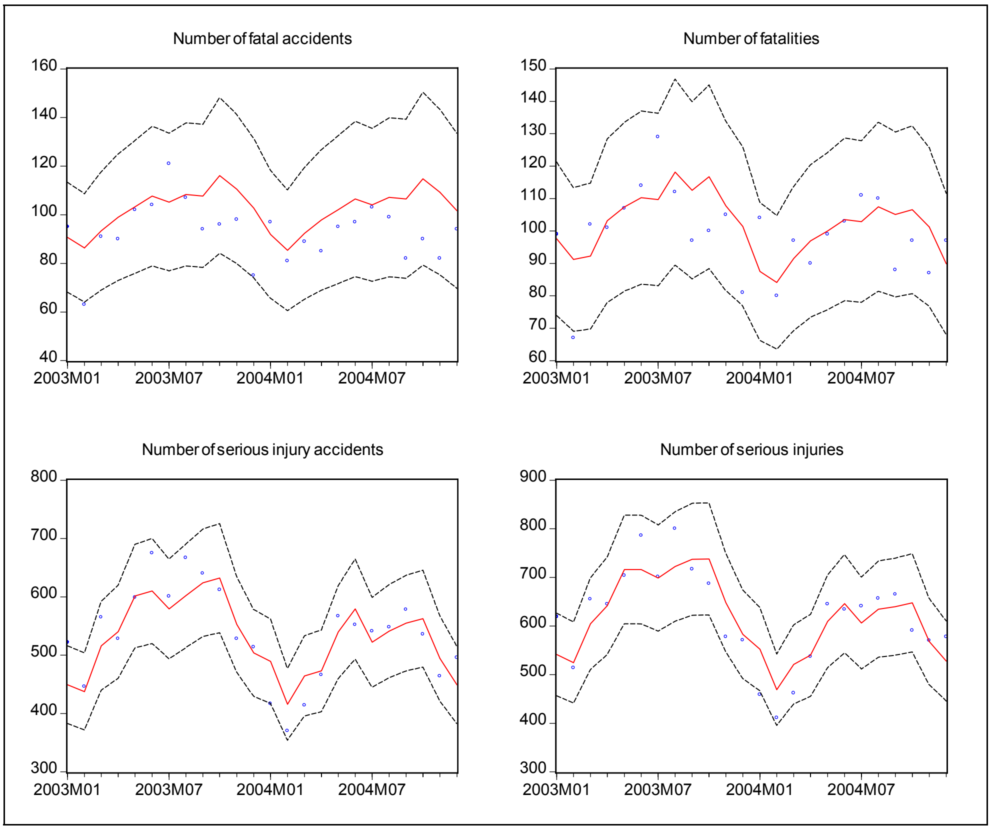

FIGURE 34: ARIMA predictions for 2003 and 2004

6.2.4 State space model for Belgian road safety data

Apart from the fact that the models can be seen as exponentially weighted moving averages of the available data, the ARIMA models do not provide any insight in the underlying patterns of the data. Typically, an ARIMA model filters away the trend and the seasonal component to obtain a series that can be modelled as a combination of moving average and autoregressive terms. Given the high degree of seasonality and the clear trends in accidents and victims, it might be instructive to explicitly model these components. State space methodology and unobserved components models are useful in this respect. This approach explicitly models the trend and the seasonal instead of filtering them away.

The four series that were modelled by means of ARIMA structures in the previous section will now be analysed with a state space model. For every model, first a stochastic level, slope and seasonal is assumed. If the variance of the corresponding component is zero, the component is fixed. If the component is at the same time not statistically significant, it is dropped from the model. An autoregressive component is also added. Further, the model residuals are evaluated and tested for serial correlation and normality. For serial correlation, the Box-Ljung Q-statistic is used (see for example equation 14 for the specification). To test for normality, the Bowman-Shenton test is used. This test takes into account the skewness and the kurtosis of the distribution of the residuals and is tested against a $\chi^2(2)$ distribution (Koopman et al., 2000). If necessary, intervention variables are defined to render the residuals normal. All this is, of course, an iterative process, of which only the final results will be shown here. Model estimation is done in STAMP 6.21.

TABLE 17: Diagnostic checking for state space models

	NACCKIL	NPERKIL	NACCSI	NPERSI
Q_6	1.927 (0.382)	2.281 (0.320)	9.730 (0.008)	9.433 (0.009)
Q_{12}	5.777 (0.672)	5.291 (0.726)	18.241 (0.020)	21.515 (0.006)
AIC	-4.338	-4.363	-5.543	-5.396
AIC*	-1.257	-1.257	-2.363	-2.209
BIC	-4.154	-4.169	-5.325	-5.167
BS-test	1.481 (0.477)	0.915 (0.633)	3.846 (0.146)	3.925 (0.141)
MAPE	9.969	9.450	7.159	7.642
Theil IC	0.059 (0.869)	0.057 (0.924)	0.044 (0.925)	0.045 (0.939)

** Calculated AIC value, to compare with TABLE 15*

In TABLE 17, the model diagnostics are shown. The Q_6 and Q_{12} statistics test for serial correlation up to order 6 and 12. For the models related to the fatalities, there does not seem to be a problem, but this cannot be said for the serious injury outcomes. Clearly, because of the correlation left in the residuals, the state space model is not capturing all the dynamics of the series. The ACF for these series (not shown here) indicate high correlation at order 4, 5 and 10. Perhaps this is caused by some outliers. However, as can be seen from TABLE 18, the serious injury series are already corrected by five distinct outliers. This results in normally distributed residuals (see the BS normality test in TABLE 17), but does not solve the correlation problem. A possible treatment would be to include lags of the dependent variable as explanatory factors in the model, creating a distributed lag model with a state space structure. In this analysis, however, the model is kept as such.

The AIC values in TABLE 17 are not completely comparable to those obtained for the ARIMA models, because of differences in the calculations. To make both AIC values comparable, the log-likelihood obtained in the STAMP software should be corrected by a constant equal to (Durbin & Koopman, 2001):

$$-\frac{n}{2}\log\left(2\pi\right)-\frac{n-1}{2} \tag{51}$$

Here, *n* is equal to the number of observation used in the log-likelihood calculations. This new value is subsequently imputed in the AIC definition given in Equation 18. The new AIC values (AIC^*) are now in the same order of magnitude and can be used to compare the state space models with the ARIMA models. Clearly, the differences among the models are small. The ARIMA models for fatalities are slightly better than their state space counterparts, and for the serious injuries the models have a comparable fit. Further, the MAPE and Theil IC statistics show the quality of the predictions. The ARIMA and state space models perform equally well for the fatality outcomes, but the ARIMA predictions for the serious injury outcomes are slightly better.

TABLE 18: Intervention and AR variables

	NACCKIL	**NPERKIL**	**NACCSI**	**NPERSI**
JUN73	-	-	-	-0.737 (0.000)
JAN79	-0.839 (0.000)	-0.722 (0.000)	-0.585 (0.000)	-0.575 (0.000)
JAN85	-0.673 (0.000)	-0.673 (0.000)	-0.324 (0.000)	-0.319 (0.000)
AUG99	-	-	-0.234 (0.000)	-0.214 (0.001)
OCT02*	-	-	-0.252 (0.000)	-0.262 (0.000)
AR(1)	-0.092	-0.096	-0.008	0.001

** The OCT02 intervention variable corrects the level.*

TABLE 19 shows the values taken by the levels and the slopes in the four models at the end of the sample. These values are in logarithms. Note that the slope component was not significantly different from zero, and was therefore dropped from the model. Returning to the original scale, it is possible (for the models with a significant slope component) to calculate a growth rate at the end of the sample. It can be seen that the yearly reduction in persons killed is smaller than for the serious injury accidents and victims.

TABLE 19: Final state diagnostics

	NACCKIL	NPERKIL	NACCSI	NPERSI
Level	4.646 (0.000)	4.697 (0.000)	6.121 (0.000)	6.304 (0.000)
Slope	n.a.	-0.002 (0.003)	-0.003 (0.000)	-0.003 (0.000)
Growth rate	n.a.	-2.633%	-3.141%	-3.300%

All models in this study were estimated with a fixed seasonal component. This was done so because the variance of the seasonal was estimated to be zero. The fixed seasonal pattern for accidents (left) and for victims (right) is shown for two years in FIGURE 35.

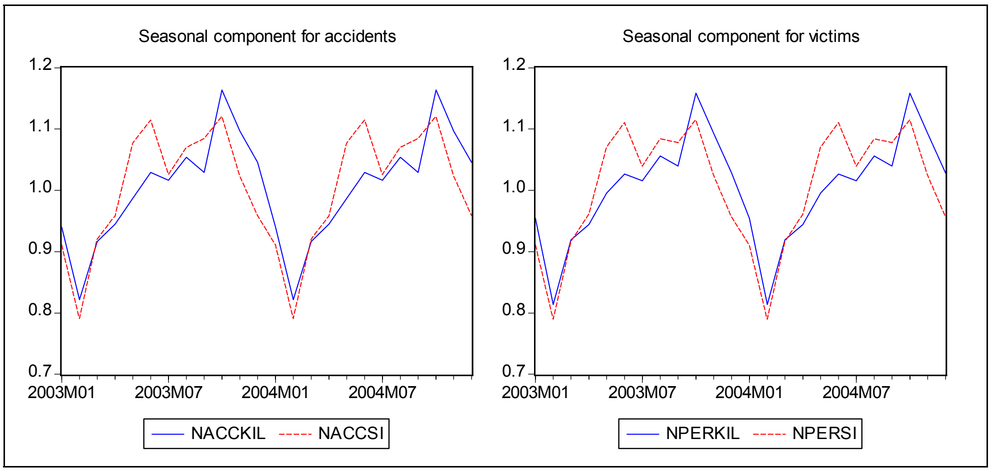

FIGURE 35: Seasonal pattern for accidents and victims

As one can expect, there is a high similarity between the curves for accidents and victims. The distribution over the months, however, depends on the outcome modelled. That is, the pattern is different for fatalities compared to serious injuries. For both outcomes, the month of February seems to be the least dangerous. The number of fatalities is clearly highest in October, while the number of serious injuries is high in June and in October.

Another interesting output from the state space model is the predicted trend for the dependent variables. As explained before, the trend does not have to be deterministic, but can move over time stochastically. Also, it is not necessary to filter the trend away. FIGURE 36 shows the estimated trends. As the seasonal and random fluctuations are separated from the trend, the graphs show a smooth picture of the evolution in road safety. The overall is much more visible from these graphs compared to the original data in FIGURE 33.

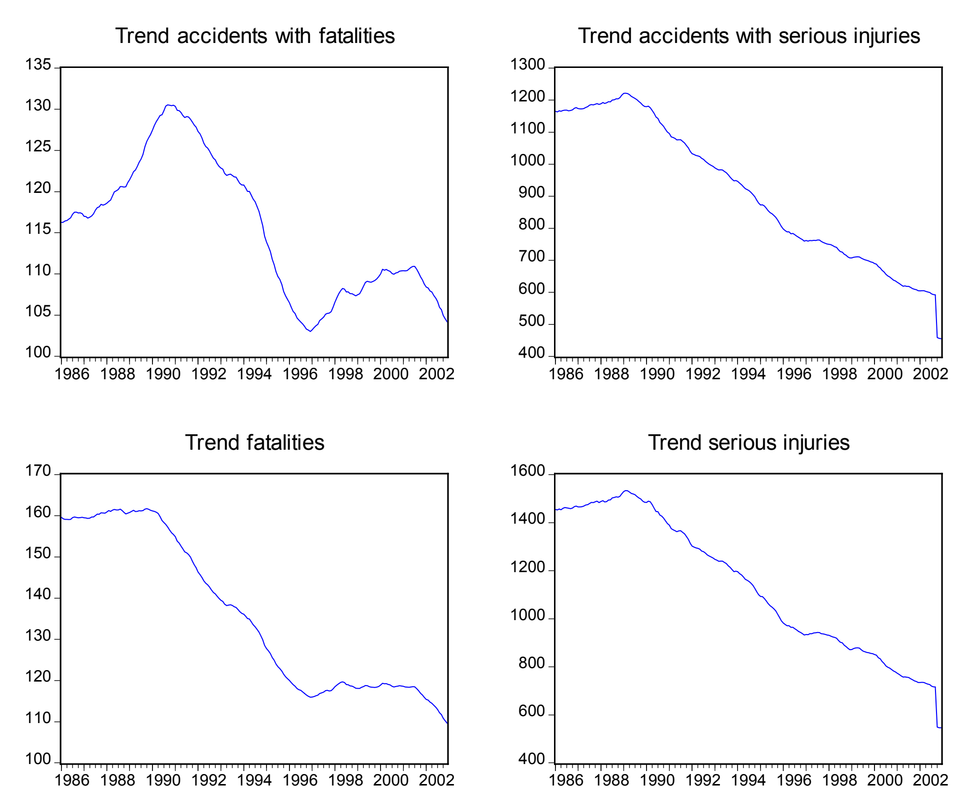

FIGURE 36: Trends for accidents and victims

To visually compare the predictions of the state space models with those obtained with the ARIMA models, FIGURE 37 shows the predicted values for the two out-of-sample periods. The graphs confirm the high similarity between the models. However, the discussion showed that state space models allow more insight in the descriptive components than the classical ARIMA models.

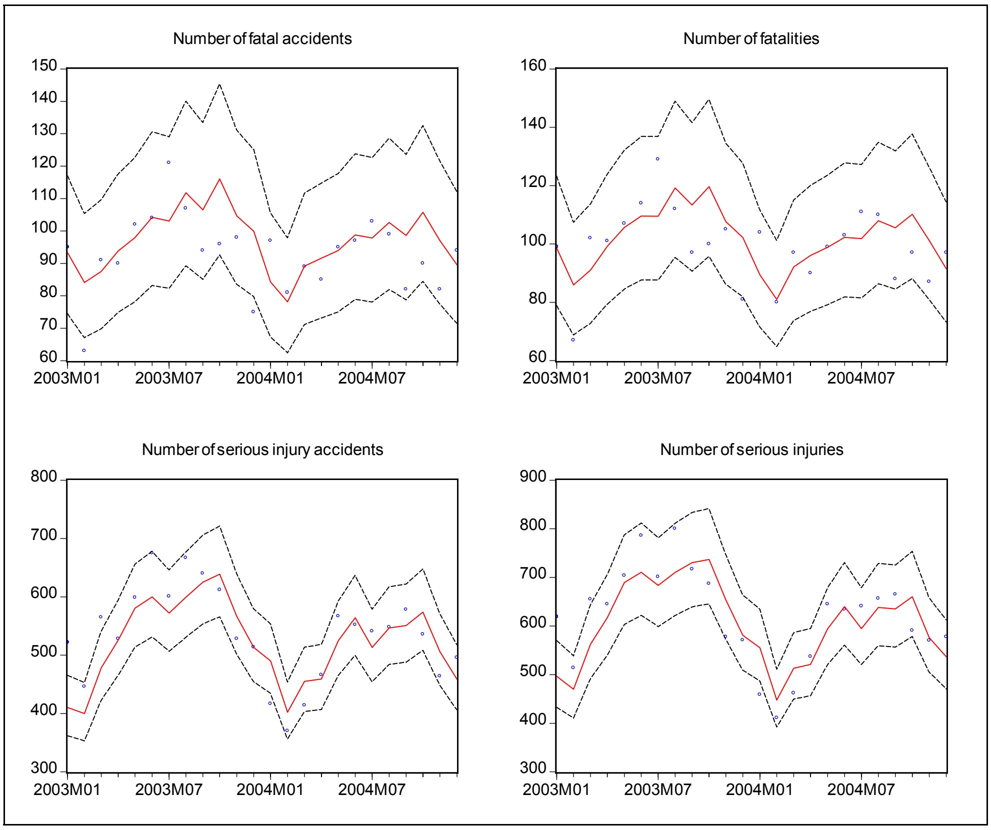

FIGURE 37: State space predictions for 2003 and 2004

6.3 Explanatory models

Whereas the descriptive models are useful to predict the future level of road safety and, in the case of state space models, to show the underlying trend and seasonal pattern, explanatory models include some variables that, at an aggregated level, might influence road safety. In this section, the ARIMA models discussed above are extended with explanatory variables. First, some background is given on how explanatory variables can be added to the ARIMA framework. Second, an explanatory model is given with calendar variables. Third, the model further extended by an exposure measure and other explanatory variables.

Note that the same procedure could be followed with state space models. This framework is at least equally flexible in including explanatory variables as the ARIMA models. The comparison of both approaches in terms of explanatory power for road safety models is an interesting topic further research, for which a start was already given in (Hermans et al., 2006b).

6.3.1 Extending ARIMA models with explanatory variables

An ARIMA model can be used in combination with classical regression. In this case, the regression model is fitted to capture the relation between the dependent and one or more independent variables, while an ARIMA structure is imposed on the residuals of the regression model. The multiple regression model can be written as:

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \ldots + \beta_k X_{k,t} + N_t \tag{52}$$

Here, Y_t is the t-th observation of the dependent variable and $X_{1,t}$, …, $X_{k,t}$ are the corresponding observations of the explanatory variables. The parameters β_0, β_1, β_2, …, β_k are fixed but unknown, and N_t is the unknown error term, which is assumed to be normally distributed. Using classical optimisation techniques, estimates for the unknown parameters are obtained. If the estimated values for β_0, β_1, β_2, …, β_k are given by b_0, b_1, b_2, …, b_k, then the dependent variable is estimated as:

$$Y_{est,t} = b_0 + b_1 X_{1,t} + b_2 X_{2,t} + \ldots + b_k X_{k,t} \tag{53}$$

The estimate $N_{est,t}$ for the error term N_t is calculated as the difference between the observed and predicted value of the dependent variable: $N_{est,t} = Y_t - Y_{est,t}$.

In the classical regression model, assumptions are made about heteroskedasticity, autocorrelation and normality. Typically, the autocorrelation assumption is likely to be violated in regression models with time series data. In a regression with autocorrelated errors, the errors will probably contain information that is not captured by the explanatory variables, and it is necessary to extract this information to finally end up with uncorrelated ("white noise") residuals. Autocorrelation can be taken into account by including an ARIMA model for the residuals of the regression equation.

As an example, assume a regression model with one explanatory variable, denoted as:

$$Y_t = \beta_0 + \beta_1 X_{1,t} + N_t \tag{54}$$

Suppose further that the error terms are autocorrelated, and that they can be appropriately described by an ARMA(1,1) process. This model can then be written as:

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 X_{1,t} + N_t \\ (1-\phi_1 B)N_t &+ (1-\theta_1 B)a_t \end{aligned} \tag{55}$$

and a_t is assumed to be white noise. Substituting the correction for the error term into the regression equation gives:

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \frac{(1-\theta_1 B)}{(1-\phi_1 B)} a_t \tag{56}$$

When the residuals of the regression equation are not stationary, then, as in the case of the classical ARMA models, they should be differenced to make them stationary. In this case, according to (Pankratz, 1991), all corresponding series (both the dependent and the explanatory variables) should be differenced. This can be seen from the small regression example. Differencing the error terms twice results in the following expression, with the ARMA(1,1) model for the differenced error terms:

$$\nabla_{12}\nabla N_t = \frac{(1-\theta_1 B)}{(1-\phi_1 B)} a_t \Leftrightarrow N_t = \frac{(1-\theta_1 B)}{\nabla_{12}\nabla(1-\phi_1 B)} a_t \tag{57}$$

Substituting back this expression into the regression equation gives:

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \frac{(1-\theta_1 B)}{\nabla_{12}\nabla(1-\phi_1 B)} a_t \tag{58}$$

or, equivalently,

$$\nabla_{12}\nabla Y_t = \beta_0' + \beta_1 \nabla_{12}\nabla X_{1,t} + \frac{(1-\theta_1 B)}{(1-\phi_1 B)} a_t \tag{59}$$

The intercept is now possibly different, but the (theoretical) regression coefficient β_1 is not affected by the differencing operation. However, its estimated value may differ slightly, since the estimation is done on different (although related) time series.

6.3.2 A model with calendar variables

In a first explanatory model, only variables related to the calendar are considered. Given the large number of possible explanatory factors in road safety models, the gathering of data requires a huge effort. The main condition to develop explanatory models is the availability of data. Moreover, if the dependent variables should be predicted with an explanatory model, future values of the explanatory factors are necessary. This implies another set of predictions that must be made beforehand. This is certainly an advantage of the descriptive models that were developed in the previous section. These models describe a time series in terms of the general trend and a possible seasonal pattern or autoregressive and moving average components, without providing any explanatory power.

On the other hand, some simple variables can be found that provide insight in the series and that are always available, for the past, the present and the future. These are variables that are related to the seasons and the calendar. It is not unrealistic to assume that these variables can help in understanding road safety time series. First, there is a seasonal pattern present in accident data. Some months always have a higher number of road crashes and victims than others. This is also related to the length of the month. It is to be expected that a 28-day month (February) will have a lower number of road crashes than a 31-day month, given the almost 10% difference in the length of the time period. Indicator variables for the months can capture these patterns. Also, it is known that the exposure to crashes is higher in some months compared to others, like for example during holidays. These peak moments in traffic can explain the number of crashes and victims during a given month. By including variables that reflect peak exposure, these effects can partly be accounted for. Second, given the problem of weekend crashes in Belgium, it is to be expected that crash counts are higher in months with more weekend days. Third, the planning of official holiday periods can influence the exposure in a month, and thus the number of crashes. Easter holiday can shift between March and April, and starting weekends of holiday periods always lead to a different kind of traffic. Based on the planning of official holiday periods, it is possible to foresee the weekends with special travel in a year. Fourth, the calendar variables (like the number of weekdays and weekend days in a month) are known for every year in the future. These properties of calendar variables make them quite appealing to practitioners, because they enrich the model and allow predictions without a heavy effort of data collection and cleaning.

In this section, some explanatory models are developed in which all independent variables are based on the calendar. This offers the undeniable advantage of availability of the data. Because an ARIMA model filters away the trend and the seasonal pattern, these components are not explicitly included. Including seasonal dummies would introduce a fixed seasonal, which is not a bad choice as such, but might unnecessarily increase the number of parameters in the model. In the model, the trading day and special day variables are included, together with an indicator for February months in leap years and level shift variables for the laws of June 1975 (seatbelt law), January 1991 (mandatory seat belt use in the rear seats), January 1992 (new speed limits), December 1994 (higher alcohol fines) and January 1998 (automatic speed cameras at intersections). These variables were introduced in Chapter 4. Also, in order to obtain normally distributed residuals, correction variables (0/1) are added to the models. Typically, these "interventions" can be related to severe weather conditions like in January 1979 and January 1985. For the serious injury outcomes, also August 1999, October 2002 and November 2002 were added, however without a clear interpretation. Fact is that these periods show a residual that is clearly smaller or larger than average. The diagnostic results of the models are shown in TABLE 20.

TABLE 20: Diagnostic checking for Regression-ARIMA models

	NACCKIL	NPERKIL	NACCSI	NPERSI
Q_6	1.930 (0.587)	1.622 (0.654)	3.666 (0.160)	2.824 (0.244)
Q_{12}	8.514 (0.483)	7.745 (0.560)	9.713 (0.286)	7.276 (0.507)
ARCH-LM	0.033 (0.856)	1.342 (0.247)	1.765 (0.184)	2.180 (0.140)
JB	1.604 (0.448)	0.457 (0.796)	3.892 (0.143)	2.330 (0.312)
AIC	-1.580	-1.636	-2.742	-2.655
BIC	-1.436	-1.491	-2.553	-2.464
MAPE	9.410	9.821	6.504	7.352
Theil IC	0.055 (0.788)	0.055 (0.801)	0.035 (0.887)	0.041 (0.900)

Like the ARIMA models (see TABLE 15) and some of the state space models (see TABLE 17), also the regression models with ARIMA errors are valid in a statistical sense, as can be seen from the Q-statistics in TABLE 20. Based on the AIC values, the regression models are slightly better in terms of model fit. For the prediction accuracy, however, the results are not so different. According to the MAPE, the ARIMA models are best for the number of serious injuries and the corresponding number of crashes. For the number of fatalities, the state space model has the lowest MAPE. The Theil IC is lower for all regression models, except for the seriously injured persons. The differences are, however, rather small. Note that the highest values for the covariance proportions of the Theil IC values are obtained for the state space models. The ARCH-LM test for heteroskedastic residuals in TABLE 20 is the same as was introduced before, and shows satisfactory results. To test for normality, the Jarque-Bera (JB) statistic is used, which has a $\chi^2(2)$ distribution under the null hypothesis of normally distributed errors. For none of the models, the normality assumption can be rejected.

The advantage of the regression models lies in the information provided by the calendar variables. TABLE 21 shows the most important results (*p*-values between brackets). Only the variables related to the calendar and the laws are shown. The autoregressive and moving average results are similar to those obtained in the pure ARIMA models.

TABLE 21: Parameter estimates for Regression-ARIMA models

	NACCKIL	**NPERKIL**	**NACCSI**	**NPERSI**
Trading Day	-0.0028 (0.130)	-0.0020 (0.269)	-0.0034 (0.001)	-0.0051 (0.000)
Special Day	-0.0006 (0.887)	0.0018 (0.636)	-0.0043 (0.0407)	-0.0039 (0.072)
Leap Year	0.0809 (0.058)	0.1173 (0.005)	0.0406 (0.079)	0.0429 (0.075)
Law0675	-0.1771 (0.003)	-0.1770 (0.002)	-0.0858 (0.028)	-0.1372 (0.001)
Law0191	-0.0153 (0.792)	-0.0265 (0.639)	-0.0214 (0.583)	-0.0221 (0.593)
Law0192	-0.0888 (0.122)	-0.1247 (0.028)	-0.0495 (0.199)	-0.0744 (0.071)
Law1294	-0.2024 (0.001)	-0.1931 (0.001)	-0.0803 (0.037)	-0.0699 (0.091)
Law0198	0.1050 (0.070)	0.1103 (0.053)	0.0086 (0.823)	0.0084 (0.839)

The trading day variable is significant for the serious injury outcomes. Since the sign of the variable is negative, months with more weekend days may be more dangerous than months with more weekdays. This shows the influence of the calendar composition in terms of weekdays and weekend days on the number of persons seriously injured, which confirms the expectations for Belgium. Weekend crashes are frequently observed, mostly with serious consequences. The trading day variable can be used to quantify the number of fatalities expected from an extra weekend day in the month. As an example, compare the months of August in the years 1997 and 2000. In 1997, 21 weekdays and 10 weekend days are observed. For this month, the trading day variable TD_{AUG97} equals -4. The same month in 2000 has 23 weekdays and only 8 weekend days. Therefore, TD_{AUG00}

equals 3. Given a parameter estimate of -0.0051 for the model NPERSI, the effect on the dependent variable is 0.0205 for August 1997, and -0.0154 for August 2000. Applying the exponential function results in an increase of persons seriously injured of exp(0.0205)-1=0.0207 or 2.07% for August 1997, and a decrease of 1-exp(-0.0154)=0.0153 or 1.53% for August 2000. A similar reasoning applies to the number of serious injury accidents. Note that this is only the effect attributable to the trading day pattern. In general, comparing two months with 9 and 10 weekend days respectively, results in a global increase in serious injuries of 1.8% for the additional weekend day, all other things being equal. Given the high monthly number of victims in Belgium, this percentage is quite considerable. The trading day variable is an interesting instrument for policy makers. The models allow measuring the number of injuries that can be expected based on the calendar structure. In a month with more weekend days, safety campaigns can be directed towards the group of people that is likely to be on the road during weekends and involved in a crash. According to the statistics, mainly young drivers die in fatal weekend accidents. In 2002, 65% of the Belgian drivers killed in a weekend crash were between 18 and 34 years old (Jolly, 2005). This group of road users is obviously the first target for weekend-related road safety campaigns.

The special day variable is significant for the serious injury outcomes, with a negative sign. For example, if a month counts one extra special day, as defined in Chapter 4, the number of serious injuries decreases by 1-exp(-0.0039)=0.39%, all other things being equal. This may seem counterintuitive at first sight, but can easily be explained. In periods of special traffic, caused by public or national holidays and Christian holy days, on the one hand large concentrations of traffic can be found, leading to reduced speed, and on the other hand the overall level of traffic is lower on these days.

