## Road safety in bits and pieces

For a better understanding of the development of the number of road fatalities


Henk Stipdonk

# Road safety in bits and pieces 

For a better understanding of the development of the number of road fatalities

# Road safety in bits and pieces 

For a better understanding of the development of the number of road fatalities

## Proefschrift

Ter verkrijging van de graad van doctor aan de Technische Universiteit Delft; op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben, voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
vrijdag 21 juni om 10.00 uur
door

Hendrik Leo STIPDONK
doctorandus in de natuurkunde
geboren te Leiden

Dit proefschrift is goedgekeurd door de promotoren
Prof. dr. B.J.M. Ale
Prof. dr. G.P. van Wee

Samenstelling promotiecommissie:

Rector Magnificus
Prof. dr. B.J.M. Ale
Prof. dr. G.P. van Wee
Prof. ir. F.C.M. Wegman
Prof. dr. G. Wets
Prof. dr. M.A.B. Postema
Prof. dr. ir. S.P. Hoogendoorn
Prof. dr. ir. J.H. van Schuppen
voorzitter
Technische Universiteit Delft, promotor
Technische Universiteit Delft, promotor
Technische Universiteit Delft
Universiteit Hasselt, België
Universiteit van Bergen, Noorwegen
Technische Universiteit Delft
Technische Universiteit Delft

Dit proefschrift is tot stand gekomen met steun van de Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.

## SWOV-Dissertatiereeks

Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV
Postbus 1090
2260 BB Leidschendam
E: info@swov.nl
I: www.swov.nl

ISBN: 978-90-73946-11-8
© 2013 Henk Stipdonk
Omslagillustratie: foto Jochem van Engers, naar een illustratie van Charlotte Dematons in "Nederland", © Lemniscaat.

Alle rechten zijn voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen of openbaar gemaakt op welke wijze dan ook zonder voorafgaande schriftelijke toestemming van de auteur

Voor mijn kinderen en kleinkinderen In de hoop dat ze altijd veilig zullen reizen

## Preface

Curiosity about the factors which explain the development in the annual number of road fatalities was the drive to write this thesis. This number has shown a strong increase in the Netherlands and many other countries since 1950 and a subsequent slow decrease since about 1972. Where does the peak in the number of fatalities about 1972 in many high income countries stem from? Why did the number of road fatalities in the Netherlands decrease by 20\% between 2003 and 2004? And although the annual number of fatalities seems to decrease gradually, there are also strong fluctuations superimposed on this gradual decrease. These fluctuations are too large and numerous to accept an explanation based on chance alone.

The road safety researcher wants to understand in order to be able to provide the best possible advice on road safety policy. Road crashes are a serious and notorious societal problem, which requires perpetual efforts to combat. Policies aim at reducing the future number of fatalities and therefore the researcher wants to forecast the effects of possible policies on the expected future number.

Since 1962 the Dutch scientific road safety research institute SWOV (Stichting Wetenschappelijk Onderzoek Verkeersveiligheid), has been running research to improve safety. One of the aims of the road safety assessment department of SWOV is to model road safety to help understand and forecast the number of fatalities or serious injuries. Such a model is to describe how the expected number of road casualties depends on the relevant factors in mathematical terms. Mathematical models are a common instrument to understand the phenomena surrounding us. In chemistry or physics, such models represent meaningful relations between quantities of interest. Think for example of Newton's gravitation law which describes how the gravitational attraction between two objects depends on their masses and the distance between them. Once masses and distance are known, the mutual attraction force can be calculated.

Newton's gravitation model is relatively straightforward. A deterministic model which describes the number of road fatalities, if it can be developed at all, would be extremely complex. Society is a collection of many people, all making their decisions in traffic, which are, of course, impossible to model. By making these decisions, people influence the safety level. This involves
factors such as the number of roundabouts or the amount of distance travelled under the influence of alcohol. While time goes by, these factors slowly (or perhaps suddenly) change, either because of specific policies or otherwise.

This thesis focuses on the role of distance travelled and its variations in the risk of a fatal crash. Distance travelled is widely understood as an important determinant of road safety. Driving, riding or walking along a road means that one is exposed to the risk of a road crash. On the other hand, road users induce risk to other road users when driving or riding. My goal was to find out how the influence of the amount and properties of distance travelled can be modelled to help understand and predict the development of the number of fatalities.

Can such a model be made? And if not, what may be a feasible compromise? At SWOV we explored the possibilities. Experts in the areas of statistical road safety, driving and riding behaviour, infrastructure and vehicle safety together tried to find ways to catch the developments in mathematical formulas. We made some small steps forward, some of which are described in this thesis. But we still have a long way to go and it isn't even certain that the way we calculate the expected number of fatalities now is actually better than the way we did this ten years ago. At least a practical justification of our approach is when policy makers have faith in the results.

Many have contributed to this thesis. First of all I want to mention Frits Bijleveld, who carried out a very large amount of analyses, mostly in Mathematica, SAS or R, resulting in a wide range of fits, parameters, forecasts and beautiful graphs like those presented in Chapter 4. Without his creativity and support this thesis could not have been written. I am also grateful to Ellen Berends, with whom I wrote Chapter 3, in spare time mostly (both hers and mine). After having written this chapter, the idea for this thesis was born. Martine Reurings was of great help to provide a logical structure to Chapter 6, and Yvette van Norden, Jacques Commandeur and Paul Wesemann have all contributed considerably, especially to Chapters 4 and 5. I am very grateful to Jochem van Engers for the professional design of the symbolic figures in Chapter 2, and the many hours spent to produce them. His designing skills were essential to explain my thoughts to the reader. On top of that, he made the photograph that adorns the front cover of this thesis. Finally, the many corrections to my English made by Marijke, my dear wife, have considerably improved the coherence of this thesis,
especially Chapters 1, 2 and 7 . She also created the title and the idea for the cover illustration of this thesis. Without her I would be lost.

Other colleagues at SWOV have been of great help to get me acquainted to the vast amount of knowledge, research skills and data sources available at SWOV. First of all, I must thank Niels Bos, whose skills are vital for SWOV. He helped me find my way in SWOV's web based databases. Thanks to Niels these databases contain less errors now than at the time SWOV received them, because he never believes