The leap year is significant for all models, with a positive sign. For example, the number of fatalities is, on average, exp(0.1173)-1=0.1245 or 12.45% higher compared to other (February) months. Note that this is a separate effect of the leap month, because the seasonal variation already has been removed from the data. It is expected that adding a leap variable corrects for a pattern that is not captured by the seasonal.

The major traffic laws of the last thirty years show mixed effects on road safety. First, the law of June 1975 on seat belt use in the front seats significantly reduces all road safety outcomes, and can be seen as a major step forward in road safety. Contrary to this law, the introduction of mandatory seat belt use in the rear seats (January 1991), together with laws to improve the safety of vulnerable road users, is not significant. This may be explained in several ways. First, the number of persons travelling on rear seats is smaller compared to the front seats. Second, and perhaps more important, according to attitude measurements in 2003 and 2004 (Silverans et al., 2005), the level of seat belt wearing in the rear seats is significantly lower than in the front seats. According to the study, only 40.9% of the rear seat passengers report systematic seat belt use, compared to 62.7% and 67.8% as a driver or front seat passenger respectively. Also, 42.5% of the surveyed rear seat passengers rarely if ever used a seat belt. Given the fact that self-reported attitude measurements usually overestimate the observations, it is expected that the true level of seat belt wearing is even lower. Third, due to the fact that 5 different measures were introduced at the same time (see Chapter 4 for more details), this law is not easily tested in an aggregated model. The introduction of higher alcohol fines (in December 1994) was successful in reducing the level of accidents and victims. The law on new speed limits did not decrease the number of accidents, but resulted in a reduction in the number of persons killed and the number of serious injuries. The installation of automatic speed cameras is significant for the fatal accidents and the number of fatalities, but has an unexpected sign, leading to more accidents and victims. Perhaps this law is different in nature, in the sense that it is a local measure, not generally applicable to all types of roads. Moreover, the introduction of the cameras was not uniformly distributed over the country, with a high density in Flanders and almost no cameras in the Walloon region. Also, the cameras were introduced gradually, which may hamper the measurement of the effect by means of a level shift dummy variable. Finally, note that the introduction of the law is almost at the end of the test data set, leaving a rather short after-law period.

The models developed in this section illustrate some interesting aspects of explanatory model building. First, they provide a road safety researcher with useful insights on the tested variables. Second, an explanatory model is not necessarily better in terms of prediction. Even a simple model can provide acceptable forecasts, without the effort of gathering data. Using calendar variables, however, offers the opportunity of getting (limited) insights in the evolutions created by the calendar constitution, without the need for gathering explanatory data and creating models with them.

6.3.3 Adding variables and the multicollinearity issue

In the previous models, the number of explanatory variables was deliberately kept quite low. This parsimonious approach towards model building has the advantage of keeping the level of multicollinearity reasonably low. On the other hand, these models are also limited from the explanatory point of view. If more variables are introduced in the model, the multicollinearity issue may rise. This section provides some background on the issue of multicollinearity, as an introduction to the explanatory models that will follow.

6.3.3.1 The nature of the multicollinearity problem

In almost all classical regression models, and variants thereof, the inclusion of too many (correlated) variables can create multicollinearity. In the presence of perfect multicollinearity, at least one of the columns in the matrix of independent X variables can be expressed as a linear combination of (some of) the other columns. In this case, the matrix $X'X$ is singular and cannot be inverted, hence the model is not identified and the parameters cannot be estimated. Similar problems can arise when some of the X variables are near-collinear. In these situations, the degree of multicollinearity is so high that, although $X'X$ is not strictly singular, numerical problems are encountered when one tries to invert it. In data sets with non-experimental data, often situations are found where (linear combinations of) the independent variables are to some extent correlated, although the correlations are not nearly 100 % and the $X'X$ matrix is perfectly invertible.

Multicollinearity among the predictor variables does not inhibit the ability to obtain a good model fit or to make inferences about mean responses (Neter et al., 1996). In fact, multicollinearity does not lead to bias, but to a reduced precision for certain (linear combinations of) parameter estimates. When variables are collinear, the estimated regression coefficients tend to have large sampling variability and certain standard errors will become large. As a result, only imprecise information is available in the model about the true regression coefficients. Because of the large variability in the coefficients, estimated parameters may be insignificant, even though a statistical relation exists between the response variable and the predictors. Further, due to the fact that collinear predictors typically move together, the common interpretation of a regression coefficient measuring the change in the expected value of the dependent variable when the predictor variable is increased by one unit, ceteris paribus, is not fully applicable. In the case of collinear explanatory variables, the regression coefficient of any variable in the model depends on the other variables that are in the model and those that are left out. The regression coefficients then only measure a partial effect, given the other (correlated) variables in the model. Moreover, in these cases adding or deleting a predictor variable will change the regression coefficients.

Multicollinearity is an intrinsic property of non-experimental data (OECD, 1997b), like road safety, economic or socio-demographic series, for which controlled experiments are not possible. However, one should realise that econometric techniques are brought into play precisely to analyse this type of data. As multicollinearity is an inescapable aspect of non-experimental data, using econometric models to analyse them is a logical choice. Multiple regression techniques are used because non-experimental data are collinear. Regression techniques are able to estimate partial effects of the independent variables and at the same time provide an estimate of its precision (variance and covariance estimates). Each single parameter estimate is then interpretable as the marginal change in the dependent variable following a marginal change in that particular independent variable, given that all other variables in the model remain constant. Of course, this is not possible in cases of perfect multicollinearity and, in all other cases, only with a precision determined by the (less than perfect) multicollinearity among the predictors.

Given the "natural" presence of multicollinearity in non-experimental studies, the data selection process and the model specification should in the first place be guided by theoretical insights in the research domain. Then, the researcher can formulate a set of hypothetical experiments, the outcome of which he would like to measure within the model (such as, for example, the effect on the number of accidents of a 10 % higher fuel price). These two starting points define, in principle, the variables that will be part of the model.

Of course, the researcher is not always free in choosing the data for the model, as he is often restricted to the available or easily accessible data sources. Also, this approach towards data selection does not mean, that multicollinearity is not important or should not be detected. It is not always straightforward to get a complete view on the multicollinearity structures in the data by only looking at the larger standard errors. In the next section, some techniques are presented to check for multicollinearity among the variables.

6.3.3.2 Checking for multicollinearity

Several techniques can be used to assess the level of multicollinearity. In general, the following phenomena typically may arise when multicollinearity is present:

- The simple correlation coefficients between pairs of predictor variables are high;
- The regression coefficients of important variables turn out to have an unexpected sign or to be insignificant;
- The regression coefficients have wide confidence intervals due to their large standard deviations;
- Insignificant results may arise for the estimated regression coefficients of important predictor variables;
- Adding or deleting an explanatory variable or an observation changes the estimated regression coefficients, in magnitude and even in sign.

Apart from these informal diagnostics, formal methods exist to detect the presence of multicollinearity. Contrary to the informal detection methods, they identify the nature of the multicollinearity and provide a quantitative tool to measure its impact. Also, the behaviour observed with the informal diagnostics may sometimes be present without multicollinearity. In this text, the method of variance components, based on a decomposition of the matrix product $X'X$ of explanatory variables, is used to detect the multicollinearity problem and to point out the independent variables that are linearly dependent (Liem et al., 2000). This procedure is especially useful in the case where $X'X$ turns out not to be invertible.

To set up the variance decomposition, the columns of the matrix X are scaled to unit length (denoted $X_{(u)}$), but not centred around their sample means, because centring obscures dependencies that involve the constant term. The sum of the eigenvalues, denoted k, is equal to the trace and each diagonal element of $X_{(u)}'X_{(u)}$ is 1:

$$\sum_{j=1}^{k} \gamma_j = tr(X'_{(u)} X_{(u)}) = k \qquad (60)$$

where γ_j are the eigenvalues of $X_{(u)}'X_{(u)}$. Small eigenvalues γ_j indicate the presence of multicollinearity. To judge what "small" is, their magnitude can be compared with k. Another possibility is to compare them to γ_{max}, which is the largest eigenvalue. Therefore, a condition index η_j is defined as:

$$\eta_j = \sqrt{\gamma_{\max} / \gamma_j} \tag{61}$$

To determine the variables that are involved in the dependency, a decomposition of the variance of the estimated regression coefficients b_k is performed. To this end, the matrix with variances and covariances of the parameters, using the scaled data, is written as:

$$Var\left(b_{(u)}\right) = s^2 \left(X_{(u)}^{'} X_{(u)}\right)^{-1} \tag{62}$$

Since $X_{(u)}'X_{(u)}$ is symmetric, it can be written as:

$$X_{(u)}^{'} X_{(u)} = PD_{\gamma}P^{'} \tag{63}$$

where D_γ = diag(γ_1, …, γ_k) and

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1,k} \\ p_{21} & p_{22} & \cdots & p_{2,k} \\ \cdots & \cdots & \cdots & \cdots \\ p_{k,1} & p_{k,2} & \cdots & p_{k,k} \end{bmatrix} \tag{64}$$

The matrix P is an orthogonal matrix with elements p_{ij} (i = 1,…,k; j = 1,…,k), and the columns of P are the eigenvectors for the corresponding eigenvalues. Given these components, one can rewrite the estimated variance of $b_{(u)}$ as follows:

$$s^2\left(b_{(u)}\right) = s^2\left(X_{(u)}^{'} X_{(u)}\right)^{-1} = s^2 PD_{\gamma}^{-1}P^{'} \tag{65}$$

and for one parameter b_j, the estimated variance is written as:

$$s^2\left(b_{(u),j}\right) = s^2 \sum_{i=1}^{k} \gamma_i^{-1} p_{ji}^2 \tag{66}$$

This expression allows considering the variance of $b_{(u),j}$ as being composed of k components. Each $\gamma_i^{-1} p_{ji}^2$ (i = 1, …, k) is a variance component of the variance of $b_{(u),j}$. The variance proportions are then defined as:

$$\varphi_{ij} = \frac{\gamma_i^{-1} p_{ji}^2}{\sum_{i=1}^{k-1} \gamma_i^{-1} p_{ji}^2} \tag{67}$$

To determine the variables that are involved in the dependencies, one can look at the eigenvectors of $X_{(u)}'X_{(u)}$ corresponding to small eigenvalues γ_j. Then, the variance components of $b_{(u),j}$ indicate the factors that cause the high variance. A table of variance-decomposition proportions can be constructed, with increasing values of the eigenvalues in the first column.

TABLE 22: Condition Index and Variance Decomposition Proportions

Eigenvalue	**Condition Index**	**Variance Decomposition Proportions**			
		***Var*(b_1)**	***Var*(b_2)**	**...**	***Var*(b_k)**
γ_1 (γ_{min})	η_1	φ_{11}	φ_{12}	...	φ_{1k}
γ_2	η_2	φ_{21}	φ_{22}	...	φ_{2k}
...	...	...	...	...	...
γ_k (γ_{max})	η_k	φ_{k1}	φ_{k2}	...	φ_{kk}

From this table, the number of interdependencies appears. A value for the condition index η_i of more than 30 is considered as an indicator of strong multicollinearity, while associated values for φ_{ik} in excess of 0.5 indicate the variables X_k involved in the collinear relations (Liem et al., 2000). Note that in each column the variance proportions add up to 1. To infer multicollinearity, one has to identify the rows with small eigenvalues. In these rows, the small eigenvalue is caused by the relationship between at least two independent variables.

6.3.3.3 Remedial Measures

As mentioned before, the presence of multicollinearity is not affecting the usefulness of a fitted model to make predictions. However, if the main purpose of the model is to assess the impact of various explanatory factors on the dependent variable, one should at least be aware of the level of multicollinearity in the data.

1. A first remedial measure might be to centre the predictor variables around their means. This typically reduces the multicollinearity among first-order

(linear) and higher order (quadratic,...) terms in a polynomial regression. Also other transformations of predictor variables (like the Box-Cox transformations proposed in the DRAG models) can sometimes be useful in reducing multicollinearity.

2. One or more explanatory variables may be dropped from the model. As a consequence, no direct information will be obtained for some of the predictor variables. Moreover, it is possible that the remaining regression coefficients in the model are affected by the correlated variable that is dropped from the model (this is called "omitted variable bias").

3. Sometimes, adding new or other cases may break down the pattern of multicollinearity. However, this option is typically not available in time series, because the added cases will probably show the same correlation patterns as the original cases in the model.

4. Another possibility is to construct composite indexes based on highly correlated variables. The method of principal components allows one to construct completely uncorrelated indexes that can be used as predictors in a regression model. It may be difficult, however, to interpret the created indexes, which limits the usefulness of this strategy in the construction of explanatory models. In a road safety context, principal components are for example used in (Christens, 2003).

5. Ridge regression is a modification of the method of least squares to overcome serious multicollinearity problems. The idea is that by allowing a certain level of bias in the estimators, a more precise estimator may be obtained. This (slightly biased) estimator may have a higher probability of being close to the true parameter value. An example of ridge regression on accident data is presented in (Christens, 2003).

Whatever procedure is used to deal with (extreme) multicollinearity, one has to keep in mind that a certain level of multicollinearity is inevitable and at the same time acceptable. As pointed out before, the important issue is that the selection of variables should be based on theoretical arguments and research objectives. In the studies presented in the subsequent section, classes of possible explanatory variables (laws, economic, climatologic,...) were selected to determine the variables that will be included in the models. Within these categories, a selection of variables should be made. If the dataset is reasonably large, the selection should be done carefully to make sure that the model can be estimated and to obtain stable parameter estimates. If the addition or deletion of another variable does not change the other estimates in the model, then one can be confident that the level of multicollinearity does not harm the estimation process. This involves a manually stepwise selection of variables. However, also the condition index procedure presented above can provide insights in the nature of possible collinear relations. In a more classical approach, the multicollinearity among a set of variables is assessed by selecting those variables for which the condition indices stay at an acceptable level. This procedure may imply a larger omitted variable bias, but gives a stronger indication that multicollinearity will not too heavily influence the parameter estimates. This approach towards the problem is used in the next section.

6.3.4 An explanatory model for Belgium

6.3.4.1 Introduction

While the approach with calendar variables, presented in the previous section, is appealing because of the low data needs and the instructive effects obtained from the results, there are compelling arguments to introduce a measure of exposure and other explanatory variables in a road safety model. In this section, an explanatory model for road safety in Belgium is developed. In particular, this model extends the examples with calendar variables to full models containing an exposure measure, economic indicators, weather variables, road safety laws and calendar variables. If necessary, a set of correction variables is included in the model to account for outliers and to obtain normally distributed residuals that satisfy the model assumptions. The exposure variable is the monthly indicator that has been created in Chapter 4, based on fuel deliveries, average fuel economy and the structure of the vehicle park. Because of the shorter time window for the exposure variable, the models in this section are developed for the period January 1986 up to December 2002. The predictive power is again assessed for the 24 observations in 2003 and 2004. The dependent variables are the number of injury accidents, the number of fatalities, the number of serious injuries and, notwithstanding the remarks made in Chapter 4, the number of light injuries.

TABLE 23 contains a description of the variables used in the models. For a full description of the variables, the reader is referred to Chapter 4. A correction variable for outliers is always indicated as the month followed by the year. For example, an outlier in August 1999 is indicated as AUG99. Autoregressive and moving average terms are written as AR or MA, followed by the order of the term, like AR1 for a first-order autoregressive term.

TABLE 23: Variable overview

Variable	**Explanation**
LNACC	Total number of injury accidents (log)
LNPERKIL	Number of persons killed (log)
LNPERSI	Number of persons serious injured (log)
LNPERLI	Number of persons lightly injured (log)
LVEHKM	Number of kilometres driven (exposure measure, log)
LCARSDRP	Number of cars per driving population (log)
LPFUELKM	Average fuel price per kilometre driven (log)
LUNEMDEGR	Degree of unemployment (log)
LTEMPF	Average temperature, expressed in Fahrenheit (log)
LQUAPREC	Quantity of precipitation, in millimetres (log)
NDAYSFROST	Number of days with frost
NDAYSSNOW	Number of days with snow
NDAYSPREC	Number of days with precipitation
LAW0191	Introduction of seatbelt law in back seats
LAW0192	Introduction of new speed limits (50 and 90 km/h)
LAW1294	Lower alcohol level allowed + higher fines
LAW0198	Installation of speed cameras at intersections
SPECIALDAY	Variable indicating special days in a month
TRADINGDAY	Composite trading day variable

After checking for multicollinearity among the explanatory variables by means of the condition index procedure described in section 6.3.3 of this chapter, the set of variables in TABLE 23 was retained. Because a table of condition indices would not be very informative, it is just mentioned that the highest condition index is 6.5 for the variables used in the road safety equations (the same variables are used in all equations) and 5.6 for the variables used in the exposure equation, meaning that there is no indication of severe multicollinearity. Knowing that autoregressive terms usually reduce the level of multicollinearity in a model (Dagenais et al., 1987), the problem is hereby assumed to be treated. In the next sections, some details on the model selection and testing procedures are given and the results of the models are presented.

6.3.4.2 Model formulation and assumptions

Contrary to the models developed in the previous sections, no differencing was needed to get stationary residuals. By adding more variables to the model, there is less information remaining in the residuals, and the stationarity assumption is often satisfied without any modification of the dependent and independent variables. The model residuals were tested on this property using the ADF test and the hypothesis of a unit root was rejected in favour of the stationarity assumption. The properties of the residuals were further assessed by testing for serial correlation, normality and heteroskedasticity. Deviances from normality are often caused by extreme values in the data. If this is the case, dummy correction variables are added to the model. Serial correlation is treated by including an AR and/or MA structure on the residuals. Heteroskedasticity is explicitly modelled in a GARCH (Generalised Autoregressive Conditional Heteroskedasticity) structure for the residuals. In general, the models have the following form:

$$\begin{aligned} \log(y_t) &= \sum_{j=1}^{k} \beta_{j,t} x_{j,t} + a_t \\ \Phi(B) a_t &= \Theta(B) \varepsilon_t \\ \sigma_t^2 &= \omega + \alpha a_{t-1}^2 + \lambda \sigma_{t-1}^2 \end{aligned} \qquad (68)$$

In the first equation, y_t indicates the variable that is modelled: the number of injury accidents, the fatalities, the serious injuries, the light injuries or the level of exposure. The right-hand side of the expression contains the explanatory variables with their associated parameter to be estimated. The residual term in this equation, a_t, may be autocorrelated and/or heteroskedastic. The second equation formulates an ARMA model for this residual term to reduce the level of serial correlation, and is similar to the structures that have been estimated earlier in this chapter. The third equation, used to model heteroskedasticity, is called a GARCH structure. It specifies that the variance of the dependent variable, σ_t^2, is modelled as a function of the lag of the squared residual and lag of the variance. The GARCH structure helps in obtaining more accurate confidence intervals and more efficient estimators, and is therefore added to the models.

In TABLE 24, an overview is given of the variables that are included in each of the models. Most explanatory variables are tested in every model, but this is not the case for the proportion of cars (only tested for exposure), the laws (not tested for exposure) and the correction variables. Also, seasonal dummies for

February (FEB) and October (OCT) were tested for all models, but they were dropped whenever they turned out to be insignificant.

TABLE 24: Variable overview per model

Variable	LVEHKM	LNACC	LNPERKIL	LNPERSI	LNPERLI
LVEHKM		X	X	X	X
LCARSDRP	X				
LPFUELKM	X	X	X	X	X
LUNEMDEGR	X	X	X	X	X
LTEMPF	X	X	X	X	X
LQUAPREC	X	X	X	X	X
NDAYSFROST	X	X	X	X	X
NDAYSSNOW	X	X	X	X	X
NDAYSPREC	X	X	X	X	X
LAW0191		X	X	X	X
LAW0192		X	X	X	X
LAW1294		X	X	X	X
LAW0198		X	X	X	X
SPECIALDAY	X	X	X	X	X
TRADINGDAY	X	X	X	X	X
FEB97		X		X	X
NOV02		X			
JUL93	X				
FEB			X	X	X
OCT			X		X
AR(1)		X		X	X
AR(2)	X			X	
AR(5)		X			
AR(6)	X				
AR(12)	X	X			X

To verify the assumptions, some statistical tests are introduced. Apart from the Box-Ljung Q-statistic to test for serial correlation, the ARCH-LM test is again used to test for autoregressive conditional heteroskedasticity in the residuals (Engle, 1982). To test for normality, the Jarque-Bera (JB) statistic is used. The diagnostics are shown in TABLE 25. It seems that the crucial assumptions are met in all models. Serial correlation is at an acceptable level (other lags than 6 and 12 were also tested), the normality assumption cannot be rejected and due to the GARCH structure, it can be assumed that the residuals are homoskedastic.

TABLE 25: Model diagnostics

	LNACC	**LNPERKIL**	**LNPERSI**	**LNPERLI**	**LVEHKM**
Q_6	4.189 (0.242)	8.484 (0.205)	7.753 (0.101)	6.012 (0.198)	3.008 (0.390)
Q_{12}	7.279 (0.608)	12.276 (0.424)	13.352 (0.205)	6.702 (0.753)	12.202 (0.202)
ARCH-LM	0.096 (0.757)	0.258 (0.612)	0.242 (0.623)	0.021 (0.884)	0.123 (0.725)
JB	0.871 (0.647)	2.390 (0.303)	1.127 (0.569)	0.863 (0.650)	1.301 (0.522)
MAPE	6.252	22.915	18.085	5.324	5.297
Theil IC	0.035 (0.711)	0.103 (0.082)	0.092 (0.313)	0.033 (0.696)	0.031 (0.597)

The table also shows the predictive performance of the models. Because of some differences in the dependent variables in comparison with the descriptive ARIMA and state space models developed before, the quality of the predictions cannot be assessed in all detail. However, for the number of fatalities and serious injuries, the comparison can be made. Clearly, the explanatory models have a much higher MAPE value than the descriptive models. For the other models, the MAPE seems reasonable. From the Theil IC values, it can be seen that, although the values in itself are quite low, the covariance proportion is much lower for the explanatory models. This is partly explained by the uncertainty that is introduced by the use of explanatory variables. Therefore, it is strongly advised against using explanatory models unless the parameter estimates of the covariates are at interest.

TABLE 26 shows the extra estimated parameters that are introduced to satisfy some of the model assumptions, like normality. Two outliers recurring in the accidents and victims models are FEB97 and NOV02, while JUL93 is a correction variable for the measure of exposure. Unfortunately, it cannot be readily explained why these periods exhibit an exceptional value. Of course, there is always the possibility of a registration error, and then it would indeed be logical that for the number of persons killed no correction is necessary, as one can assume the most complete and accurate registration for this road safety outcome. Further, the parameter for the months of February and October are significant in the equations for the victims. These months are known to have the lowest and highest number of victims, respectively. The effect is stronger for the persons killed, were on average, 8% less fatalities are counted in February and almost 10% more in October.

TABLE 26: Correction parameters

	LNACC	LNPERKIL	LNPERSI	LNPERLI	LVEHKM
FEB97	-0.355 (0.012)	-	-0.262 (0.005)	-0.327 (0.000)	-
NOV02	-0.236 (0.013)	-	-	-	-
JUL93	-	-	-	-	0.272 (0.000)
FEB	-	-0.083 (0.012)	-0.066 (0.002)	-0.050 (0.054)	-
OCT	-	0.099 (0.007)	-	0.061 (0.002)	-

TABLE 27 shows the parameter estimates needed to achieve independence of the residuals. As can be seen, autoregressive corrections were sufficient to obtain uncorrelated residuals.

TABLE 27: Autoregressive parameters

	LNACC	LNPERKIL	LNPERSI	LNPERLI	LVEHKM
AR(1)	0.454 (0.000)	-	0.527 (0.000)	0.465 (0.000)	-
AR(2)	-	-	0.232 (0.002)	-	0.180 (0.003)
AR(5)	0.290 (0.000)	-	-	-	-
AR(6)					0.113 (0.066)
AR(12)	0.398 (0.000)	-	-	0.297 (0.000)	0.613 (0.000)

Note that typically an order 1 or 12 coefficient is needed, and for some models also other orders are added. While first and twelfth order corrections can be explained by the nature of the monthly data, this is less straightforward for the orders in between. Often, these orders are initiated by outliers in the data. However, the GARCH structure will also partly take into account the possible outliers. In any case, it is assumed that, even when the extra orders are caused by outliers, a correction is preferred to large residuals that do not satisfy the assumptions.

Finally, TABLE 28 shows the GARCH structure estimated for the models. For the number of persons killed and seriously injured, only the GARCH parameter λ is significant. This indicates that the current period's variance is mainly determined by the last period's variance. For the number of accidents and the number of persons lightly injured, also the parameter α is significant, meaning that the information about the volatility of the series in the previous period influences the variance in the current period. Also for the number of kilometres, all GARCH terms are highly significant.

TABLE 28: GARCH parameters

	LNACC	LNPERKIL	LNPERSI	LNPERLI	LVEHKM
ω	0.000 (0.085)	0.002 (0.478)	0.000 (0.952)	0.007 (0.000)	0.000 (0.000)
α	-0.079 (0.012)	-0.063 (0.402)	-0.025 (0.334)	0.301 (0.011)	-0.128 (0.000)
λ	1.032 (0.000)	0.830 (0.006)	1.030 (0.000)	-0.591 (0.015)	1.018 (0.000)

6.3.4.3 Parameter estimates for explanatory variables

TABLE 29 contains the parameter estimates of the models. Parameter estimates are grouped according to their meaning: the laws, weather variables, economic indicators, calendar structure and vehicle variables. For all variables, the estimated parameter is given, with the *p*-value between brackets. For the variables in the log space, also the indirect effect is given (in italics). This is the effect of an explanatory variable on the dependent variable that is incurred via the exposure measure. When both dependent and independent variables are in the log space, then the parameter estimate equals the (constant) elasticity. The indirect effect of a variable *x* on *y*, denoted $\varepsilon^*_{y,x}$, is then equal to the elasticity of *y* with respect to changes in exposure, $\varepsilon_{y,vehkm}$, multiplied by the elasticity of the exposure variable with respect to changes in *x*, or $\varepsilon_{vehkm,x}$. More formally, this can be written as follows:

$$\varepsilon^*_{y,x} = \varepsilon_{y,vehkm} \cdot \varepsilon_{vehkm,x} \tag{69}$$

Note that also for the other variables, one could calculate a similar concept, as is for example shown in (Gaudry, 1984, revised 2002). To keep the model interpretation clear in the given context, this is not done here.

Law variables

In the table, first the results for the law variables (not tested for LVEHKM) are shown. The mandatory seat belt law in the back seat is significant for the serious injuries, and only moderately significant for the number of accidents in general. These results are in line with those obtained in the model with calendar variables. The speed laws are significant, leading to a 13% (LNPERLI) to 19% (LNPERSI) reduction in accidents and victims. Another favourable effect can be noted for the laws and fines on alcohol. This law seems to be very useful in reducing both accidents and victims. Note, however, that the percentage reduction is about 50% smaller for the light injuries. This underlines the hypothesis that drunken drivers do frequently provoke serious or fatal accidents.

Finally, the law on automatic speed cameras at intersections has a significant negative sign in the serious injury equation. The absence of further effects may be attributed to the reasons that were already given before in section 6.3.2. Also, it is still assumed that probably a compound effect of different laws is measured by the dummy variable and that a disaggregated approach towards this law might give more detailed results. An example of such a disaggregated analysis can be found in (Nuyts, 2004). In this study, the effect of automatic cameras is investigated for various locations in Antwerp. For the number of accidents (including crashes with material damage only), the parameter estimate had a negative sign, although the effect was not significant. A significant decrease was further found for accidents with persons killed or injured. According to the author, this underlines the hypothesis that the cameras will mainly reduce severe crashes. Therefore, although the results are not fully comparable with the study at hand (no effects on the number of victims were estimated, and the study is local and cross-sectional), some parallel results are found. The change in significance for the number of fatalities, compared to the regression-ARIMA model, may partly be due to the different time horizon in combination with the rather limited after-period. Obviously, a negative sign for the parameter estimates, as obtained for the seriously injured persons, is more in line with common expectations, but still this issue needs further clarification.

It is assumed that the introduction of a law results in a sudden and permanent increase or decrease in the dependent variable. For example, the introduction of the new speed limits resulted in a 1-exp(-0.160)=0.1479 or 14.79% reduction of the number of fatalities, ceteris paribus. This assumption of a "step-based intervention" is not always a natural one. Moreover, it is not possible to isolate the effect of a single measure when several regulations are put into practice at the same or at a nearby moment in time. The significant impact of laws and regulations may be better described as "something changed at that time", instead of attributing the whole effect to the law itself. Nevertheless it makes sense to test whether these changes are indeed substantial.

TABLE 29: Parameter estimates for the explanatory variables

	LNACC	LNPERKIL	LNPERSI	LNPERLI	LVEHKM
Law variables					
LAW0191	-0.056 (0.126)	-0.053 (0.244)	-0.084 (0.037)	-0.039 (0.301)	-
LAW0192	-0.154 (0.000)	-0.160 (0.001)	-0.188 (0.000)	-0.131 (0.002)	-
LAW1294	-0.081 (0.025)	-0.229 (0.000)	-0.248 (0.000)	-0.123 (0.001)	-
LAW0198	0.044 (0.450)	0.020 (0.462)	-0.191 (0.000)	-0.031 (0.342)	-
Weather variables					
LTEMPF	0.462 (0.000)	0.342 (0.000)	0.410 (0.000)	0.388 (0.000)	0.153 (0.001)
	0.080	*0.052*	*0.055*	*0.090*	
LQUAPREC	0.019 (0.037)	0.021 (0.226)	0.014 (0.213)	0.034 (0.001)	0.012 (0.169)
	0.006	*0.004*	*0.004*	*0.007*	
NDAYSFROST	0.001 (0.764)	0.003 (0.251)	-0.001 (0.709)	0.001 (0.677)	0.000 (0.783)
NDAYSSNOW	0.003 (0.017)	0.001 (0.651)	0.000 (0.794)	0.002 (0.233)	0.004 (0.002)
NDAYSPREC	0.001 (0.439)	-0.001 (0.633)	-0.001 (0.494)	0.001 (0.717)	-0.003 (0.038)
Economic variables					
LPFUELKM	0.871 (0.000)	0.269 (0.034)	1.354 (0.000)	0.786 (0.000)	-0.178 (0.019)
	-0.093	*-0.061*	*-0.064*	*-0.105*	
LUNEMDEGR	0.479 (0.000)	0.178 (0.003)	0.396 (0.002)	0.406 (0.000)	0.028 (0.604)
	0.015	*0.010*	*0.010*	*0.016*	
Calendar variables					
SPECIALDAY	-0.006 (0.020)	0.003 (0.380)	-0.002 (0.389)	-0.002 (0.514)	-0.002 (0.515)
TRADINGDAY	-0.004 (0.008)	-0.004 (0.259)	-0.007 (0.000)	-0.005 (0.010)	0.006 (0.001)
Vehicle variables					
LCARSDRP	-	-	-	-	2.079 (0.000)
LVEHKM	0.520 (0.000)	0.341 (0.000)	0.357 (0.000)	0.590 (0.000)	-

<u>Weather variables</u>

The weather variables seem to influence the number of accidents and victims in quite different ways. A higher temperature increases the number of all types of victims and the number of accidents, as well as the level of exposure. A 1% increase in average temperature may lead to increases of 0.34% (persons killed) to 0.46% (number of accidents). This positive sign is in line with results obtained in (Gaudry et al., 1995) and (Blum & Gaudry, 2000). Also note that the indirect effects are positive, increasing the total effect of the variable on the road safety outcomes. The quantity of precipitation increases the number of accidents and the persons lightly injured. This may indicate a limited level of risk compensation. The number of accidents increases, but due to lower speeds they lead only to minor injuries. Also, a month with more rainy days has less vehicle kilometres, on average, but the effect of this variable on the road safety outcomes is not significant.

The number of days with frost has no impact whatsoever on the number of accidents and victims or the level of exposure. The number of days with snow has an effect on the total number of accidents, probably most related to the number of light injuries. Perhaps the winter conditions in Belgium are not that severe that an increase in accidents should be expected. In countries where snow is more prevalent, the effect on traffic safety is comparable to that of frost. Moreover, the government invests a lot of money each year to scatter salt on icy roads, probably reducing possible effects. Also, because heavy snow is quite rare in Belgium, it might be that road users compensate for the higher risk. They probably adjust their driving habits more than in normal weather conditions. A strange result is obtained, however, for the level of exposure, as snow seems to increase the number of kilometres driven. This may be related to the construction of the variable, in which the fuel deliveries play an important role. It might be interesting to investigate the relation between weather conditions and the delivery and consumption of fuel.

Economic variables

The next lines in the table show the estimated parameters for the fuel price per kilometre and the degree of unemployment. The results for unemployment are not in line with the literature. In the current models, the effect has a positive sign and is strongly significant for the road safety outcomes, while there is no effect on exposure. In many other models, the effect is either not found or strongly negative. It might well be that differences in the economic systems between countries result in diverging results.

Also, fuel prices are usually expected to reduce the level of accidents and victims, because of a reduced demand for travel and possibly lower (less energy consuming) speed. However, it is possible that a rise in fuel prices reduces the transport demand, thereby also reducing traffic problems and increasing speeds. The reduction in traffic is supported by the negative sign obtained in the equation for the level of exposure. A 1% increase in the price per kilometre reduces the number of kilometres driven by 0.18%. This, in turn, results in negative indirect effects for the number of accidents and victims, but these are not strong enough to turn the strong direct impacts.

Calendar variables

As for the calendar variables, the special day variable has only an effect on the number of accidents, and not on the number of victims or the level of exposure. Perhaps this variable is too general to measure any specific effect. However, the sign is in line with expectation, in the sense that special days are usually synonymous with less traffic, leading to fewer accidents.

The estimated parameters for the trading day variable are comparable to those obtained earlier. In particular, the number of fatalities is not influenced, contrary to the number of persons with serious or light injuries. This variable is a clear illustration of the problem of weekend accidents in Belgium. Further, the trading day pattern increases the level of exposure. That is, more kilometres are driven in months with more weekdays, which is an acceptable assumption.

Vehicle variables

The vehicle variables in the table are the number of personal cars per capita of the population at driving age (LCARSDRP) and the exposure variable itself (LVEHKM). The variable LCARSDRP is only used to explain the level of exposure, and is highly significant in this equation. More specifically, a 1% increase in cars per driving population leads to a 2% increase in kilometres driven. Clearly, the cars are responsible for the largest part of transport.

TABLE 30: Parameter estimates for exposure

	Coefficient	Std. Error	t-Statistic	Prob.	Wald stat.
LNACC	0.5200	0.0506	10.2816	0.0000	90.0475 (0.0000)
LNPERKIL	0.3406	0.0433	7.8641	0.0000	231.8504 (0.0000)
LNPERSI	0.3573	0.0526	6.7882	0.0000	149.0907 (0.0000)
LNPERLI	0.5904	0.0514	11.4847	0.0000	63.4597 (0.0000)

The estimated parameters for exposure are shown in TABLE 29, and are reprinted in TABLE 30 in more detail. As can be seen, exposure is highly significant in all models, with a positive sign. Compared to most macro-models in the literature, however, the magnitude of the parameters is low. Although the parameters are significantly different from zero, it is instructive to test whether they are different from one. The Wald statistics (see TABLE 30), used to test the hypothesis that the exposure parameters are equal to one, show that, for each road safety outcome, the effect of exposure is less than proportional.

On the one hand, the significance of the parameters is reassuring, because it is in line with expectations and previous research. On the other hand, one might wonder why the values are so small. There are some possible explanations for this result. First, as the level of exposure increases, it is not illogical to assume that also its effect on road safety will alter. If the parameter estimates are far below one, then a 1% increase in exposure will result in a less than 1% increase in accidents and victims. Increases in exposure are nowadays slower than in the past, in line with the sigmoid trends that can usually be seen in exposure data. Second, the learning effect in road safety might also be present in the context of exposure. That is, people get used to higher levels of traffic. Per extra kilometre driven, the marginal increase in accidents will become smaller. Similar results are found in (Elvik, 2006). Here, the author introduces some "laws of accident causation, one of them being the "universal law of learning". Road user accident rate per kilometre travelled declines as the number of kilometres travelled increases. Also, driving long annual distances is associated with more success in avoiding accidents. Clearly, it might be interesting to further investigate Elvik's law in a time series context, as the results of the models at hand support his hypothesis. Third, the level of congestion is quite high in Belgium. Especially the last years, this problem is even larger than ever. A higher level of congestion is expected to increase the level of road safety. This would imply an inverse U-shaped relationship between exposure and road safety. That is, at lower levels of exposure, an increase in kilometres driven will more than proportionally increase the number of accidents and victims, up to a certain saturation level. Then, additional cars may slow down traffic to such an extent that the number of fatalities will start to decrease. In other words, at low levels of traffic, exposure surely increases the number of accidents, while at high levels the marginal effect is getting smaller. Introducing congestion in road safety modelling is, in essence, recognising the fact that possible turning points are present in the relation between exposure and road safety.

To verify this assumption, a small stylised exercise is performed. Instead of modelling a linear relation between the logarithms of the exposure measure and the road safety outcomes, as was done before, a quadratic effect is assumed. Parabolas are the most simple examples of turning point curves. They are, of course, completely symmetric around the optimum, and therefore probably not very well suited to show the real relationship, but they are useful to get the idea. When a quadratic term is added in the previous models, the results shown in TABLE 31 are obtained (using *x* as an indication of the exposure variable).

TABLE 31: Quadratic exposure: stylised exercise

	Equation	Solutions	Maximum
LNACC	$-0.0979x^2 + 1.6803x$	0; 17.1634	8.5817
LNPERKIL	$-0.0522x^2 + 0.9020x$	0; 17.2797	8.6398
LNPERSI	$-0.1023x^2 + 1.6757x$	0; 16.3803	8.1901
LNPERLI	$-0.0968x^2 + 1.7590x$	0; 18.1715	9.0857

From the signs of the coefficients, it is seen that the curves have indeed a maximum. That is, it starts from zero, then increases and finally decreases to the second solution. Because no intercept is estimated in the model, zero is always a solution, which is the road safety outcome associated with no traffic. The other solution is the theoretical level of traffic that would imply zero accidents and victims, that is, in an extreme congestion situation. From the data set used in the full models (1986-2002), it is derived that the minimal exposure level (in logs) is 8.3372, while the maximum is 9.0497. These values are at the borders of the exposure values in the study. Then, it is interesting to compare the range of exposure values with the maximum of the stylised curves. This is done in FIGURE 38. On the horizontal axis, the level of exposure is shown, while the vertical axes show the road safety outcomes analysed above. The bold lines on the parabolas indicate the exposure range in the data set. For the number of accidents and the number of persons killed, the maximum of the parabola lies in the range of observed exposure values. For the serious injuries, the observed values are to be found in the decreasing part of the curve, while for the lightly injured victims the exposure values are in the increasing part.

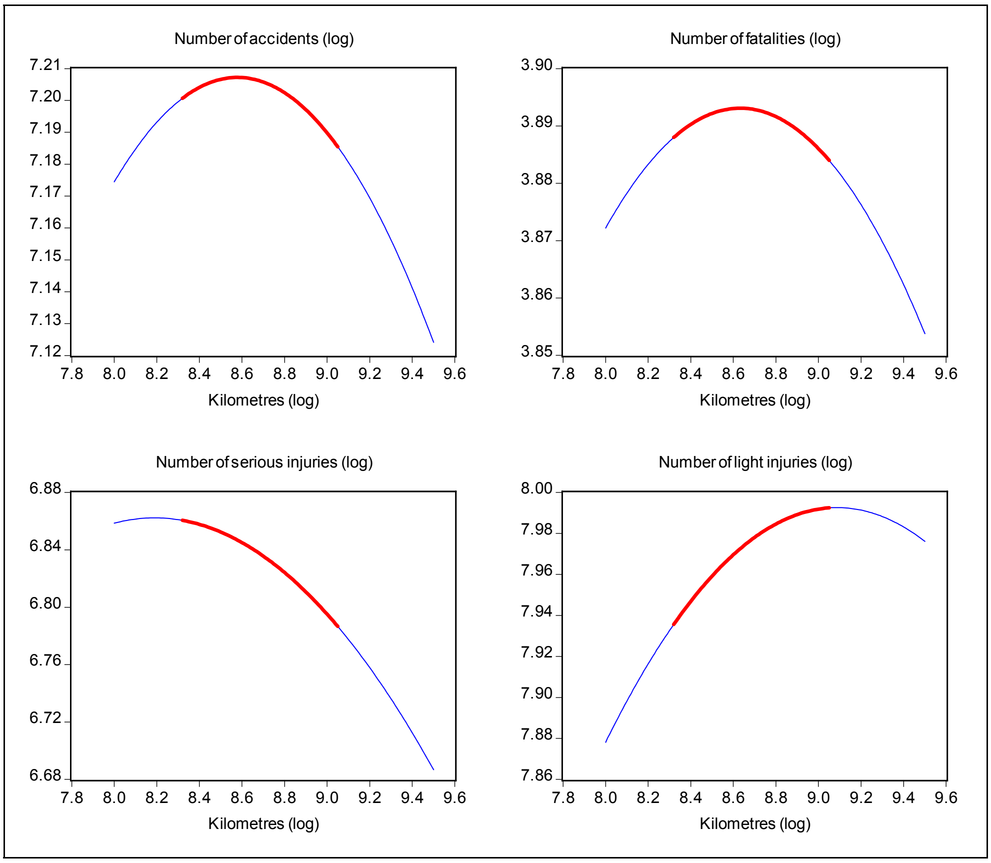

FIGURE 38: Stylised parabolas

This observation is completely in line with the congestion assumption: the number of serious injuries and victims will be lower because of lower speeds, but the number of persons lightly injured is not expected to decrease. Indeed, a congestion situation may lead to more light injuries, but it probably reduces the number of fatalities and severe injuries. Although not completely comparable, this result is also reflected in the parameter estimates in the previous models. Recall that the values were all positive, and that the estimates were 0.34 for fatalities, 0.36 for serious injuries and 0.59 for light injuries. Clearly, the largest positive effect is obtained for the number of light injuries, which are also completely in the increasing part of the parabola. Although this is a stylised study (the real relation will probably not be parabolic in nature and perhaps the functional approach can be questioned), these results are a plea for further investigation of the true nature of the relationship between exposure and risk.

6.4 Conclusion

This chapter has shown some applications of regression and time series modelling on monthly road safety data. Apart from the classical ARIMA and state space models, used to describe and forecast road safety outcomes, the framework is easily extended to include explanatory variables. Although the calendar variables, tested in a first small explanatory model, are instructive and provide interesting results, they are of course limited in their explanatory power. Clearly, a calendar composition cannot "explain" road safety, it can only point out patterns related to how time evolves. Therefore, they can better be considered as descriptive models than as explanatory models.

The introduction of explanatory variables widens the possibilities of the models, but at the same time increases the complexity of the modelling process. In the first place, explanatory models are quite time consuming at the data pre-processing stage. In particular, getting data that is suited for road safety macro modelling is not an easy task. For example, no exposure measure was available on a monthly basis and it had to be constructed from other available data sources. The availability of monthly data of fuel sales was decisive in the choice of the analysis period. These data were available from 1986 onwards, so older road safety data could not be used. This is probably a problem that is faced in every country when explanatory road safety models are developed. Second, models with explanatory variables are also less suited for prediction because of the necessity of having estimates of future values of the explanatory variables and the related problems of increased observation error when predicting with uncertain independent variables. Third, the results obtained from the models may differ from country to country. On the one hand, this pinpoints the differences between countries, but on the other hand, it shows the problem of comparing variables over countries that are developed in a completely different way. Also, the analysis periods of different models can differ considerably, leading to different results because of fundamental changes in the road system over time.

Because of these difficulties, it is very important to develop models that are statistically sound and in line with the model assumptions. Therefore, model diagnostics were presented regularly in this chapter. Also, it highlights the need for a unified framework on road safety modelling. The DRAG family of models is an example of a proposed structure that is followed by several countries. The interpretation in terms of elasticities and a recurring model structure facilitate the comparison of results. This approach, however, is quite complex and demanding in terms of development.

Another interesting point discussed in this chapter is the special role of exposure in road safety models. In the presented models, exposure was measured by the created variable for the number of kilometres driven. The variable was significant, with a positive sign, but with a less than proportional effect on road safety. Clearly, the effect of exposure is determined by a number of choices. First, the time window and the length of the series will show the relations in that period. A 30 years period of monthly observations from 1951 to 1980 will probably show other relations than a 20 years period from 1981 to 2000. It is to be expected that the relation between exposure and the number of accidents and victims changed over time. This is easily seen from the basic models developed by Oppe and his successors (see for example Chapter 5). If a logistic curve is fitted for exposure, then the marginal changes in exposure over time are getting smaller and smaller. A longer series may capture more patterns than a shorter one, and the pattern that is found depends on the period considered. This might explain the differences between the models in this chapter and previous models found in the literature. However, the signs are in the expected direction, and there are compelling reasons to assume that the road system and exposure are changing over time.

It goes without saying that the effect of exposure also depends on the measure used. Exposure in terms of population yields different relations than the number of kilometres driven. Also the quality of the measure is of primary importance. Often, this can only be assessed by checking the results obtained with expectations. However, as pointed out before, these expectations will change over time, and models will not be comparable in every respect. This is, of course, also true for the exposure measure used in this chapter. The results are reasonable, but further research should indicate if the variable indeed measures the exposure to the risk.

In the light of these remarks, it might be tempting to drop the exposure variable from the models. If the parameter estimate is getting closer and closer to zero, one can question its added value for the model. This is a difficult question, with both theoretical and practical implications. First, theoretically, it is interesting to show how the relation between road safety and exposure is evolving. This is a strong argument for keeping the variable in the model, even when it is not significant. Then, the question of importance of the variable shifts from "significant or not" to "proportional or not". Therefore, the exposure variable was often tested for a significant deviation from one instead of from zero. A parameter estimate that is significantly different from one means that the proportionality assumption is not satisfied, in either direction. A parameter that is not significantly different from zero indicates a less than proportional relationship. Dropping the variable therefore means that the relation between risk and fatalities, as determined by exposure, cannot be established. It is instructive to know the nature of this relation, even when the parameter is close to zero. Also, given the basic relationship between the variables, stating that the road safety trend is the result of changes in exposure and risk, the theoretical framework requires an exposure variable. Therefore, it is actually undesirable to fix the parameter at a particular value or drop the variable from the equation (which amounts to fixing it at zero).

Practically, there may be circumstances where dropping (or not including) the exposure measure will not invalidate the results of an explanatory model. Consider, in the given context, a small model. Starting from the "extended" relationship between fatalities and risk, namely $F_t = V_t^{\eta} \times R_t$, and further assuming a log-linear relationship, such that $\log(F_t) = \eta \log(V_t) + \log(R_t)$, it can be seen that the explanatory variables in the models developed in this chapter are used to explain the risk part of this expression, while the parameter of exposure measures the proportionality of the relation between risk and fatalities. A parameter estimate for exposure that is not significantly different from zero is in fact an indication of a relation that is not proportional. That is, the hypothesis that η equals one is probably rejected. In that case, statistical theory normally implies that the variable can be dropped from the model. By dropping the exposure variable, it is assumed that $\eta = 0$, or, in other words, that the relation is not proportional. If the proportionality is indeed not present, dropping the exposure variable from the model would not harm the results. Along the same lines of reasoning, if the exposure parameter is set equal to 1, the proportionality is explicitly assumed. In short, there are situations in which the absence of an exposure variable (that is, when exposure data are not available) will not invalidate the results, but if exposure data is available, it is instructive to test it. Rather than estimating the model without an exposure variable, one should strive to improve and extend the data sources so as to enhance the efficiency of statistical inference in these models. Removing the crucial variable of exposure will not contribute to new insights in the safety-exposure-risk relation.

Furthermore, it becomes increasingly interesting to investigate other kinds of relations between exposure and road safety, as was demonstrated with the stylised parabola example. This topic, introduced by Gaudry & Lassarre (2000), certainly needs further research, as it is both an opportunity to explore the true relations as well as a risk of over-fitting the data. Indeed, adding turning points to a relation inevitably leads to a higher number of parameters to be estimated, which should be balanced against the number of observations. Moreover, over-fitted relations are known to provide a seemingly perfect fit to the sample data, but they often produce less accurate predictions. In any case, given the changing patterns in road safety and the importance of exposure in this story, further research on this topic is crucial. It is a critical success factor for future road safety research.

Chapter 7 Disaggregated models for yearly exposure and risk

7.1 Introduction

This chapter introduces some models developed for Belgian disaggregated road safety data. In (COST 329, 2004), disaggregated models are defined as "models in which the response (dependent) variable comprises a sub-group of the total, aggregated number of accidents or their consequences". Sometimes, the term "disaggregated" is used for models in which data pertaining to individual units or decision makers (cars, travellers, households, firms, etc.) are analysed. The models presented here could therefore rather be named "accident or casualty subset models". Although they are clearly less aggregated than the models in the previous chapters, they still consider total numbers of accidents or victims, aggregated in time (per year), but now for a subset of the whole road safety system. Typically, roads are divided into highways, rural roads and urban roads, road users are analysed according to their age or gender, or crashes are studied in terms of the vehicle types involved. Disaggregated models are useful in addition to the aggregated models presented in the previous chapters. In the first place, road safety models are meant to support decision- and policy makers in their analysis of road safety developments, especially when setting road safety targets and developing road safety programmes. While an aggregate model will typically be used for the description and forecasting of general trends in road safety on a high level, they are less suited for analysing parts of the transport system or subgroups of road users. For example, changes in the proportion of young and old road users may affect the evolution in road safety. It is not possible to derive the effects of these changes from an aggregated model. Therefore, it is necessary to analyse the trends in road safety at a lower level of aggregation.

Studies at a lower aggregation level should enhance the insights in the road safety trends for parts of the road system. However, the lower the level of aggregation, the more difficult it becomes to obtain the data needed for the analysis. While aggregated statistics on road safety and exposure are generally available, they are mostly not easily found for subgroups of road users. Also, due to the smaller counts on a lower aggregation level, the data might be less accurate and more advanced statistical techniques are needed to adequately model the problem. The higher data requirements, especially in a time series context, can probably explain the relatively low number of disaggregated time series studies.

The major levels of disaggregation that are considered in a time series context are transport mode, road type and road user characteristics (age and gender). Depending on the data available, these subgroups can be crossed-over, by looking for example at the fatalities of female road users on highways or children injured in the urban area. It is clear that with more restricted groups of road users, one runs faster into data problems.

In (COST 329, 2004), an overview of several modelling approaches towards disaggregated road safety modelling is given. Broughton (1994) computed trends for casualty rates per type of road user. On the basis of population forecasts, these trends are further disaggregated according to age and gender. In (Bijleveld & Oppe, 1999), the number of fatalities and rates for combinations of transport mode and age, and separately for types of road, are analysed. They introduced the unobserved components model in this type of analysis. Later, (Bijleveld, Commandeur, Koopman et al., 2005) developed a multivariate nonlinear time series model for the analysis of traffic volumes and road casualties inside and outside urban areas.

Greibe (1999) gives a description of the number of killed and seriously injured cyclists on urban roads using a log-linear model including variables on bicycle and motor vehicle traffic and variables on general safety improvements. Pedersen (1998) uses log-linear splines to detect abrupt changes in the development of casualties. According to the author, this approach can be applied on every level of aggregation. This model, however, is based on accident data only, which is not unusual as disaggregated exposure data are often unavailable. Also some of the DRAG models, introduced in Chapter 2, can be seen as disaggregated models, in the sense that they consider aspects of specific groups of road users, like for example in (Fridstrøm, 2000). Note that these applications are different from the classical black-spot analysis or the before-after studies popularized by Hauer (1997), in the sense that they always include a trend component. That is, interest is mainly in the trends over time and, if the data allows it, in explaining certain developments.

This chapter shows four possible analyses of disaggregated Belgian road safety data. The numbers are always yearly totals, which means that no seasonal effects are included. All applications are concerned with the relation between the number of fatalities and the level of exposure. In a first study, an ARMA regression model is presented to analyse the risk of road users depending on their age and gender. In a second study, the road risk per type of road is investigated. This is done by means of a state space model that allows extracting a trend of the risk over time for each type of road. Next, the road risk per type of motorised road user is studied. Finally, the number of fatalities in 2-by-2 crashes with cars, trucks and motorcycles is modelled.

7.2 Analysis of road risk for age and gender categories

In the literature, it is often found that the risk of road users varies with their age and gender. According to Evans (2004), the number of fatalities per driver age in the US shows a peak at the age of 19 years old. The number of fatalities then steadily decreases with age. However, one of the reasons for this decline is that there are (still) fewer people of older age in the population. Correcting for the population rate shows the typical groups of vulnerable road users, namely the youngsters and the elderly.

For Belgium, data are available on the number of fatalities per age and gender category, as well as on the magnitude of the population for these groups. In FIGURE 39, the number of fatalities is shown for each age-gender category. Data are available for the period 1973-2001. For the age variable, five groups are made according to the availability of the data: (1) 0-14 years, (2) 15-24 years, (3) 25-44 years, (4) 45-64 years and (5) 65 years and more.

It is clear from the graphs that the evolution in time of the fatalities per age and gender category shows some very specific properties. First, the number of fatalities is higher for males in every age category. This is true for the complete time window of the study. However, the differences get smaller for all categories, except for the 25-44 years old victims. Second, the number of fatalities is decreasing over time, except for the 25-44 years old victims. For the latter group, the level of fatalities hardly changed. It is clear that the number of fatalities changes with age and gender. In this section, these relations will be quantified in classical exponential risk models, corrected for serial correlation.

The models that will be used to analyse the number of fatalities for each age and gender combination can be written as:

$$\begin{aligned}
&\log(F_t) = \log(P_t) + (\alpha_0 + \alpha_1 t) + u_t \\
&\Phi(B)u_t = \varepsilon_t \\
&\varepsilon_t \sim NID(0, \sigma^2)
\end{aligned} \quad (70)$$

The dependent variable is the log of the number of fatalities (F_t) for each age-gender combination (5 age categories and 2 gender categories). The index t specifies the year of the observation (t = 1973,..., 2001).

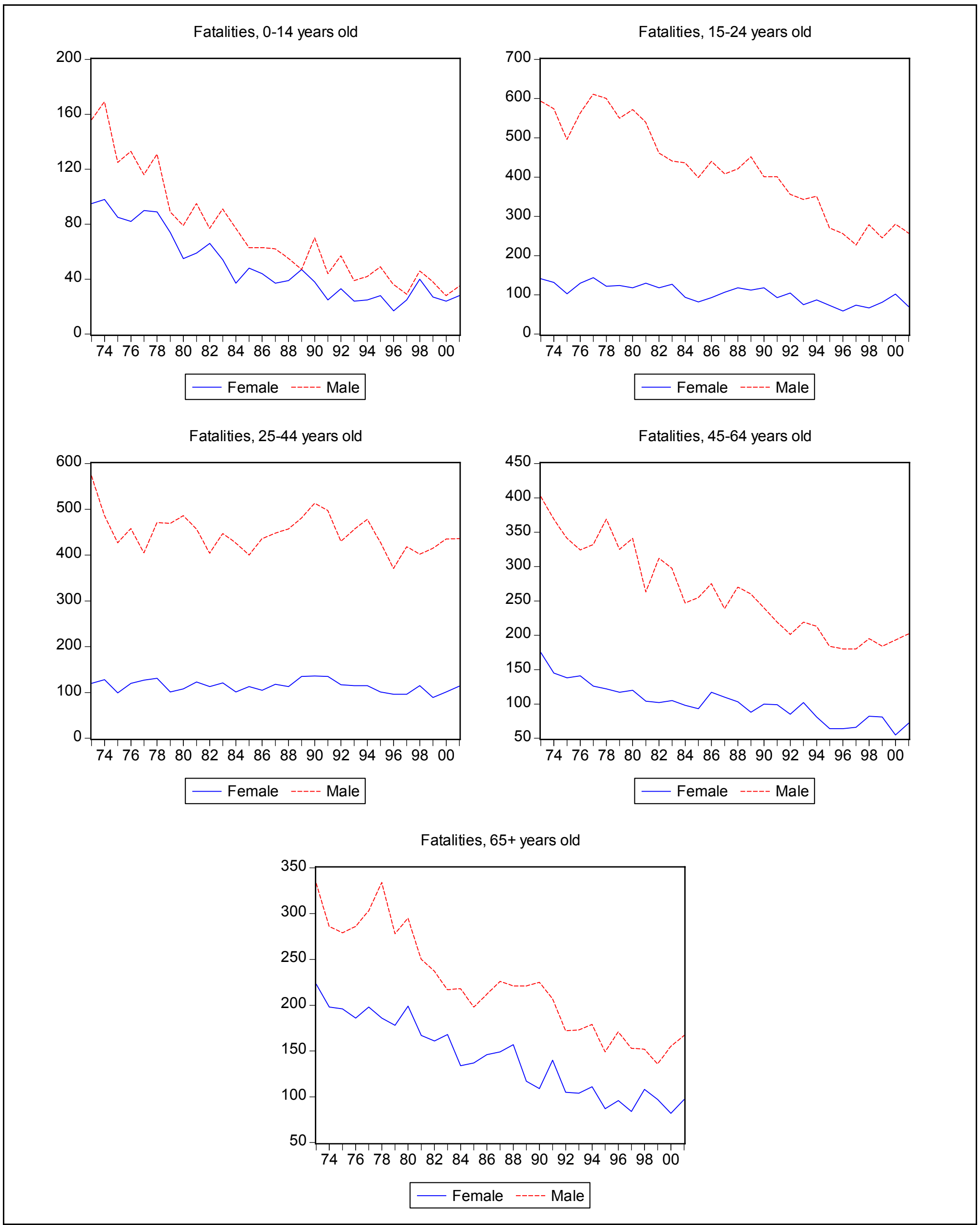

FIGURE 39: Annual number of fatalities by age and gender

The log of the population (P_t) is used as an offset to correct for exposure. As discussed in Chapter 4, there is no other exposure data available for age-gender combinations of road users. Therefore, population data will be used instead. For the risk, an exponential curve is assumed. The residuals u_t in the first equation are possibly autocorrelated. The second equation defines an autoregressive structure for the error term, whenever this is necessary to obtain normally distributed white noise residuals ε_t (see the third equation). Usually 0, 1 or 2 autoregressive terms are needed, but these are not reported here. The Q-statistics for the final models did not indicate any problem of autocorrelation. The model is estimated for each age-gender combination, leading to 10 separate equations. The results are shown in TABLE 32.

TABLE 32: Results for the age-gender models

AGE	GENDER	Param.	Coefficient	Std. Error	t-Statistic	Prob.
0-14	**Male**	α_0	95.9226	3.3651	28.5048	0.0000
		α_1	-0.0496	0.0017	-29.3155	0.0000
	Female	α_0	90.9531	8.7076	10.4452	0.0000
		α_1	-0.0473	0.0044	-10.7961	0.0000
15-24	**Male**	α_0	46.1213	6.8046	6.7779	0.0000
		α_1	-0.0235	0.0034	-6.8690	0.0000
	Female	α_0	26.9087	5.2746	5.1016	0.0000
		α_1	-0.0145	0.0027	-5.4720	0.0000
25-44	**Male**	α_0	18.5408	4.6687	3.9713	0.0005
		α_1	-0.0099	0.0023	-4.2275	0.0003
	Female	α_0	20.6615	3.7985	5.4394	0.0000
		α_1	-0.0117	0.0019	-6.1035	0.0000
45-64	**Male**	α_0	61.2717	3.4335	17.8453	0.0000
		α_1	-0.0316	0.0017	-18.2765	0.0000
	Female	α_0	58.7781	5.4830	10.7201	0.0000
		α_1	-0.0308	0.0028	-11.1715	0.0000
65+	**Male**	α_0	73.7008	7.5462	9.7666	0.0000
		α_1	-0.0376	0.0038	-9.9061	0.0000
	Female	α_0	82.2240	4.3730	18.8028	0.0000
		α_1	-0.0423	0.0022	-19.2321	0.0000

It is seen from the table that both the intercept (risk level for $t = 0$) and the parameter for the trend are significant in all models. For each gender category, the intercepts show a U-shaped curve over the age categories, with the highest values for the younger and older road users. The same shape is found by (Evans, 2004) for population fatality rates of drivers between 20 and 80 years old, although his analysis was not over time. This is further illustrated in FIGURE 40.

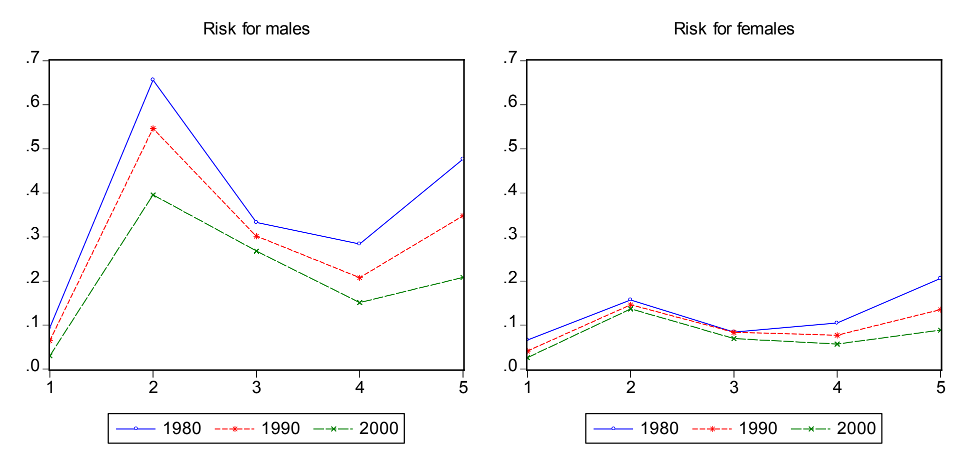

FIGURE 40: Risk comparison over time

The graphs show the estimated risk for male and female road users for each age category for the years 1980, 1990 and 2000, calculated from the 10 models. As the age of 20 years old belongs to the second category, a U-shaped relation, comparable to that found by Evans (2004), appears again. It is also seen that the risk is generally lower for females than for males, although the difference is quite small in the first age category. Further, the three curves in each graph indicate the reduction in risk over time. Clearly, the decrease is not equally large for every age-gender combination, which underlines the importance of targeted road safety programs and campaigns.

The yearly reduction in risk for the different age-gender combination can be read from the α_1 parameters (see TABLE 32). This parameter learns that, for example in the group of male road users between 15 and 24 years old, the average decrease in risk equals 1-exp(-0.0235) or 2.32%. The parameter estimates for all groups are shown in absolute values in FIGURE 41. A high value indicates a higher reduction in risk over time. Again, a U-shaped relation is found, which starts from the first age category, indicating a higher risk reduction per year for younger and older people. Clearly, the second age group has a high risk (see FIGURE 40), with a relatively low yearly reduction compared to the other groups. Also note the difference in reduction between males and females. The reduction is larger for males in the first two categories, but is smaller for the oldest road users.

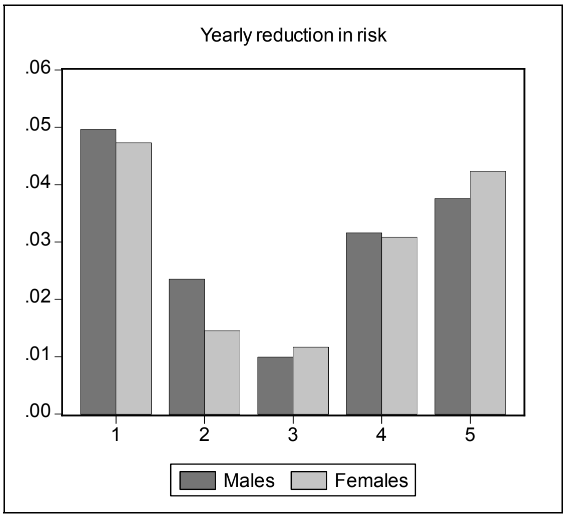

FIGURE 41: Yearly reduction in risk for males and females

Although the models presented here are quite simple, they offer the possibility of describing the differences in fatalities and risk between various age-gender categories, together with their trend over time. Of course, using population data as a measure of exposure will certainly not give the same results as when the number of kilometres driven is used. Typically, one can expect that older people drive less than younger ones, although their share in the total population is rising. However, (Evans, 2004) found U-shaped curves for driver fatalities per billion kilometre, which are comparable to the shapes in the population fatality rates. This confirms the fact that it is useful to include population data as an offset, to correct for the differences between the age groups over time.

7.3 Analysis of road risk per type of road user

One of the basic distinctions that can be made in disaggregated road safety research is according to the type of road user. In this section, an analysis is presented of the developments in time of the number of persons killed or seriously injured for three types of motorised road users, namely the cars, the trucks and the motorcycles. The car is by far the most popular means of transport in Belgium. During the last decennium, however, motorcycles gained popularity and there has been a steady increase in the number of motorcycles on the road. Also, the transportation sector has been growing, resulting in more freight transport on the road. FIGURE 42 shows the number of killed or seriously injured victims for these means of transport for the years 1973-2001, together with the number of vehicle kilometres, introduced in Chapter 4. Each observation contains the number of killed or seriously injured persons counted in the given mode of transport. For example, if a crash leads to three victims, say two in a car and one on a motorcycle, then these numbers are added to the counts for the respective groups of road users (namely two car victims and one motorcycle victim). This also means that victims of other transport modes (bicyclists, pedestrians, etc.) are not considered here.

While the number of victims decreased clearly for the cars and the motorcycles, this is not the case for the trucks. For this transport mode, no decreasing trend can be observed. Also, the decrease in motorcycle victims seems to stop in the early nineties, which is exactly the period where the number of vehicle kilometres for this group starts increasing at a much faster rate.

The data will be analysed in a multivariate state space model. The number of kilometres driven is considered as an explanatory variable, and is therefore assumed to be known without error. Moreover, a parameter is estimated for this variable, such that the proportionality assumption of exposure and the number of victims can be verified. Also, the data are analysed in logs, as this follows the rationale of a multiplicative model for exposure and risk on the original scale.

The estimation is done in the STAMP 6.21 software (Koopman et al., 2000). The multivariate model can be written as:

$$
\begin{aligned}
y_t &= \mu_t + \beta_t x_t + \varepsilon_t\,, & \varepsilon_t &\sim NID(0, \Sigma_\varepsilon) \\
\mu_t &= \mu_{t-1} + \nu_{t-1} + \eta_t\,, & \eta_t &\sim NID(0, \Sigma_\eta) \\
\nu_t &= \nu_{t-1} + \zeta_t\,, & \zeta_t &\sim NID(0, \Sigma_\zeta)
\end{aligned}
\tag{71}
$$

In this formulation, y_t is a vector containing the observed values for the number of victims for the three types of road users and x_t contains the three series of kilometres driven. Because each series is concerned with one road user type, the explanatory matrix is diagonal, such that the number of victims of each road user type is related to its own number of kilometres driven. The unobserved components vectors μ_t and ν_t contain the levels and the slopes, respectively. Finally, Σ_ε, Σ_η and Σ_ζ are the variance matrices for the observation equation and the two state equations, while ε_t, η_t and ζ_t are multivariate, normally and independently distributed (NID) disturbances.

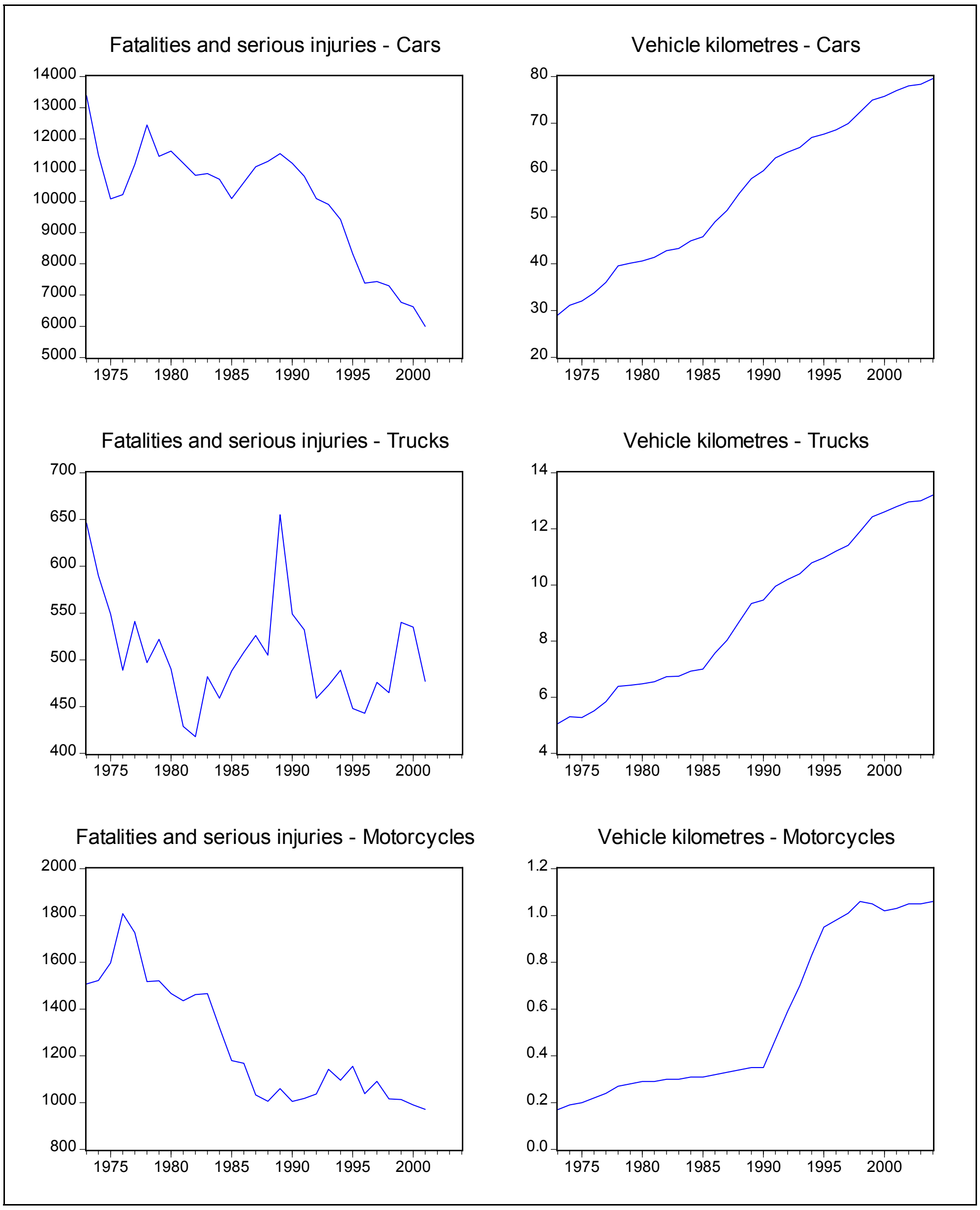

FIGURE 42: Road safety data per type of motorised road user

A first model includes a separate level and slope component for each transport mode. The AIC values for this model are -4.2259 for cars, -3.2052 for trucks and -4.0479 for motorcycles. After inspecting the covariance matrices of the level and slope components, it turns out that a common factor model might give a better fit. Setting the rank of the variance matrix for the slope component equal to 1 reduces the AIC values to -4.4405, -3.4692 and -4.2845, respectively. Further, the Q-statistics did not indicate any problem of autocorrelation. The common factor model was therefore retained.

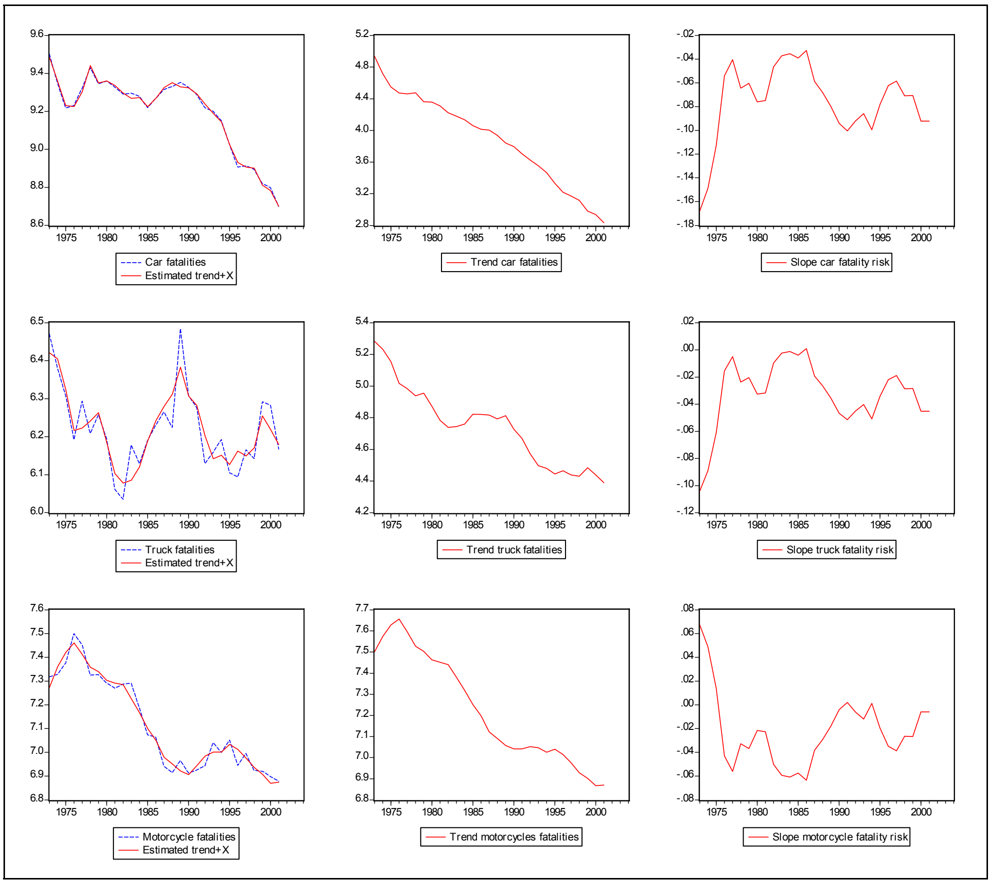

FIGURE 43: Fitted values, trends and slopes for transport modes

FIGURE 43 shows the fitted curves for the three series of victims (first column), together with an estimate of the trend (second column) and the slope curves (third column). As can be seen from the first column, the model seems to fit the data fairly well, although the fluctuations in the truck victims are quite irregular. The trend curves in the middle column in FIGURE 43 are corrected for the vehicle kilometres and are, as such, an indication of the risk for each type of road user. The risk is decreasing almost linearly for the cars, but shows some stagnation for the trucks and the motorcycles. At the end of the eighties, the risk curve is almost flat for the trucks, and the same occurs for the motorcycles in the early nineties. As for the slopes, it is seen that the rate of decrease in risk is getting smaller near the end for all types of road users. At the end of the sample period, there is a decrease in risk of 9% for cars and 4.5% for trucks, but only 0.59% for motorcycles. In fact, as can be seen from the estimations of the final state coefficients in TABLE 33, at the end of the series the slope is only significantly different from zero for the cars. This means that the reduction in risk for trucks and motorcycles is almost zero around the year 2000.

TABLE 33: Estimated coefficients

	Transport mode	Coefficient	R.m.s.e.	t-value	p-value
Level	**Cars**	2.8329	1.4841	1.9089	0.0670
	Trucks	4.3891	1.2888	3.4055	0.0021
	Motorcycles	6.8703	0.0209	329.39	0.0000
Slope	**Cars**	-0.0922	0.0356	-2.5895	0.0153
	Trucks	-0.0452	0.0322	-1.4024	0.1722
	Motorcycles	-0.0059	0.0330	-0.1788	0.8595
Veh-km	**Cars**	1.3508	0.3410	3.9618	0.0005
	Trucks	0.7025	0.5031	1.3965	0.1739
	Motorcycles	0.1290	0.0660	1.9538	0.0612

TABLE 33 also shows the estimated coefficients for the exposure measures in each equation. Note that the cars exposure only occurs in the cars equation, the trucks exposure in the trucks equation, and so on. For the cars and the motorcycles, a significant value (compared to zero) is found. For cars, the increase in the number of victims is more than proportional (1.34%), while it is less than proportional for trucks and motorcycles. A 1% increase in kilometres driven by motorcycles results in a 0.13% increase in the number of killed or seriously injured motorcyclists.

Given the risk functions obtained from the model, it is interesting to look at the relative risks of the three modes of transport. As the risk levels are modelled in logs, the difference between, for example, the car risk and the truck risk is equal to the log of the ratio of both risk curves:

$$\text{Log Relative Risk}_{\text{Cars}} = \log\left(Risk_{Cars}\right) - \log\left(Risk_{Trucks}\right) = \log\left(\frac{Risk_{Cars}}{Risk_{Trucks}}\right) \qquad (72)$$

This is shown in FIGURE 44 on the left graph. Since the risk components of cars and motorcycles are compared to the risk component of trucks, the latter is equal to zero for all time periods. Compared to the risk of trucks, the risk of motorcycles has been much higher since the early seventies, while the risk of cars is lower ever since. Also, compared to trucks, the risk of motorcycles seems not to be declining, as the risk of cars does. It is striking that, after a decrease in relative risk for motorcycles in the eighties, it is increasing again in the nineties, that is, together with the strong increase in kilometres driven by motorcycles. Although they travel the lowest number of vehicle kilometres, the motorcyclists have a risk that is between 2 and 3 times the risk of trucks.

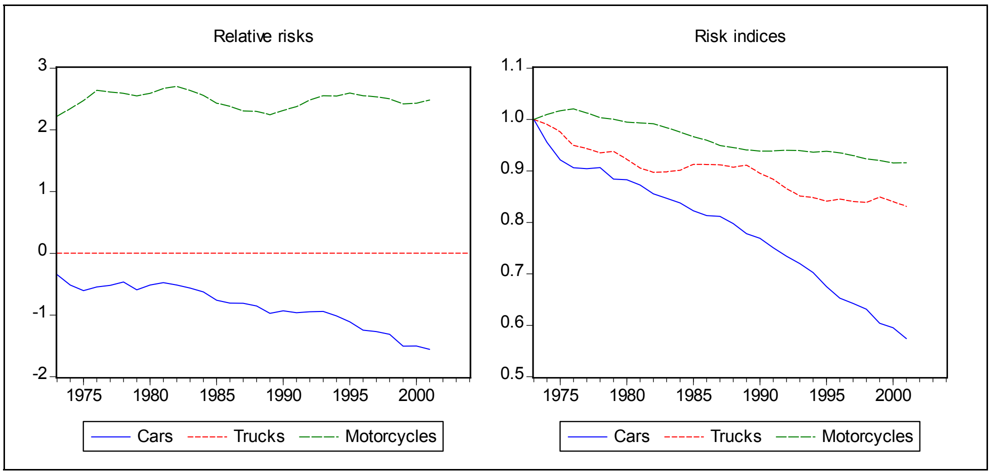

FIGURE 44: Relative risk and risk indices

An alternative view is given in the right graph of FIGURE 44. Here, a risk index is calculated by putting the risk component for the year 1973 equal to 1 for all transport modes. The curves then show how the trends change over time, relative to one another. It is seen that cars have indeed the strongest reduction in risk compared to 1973, while the decrease is less pronounced for trucks and motorcycles.

This type of analysis shows the importance of a disaggregated time series analysis, next to the aggregated models. The example focuses on three motorized transport modes, because for these an exposure measure was available. However, it should be clear that this analysis is easily extended to other types of transport, provided that an exposure measure can be included. One drawback of the data used in this study is the fact that it does not take into account the parties that are involved in an accident. The number of car victims contains the victims in car-car accidents, but also car-truck accidents, car-cyclists accidents and so on. The next section extends this topic by an analysis of two-sided crashes with motorised modes of transport.

7.4 Analysis of road risk in two-sided crashes

In the official Belgian road safety statistics, some matrices can be found with data on the number of accidents and victims in two-sided accidents. They represent the victims that are counted among the main parties involved in the crash. This section investigates the evolution over time of the number of fatalities and serious injuries in crashes where the main road users involved are cars, trucks or motorcycles. More specifically, time series of victims in collisions of a car with a car (C-C), a car with a motorcycle (C-M), a car with a truck (C-T) and a truck with a motorcycle (T-M) are studied. Again, the choice is determined by the availability of both crash and exposure data. For non-motorised modes of transport, the exposure data is not available over time. Also, the crashes of two motorcycles and two trucks are not studied. The objective is to gain insight in the risk functions of the analysed crash types.

The data that is available for this analysis is, unfortunately, not complete. That is, data were found for the years 1973-1989 and 1994-2001, with four missing values in each series of victims. The series of vehicle kilometres are complete. To deal with the missing value problem, state space models are used. Indeed, as will become clear below, the Kalman filter, that is used to provide optimal estimates of the current state of a dynamic system (Chatfield, 2004), is well suited to estimate missing values in a series. Details can be found in (Harvey, 1989) and in (Durbin & Koopman, 2001). More specifically, the models presented here are based on the multivariate latent risk models discussed in (Bijleveld, Commandeur, Gould et al., 2005). However, whereas these models assume a proportional relation between fatalities and risk, extra parameters are estimated here to test this assumption. For example, for the case of car-truck victims, one can write the model as:

$$
\begin{aligned}
&\log(V_{cars,t}) = \mu_{cars,t} + \varepsilon_{cars,t} \\
&\mu_{cars,t} = \mu_{cars,t-1} + \nu_{cars,t-1} + \eta_{cars,t} \\
&\nu_{cars,t} = \nu_{cars,t-1} + \zeta_{cars,t} \\
&\log(V_{trucks,t}) = \mu_{trucks,t} + \varepsilon_{trucks,t} \\
&\mu_{trucks,t} = \mu_{trucks,t-1} + \nu_{trucks,t-1} + \eta_{trucks,t} \\
&\nu_{trucks,t} = \nu_{trucks,t-1} + \zeta_{trucks,t} \\
&\log(F_t) = \alpha_{cars}\mu_{cars,t} + \alpha_{trucks}\mu_{trucks,t} + \mu_{fat,t} + \varepsilon_{fat,t} \\
&\mu_{fat,t} = \mu_{fat,t-1} + \nu_{fat,t-1} + \eta_{fat,t} \\
&\nu_{fat,t} = \nu_{fat,t-1} + \zeta_{fat,t}
\end{aligned}
\qquad (73)
$$

The first three equations are respectively, the observation equation, the level and the slope for the car exposure, V_{cars}. The same structure is then repeated for the truck exposure, V_{truck}. The last three equations are for the fatalities and serious injuries in crashes between cars and trucks, denoted F_t. The victims are modelled as a function of the unobserved levels of car and truck kilometres and the level of the risk, denoted $\mu_{fat,t}$. The parameters α_{cars} and α_{trucks} measure the effect of the exposure levels in the equation. Note that all quantities are in logs, which means that, again, the multiplicative structure between exposure and risk is respected. The last term in each equation is the mutually uncorrelated, normally distributed disturbance term, denoted $\varepsilon_{i,t}$ for the observation equations, $\eta_{i,t}$ for the level equations and $\zeta_{i,t}$ for the slope equations (i = cars, trucks), each with a mean zero and a specific variance. When studying crashes between cars and cars, the second measure of exposure is dropped, leaving six instead of nine equations to be estimated. Also, if the variance of the level or slope equation is estimated to be equal to zero, the corresponding component is assumed to be fixed. The series are shown in FIGURE 45. The number of vehicle kilometres for the three transport modes were already shown in FIGURE 9 (Chapter 4).

The models were estimated in Eviews 5.1 (QMS, 2004). The Q-statistics for the first 12 orders did not indicate the presence of autocorrelation. In the model for car-car crashes, the level of the crashes was fixed. This was also the case for the slope of the crashes in the other models (C-M, C-T, T-M). These slopes are interesting, as they show the yearly decrease in risk for each group of crashes. The results are shown in TABLE 34.

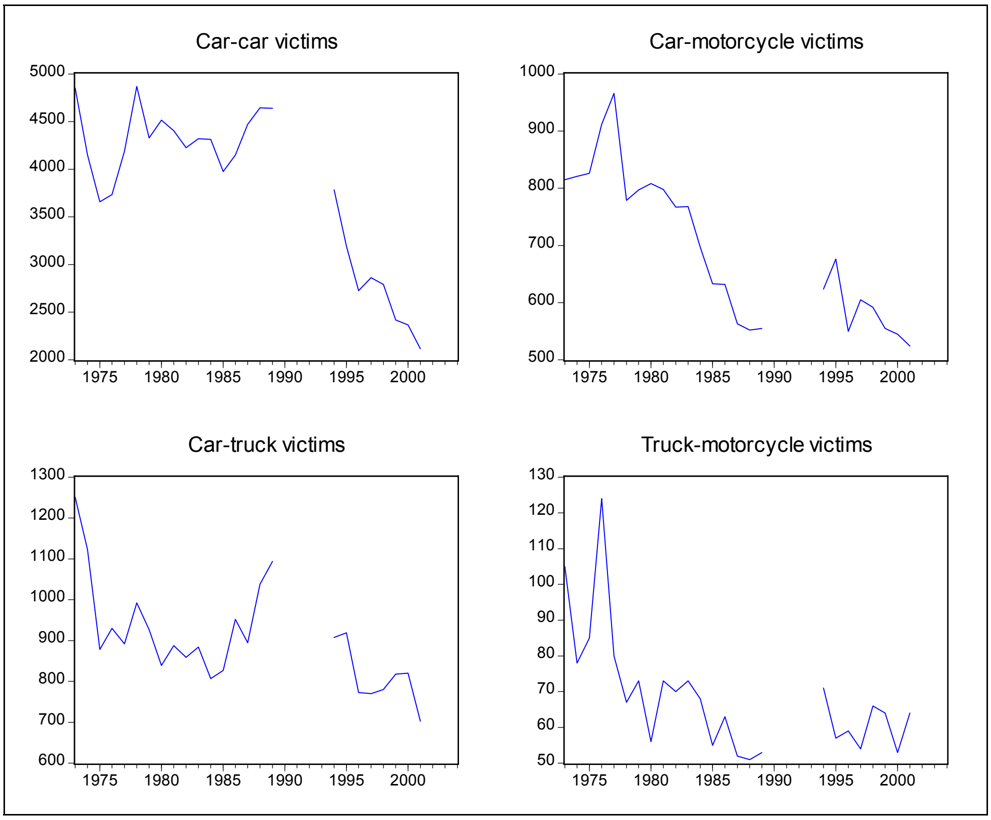

FIGURE 45: Fatalities and serious injuries in two-sided crashes

For the car-car crashes, a yearly decrease of about 12% is registered, while this is much lower for the other types. The risk of crashes between cars and motorcycles and cars and trucks is estimated to decrease at a rate of 3% and 6% respectively. A remarkable result is found for the truck-motorcycle victims. Here, the slope is fixed and not significantly different from zero. Therefore, it could be dropped from the model. This indicates that the rate of reduction in the number of these fatalities is to be neglected.

A second output from the model is the estimation of the exposure parameters. These are shown in TABLE 35. Again some interesting findings should be highlighted. First, apart from the exposure for motorcycles, all parameter estimates are highly significant. For example, the car exposure shows a more than proportional effect on the number of victims in model C-C.

TABLE 34: Estimated slopes for the risk in two-sided crashes

Model	Coefficient	R.m.s.e.	z-statistic	Prob.
C-C	-0.1204	0.0477	-2.5254	0.0116
C-M	-0.0332	0.0049	-6.8111	0.0000
C-T	-0.0622	0.0100	-6.2520	0.0000
T-M	0.0063	0.0103	0.6151	0.5385

Second, when the two parties in the crash are different, one exposure parameter is positive, while the other is negative. This is, at first sight, a strange result. However, the measures in the model are on the log scale, so that, on the original (multiplicative) scale, the exposure measures are raised to a power equal to their respective parameter estimates. If one of the two parameters is negative, it indicates a difference in logs, or the log of a ratio of the exposure levels. The number of victims is therefore determined by the ratio of the (trend of the) two exposure measures, rather than by their separate levels. For example, in the case of the C-T model, the number of victims would be determined by the ratio of truck kilometres divided by car kilometres (each raised to a specific power). If the ratio is higher (more truck kilometres for a given number of car kilometres), then the number of victims will rise. If the number of car kilometres is higher, then the ratio is lower (less truck kilometres for a given number of car kilometres) and the number of victims will fall. The way in which the ratio behaves is determined by the magnitude of the parameter estimates. The same can be said for the C-M and T-M models.

Given the importance of the parameters, it might be instructive to investigate them further. In particular, one can test the proportionality assumption of the exposure measures relative to the number of victims. Also, for the models with two exposure measures, the hypothesis can be tested whether the two parameters are equal with opposite sign, meaning that the relation between the number of victims and the risk is determined by the ratio of the two exposure measures raised to only one power.

TABLE 35: Estimated coefficients for the exposure parameters

Model	Parameter	Coefficient	Std. Error	z-statistic	Prob.
C-C	α_{cars}	2.2631	0.2150	10.5264	0.0000
C-M	α_{cars}	-0.3451	0.0911	-3.7869	0.0002
	$\alpha_{motorcycles}$	0.4282	0.3230	1.3257	0.1849
C-T	α_{cars}	-2.3658	0.4456	-5.3097	0.0000
	α_{trucks}	3.8120	0.5061	7.5324	0.0000
T-M	α_{trucks}	-1.3859	0.2343	-5.9140	0.0000
	$\alpha_{motorcycles}$	0.3601	0.2539	1.4184	0.1561

To test these hypotheses, the Wald test for coefficient restrictions is used (QMS, 2004). TABLE 36 shows the results for the four models.

TABLE 36: Wald statistics for exposure measures

Model	Test	Chi-square	Prob.
C-C	$\alpha_{cars} = 1$	34.5164	0.0000
C-M	$\alpha_{cars} = 1$	217.8743	0.0000
	$\alpha_{motorcycles} = 1$	3.1333	0.0767
	$\alpha_{cars} = -\alpha_{motorcycles}$	0.0610	0.8049
C-T	$\alpha_{cars} = 1$	57.0633	0.0000
	$\alpha_{trucks} = 1$	30.8735	0.0000
	$\alpha_{cars} = -\alpha_{trucks}$	4.0204	0.0450
T-M	$\alpha_{trucks} = 1$	103.6590	0.0000
	$\alpha_{motorcycles} = 1$	6.3519	0.0117
	$\alpha_{trucks} = -\alpha_{motorcycles}$	7.7086	0.0055

It is clear that none of the parameter estimates indicate proportionality. Also, for the models in which trucks are involved (C-T and C-M), the hypothesis of only one parameter for the ratio is rejected. This is not true for the C-M model. Here, the hypothesis that the car and motorcycle exposure have the same coefficient in absolute value is not rejected.

Each model further estimates a risk component for the number of victims. In FIGURE 46, the risk indices, showing the risk for each time period, scaled to the first year of observation are plotted.

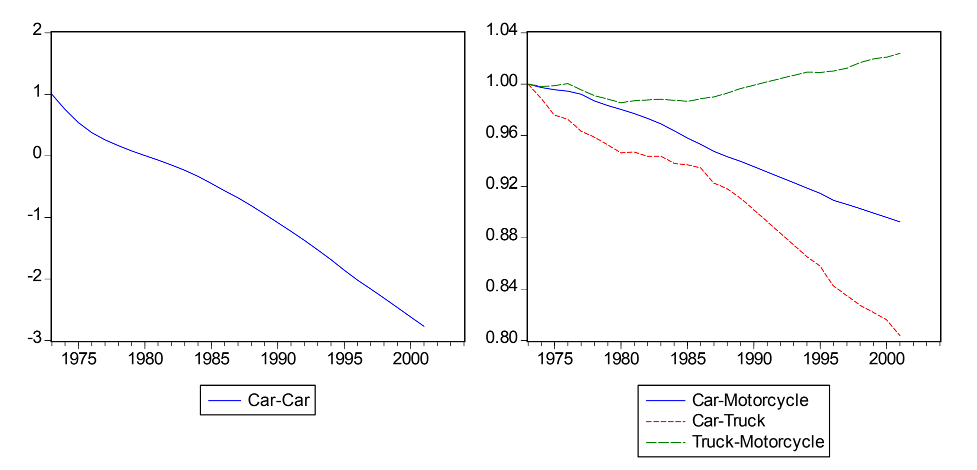

FIGURE 46: Risk indices for two-sided accidents

The left graph separately shows the index for car-car accidents, as this value is decreasing much faster than for the other two-sided accidents. The latter are shown in the right graph. Here, the decrease is less strong. The number of victims in car-trucks accidents shows the fastest decline, while the truck-motorcycle risk index is still close to one, even after a period of almost 30 years. From the late eighties onwards, the risk is increasing. Clearly, motorcycles should be counted as vulnerable road users. In fact, the risk curves nicely confirm the trends in Belgian road safety policy, where attention is given to these modes of transport in terms of road safety. Campaigns are organised to make the motorcycle visible in traffic, blind spot mirrors are introduced to reduce the number of fatalities in crashes between trucks and vulnerable road users, and proposals for lower speed limits for trucks have been formulated. These models show that the extra attention is indeed necessary.

The last interesting outputs from the models are the curves fitted to the observed values. These are shown in FIGURE 47. The full lines are the fitted values, together with the (dashed) 95% confidence interval. The irregular dashed lines show the observed series. Note how the Kalman filter fits the missing values of the series in a way that is, at least with the naked eye, quite reasonable. Also, the missing values are surrounded by confidence bounds that are slightly larger, although this is only visible for the car-truck victims. This is also the most fluctuating curve, and since the smoothed estimates take into account all observed values, a higher uncertainty is found. For the motorcycle graphs (right column), the missing values are in the period where the number of kilometres driven by motorcycles shows a steady increase. This evolution is nicely picked up in the number of victims, which is also increasing from 1990 onwards.

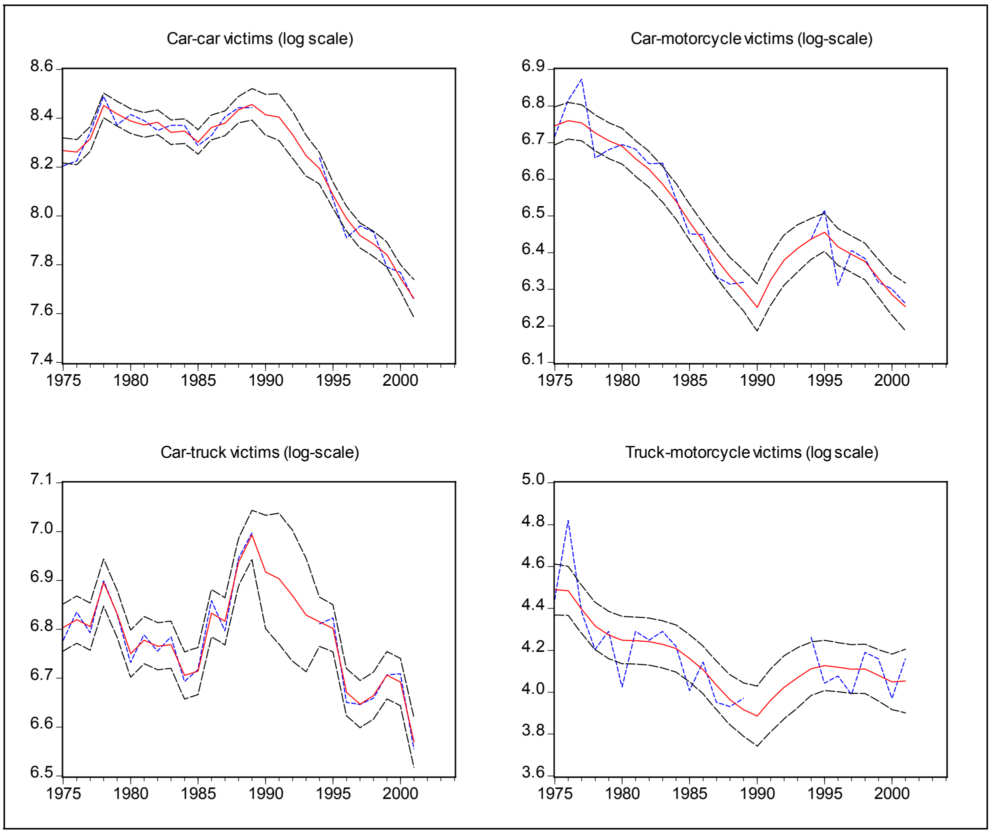

FIGURE 47: Fitted values for victims in two-sided accidents

It might be argued that some extreme values can be accounted for in the analysis by including intervention variables. However, there is no real explanation for these values and, given the low level of aggregation, they are quite probably caused by mere statistical fluctuations. Including interventions would in that case over-fit the model and unnecessarily reduce the number of degrees of freedom. This is not desirable, given the limited number of observations used in the study.

This section clearly illustrates the merits of a disaggregated approach towards different modes of transport in two-sided accidents. The higher risk of trucks and motorcycles was found in the data and the results were quantified by means of the very flexible state space models. However, the large fluctuations in the disaggregated series make it less easy to model the trends in the data, but nevertheless interesting insights are obtained. An extension would be to model the curves all together in a multivariate framework. This approach, however, is computationally very demanding, but would allow for a fine-tuning in the way in which the variances of the equations are related to one another. Also, other modes of transport could be added, provided, of course, that the appropriate data is available.

7.5 Analysis of road risk per type of road

Apart from the distinction that can be made between several types of road users, as demonstrated in the previous sections, one can also disaggregate road safety figures according to the type of the underlying road network. In the subsequent analysis, three types of roads are distinguished: highways, provincial roads (national and regional) and local roads. For these types of roads, official statistics are available on the number of accidents and victims as well as on the number of vehicle kilometres. Because of the differences between the roads (physical characteristics, number of lanes, speed limits, etc.), it is expected that the risk curves will behave differently. This aspect will be investigated in this section by means of a multivariate state space analysis, comparable to the one that has been used in section 7.3.

Apart from the exposure measure, however, this section also tests the significance of two major road safety interventions that are related to different types of road. The first variable, LAW0192, tests the introduction of the law on speed limits in January 1992. The law is concerned with the speed limits of 50 km/h in urban areas and 90 km/h on provincial roads. The second intervention, called LAW0198, measures the effect of the installation of speed cameras starting from January 1998. These two laws are expected to influence road safety on urban and provincial roads. Note, however, that speed cameras were installed shortly after the introduction of new right-of-way rules for roundabouts (October 1997). Probably this variable will measure a compound effect, which is inevitable with dummy intervention variables for yearly data.

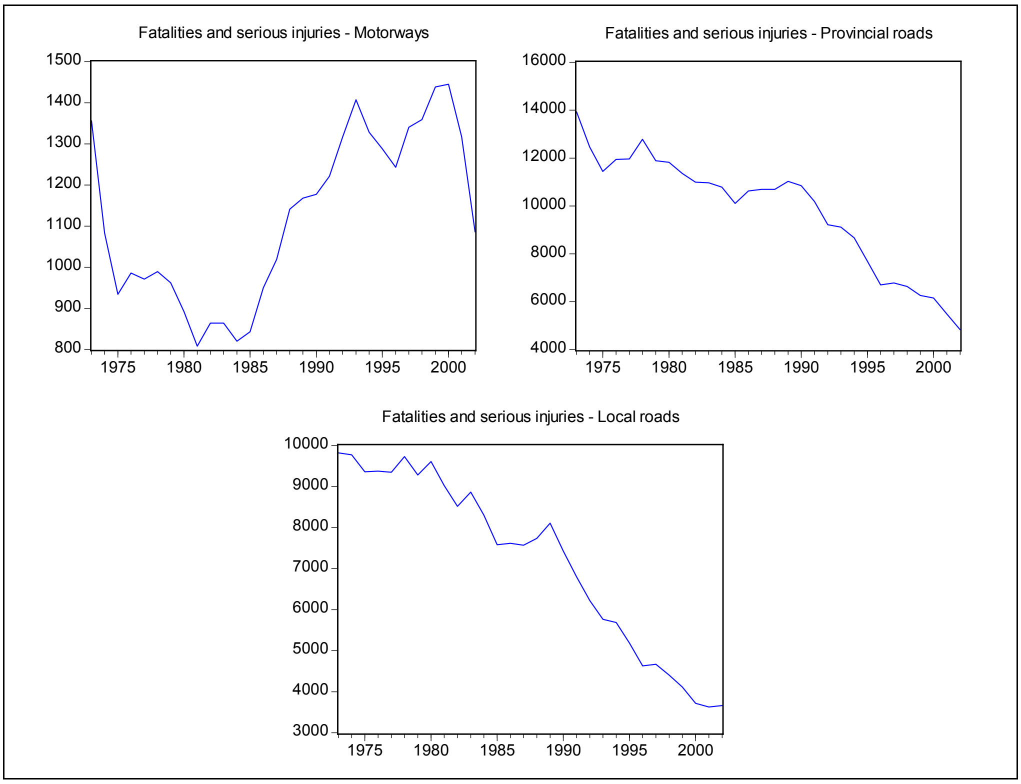

FIGURE 48: Fatalities and serious injuries per type of road

Data are used for the years 1973-2002. The statistics for 2003 and 2004 are also available, but they show a very unlikely pattern, probably caused by changes in the registration procedure and by shifting some roads from one category to another. FIGURE 48 shows the data on the number of persons killed or seriously injured for each type of road. It is immediately clear that the number of victims for motorways follows a different pattern from that observed for the provincial and local roads. From the eighties onwards, the trend on motorways is upwards, and turns only in the late nineties. However, the number of fatalities and serious injuries on motorways is, in absolute values, much lower than for the other two types of road.

For the model formulation, reference is made to Equation 71. The model tested here is identical, except for the explanatory variables. Here, the number of vehicle kilometres per type of road is used, together with the two tested laws. Note again that, as in the study per type of road user, the number of vehicle kilometres is only included in the equation for the corresponding type of road.

That is, the number of kilometres driven on motorways is included in the equation for motorways, and so on.

First, a full model in STAMP 6.21 (Koopman et al., 2000) was estimated, with a separate level and slope for each type of road. After inspection of the correlations between the level and slope disturbances respectively, it was decided to reduce the rank of the variance matrix for the level to 1 and for the slope to 2. The Q-statistics were satisfactory for the motorways and provincial roads, but indicated some serial correlation for the local roads. Inspecting the residuals shows an outlying value for the year 1989. Therefore, an intervention variable (Y1989) was added to the observation equation of the local road fatalities, which reduced significantly the level of autocorrelation. The values of the final state vector are shown in TABLE 37. The slope estimates represent the growth rates of the risk at the end of the estimation period. The model shows a reduction in the risk of 18% per year for the motorways, 13.5% for provincial roads and about 6% for local roads. For the motorways, the high value is probably caused by the enormous drop in victims in the last two years of the analysis period.

TABLE 37: Estimated coefficients for the final state parameters

	Road type	Coefficient	R.m.s.e.	t-value	p-value
Level	**Motorways**	6.5638	0.7692	8.5337	0.0000
	Provincial roads	7.1863	0.8267	8.6923	0.0000
	Local roads	6.9534	0.5348	13.002	0.0000
Slope	**Motorways**	-0.1808	0.0769	-2.3532	0.0259
	Provincial roads	-0.1349	0.0369	-3.6502	0.0011
	Local roads	-0.0607	0.0117	-5.2043	0.0000

The parameter estimates for the exposure measure and the intervention variables are shown in TABLE 38. For the exposure measure, the number of vehicle kilometres per type of road, it can be seen from the table that all parameter estimates are positive. The coefficient is not significant for motorways, while it is slightly significant for provincial roads and highly significant for local roads. A 1% increase in kilometres driven on motorways will result in a 0.13% increase in fatalities and serious injuries. On provincial roads, the increase is about 0.36% and for local roads it is 0.43%. This indicates that the effect of exposure is larger for more local roads compared to motorways.

TABLE 38: Estimated coefficients for the explanatory parameters

	Road type	Coefficient	R.m.s.e.	t-value	p-value
Veh-km	**Motorways**	0.1269	0.2227	0.5700	0.5733
	Provincial roads	0.3654	0.2240	1.6315	0.1140
	Local roads	0.4268	0.1752	2.4362	0.0215
LAW0192	**Motorways**	0.0323	0.0614	0.5258	0.6032
	Provincial roads	-0.0662	0.0419	-1.5814	0.1250
	Local roads	-0.0851	0.0306	-2.7789	0.0096
LAW0198	**Motorways**	-0.0448	0.0618	-0.7254	0.4742
	Provincial roads	0.0142	0.0422	0.3363	0.7392
	Local roads	-0.0115	0.0308	-0.3746	0.7108
Y1989	**Local roads**	0.1102	0.0308	3.5811	0.0013

An instructive result is obtained for the interventions. As expected, the law on speed limits is significant for local roads and, to a lesser degree, for provincial roads, but not on motorways. Indeed, this law only regulates the speed limits for these categories of roads. Also, the effect is larger for local roads compared to provincial roads. The introduction of the new speed limits resulted, on average, in a decrease of victims of 6.6% on provincial roads and 8.5% on local roads. On the other hand, the installation of speed cameras and the introduction of the new right-of-way rules for roundabouts are not significant for any of the road types. Probably, these measures, compared to the speed limits, have a very concentrated influence, with high variability from one location to another. It is then very difficult to assess the effect of a measure on data that are aggregated in time. Perhaps a before-and-after study on certain locations might shed a different light on these results.

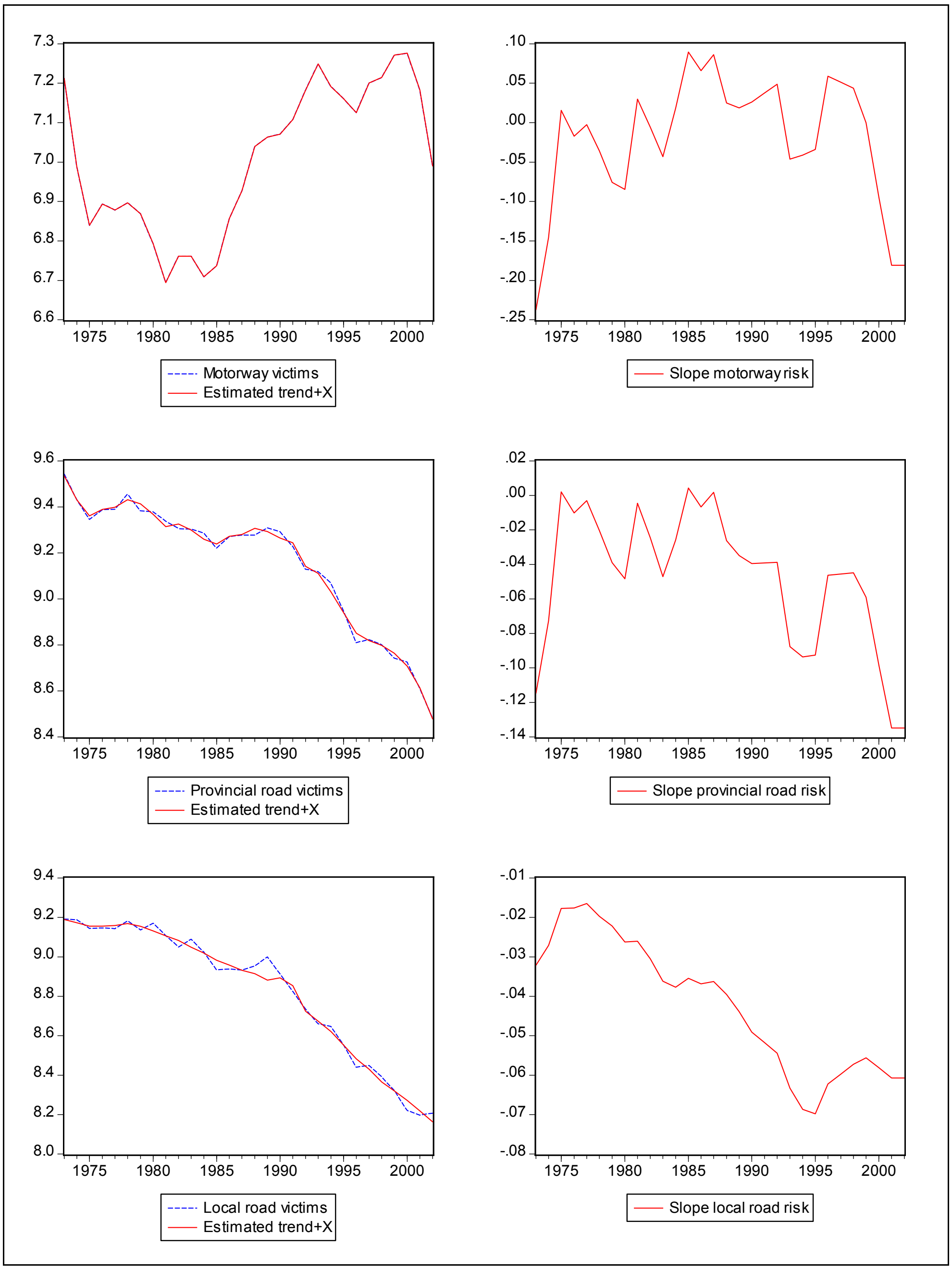

FIGURE 49: Fitted values plus trends and slopes for road types

The graphs for the fitted values plus trends and slopes are shown in FIGURE 49. The slopes show a very irregular pattern. For motorways and provincial roads, the slope is around zero for many periods, indicating stagnation in the trend of the number of victims. However, the final absolute values of these slopes are large. For the local roads, the slope values are smaller in absolute values, but they are always negative, indicating a systematic reduction in the number of victims. Around the year 1991, the changes in the trend and the slope, caused by the speed law, are visible, but the slope curves indicate a smaller decrease already some years later.

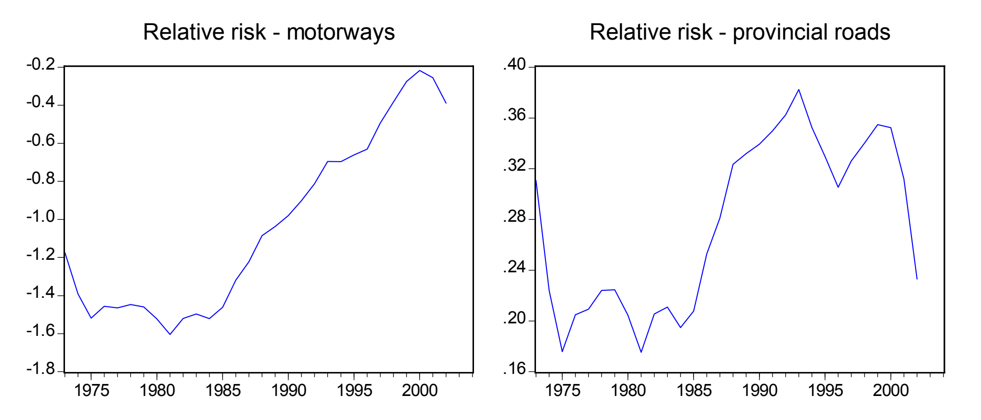

FIGURE 50: Relative risks for motorways and provincial roads

In FIGURE 50, the relative risks for motorways and provincial roads compared to local roads are shown. The relative risk is the evolution of the risk of motorways and provincial roads compared to the risk of local roads. A negative value indicates a lower risk than the reference type of road (in this graph the local roads). For example, at its lowest point, the relative risk for motorways is -1.60, indicating a risk that is 1-exp(-1.60)=0.7981 or almost 80% lower than the risk on local roads. The risk of motorways is always lower than the reference line, but the difference gets smaller over time. In 2000, the difference with local roads is reduced to 20%. The risk of provincial roads is higher for the whole analysis period, and the difference between this type of road and the local roads fluctuates in almost the same way as the motorways until the early nineties.

The risk index (FIGURE 51) shows the evolution of the risk for each type of road, rescaled to one in the first year of analysis. It indicates again the comparable behaviour of the risk curves for provincial and local roads, and the deviating pattern for motorways. Although the relative risk points out that motorways may be considered safer than the other types of road, the risk index does not show a monotonically decreasing risk pattern.

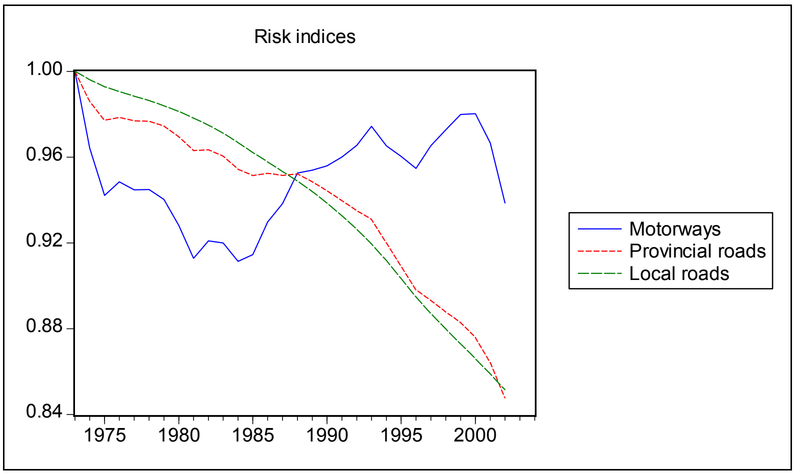

FIGURE 51: Risk index for motorways, provincial roads and local roads

The three types of road have their own characteristics, but it seems that the decrease in victims and risks during the last decennia can be attributed more to the local and provincial roads rather than to the motorways. The latter road type does not show the decreasing pattern that is also present in the global number of victims.

7.6 Conclusion

This chapter presented some possible models that can be developed on Belgian disaggregated road safety data. The number of victims was disaggregated according to road user properties (age and gender), motorised road user type and road type. The results show that the various groups in the studies often have very specific characteristics, which are worth investigating separately. Various road safety measures are only oriented towards one or more subsections of the transport system, and testing these measures on a highly aggregated level might obscure the nature of the effects.

The field of disaggregated time series models in road safety research seems to be scarcely out of the egg. However, the merits of this approach are clear. A disaggregated analysis is complementary to the aggregated models in the sense that the trends for specific parts of the transport system can be analysed separately, and the specific countermeasures, introduced to enhance road safety for a specific group of road users, can be benchmarked. The main problem, however, is the availability of relevant data. The lower the level of aggregation, the more difficult it is to find valid exposure data. For Belgium, exposure data per road (user) type are only available on a yearly basis for motorised means of transport. For age and gender categories, population data seemed to be a valid alternative to correct for exposure, but it has disadvantages. For example, changes in population will, for some age and gender categories (especially for older and younger persons), diverge from the number of kilometres driven by these groups of road users. Moreover, it is currently completely impossible to derive data on combinations of disaggregated characteristics, like for example the kilometres driven by female bicyclists on local roads. It goes without saying that the data determines the possibilities and limitations of the models.

This chapter can therefore be seen as a strong plea in favour of the collection of disaggregated time series data in road safety. The analyses in this chapter demonstrate the possibilities of this approach, but they show, at the same time and equally obvious, their limitations. The models are, both deliberately and out of (sheer) necessity, kept very simple. Although there are, consequently, restrictions to the applicability of the results, they make optimal use of the available data. Moreover, it is clear that disaggregated series are related to one another in many ways, and it is probably more important here than for disaggregated models to analyse them together. For example, if more people use their bike, less people will drive and vice versa. These changes in the transport system will probably not affect the aggregated time series, but they have an influence on the disaggregated results. Consequently, more advanced statistical techniques are needed to analyse the available data.

There are two other remarks that should be made in the context of disaggregated models. First, although the models presented here are disaggregated in terms of road user characteristics (age and gender), road user type or road type, the observations are still aggregated in time. That is, all models use yearly aggregations of disaggregated road safety figures. For example, one considers the number of fatalities on local roads per year, taking together all the fatalities on all local roads at that time. The high level of aggregation in time may imply that certain effects are difficult to assess. Therefore, these models are complementary to other, mostly cross-sectional models, and the results should be taken for what they are worth. Second, the studies in this chapter contain only disaggregated data series for motorised transport. The reason is that these series are all based on the official traffic counts from the Ministry of Mobility and Transport. For other means of transport (bicyclists, pedestrians, etc.), no such series are available and, consequently, no time series analyses can be carried out. However, given the attempts of the government to promote sustainable means of transport, it would be instructive to analyse road risk and measures of exposure for these groups of road users as well. The next chapter presents some cross-sectional descriptive studies for other modes of transport.

Chapter 8 Disaggregated analysis in Flanders

8.1 Introduction

In the previous chapters, the relation between exposure and road safety was extensively analysed with a number of models. Aggregated models were developed on yearly data to obtain predictions for road safety and to assess the long run objectives that were set by the government. Also, disaggregated models were applied to each type of road and to various types of road users. Using monthly data, descriptive and explanatory models were built to make short term predictions that take into account the seasonal fluctuation in the data. It is clear that the combination of exposure data and road safety outcomes can lead to interesting conclusions.

However, the main difficulty when studying the relationship between mobility and traffic safety is the availability of appropriate data. If more disaggregated data on road safety and exposure were available, it would be possible to analyse other types of road users (like pedestrians or bicyclists), or various kinds of accidents. The data for this kind of models is, unfortunately, not available in Belgium to analyse in a time series setting. In Flanders, a travel survey is the only source of cross-sectional data on disaggregated exposure. Therefore, in this chapter, a relatively simple approach is introduced to enhance the insight in the relation between road safety and mobility on a cross-sectional basis. More specifically, some exploratory analyses are presented in which mobility and traffic safety data are combined, using Flemish travel survey data (Zwerts & Nuyts, 2004) and accident data for the year 2000 to study the factors of mobility that determine traffic safety. As was done in the previous chapters, the analysis starts from the basic relation between the number of accidents, the number of victims and a measure of exposure.

The main issue, compared to the previous chapters, is that results can be derived for non-motorized means of transport in a cross-sectional setting. In particular, the relationship between exposure and risk can be analysed for vulnerable road users. It is well-known that different user groups (in terms of transport mode, age and gender) can have different patterns of traffic and a corresponding level of risk and exposure. In a time series context, this was already shown in Chapter 7. When sustainable transport modes are promoted, policy makers have to make sure that these are safe and useful for the target group. It is important to find out whether a transport means is unsafe for a given user group because of a high level of exposure, or rather because of a higher level of risk. A decomposition using travel survey data will be presented. Note, however, that other methods for deriving a risk indicator are available, as demonstrated for example in (de Leur & Sayed, 2003).

For the studies in this chapter, various sources of data are needed. First, the Flemish accident data for the year 2000 will be used, containing all information from the crash report, including gender and age of the victims. Second, data on the mobility of road users are needed. As detailed mobility data are desirable for this study, the Flemish travel survey 2000-2001 (Zwerts & Nuyts, 2004), introduced in Chapter 4, will be used. Given the accident statistics and exposure measures, it seems that all necessary ingredients for a traffic safety analysis are at hand. However, there are still some remarks. First, mobility and accident information are two different data sources. Usually there are no mobility data available related to the exact place and time of an accident. Therefore it is not always possible to find a clear link between both. Second, travel surveys consist of samples of road users, while accident counts are observed statistics. Both sources have their own problems and limitations. For surveys, it is sometimes difficult to guarantee that all groups of road users are present in the same proportions as in the population. Accident counts are subject to problems of under-registration and wrong or incomplete accident information. Third, travel surveys are conducted with the objective of gaining insight in the travel habits of citizens, and rarely if ever to increase (the knowledge on) traffic safety. Some questions are irrelevant for traffic safety research, and other important questions are not asked.

8.2 Travel surveys and road crashes

In this section, the decomposition of fatalities in exposure and risk is applied on cross-sectional data for the year 2000. For example, if one is interested in the distribution of fatalities over age and gender, it is possible to calculate a measure of exposure and an indicator of risk for each age-gender category. The product of exposure and risk will again provide an estimate of the fatalities. Note that a similar analysis was presented in Chapter 7, section 7.2. This approach can be useful to extend the insight in the accident generating process for a given age category. A similar study was presented in (de Leur & Sayed, 2003) and in (Toomath & White, 1982), where information on driving patterns is used in association with reported injury crashes to determine exposure-adjusted accident rates by age and gender group.

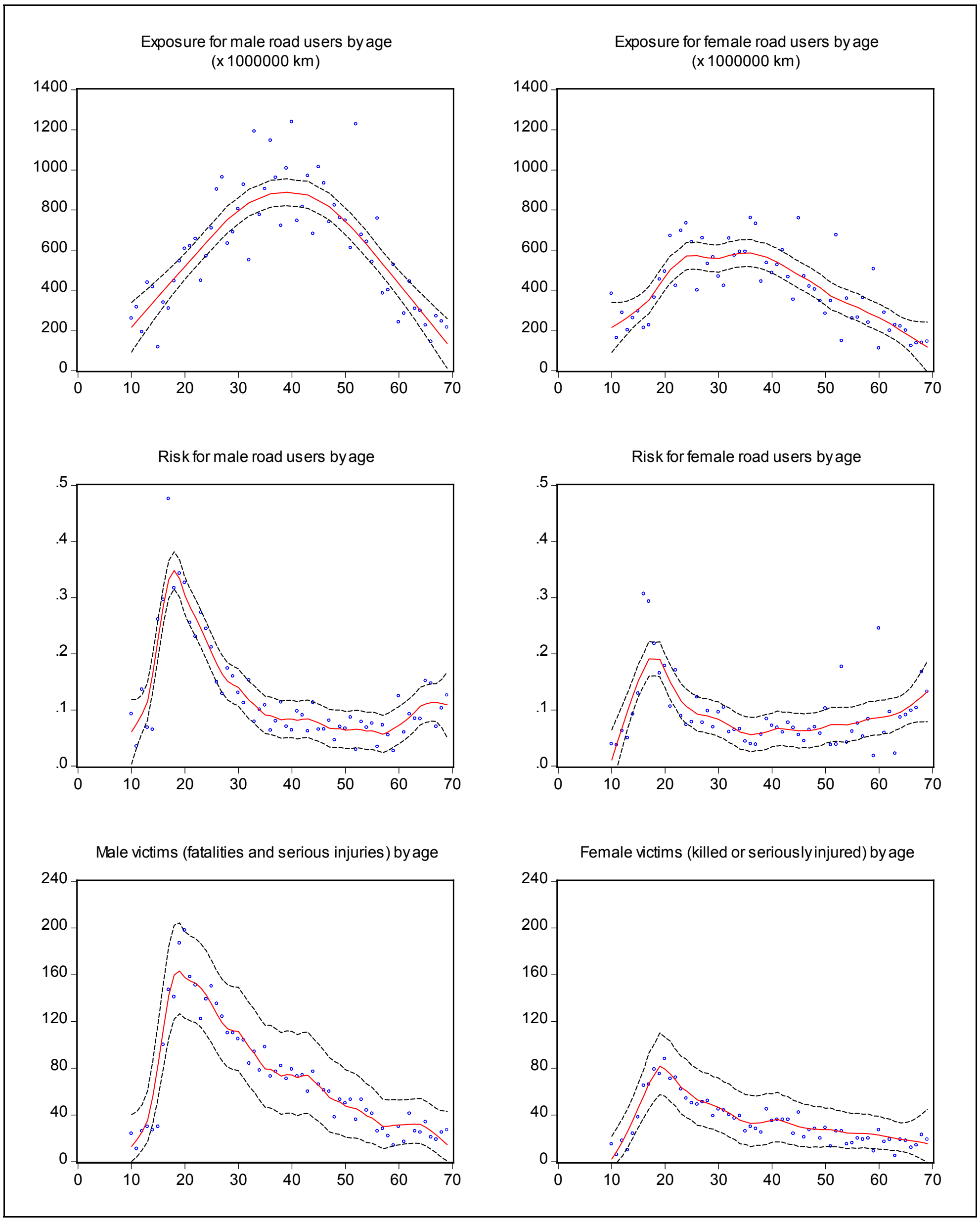

FIGURE 52: Decomposition of fatalities in risk and exposure

Using the data of travel surveys for Flanders (Zwerts & Nuyts, 2004), a measure of exposure can be created for different age and gender groups for the year 2000. This measure is based on the reported average amount of travel, which has been extrapolated using official statistics on the segmentation of the age and gender groups in the population. Together with the road accident records for the same period, a risk indicator is constructed as the ratio of the number of fatalities and serious injuries and the level of exposure. By doing so, a decomposition of fatalities in exposure and risk is obtained. For example, the product of risk and exposure for male persons of 25 years old should result in a corresponding number of victims. In comparison with a decomposition over time, the curve-fitting exercise is more complex. It is expected that exposure will be higher for the working category of people, and lower for younger and older persons. Also the risk will not be continuously decreasing, but will be higher for younger and older people and for vulnerable road users in general. Therefore, the nonparametric LOESS method for estimating regression surfaces is used (Cleveland et al., 1988; Cleveland & Grosse, 1991). This method can be used for situations in which no suitable parametric form of the regression surface can be found.

Assume that a dependent variable y and an independent variable x are related by $y = g(x) + \varepsilon$, where g is the regression function and ε is the random error, then for a given value x_i of the independent variable, g can be locally approximated by fitting a regression surface to the data points within a chosen neighbourhood of the point x_i. The method of weighted least squares is used to fit linear or quadratic functions at the centres of the neighbourhoods. Each neighbourhood contains a fraction of the data, which is determined by the smoothing parameter. Data points in a given local neighbourhood are weighted by a smooth decreasing function of their distance from the centre of the neighbourhood (Cohen, 1999). The selection of the smoothing parameter determines the fit of the model. If this parameter is too low, the data is over-fitted. If it is too high, an overly smooth fit is obtained, loosing essential features of the data.

TABLE 39: Model characteristics for exposure and risk

Group	**Equation**	**AIC**	**Smooth**
Male	Exposure	11.1995	0.5250
	Risk	-5.1154	0.1417
Female	Exposure	10.6536	0.2750
	Risk	-4.9823	0.2250

In the study, the smoothing parameter is chosen to minimize AICC, a bias corrected AIC criterion (Hurvich et al., 1998), which is a trade-off between smoothness of the fit and complexity of the model. The LOESS model fitting is done in SAS (SAS Institute Inc., 2004b). The dependent variables are exposure and risk, for male and female road users respectively. The independent variable is the age of the road users, starting at 10 and ending at 69 (due to data restrictions in the travel survey). TABLE 39 contains the parameter estimates for exposure and risk, for both male and female road users.

FIGURE 52 shows a decomposition of the number of fatalities in risk and exposure, for each age and gender combination. The curves give an acceptable mental fit to the data, due to the optimal choice of the smoothing parameter. For exposure, it is reasonable to assume that both younger and older people are less exposed to risk, because of their lower frequency of travel. The estimate for the victims is again the product of the estimates for exposure and risk. Some interesting insights can be gained from this decomposition.

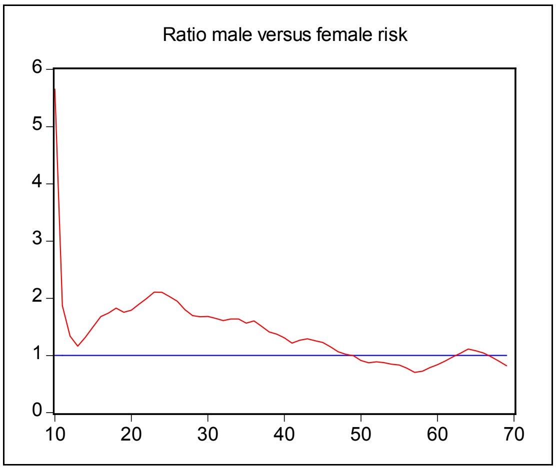

FIGURE 53: Ratio of male versus female risk

First, persons between 15 and 25 years old are high risk road users. The risk stays at more or less the same level from age 40 up to age 60 and then starts going up for the elderly. Even if their exposure is lower, they have a higher probability of being killed or seriously injured. These risk pictures clearly show how vulnerable the younger and older road users are. It is easily seen that these results are more or less in line with what was found in Chapter 7, where a similar study was done in a time series context. However, whereas the previous model used population data as a proxy for exposure, an estimated number of kilometres travelled for the year 2000 is used here. Second, the highest exposure is found for people between 30 and 50 years old. This is the class of the working people, who are at the same time socially active. Their activities result in a higher number of kilometres driven. For this group of persons, the number of victims is more determined by exposure than by risk.

Third, the risk and exposure for female road users is lower than for male road users of the same age. This indicates that women are less frequent road users, and if they are, their probability of being killed or seriously injured in an accident is lower. However, this difference is getting smaller with age. In FIGURE 53, the ratio of male to female risk is shown. On the reference line, the ratio equals one, indicating an equal risk for males and females. It seems that male road users have a higher risk up to an age between 40 and 50, and afterwards risks seem to be quite similar.

8.3 Modal split and road crashes

Another advantage of travel surveys is that detailed information on travelling choices is available, like for example the modal split, for each age category. It is expected that younger people will travel more as a car passenger, while at a later age they will go by bike or drive a car themselves. This information can also be linked with the number of victims of the various modes of transport. In FIGURE 54, the risk, the exposure and the number of victims for car drivers, car passengers, bicyclists and pedestrians are shown, categorised by age. Because of the higher level of detail, and in order to reduce sampling errors, categories of ages instead of the ages themselves are used. The victims are again the number of persons killed or seriously injured, while exposure is the number of kilometres travelled using a specific mode of transport.

Some interesting results are found. First, car drivers have the highest number of victims. Exposure is largest for drivers between 25 and 54 years old. The risk is higher for young drivers, but decreases with age. This is perhaps a kind of learning effect, showing that older drivers are more experienced than younger ones. On the other hand, elderly people show a higher risk. Although studies have shown that specific driver performance skills decline with increasing age (Warshawsky-Livne & Shinar, 2002), indicating that they become less proficient in driving a car, it is reasonable to assume that the higher risk of the elderly drivers stems from their reduced ability to survive injury crashes. As explained in (Evans, 2004), the risk of being killed in a crash is higher for older than for younger drivers. Older drivers involved in a crash are more likely than younger drivers to suffer serious injury or death. The number of victims is higher for younger people, mainly because of the higher risk. For the working category, the number of victims is more determined by the higher level of exposure.

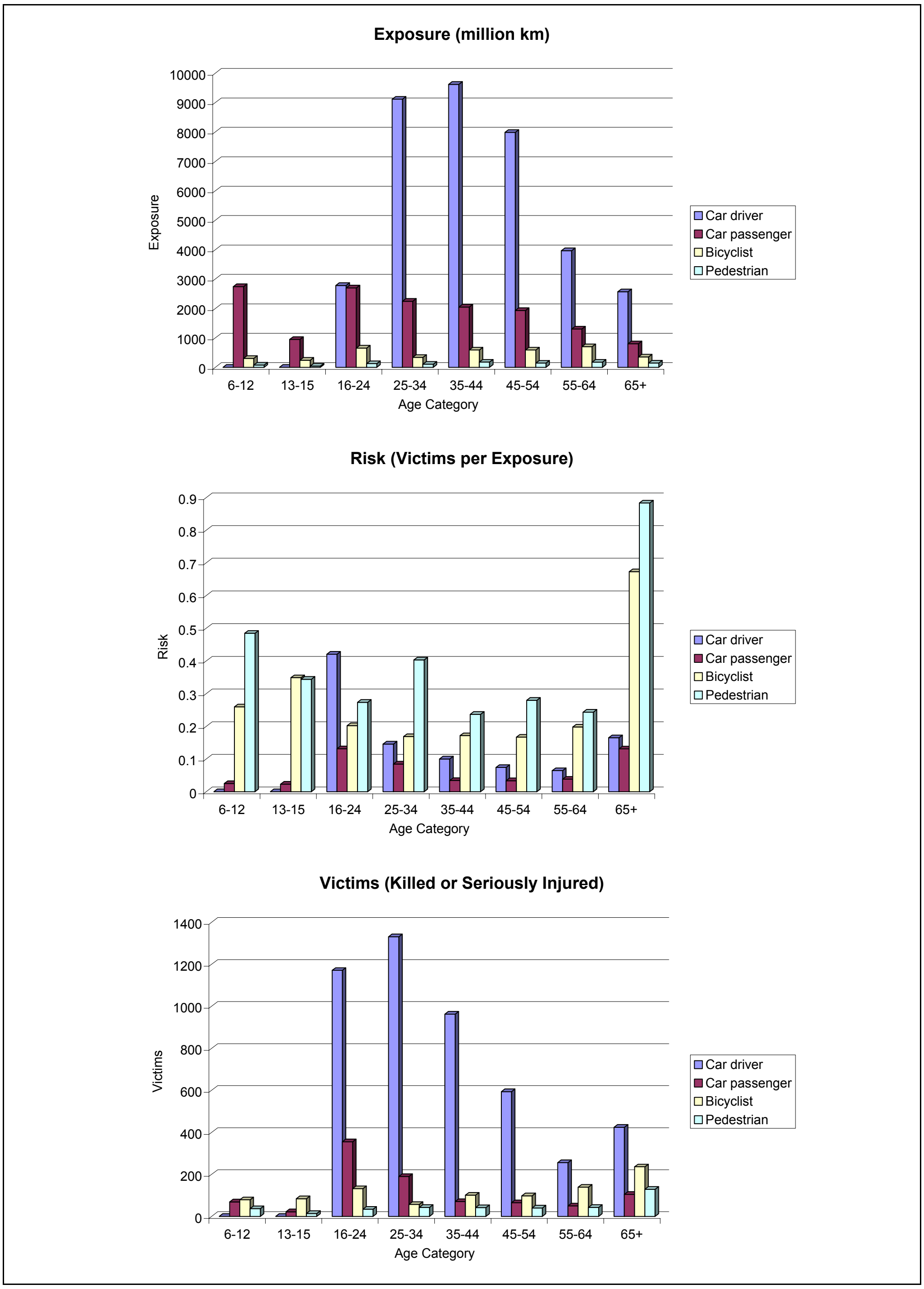

FIGURE 54: Decomposition by modal split

Second, car passengers show high exposure in the youngest age category. Very young children are mostly taken by car to their social or educational activities. The children of age 13-15 are less often car passengers. They probably prefer the bike or go on foot. From the next age category on, car passengers are again more frequent, but their level decreases with age. The risk for car passengers is highest for young (16-24 years old) and elderly people. Many young people travel together and passengers of this age category are probably accompanied (and driven) by peers. The highest number of car passenger victims is found among the youngsters between 16 and 24, mainly because of the higher risk. Third, bicyclists and pedestrians show a relatively low level of exposure over all ages. The number of victims is especially high for older bicyclists and pedestrians. Their number is still relatively low compared to car users, but the risk of vulnerable road users is remarkably high. Evans (2004) showed that, after the age of 60, the risk of pedestrian death per person increases steeply, and then again declines, likely reflecting reduced walking. The increasing involvement of elderly people in pedestrian crashes may reflect the decreasing perceptual skills and agility.

These findings are interesting for policy makers who should promote bicycle use and walking as examples of sustainable means of transport. As long as the risk of these road users is high, these modes of transport may be less attractive than any other alternative.

8.4 Conclusion

For many traffic safety researchers, the main problem is finding the right data sources. The indicators needed for traffic safety analysis are exposure, risk and the number of victims or crashes. In many countries, these data are not always available in a format that is needed for specific modelling purposes. Especially exposure data is mostly not gathered with the objective of analyzing traffic safety. In Flanders, the same problems arise. In this chapter, some examples of the use of travel survey data for the analysis of traffic safety were presented.

A decomposition into exposure and risk, by age and gender category, shows the impact of the main traffic safety indicators for road users of different age-gender combinations. For each combination, the number of victims can be explained as the result of exposure to risk and the risk itself. The models also show which factor is most important. Typically for age categories, it is possible to detect the high-risk groups. Also, exposure measures based on travel surveys are mostly more detailed than the exposure measures based on traffic counts. They allow splitting up the analysis by socio-demographic characteristics like age and gender. As shown in this chapter, they are better suited to point out the safety differences in the various modes of transport.

On the other hand, the approach has some limitations. First, the travel survey data are based on a sample of road users, who registered their travel behaviour for a few days. Sampling errors may influence the results, and for some subcategories by age and gender, the number of respondents may be too low to allow valid extrapolations. Second, the sample of the travel survey is by no means in accordance with the observed crashes and victims. It is quite possible that there were no crashes registered among the road users in the travel survey sample. Instead, we can only match the extrapolated sample data of the travel surveys with the observed accident counts. It is then implicitly assumed that the exposure, calculated from the travel surveys, is also representative for the exposure that (partly) causes road crashes. Third, the availability of travel survey data determines the possibilities for analysis of the traffic safety situation. If trends in the number of victims are to be studied over a long period of time, we have to make sure that the travel survey is conducted on a regular basis. In Flanders, travel surveys are only available for 2 separate years, namely 1995 (Hajnal & Miermans, 1996) and 2000 (Zwerts & Nuyts, 2004), which excludes the possibility of a reliable evolutionary study. For Belgium, only one travel survey study is known (Hubert & Toint, 2002). Given the interesting conclusions that can be drawn on the basis of this kind of analysis, there is a clear incentive to extend the frequency of conducting travel surveys in Flanders, as it is done in other countries.

The use of travel survey data, which are typically mobility-related, can greatly improve the knowledge of the relationship between mobility and traffic safety. Given the importance of exposure in traffic safety studies, this does not come as a surprise. If mobility data are gathered on a regular basis, and directed towards traffic safety, these data can provide useful insights that would remain hidden if only less detailed information is used. For policy makers, this information can steer their campaigns, and determine which kind of transport mode should be made more safe or more attractive for specific groups of road users.

Chapter 9 Final conclusions

9.1 Looking back…

Probably every fundamental research effort that comes to an end is, in several ways, restricted in its practical use. Typically a large amount of knowledge on a very specific problem is generated and gathered. Too often, the relevance of such a piece of work is obscured by the lack of a larger context in which the scientific progress gets its meaning. In road safety research, the relevance and the final objectives are obvious. Given the demand for a transport system, initiated by economic, social and cultural activities that are to be accomplished, road safety research should contribute to the reduction of the number of accidents and victims, which are the negative by-products of the system. This is a huge task, but it is hoped that every piece of work in this domain will enhance the knowledge on road safety and reduce the magnitude of the problem. This is a never-ending story, because the level of road safety does not remain constant until the research is done.

In this manuscript, yet another contribution to the field of road safety research is presented. The purpose was to provide an insight in the relation between road safety, exposure and risk in Belgium, using data at various aggregation levels. With this topic, the author takes a bird's-eye view to the problem. Instead of investigating accidents and victims on a one-per-one basis, road safety data are aggregated per time unit (per month and per year), and groups of subjects are analysed. As a consequence, the models developed in this work should be considered as strategic devices to get insight in the way road safety evolves.

In Chapter 1, a motivating introduction to the topic was given. Belgium is still a poor student as it comes to road safety. Governments and road safety institutes are concerned with the problem and try to introduce measures that eventually lead to a reduction in the number of victims and accidents. Clearly, at a strategic level, a policy maker needs high-level information that shows general trends in road safety. The models presented in this work are very well suited for this task. A time series approach is taken, such that trends can become clear and interventions can be evaluated. Chapter 2 provided an overview of this kind of models. Time series road safety models are not new, but they have never been applied in a structured way to the Belgian road safety situation.

In very general terms, the connecting thread between the various chapters in this work is the basic relation that decomposes road safety (in terms of accidents or victims) in a risk component and a corresponding measure of exposure to the risk. The two basic sources of information that are needed for this kind of analysis are the accident data and information on the level of exposure. While the governments of several countries significantly improved upon the quality of their accident databases, it is found that exposure measures are far less available in a format and with a content that is useful in road safety research. Therefore, Chapter 3 offered an overview of possible measures of exposure, and indicated how they can be obtained and used. Because monthly exposure data were not available for Belgium, an exposure measure was created in Chapter 4, on the basis of fuel sales, average fuel efficiency and the vehicle park. Also other sources of data were introduced in the same chapter, including an in-depth discussion of Belgian accident data, an overview of the available exposure data (on a yearly basis) and a set of possible explanatory variables.

Next, in Chapter 5 some highly aggregated road safety models were developed for Belgium. On the basis of the yearly number of fatalities and the number of kilometres driven, the models provided an insight in the interaction between road safety, risk and exposure. Starting from the famous Oppe models, some extensions were demonstrated. First, the logistic curve for exposure was replaced by the more flexible Richards curve, and a constant was added to the exponential risk curve. Second, using alternative risk models, the basic relationship was extended with a parameter that allows for non-proportionality in the exposure-risk relation, and some road safety interventions were tested. Globally, the models did not show any reason to reject the proportionality for the yearly data. Further, depending on the model considered, the road safety interventions concerning seat belt use, speeding and alcohol indicated significant reductions in the number of victims. In this chapter, also the flexible state space models were introduced. This approach is based on the idea that several components in the classical exposure-risk relation are stochastic and unobserved, recognising that these variables are measured with uncertainty. In a first setting, where exposure was treated as an explanatory variable, the proportionality assumption could again not be rejected. In a second state space model, a latent risk time series model, exposure was introduced as a latent variable. In general, the stochastic approach did not lead to very different results compared to the previous models, but it provided a framework that is more naturally related to the basic idea behind the relation between exposure and risk. The different models tested in this chapter resulted in very diverging predictions of the number of fatalities in 2010, although the more elaborated models lead to quite sensible trends.

Chapter 6 introduced models on monthly data. The difference in frequency (monthly compared to yearly data) offers the advantage of a larger data set, which makes it possible to include more explanatory variables in the models. Also, monthly data show specific seasonal patterns, which are interesting from a descriptive point of view. Both ARIMA and state space descriptive models were developed, leading to quite comparable results in terms of predictions. However, the state space approach is more “descriptive” than ARIMA models, in the sense that trends and seasonal components can be easily extracted. The explanatory model with calendar variables slightly improved the fit, but resulted in a forecasting accuracy that can be compared with the ARIMA models. However, even the simple calendar variables like a trading day pattern or the leap years can give an added value to the models. The laws on seat belt use, speed and alcohol gave significant and reassuring results. This was not the case for the law on seat belt use in the rear seats and the installation of automatic speed cameras. Probably these laws are different in nature and are therefore hard to test in an aggregated model.

In a second explanatory model, apart from the laws and the calendar variables, also the effect of the weather, economic conditions and exposure were tested. Further, a model for the number of light injuries was added and exposure was modelled in a separate equation. This allows measuring both the direct and indirect (via exposure) effects of some explanatory variables. Again, the model is a regression-ARIMA model, extended with a GARCH structure on the residuals to account for possibly unequal error variances. The results for the laws on seat belt use in the rear seats, alcohol and automatic cameras were comparable with those obtained earlier. As for the weather, higher temperatures and more precipitation seem to reduce safety, while no effects for frost are found. The number of days with snow increases the number of accidents. Further, the relation between the weather and exposure is not according to expectations, and probably indicates an influence of the way the exposure measure was developed. For the economic conditions, both a higher fuel price and a higher level of unemployment reduce road safety. This is not in line with literature, but it is not unthinkable that these relations are influenced by the economic structure of a country (especially for unemployment) or by the complex role of exposure (especially for fuel price). Indeed, some effects seem to reflect the changes in the relation between exposure and road safety. When comparing the outcomes of these models with the previous ones, it is important to note that the series in the last models are shorter (because of the availability of exposure data), which can undoubtedly influence the results.

The most important results in these models were again obtained for exposure. The relation between exposure and road safety deviated from proportionality. Also, there is a difference between the results for the number of fatalities compared to serious or light injuries. This was further investigated in a stylised parabolic experiment. For each kind of victims a quadratic relation with exposure was fitted. Whereas the number of light injuries is still increasing with traffic, it seems that a higher level of exposure may lead to reductions in the number of fatalities and serious injuries. The insights from this chapter bring up the topic of a turning point in the relation between exposure and road safety, which is closely related to the congestion problem. Further, if the relation between exposure and road safety is less than proportional, the effect of modelling road safety without this key variable will probably not harm the other results obtained in the model. However, it is always preferred to keep this variable in the model, provided that it is available.

Apart from the very general models that were developed in Chapter 5 and Chapter 6, a more disaggregated or “subset” approach was presented in Chapter 7. That is, instead of looking at road safety for the whole country, types of roads and groups of road users are considered. Note that these models remain highly aggregated in time, as they are developed on yearly data. Disaggregation was done by age and gender and by type of road user (car, truck, motorcycle). Further, models were developed according to road type (motorways, provincial roads and local roads) and for two-sided accidents.

In the analysis of fatalities per age and gender category, it was found that road risk is changing over the age groups according to a U-shaped curve, and that men generally have higher risk than women. Further, the risk is decreasing over time, but not at the same rate for all age-gender groups. The highest yearly reduction in risk is found for the oldest and youngest road users.

In the analysis per type of road user, the relative risk curves and the risk indices clearly show that trucks and motorcycles have higher risks than cars. This information is useful for policy makers, as it helps them to recognize that road safety policies are not equally effective for all groups of road users. It is especially striking that almost no reduction in risk is observed for motorcycles over time.

The less than proportional effects found for the truck and motorcycle exposure are further clarified in the analysis of two-sided accidents. These analyses showed that two-sided accidents are more determined by the ratio of the exposure measures of the different transport modes involved than by the levels of these exposure measures. This result strongly suggests an integrated approach towards road safety, in which the safety of various transport modes can be jointly considered and evaluated. Also, the flexibility of state space models in treating incomplete series of data was nicely illustrated.

The analysis per type of road showed that provincial and local roads decreased together over time in risk, while the risk on motorways is highly irregular and does not show a clear risk-reducing pattern.

Chapter 8, finally, introduced some examples of disaggregated analyses, based on Flemish travel survey data, which are not yet possible in a time series context for Belgium. First, the decomposition of the number of victims in exposure and risk was developed in a cross-sectional setting, per age and gender for the year 2000. In a second study, the risk, the exposure and the number of victims for car drivers, car passengers, bicyclists and pedestrians were analysed per age category. Although the models in this chapter are more qualitative in nature compared to the previous applications, they show some nice examples of highly disaggregated analyses. Moreover, they prove the potential of the use of travel survey data in road safety research, especially for the analysis of vulnerable road users.

9.2 Looking forward…

The analysis of road safety in Belgium, in relation with exposure measures at various levels of aggregation, revealed some useful insights for policy makers and clearly illustrated the added value of macro models in road safety research. Many of these models can be seen as triggers for further research in this area. Of course, apart from the insights obtained from the models developed in this manuscript, the results should also be seen as a plea for a more structured and consistent road safety data gathering in Belgium. Data on road safety, exposure and, if applicable, explanatory variables should be gathered in a consistent way and at various levels of aggregation in time and space.

The main challenges that were found in this text are at the core of the relation between exposure, risk and road safety. This relation is not unambiguous, shows different patterns for subparts of the road system and seems to be changing over time. The (more than) proportionality property, that used to be found in this relation, is not always present anymore.

First, the results of the analysis of road safety and exposure may depend on the length of the time series that is analysed and, perhaps more important, on the period or time window that is considered. The changes in the relation between exposure and road safety introduce the issue of turning points or, more generally, of flexible form relationships. Indeed, depending on the series that is analysed in a specific time window, the effects on road safety of exposure (and even other variables in the model) may change. Also, a change of the level of aggregation in time (yearly versus monthly data) can shed another light on the problem. This topic needs further clarification.

Second, it is clear that the accepted properties of the relation between exposure and risk are changing. As traffic continues to grow over time, it is expected that in the future increasing traffic will reduce the number of accidents and victims. It will be very hard, then, to find an equilibrium between transport demand and supply in which the number of accidents and victims can be controlled. The results from previous research will probably not be valid anymore and models should be adapted or re-fitted in order to take these changes and new dimensions into account.

Third, the relation between road safety, exposure and risk depends on the level of aggregation. Indeed, the disaggregation of road safety according to age and gender of the road users, transport modes, types of roads and types of accidents provided useful insights that can improve road safety policies. Such an analysis would also point out very diverging results over different countries. In Belgium, for example, two types of accidents are frequently observed: one-sided fatal accidents (typically at night) and accidents in which many road users, especially trucks, are involved. Both situations can result in the same number and kind of victims, although the crucial role of exposure in both types of accidents can be very different. This topic includes modelling and data gathering challenges.

Fourth, given the importance of exposure in road safety research, continuous attention should be given to the quality and the meaning of the various exposure measures available. Every result concerning the road risk or its proportionality towards the number of accidents or victims depends on the exposure measure used. Although the approaches with unobserved components explicitly recognise the (lack of) quality of the exposure measure, this does not protect the results from the influence of invalid data. Fortunately, many countries and governmental institutions are spending time and money on the construction of valid exposure measures. Better exposure measures will also improve the knowledge about the true impact on road safety.

Fifth, when using explanatory models, it is expected that the changing relation between exposure and risk will also affect the parameter estimates obtained for the other covariates in the model. Typically for the variables that have an impact on both the level of exposure and the level of road safety (the weather, economic conditions, etc.), different results might be obtained, changing at the same time the elasticities of direct and indirect effects.

Sixth, in line with the previous comments and also in relation with explanatory models, it might be interesting to consider the combination of established model structures like the DRAG approach and the flexibility of state space modelling. In particular, the flexibility of stochastic trend modelling may be an added value in the context of the changing impact of exposure on road safety. For example, if the effect of exposure is changing from "more than proportional" to "less than proportional", it is natural to ask how this transition took place over time and to test whether this is also present in disaggregated time series. Also, the layered structure of the DRAG approach (exposure, frequency, severity) can be quite naturally modelled in a state space formulation. This would open up the possibility of a new reference class of models that is at the same time statistically sound and flexible enough to allow turning points or changing patterns in the relation between exposure and road safety, without assuming fixed functional form relationships.

The models developed in this manuscript serve two distinct higher purposes. First, an attempt is made to unearth the properties of the Belgian road safety situation. A similar approach has never been followed in this country. Therefore, it is believed that the results have an added value for road safety research and policy in Belgium. In the second place, the approaches followed in this text join a movement in international research on the topic of macroscopic road safety models. It is hoped that the combination of descriptive and explanatory time series models, the development of studies at various levels of aggregation and the results obtained in this context for the Belgian situation also contribute to the state-of-the-art of this type of road safety research. Of course, it is still an unpretentious attempt to sharpen the questions for future research in this domain and to raise a corner of the veil of the complex issue of road safety modelling.

Samenvatting

Decennia lang werden duizenden pagina's geschreven over het verkeersveiligheidsprobleem, maar daarvan kan geen enkel woord, zelf geen enkele zin, de pijn, de zorgen en het onherstelbare verlies uitdrukken dat met verkeersongevallen gepaard gaat. Verkeersveiligheid is een wereldwijd probleem, met gevolgen voor de volksgezondheid, het sociale leven en de economische voorspoed van een land. Met wereldwijd 1.2 miljoen doden per jaar behoort de verkeersdrukte tot de belangrijkste doodsoorzaken, niet in het minst voor kinderen. Daarnaast zorgt de onveiligheid op de wegen voor enorme economische kosten. België is op het vlak van verkeersveiligheid zeker geen koploper, en noteert al meerdere jaren een aantal verkeersdoden boven het Europese gemiddelde. Verkeersveiligheid stopt niet aan de landsgrenzen, en wordt momenteel op alle beleidsniveaus aangepakt. Dit resulteerde de laatste jaren in een daling van het aantal ongevallen en slachtoffers. België zal echter nog heel wat inspanningen moeten leveren om de Europese koplopers in te halen. Hoe beter de verkeersveiligheid, des te meer inspanningen er nodig zijn om een bijkomende daling in het aantal slachtoffers te realiseren.

Er zijn verschillende benaderingen om het verkeersveiligheidsprobleem aan te pakken. In dit proefschrift wordt de verkeersveiligheid vanuit een macroscopisch perspectief bekeken. In plaats van naar individuele ongevallen of locaties te kijken, worden maandelijkse of jaarlijkse statistieken van ongevallen geanalyseerd. Dit houdt niet alleen het totale aantal doden per tijdseenheid in, maar bijvoorbeeld ook het aantal slachtoffers per leeftijd, per type weg of per type weggebruiker (en dit telkens per tijdseenheid). In elk geval is de analyse gebaseerd op een decompositie van het probleem in drie dimensies, namelijk blootstelling, risico en gevolg. Veranderingen in een van deze dimensies kan de volledige verkeersveiligheidssituatie wijzigen. Elke analyse in deze tekst zal van deze (of een gelijkaardige) assumptie vertrekken om de Belgische verkeersveiligheid in kaart te brengen. Terwijl de meeste studies voor België op vandaag gebruik maken van cross-sectionele data, worden in dit werk tijdreeksen van ongevallen en slachtoffers geanalyseerd. Dit betekent dat alle data sequentieel en met een regelmatige frequentie in de tijd werden verzameld, en dat trends in verkeersveiligheid en blootstelling kunnen geanalyseerd worden.

De belangrijkste doelstelling van dit werk bestaat erin de relaties tussen verkeersveiligheid, blootstelling en risico in een tijdsperspectief te beschrijven, te verklaren en te voorspellen voor de Belgische situatie. De overheden en

verkeersveiligheidsinstellingen zijn begaan met het probleem en voeren maatregelen in die uiteindelijk zouden moeten leiden tot een daling van het aantal slachtoffers. Op een strategisch niveau hebben de beleidsmakers dus geaggregeerde informatie nodig over de algemene evolutie in de verkeersveiligheid. De modellen die in deze tekst worden gebruikt, en die in hoofdstuk 2 worden toegelicht, zijn hiervoor geschikt. Naast een motivatie voor dit type van onderzoek wordt ook een typologie van de meest gangbare modellen voorgesteld en worden de belangrijkste resultaten uit de literatuur bestudeerd. In een derde hoofdstuk worden maten voor blootstelling in detail toegelicht. De blootstelling is immers een cruciale variabele in de analyses, en kent heel wat problemen die typisch in de context van verkeersveiligheidsonderzoek kunnen opduiken. Deze hoofdstukken vormen de inleiding tot het onderzoek in de volgende hoofdstukken.

In hoofdstuk 4 wordt een inspanning geleverd om de mogelijke bronnen van informatie (blootstelling, ongevallen en mogelijke verklarende factoren) voor het besproken type van onderzoek te identificeren en in kaart te brengen. Het gaat hier steeds om (officiële) Belgische of Vlaamse data, in de vorm van tijdreeksen. Daarnaast wordt in dit hoofdstuk een maandelijkse maat voor de blootstelling (in aantal gereden kilometers) opgesteld voor België, gebaseerd op cijfers van de leveringen van brandstof, het voertuigenpark en de verbruiksefficiëntie van voertuigen. De data die in dit hoofdstuk worden verzameld zullen in de volgende hoofdstukken geanalyseerd worden.

In hoofdstuk 5 worden trends in de Belgische verkeersveiligheid geanalyseerd op een zeer sterk geaggregeerd niveau. Aan de hand van de jaarlijkse statistieken van ongevallen en voertuigkilometers wordt de relatie tussen het aantal slachtoffers, de blootstelling en het risico onderzocht. Hiertoe worden een aantal basis- en meer geavanceerde statistische modellen toegepast, en worden de fit en het voorspellende karakter van de modellen vergeleken. Als basis wordt het klassieke "Oppe-model" toegepast op de Belgische data, waarin het aantal slachtoffers (F_t) wordt beschouwd als het product van een logistische blootstelling (V_t) en een exponentieel risico (R_t), of $F_t = V_t \times R_t$. Hiervoor worden vervolgens een aantal uitbreidingen getest. Eerst wordt de logistische curve voor blootstelling vervangen door de meer flexibele Richards curve, en wordt aan de klassieke exponentiële risicocurve een constante toegevoegd. Vervolgens worden alternatieve risicomodellen voorgesteld, die een niet-proportionele relatie tussen het aantal slachtoffers en het risico toelaten door een parameter voor de blootstelling te schatten, zodat $F_t = V_t^{\eta} \times R_t$, en worden een aantal verkeersveiligheidswetten (gordeldracht, snelheid en alcohol) getest. Verder worden in dit hoofdstuk ook state space modellen gebruikt. Deze benadering is gebaseerd op de veronderstelling dat bepaalde componenten in het klassieke Oppe-model niet observeerbaar zijn. De proportionele relatie kan op de jaarlijkse data niet worden verworpen, en de wetten voor gordeldracht, snelheid en alcohol realiseren (afhankelijk van het model) een significante daling in het aantal slachtoffers. Vervolgens worden in een multivariaat state space model zowel het risico als de blootstelling als niet geobserveerde componenten beschouwd. Deze geavanceerde modellen geven geen afwijkende resultaten ten opzichte van de eerdere modellen, maar zijn wel op een heel natuurlijke manier verbonden met de basisrelatie tussen slachtoffers, blootstelling en risico. Alle modellen worden verder ook gebruikt om een voorspelling te maken van het aantal slachtoffers in 2010. Heel uiteenlopende voorspellingen worden opgetekend, maar algemeen kan men stellen dat de geavanceerde modellen tot zeer redelijke en aanvaardbare voorspellingen leiden. Vergelijking is evenwel noodzakelijk.

In hoofdstuk 6 wordt de relatie tussen verkeersveiligheid en blootstelling geanalyseerd op een lager frequentieniveau, namelijk voor maandelijkse data. Dit heeft het voordeel dat meer observaties beschikbaar zijn, waardoor ook meer uitgebreide (verklarende) modellen kunnen worden opgesteld. Verder tonen maandelijkse data de seizoensschommelingen die typisch zijn voor verkeersongevallen en voor een aantal gerelateerde grootheden. Voor de beschrijvende modellen worden zowel ARIMA als state space technieken gebruikt, en de voorspellingen van deze modellen zijn in grote lijnen vergelijkbaar. In alle geval kan men stellen dat een state space model meer "beschrijvend" is dan een ARIMA model, omdat de onderliggende dimensies (trends en seizoenschommelingen) expliciet kunnen worden weergegeven. Een eerste verklarend regressiemodel met ARMA foutentermen model maakt enkel gebruik van kalenderdata (trading days, schrikkeljaar, invoering van wetten,…), en is dus niet veeleisend op het vlak van dataverzameling. Hoewel het verklarende karakter van deze modellen dan ook eerder beperkt is, kunnen interessante inzichten worden verworven, zoals over het fenomeen van de weekendongevallen. Ook de wetten in verband met gordeldracht, snelheid en alcohol geven hier significante en geruststellende resultaten. Dit is niet het geval voor de wetten in verband met de gordeldracht achter in de wagen en de introductie van onbemande camera's. Vermoedelijk zijn deze wetten van dusdanige aard dat ze moeilijk in een geaggregeerd model kunnen getest worden. Het voorspellend karakter van het model met kalendervariabelen is ongeveer vergelijkbaar met dat van de pure ARIMA of state space modellen.

In een tweede verklarende model worden naast de wetten en de kalendervariabelen ook de impact van het weer, de economie en de blootstelling getest. Verder wordt een model toegevoegd voor het aantal lichtwonde slachtoffers en wordt ook de blootstelling in een aparte vergelijking gemodelleerd. Dit laat toe om zowel de directe als de indirecte (via de blootstelling) effecten van een aantal verklarende variabelen te meten. Het model is opnieuw een regressiemodel met ARMA foutentermen, maar laat tevens een GARCH structuur toe om eventuele ongelijke varianties op te vangen. De resultaten voor de wetten in verband met gordeldracht achter in de wagen, snelheid, alcohol en onbemande camera's zijn vergelijkbaar met eerdere resultaten. Wat het weer betreft blijken een hogere temperatuur en meer neerslag te leiden tot meer onveiligheid, terwijl de vrieskou geen enkel effect heeft. Het aantal sneeuwdagen in een maand doet het aantal ongevallen stijgen. De relatie tussen het weer en de blootstelling in het model is niet altijd volgens verwachting, en wijst vermoedelijk op een invloed van de manier waarop de variabele voor blootstelling werd opgesteld. Bij de economische variabelen geven zowel een hogere brandstofprijs als een hogere werkloosheid aanleiding tot meer onveiligheid. Deze resultaten zijn niet volledig in overeenstemming met de literatuur, maar het is niet ondenkbaar dat ze mede worden beïnvloed door de economische structuur van een land of door de complexe impact van de blootstelling. Wanneer men deze modellen vergelijkt met eerdere modellen, dan is het tevens van belang te noteren dat de tijdreeksen in het laatste verklarende model korter zijn dan in de vorige modellen; dit kan ongetwijfeld de uitkomsten beïnvloeden.

De belangrijkste resultaten van dit model zijn die voor de bloostelling. De relatie tussen de blootstelling en de verkeersveiligheid is in deze modellen niet meer proportioneel te noemen (de parameterschatting voor blootstelling is significant kleiner dan 1). Verder is er een duidelijk verschil tussen de ernst van de afloop (doden, zwaargewonden, lichtgewonden) en het effect van blootstelling. Dit werd verder onderzocht in een gestileerd parabolisch experiment. Voor elke ernst van ongevallen werd een kwadratisch verband met de blootstelling gezocht. Hieruit blijkt enerzijds dat het aantal lichtgewonden nog steeds stijgt met de blootstelling, maar dat anderzijds de blootstelling al zodanig hoog is dat een verdere stijging op termijn wel kan leiden tot een daling in het aantal doden of zwaargewonden. Dit verwijst naar mogelijk "keerpunten" in de relatie tussen verkeersveiligheid en blootstelling, wat uiteraard sterk gerelateerd is aan het probleem van congestie. Een ander aspect van de minder dan proportionele verhouding is het belang van de blootstelling in een model voor de verkeersveiligheid. Met een aan nul grenzende elasticiteit zal het niet opnemen van de blootstelling de resultaten van een model vermoedelijk niet schaden. Uiteraard is dit enkel een optie wanneer de blootstelling niet beschikbaar is, en zelfs in die gevallen kan het moeilijk gelden als een argument om deze variabele niet te bekomen. Blootstelling is immers een sleutelvariabele in de basisrelatie tussen verkeersveiligheid, blootstelling en risico, en is op die manier onmisbaar vanuit een conceptueel standpunt.

Vermits het algemeen gekend is dat maatregelen vaak gericht zijn op de reductie van het aantal slachtoffers in een welbepaald segment van het verkeerssysteem, worden in hoofdstuk 7 een aantal "subset" modellen ontwikkeld die de relatie tussen verkeersveiligheid, blootstelling en risico analyseren voor een bepaalde groep van weggebruikers, wegen, of ongevallen. Uit de analyse van het aantal slachtoffers per leeftijd en geslacht kan men afleiden dat de risico's over de leeftijden heen een U-vormig verloop kennen, en dat mannen over het algemeen een hoger risico vertonen dan vrouwen. Het risico blijkt te dalen in de tijd, maar deze daling is niet voor elke groep gelijk. De hoogste jaarlijkse reductie in risico werd opgetekend voor de oudste en de jongste weggebruikers.

De relatieve risico's en de risico-indices die werden afgeleid in de analyse per type weggebruiker (wagen, vrachtwagen, motorfiets) tonen aan dat vrachtwagens en bromfietsen een hoger risico dan wagens vertonen. Het is ook opvallend dat het risico van motorfietsers doorheen de tijd bijna niet gedaald is. Dergelijke kennis is nuttig voor de beleidsmakers, want ze toont aan dat verkeersveiligheidsmaatregelen niet tot dezelfde resultaten leiden voor elk type weggebruiker. De minder dan proportionele effecten die werden gevonden voor de blootstelling van vrachtwagens en motorfietsen worden verder verduidelijkt in de analyse van de tweezijdige ongevallen tussen deze types van weggebruikers. Hieruit blijkt dat het aantal slachtoffers in deze ongevallen in feite wordt bepaald door een ratio van de blootstellingen (voertuigkilometer) van de respectievelijke groepen van weggebruikers. Verder illustreert deze analyse heel mooi de flexibiliteit van de state space modellen bij de analyse van onvolledige tijdreeksen. De analyse per type weg, tenslotte, toont aan dat de rijkswegen en gemeentewegen een duidelijke reductie in risico vertonen over de tijd heen, wat niet kan gezegd worden van de autosnelwegen. Hoewel het risico hier over het algemeen lager ligt, is het patroon heel grillig en zeker niet monotoon dalend.

Uit de vorige hoofdstukken is gebleken dat bepaalde types van tijdreeksanalyse niet mogelijk zijn in België omwille van het gebrek aan goede data over de blootstelling. Zo is het niet mogelijk om de risico's van zwakke (niet-gemotoriseerde) weggebruikers te analyseren, omdat hun blootstelling niet over de tijd heen werd gemeten. Daarom toont hoofdstuk 8 een aantal eenvoudige cross-sectionele studies waarin gebruik wordt gemaakt van het Vlaamse "Onderzoek Verplaatsingsgedrag" (OVG). Een eerste studie toont een decompositie van het aantal slachtoffers in blootstelling en risico voor het jaar 2000, per leeftijd en geslacht. In een tweede studie werden deze gegevens, per leeftijdscategorie, geanalyseerd voor autobestuurders, autopassagiers, fietsers en voetgangers. Hoewel deze modellen, in vergelijking met de vorige hoofdstukken, eerder kwalitatief zijn van aard, tonen ze welke interessante resultaten op een laag niveau van aggregatie kunnen worden bekomen. Ze geven overigens ook het potentieel aan van onderzoek naar verplaatsingsgedrag in studies voor de verkeersveiligheid, in het bijzonder voor de zwakke weggebruikers.

De modellen die in deze tekst worden voorgesteld geven de lezer een idee van de mogelijkheden en de beperkingen van de beschikbare data voor onderzoek naar verkeersveiligheid in België. Vermits quasi alle bestaande tijdreeksen over de Belgische blootstelling in dit werk aan bod komen, geeft het bovendien een stand van zaken van de tijdreeksmodellen die kunnen worden aangewend om de Belgische verkeersveiligheid in kaart te brengen, op verschillende niveaus van aggregatie. Met de ontwikkeling van deze modellen wil de auteur dan ook twee hogere doelen dienen. Eerst en vooral moeten de resultaten van de modellen bijdragen tot het ontrafelen van de karakteristieken van de Belgische verkeersveiligheid, ter ondersteuning van het onderzoek en het beleid in dit domein. In de tweede plaats maken de modellen, die in deze tekst werden voorgesteld, deel uit van een internationale stroming in verkeersveiligheidsonderzoek. De combinatie van beschrijvende en verklarende modellen, de ontwikkeling van studies op verschillende aggregatieniveaus en de resultaten hiervan voor de Belgische context zouden moeten bijdragen tot de state-of-the-art van dit type onderzoek. Het blijft uiteraard een pretentieloze poging om de uitdagingen in het betrokken onderzoeksdomein scherper te stellen, en zo een bescheiden bijdrage te leveren aan de vooruitgang van de verkeersveiligheid.

References

Abbas, K. A. (2004). Traffic safety assessment and development of predictive models for accidents on rural roads in Egypt. *Accident Analysis and Prevention, 36*(2), 149-163.

Adams, J. G. U. (1987). Smeed's Law: some further thoughts. *Traffic Engineering and Control, Vol 28*(2), 70-73.

Akaike, H. (1973). *Information theory and an extension of the maximum likelihood principle*. Paper presented at the Second international symposium on information theory, Budapest, Hungary.

Alsop, J., & Langley, J. (2001). Under-reporting of motor vehicle traffic crash victims in New Zealand. *Accident Analysis and Prevention, 33*(3), 353-359.

Atkins, M. S. (1979). A case study on the use of intervention analysis applied to traffic accidents. *Journal of the Operational Research Society, 30*(7), 651-659.

Bell, W. R., & Hillmer, S. (1983). Modeling time series with calendar variation. *Journal of American Statistical Association, 78*(383), 526-534.

Bell, W. R., & Martin, D. (2004, August 8-12). *Modeling time-varying trading-day effects in monthly time series*. Paper presented at the Joint Statistical Meetings, Toronto, Canada.

Bergel, R. (1992). *Utilisation de modeles ARIMA avec variables externes dans le secteur de transport*. Paper presented at the 6th World Conference of Transport Research, Lyon, France.

Bijleveld, F. D., & Commandeur, J. J. F. (2004, 27-28 May 2004). *The basic evaluation model*. Paper presented at the ICTSA meeting, INRETS, Arceuil, France.

Bijleveld, F. D., Commandeur, J. J. F., Gould, P., & Koopman, S. J. (2005). *Model-based Measurement of Latent Risk in Time Series with Applications* (Tinbergen Institute Discussion Paper TI 2005-118/4). Amsterdam, The Netherlands: Tinbergen Institute.

Bijleveld, F. D., Commandeur, J. J. F., Koopman, S. J., & van Montfort, K. (2005). *Multivariate nonlinear time series modelling of exposure and risk in road safety research*. Amsterdam: Vrije Universiteit.

Bijleveld, F. D., & Oppe, S. (1996). *Aggregated models for traffic safety*. *Research note COST 329*. Leidschendam, The Netherlands: Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.

Bijleveld, F. D., & Oppe, S. (1999). Time-ordered analysis of disaggregate road safety data, *COST-329 reader*. Leidschendam, The Netherlands: SWOV.

BIVV. (2002). *Verkeersveiligheid statistieken 2001*. Retrieved March, 20th, 2004, from the World Wide Web: http://www.bivv.be/main/PublicatieMateriaal/Statistieken.shtml

Blum, U., & Gaudry, M. (2000). The SNUS-2.5 Model for Germany. In M. Gaudry & S. Lassarre (Eds.), *Structural Road Accident Models: The International DRAG Family* (pp. 67-96). Oxford: Elsevier Science.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: forecasting and control*. San Francisco, USA: Holden-Day.

Broughton, J. (1991). Forecasting road accident casualties in Great Britain. *Accident Analysis and Prevention, 23*(5), 353-362.

Broughton, J. (1994). *Disaggregate casualty forecasting*. Great Britain: Road Safety Division, DETR, reproduced in COST-329 reader.

Cameron, M. (1997). *Review of modelling of road casualties in The Netherlands* (D-97-18). Leidschendam: SWOV.

Campbell, B. J., & Campbell, F. A. (1986). *Seat belt law experience in four foreign countries compared to the United States*. Falls Church, Virginia: AAA Foundation for Traffic Safety.

Campbell, B. J., & Campbell, F. A. (1988). Injury reduction and belt use associated with occupant restraint laws. In G. J. D. (Ed.), *Preventing automobile injury* (pp. 24-50). Dover, MA: Auburn House.

Cardoso, J. L. (2005). The use of international data on fuel sales and vehicle fleet for the estimation of yearly national traffic volumes. *Accident Analysis and Prevention, 37*, 207-215.

CARE. (2004, 09-11-2004). *CARE: Community Road Accident Database*. European Commission, Directorate-General for Energy and Transport. Retrieved, from the World Wide Web: http://europa.eu.int/comm/transport/care/

Chatfield, C. (2004). *The analysis of time series - An introduction* (6 ed.). Boca Raton, Florida: Chapman & Hall/CRC.

Christens, P. F. (2003). *Statistical modelling of traffic safety development*. Unpublished PhD, Technical university of Denmark, Kongens Lyngby, Denmark.

Cleveland, W. P., & Grupe, M. R. (1983). Modeling time series when calendar effects are present. In A. Zellner (Ed.), *Applied Time Series Analysis of Economic*

Data (pp. 57-67). Washington D.C.: U.S. Department of Commerce, U.S. Bureau of the Census.

Cleveland, W. S., Devlin, S. J., & Grosse, E. (1988). Regression by Local Fitting. *Journal of Econometrics, 37*, 87-114.

Cleveland, W. S., & Grosse, E. (1991). Computational Methods for Local Regression. *Statistics and Computing, 1*, 47-62.

Cochrane, D., & Orcutt, G. H. (1949). Application of least squares regression to relationships containing autocorrelated error terms. *Journal of American Statistical Association, 44*, 32-61.

Cohen, R. A. (1999, April 11-14). *An Introduction to PROC LOESS for Local Regression*. Paper presented at the 24th SAS Users Group International Conference, Florida, USA.

Commandeur, J. J. F. (2002). *Algemene en periodieke trends in de ontwikkeling van de verkeersveiligheid in acht ontwikkelde landen* (R-2002-17). Leidschendam, The Netherlands: Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.

Commandeur, J. J. F., & Koornstra, M. J. (2001). *Prognoses voor de verkeersveiligheid in 2010* (R-2001-9). Leidschendam, The Netherlands: Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.

Cook, P. J., & Tauchen, G. (1984). The effect of minimum drinking age legislation on youthful auto fatalities, 1970-1977. *Journal of Legal Studies, 13*, 169-190.

COST 329. (2004). *Models for traffic and safety development and interventions* (Final report of the action EUR 20913). Brussels, Belgium: Directorate General for Transport, European Commission.

D'Hondt, H. (2002). *Gezondheid - Doodsoorzaken in 1997 volgens de aangiften van overlijden*. Brussels, Belgium: Nationaal Instituut voor de Statistiek.

Dagenais, M., Gaudry, M., & Liem, T. (1987). Urban travel demand: the impact of Box-Cox transformations with nonspherical residual errors. *Transportation Research B, 21B*(6), 443-477.

De Borger, L., & De Borger, B. (1987). *Personenvervoer en energie: modelbouw en simulatie. Eindrapport opgesteld voor de Diensten van de Wetenschappelijke Politiek in het kader van het nationaal programma "Energie"*. Antwerp, Belgium: UFSIA.

de Leur, P., & Sayed, T. (2003). Development of a road safety risk index. *Transportation Research Record, 1784*(02-2814), 33-42.

De Mol, J., & Boets, S. (2003). *Optimisation of traffic accident statistics* (CP-39). Brussels, Belgium: Belgian Federal Science Policy Office (DWTC).

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association, 74*, 427-431.

Durbin, J., & Koopman, S. J. (2001). *Time series analysis by state space methods* (Vol. 24). Oxford, UK: Oxford University Press.

Economic Commission for Europe. (2005). *Statistics of Road Traffic Accidents in Europe and North America*. Geneva: United Nations.

Eisenberg, D. (2004). The mixed effects of precipitation on traffic crashes. *Accident Analysis and Prevention, 36*(4), 637-647.

Elvik, R. (2006). Laws of accident causation. *Accident Analysis and Prevention, 38*, 742-747.

Elvik, R., & Vaa, T. (2004). *Handbook of road safety measures*. Oxford, UK: Elsevier Ltd.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica, 50*, 987-1008.

ETSC. (2003). *Transport safety performance in the EU: a statistical overview*. Brussels, Belgium: European Transport Safety Council.

European Commission.*SafetyNet - Building a European Road Safety Observatory*. Retrieved, from the World Wide Web: http://safetynet.swov.nl/

European Commission. (2001). *White Paper "European transport policy for 2010: time to decide"* (COM(2001)0370). Brussels: Commission of the European Communities.

Evans, L. (1990, January 7-11). *The fraction of traffic fatalities attributed to alcohol*. Paper presented at the 69th Annual Meeting of the Transportation Research Board, Washington, DC, USA.

Evans, L. (2004). *Traffic safety*. Michigan: Bloomfield Hills.

Findley, D., Monsell, B., Bell, W., Otto, M., & Chen, B. (1998). New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program. *Journal of Business and Economic Statistics, 16*(2), 127-177.

Fournier, F., & Simard, R. (2000). The DRAG-2 Model for Quebec. In M. Gaudry & S. Lassarre (Eds.), *Structural Road Accident Models: the International DRAG Family* (pp. 37-66). Oxford: Elsevier Science.

Fowles, R., & Loeb, P. D. (1989). Speeding, coordination, and the 55-mph limit: Comment. *American Economic Review, 79*, 916-921.

Fridstrøm, L. (1997). *TRULS: an econometric model of road use, accidents and their severity*. Oslo, Norway: Institute of Transport Economics.

Fridstrøm, L. (1999). *Econometric models of road use, accidents, and road investment decisions* (Volume II. Report 457). Oslo, Norway: TØI.

Fridstrøm, L. (2000). The TRULS-1 model for Norway. In M. Gaudry & S. Lassarre (Eds.), *Structural Road Accident Models: the International DRAG Family* (pp. 97-126). Oxford: Elsevier Science.

Fridstrøm, L., Ifver, J., Ingebrigtsen, S., Kulmala, R., & Thomsen, L. K. (1995). Measuring the contribution of randomness, exposure, weather and daylight to the variation in road accident counts. *Accident Analysis and Prevention, 27*(1), 1-20.

Fridstrøm, L., & Ingebrigtsen, S. (1991). An aggregate accident model based on pooled, regional time-series data. *Accident Analysis and Prevention, 23*(5), 363-378.

Friedland, M., Trebilcock, M., & Roach, K. (1987). *Regulating traffic safety: a survey of control strategies*. Toronto: University of Toronto, Faculty of Law.

Gaudry, M. (1984, revised 2002). *DRAG, model of the Demand for Road use, Accidents and their Severity, applied in Quebec from 1956 to 1982* (AJD-17): Université de Montréal - Agora Jules Dupuit.

Gaudry, M. (1993). *Le modèle DRAG: éléments pertinents au monde du travail - une expertise exploratoire* (Publication CRT-948). Montréal, Canada: Centre de recherche sur les transports.

Gaudry, M. (2000). Multiple Levels, Damages, Forms, Moments and Variables in Road Accident Models. In M. Gaudry & S. Lassarre (Eds.), *Structural Road Accident Models: the International DRAG Family* (pp. 1-36). Oxford: Elsevier Science.

Gaudry, M., & Blum, U. (1993). Une présentation brève du modèle SNUS-1 (Straßenverkehrs-Nachfrage, Unfälle und ihre Schwere): modélisation de l'insécurité routière. *Collection Transports et Communication, 47*, 37-43.

Gaudry, M., Blum, U., & Liem, T. (2000). *Turning Box-Cox including quadratic forms in regression* (CRT-2000-20). Montréal, Canada: Centre de recherche sur les transports.

Gaudry, M., Fournier, F., & Simard, R. (1995). *DRAG-2, un modèle économétrique appliqué au kilometrage, aux accidents et à leur gravité au Québec: Synthèse des résultats.*: Société de l'assurance automobile du Québec.

Gaudry, M., & Lassarre, S. (2000). *Structural road accident models: the international DRAG family*. Oxford, UK: Elsevier Science Ltd.

Gómez, V., & Maravall, A. (1996). *Programs TRAMO and SEATS, Instructions for the User (Beta Version: September 1996)* (Documento de Trabajo no. 9628): Banco de España – Servicio de Estudios.

Greibe, P. (1999). *Cyclist safety in urban areas* (Paper prepared for COST329 report reader). Denmark: Danish Road Directorate, Ministry of Transport.

Haight, F. (1982). *Accident Analysis and Prevention, 14*(5), 335-424.

Haight, F. (1984). Why the per capita traffic fatality rate is falling. *Journal of Safety Research, 15*, 137-140.

Haight, F. (1991). *Accident Analysis and Prevention, 23*(5), 343-473.

Hajnal, I., & Miermans, W. (1996). *Onderzoek Verplaatsingsgedrag Vlaanderen. Analyseopdracht. Eindverslag*. Diepenbeek, Belgium: Provinciale Hogeschool Limburg, Departement architectuur.

Hakim, S., Shefer, D., Hakkert, A. S., & Hocherman, I. (1991). A critical review of macro models for road accidents. *Accident Analysis and Prevention, 23*(5), 379-400.

Hakkert, A. S., & McGann, A. (1996). *A comparative study of road safety in Australian states* (ARR 278). Melbourne, Australia: ARRB Transport Research Ltd.

Harris, S. (1990). The real number of road traffic accident counts in the Netherlands, a year-long survey. *Accident Analysis and Prevention, 22*(4), 371-378.

Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge, United Kingdom: Cambridge University Press.

Harvey, A. C., & Durbin, J. (1986). The effects of seat belt legislation on British road casualties: a case study in structural time series modelling. *Royal Statistical Society, A, 149*(Part 3), 187-227.

Hauer, E. (1997). *Observational before-after studies in road safety*. Oxford, UK: Elsevier Science Ltd.

Hermans, E., Wets, G., & Van den Bossche, F. (2006a, 22-26 January). *Describing the Evolution in the Number of Highway Deaths by a Decomposition in Exposure, Accident Risk and Fatal Risk.* Paper presented at the 85th Annual Meeting of the Transportation Research Board, Washington D.C., USA.

Hermans, E., Wets, G., & Van den Bossche, F. (2006b). The Frequency and Severity of Road Traffic Accidents Studied by State Space Methods. *Journal of Transportation and Statistics, 9*.

Hoxie, P. D., & Skinner, D. (1985). *The statistical analysis of socio-economic influences on three groups of high risk fatalities*. Washington, DC, USA: National Highway Traffic Safety Administration, US DOT.

Hoxie, P. D., & Skinner, D. (1987a). *Effects of Mandatory Seatbelt Use Laws on Highway Fatalities in 1985*. Washington, DC, USA: National Highway Safety Administration, US DOT.

Hoxie, P. D., & Skinner, D. (1987b). *A Statistical analysis of the effects of a uniform minimum drinking age*. Washington, DC, USA: National Highway Safety Administration, US DOT.

Hoxie, P. D., Skinner, D., & Wang, H. (1984). *Socio-economic influences on highway fatalities: an empirical investigation*. Washington, DC, USA: National Highway Traffic Safety Administration, US DOT.

Hubert, J.-P., & Toint, P. (2002). *La mobilité quotidienne des Belges*. Namur, Belgium: Presses Universitaires de Namur.

Hurvich, C. M., Simonoff, J. S., & Tsai, C. L. (1998). Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal of the Royal Statistical Society B, 60*, 271-293.

Jaeger, L. (1999). *L'évaluation du risque dans le système des transports routiers par le développement du modèle TAG* (PhD). Strasbourg: Université Louis Pasteur.

Jaeger, L., & Lassarre, S. (1997). *Pour une modélisation de l'évolution de l'insécurité routière. Estimation du kilométrage mensuel en France de 1957 à 1993: méthodologie et résultats*. (DERA no 9709). Strasbourg/Paris: Convention DRAST/INRETS.

Jaeger, L., & Lassarre, S. (2000). The TAG-1 model for France. In M. Gaudry & S. Lassarre (Eds.), *Structural Road Accident Models: The International DRAG Family* (pp. 157-184). Oxford: Elsevier Science.

James, H. (1991). Under-reporting of road traffic accidents. *Traffic Engineering and Control, 32*(12).

Johansson, P. (1996). Speed limitation and motorway casualties: a time series count data regression approach. *Accident Analysis and Prevention, 28*(1), 73-87.

Jolly, J. (2005). Weekendongevallen in cijfers. *Via Secura, 70*.

Joshua, S. C., & Garber, N. J. (1990). Estimating truck accident rates and involvements using linear and Poisson regression models. *Transportation planning and technology, 15*.

Jovanis, P. P., & Chang, H.-L. (1986). Modeling the Relationship of Accidents to Miles Traveled. *Transportation Research Record, 1068*, 42-51.

Kinet, S., Boets, S., Drevet, M., Scheers, M., Derweduwen, P., De Beuckeleer, E., Keuleers, B., Wets, G., Hannes, E., Van Hout, K., & Cuyvers, R. (2004). *Exploitatie van gegevens inzake verkeersveiligheid: eindrapport* (AG/EE/054). Brussels, Belgium: Federaal Wetenschapsbeleid.

Koopman, S. J., Harvey, A. C., Doornik, J. A., & Shephard, N. (2000). Stamp: structural time series analyser, modeller and predictor (Version 5). London, UK: Timberlake Consultants Ltd.

Kulmala, R. (1995). *Safety at three- and four-arm junctions; development and applications of accident prediction models* (Technical report 233): VTT.

Labeeuw, G. (2005). *Verkeerstellingen 2004* (28). Brussels, Belgium: Federale Overheidsdienst Mobiliteit en Vervoer, Directoraat-generaal Mobiliteit en Verkeersveiligheid, Directie Mobiliteit.

Lassarre, S. (1986). The introduction of the variables "traffic volume", "speed", and "belt-wearing" into a predictive model of the severity of accidents. *Accident Analysis and Prevention, 18*(2), 129-134.

Lassarre, S. (2001). Analysis of progress in road safety in ten European countries. *Accident Analysis and Prevention, 33*, 743-751.

Lave, C. A. (1985). Speeding coordination and the 55-mph limit. *American Economic Review, 75*, 1159-1164.

Ledolter, J., & Chan, K. S. (1996). Evaluating the impact of the 65 mph maximum speed limit on Iowa rural interstates. *The American Statistician, 50*(1), 79-85.

Lereim, I. (1984). *Traffic Accidents and Their Consequences: a study on injured road users treated at the Regional Hospital of Trondheim* (PhD. thesis). Trondheim, Norway: Trondheim University, Faculty of Medicine.

Levy, D. T., & Asch, P. (1989). Speeding, coordination, and the 55-mph limit: Comment. *American Economic Review, 79*, 913-915.

Li, L., & Kim, K. (2000). Estimating driver crash risks based on the extended Bradley-Terry model: an induced exposure method. *Journal of the Royal Statistical Society A, 163*(3), 227-240.

Liem, T., Gaudry, M., Dagenais, M., & Blum, U. (2000). The TRIO LEVEL-1.5 Algorithm for BC-GAUHESEQ regression. In M. Gaudry & S. Lassarre (Eds.), *Structural road accident models: the international DRAG family* (pp. 263-324). Oxford: Elsevier Science.

Loeb, P. D. (1987). The determinants of automobile fatalities. *Journal of Transport Economics and Policy, 21*, 279-287.

Lopez, D. G., Rosman, D. L., Jelinek, G. A., Wilkes, G. J., & Sprivulis, P. C. (2000). Complementing police road-crash records with trauma registry data — an initial evaluation. *Accident Analysis and Prevention, 32*(6), 771-777.

Maas, M. W., & Harris, S. (1984). Police recording of road accident in-patients: investigation into the completeness, representativity and reliability of police records of hospitalized traffic victims. *Accident Analysis and Prevention, 16*(3), 167-184.

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics, 11*, 601-618.

Makridakis, S., Wheelwright, S. C., & Hyndman, R. (1998). *Forecasting: methods and applications* (3 ed.): John Wiley & Sons, Inc.

McCarthy, P. (2000). The TRACS-CA Model for California. In M. Gaudry & S. Lassarre (Eds.), *Structural Road Accident Models: The International DRAG Family* (pp. 185-204). Oxford: Elsevier Science.

Michener, R., & Tighe, C. (1992). A Poisson Regression Model of Highway Fatalities. *American Economic Review Papers and Proceedings, 82*, 452-456.

Ministerie van de Vlaamse Gemeenschap. (2001). *Ontwerp Mobiliteitsplan Vlaanderen*. Ministerie van de Vlaamse Gemeenschap. Retrieved February, 9th, 2004, from the World Wide Web: http://viwc.lin.vlaanderen.be/mobiliteit

Minter, A. L. (1987). Road casualties - improvement by learning processes. *Traffic Engineering and Control, 28*(2), 74-79.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied Linear Statistical Models*: WCB/McGraw-Hill.

Nuyts, E. (2004). *Effectiviteit van onbemande camera's* (RA-2004-46). Diepenbeek, Belgium: Steunpunt verkeersveiligheid.

OECD. (1997a). *Road safety principles and models* (IRRD NO. 888815). Paris, France: Organisation for economic co-operation and development.

OECD. (1997b). *Road safety principles and models: review of descriptive, predictive, risk and accident consequence models* (OECD/GD(97)153). Paris, France: OCDE-OECD.

OECD. (1998). *IRTAD Special Report: definitions and data availability*. Bergisch Gladbach, Germany: BASt.

Oppe, S. (1989). Macroscopic models for traffic and traffic safety. *Accident Analysis and Prevention, 21*, 225-232.

Oppe, S. (1991). The development of traffic and traffic safety in six developed countries. *Accident Analysis and Prevention, 23*(5), 401-412.

Oppe, S. (2001). *Traffic safety developments in Poland* (D-2001-8). Leidschendam, The Netherlands: Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.

Pankratz, A. (1991). *Forecasting with dynamic regression models*. New York: John Wiley & Sons.

Partyka, S. C. (1984). Simple models of fatality trends using employment and population data. *Accident Analysis and Prevention, 16*, 211-222.

Pedersen, C. B. (1998). Log linear splines, *COST-329 reader*. Brussels, Belgium: European Commission, Directorate General for Transport.

Peltzman, S. (1975). The effects of automobile safety regulation. *Journal of Political Economy, 83*(4), 677-725.

Pereira, P., & Pernías-Cerrillo, J. C. (2005). *The Diffusion of Cellular Telephony in Portugal before UMTS: A Time Series Approach*. Lisbon, Portugal: Competition Authority.

QMS. (2004). *EViews 5.0 User's Guide*. Irvine CA, USA.

Quantitative Micro Software, L. (2004). *EViews 5.0 User's Guide*. Irvine CA, USA.

Raeside, R., & White, D. (2004). Predicting casualty numbers in Great Britain. *Transportation Research Record: Journal of the Transportation Research Board, 1897*, 142-147.

Recht, J. L. (1965). *Multiple regression study of the effects of safety activities on the traffic accident problem* (Technical report). USA: National Safety Council.

Rosman, D. L., & Knuiman, M. W. (1994). A comparison of hospital and police road injury data. *Accident Analysis and Prevention, 26*(2), 215-222.

Rutherford, W. H. (1987). Compulsory wearing of seat belts in the United Kingdom: the effect on patients and on fatalities. *Transport Reviews, 7*(3), 245-257.

SAS Institute Inc. (2004a). *SAS/ETS 9.1 User's Guide*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. (2004b). *SAS/STAT 9.1 User's Guide*. Cary, NC, USA: SAS Institute INC.

Scott, P. P. (1986). Modelling time series of British road accident data. *Accident Analysis and Prevention, 18*(2), 109-117.

Scuffham, P. A. (2001). *An econometric analysis of motor vehicle traffic crashes and macroeconomic factors*. Unpublished PhD thesis, University of Otago, Dunedin, New Zealand.

Scuffham, P. A. (2003). Economic factors and traffic crashes in New Zealand. *Applied Economics, 35*(2), 179-188.

Scuffham, P. A., & Langley (2002). A model of traffic crashes in New Zealand. *Accident Analysis and Prevention, 34*(5), 673-687.

Silverans, P., Drevet, M., & Vanlaar, W. (2005). *Attitudemeting verkeersveiligheid 2003-2004*. Brussel: Belgisch Instituut voor de verkeersveiligheid (BIVV).

Smeed, R. J. (1949). Some statistical aspects of road safety research. *Journal of the Royal Statistical Society A, 62*(Part I, series 4), 1-24.

Snyder, D. (1989). Speeding, coordination, and the 55-mph limit: Comment. *American Economic Review, 79*, 922-925.

Soukup, R., & Findley, D. (2000). *Detection and Modeling of Trading Day Effects.* Paper presented at the ICES.

Staten-Generaal voor de Verkeersveiligheid. (2001). http://www.wegcode.be.

Steenberghen, T., Wets, G., Thomas, I., Eckhardt, N., Casaer, F., & Wijnant, J. (2003). *Innovative Spatial Analysis Techniques for Traffic Safety*. Brussels, Belgium: Belgian Public Planning Service Science Policy.

Stewart, J. (1991). *Econometrics*. Cambridge, Great Britain: Philip Allan.

Tegnér, G., Holmberg, I. H., Loncar Lucassi, V. M., & Nilsson, C. (2000). The DRAG-Stockholm Model. In M. Gaudry & S. Lassarre (Eds.), *Structural Road*

Accident Models: The International DRAG Family (pp. 127-156). Oxford: Elsevier Science.

Tegnér, G., & Loncar Lucassi, V. M. (1997). *Time-series models for urban road traffic and accidents in Stockholm*: Paper presented at the Aalborg University, Denmark.

Theil, H. (1966). *Applied Economic Forecasting*. Amsterdam: North-Holland.

Toomath, J. B., & White, W. T. (1982). New Zealand survey of driver exposure to risk of accidents. *Accident Analysis and Prevention, 14*(5), 407-411.

Transport en Logistiek Nederland. (2002). *Licht op groen: visie van Transport en Logistiek Nederland op een effectief en efficiënt milieubeleid in het wegvervoer*. Zoetermeer, The Netherlands.

Van den Bossche, F., & Wets, G. (2003a). *Macro models in traffic safety and the DRAG family: literature review* (RA-2003-08). Diepenbeek, Belgium: Steunpunt Verkeersveiligheid.

Van den Bossche, F., & Wets, G. (2003b). *A structural road accident model for Belgium* (RA-2003-21). Diepenbeek, Belgium: Steunpunt verkeersveiligheid bij stijgende mobiliteit.

Van den Bossche, F., Wets, G., & Brijs, T. (2004). *A regression model with ARMA errors to investigate the frequency and severity of road traffic accidents*. Paper presented at the 83rd annual meeting of the Transportation Research Board, Washington, USA.

Van den Bossche, F., Wets, G., & Brijs, T. (2005a). Role of Exposure in Analysis of Road Accidents: A Belgian case study. *Transportation Research Record, 1908*, 96-103.

Van den Bossche, F., Wets, G., & Brijs, T. (2005b, October 3-5). *The use of travel survey data in road safety analysis*. Paper presented at the European Transport Conference (ETC) 2005, Strasbourg, France.

Van den Bossche, F., Wets, G., & Brijs, T. (2005c). The use of travel survey data in road safety analysis. *ETSC Yearbook 2005*, 64-75.

Van den Bossche, F., Wets, G., & Brijs, T. (2006, January 22-26). *Predicting road crashes using calendar data*. Paper presented at the 85th Annual Meeting of the Transportation Research Board, Washington D.C., USA.

Van Hout, K., Van den Bossche, F., & Daniels, S. (2004). *Data voor verkeersveiligheidsonderzoek in Vlaanderen* (SN-2004-03). Diepenbeek, Belgium: Steunpunt Verkeersveiligheid.

Van Schagen, I. (2000). *De verkeersonveiligheid in Nederland tot en met 1999: analyse van aard, omvang en ontwikkelingen* (D-2000-15): Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.

Verbeek, M. (2000). *A guide to modern econometrics*. West Sussex, England: John Wiley & Sons.

Wagenaar, A. C. (1983). *Alcohol. young drivers, and traffic accidents*. Lexington, MA: Lexington Books.

Wagenaar, A. C. (1984). The effects of macroeconomic conditions on the incidence of motor vehicle accidents. *Accident Analysis and Prevention, 16*(3), 191-205.

Warshawsky-Livne, L., & Shinar, D. (2002). Effects of uncertainty, transmission type, driver age and gender on brake reaction and movement time. *Journal of Safety Research, 33*, 117-128.

WHO. (2004). *World report on road traffic injury prevention*: World Health Organization.

Williams, A. F., & Lund, A. K. (1988). Mandatory seat belt use laws and occupant crash protection in the United States: presenting status and future prospects. In J. D. Graham (Ed.), *Preventing automobile injury* (pp. 51-89). Dover, MA: Auburn House.

Wolfe, A. C. (1982). The concept of exposure to the risk of a road traffic accident and an overview of exposure data collection methods. *Accident Analysis and Prevention, 14*(5), 337-340.

Young, A. (1965). *Estimating trading-day variation in monthly economic time series* (12). Washington D.C., USA: Bureau of the Census.

Zlatoper, T. J. (1984). Regression analysis of time series data on motor vehicle deaths in the United States. *Journal of Transport Economics and Policy, 18*, 263-274.

Zlatoper, T. J. (1989). Models explaining motor vehicle death rates in the United States. *Accident Analysis and Prevention, 21*, 125-154.

Zwerts, E., & Nuyts, E. (2004). *Onderzoek verplaatsingsgedrag Vlaanderen 2* (D/2004/3241/016). Diepenbeek, Belgium: Provinciale Hogeschool Limburg.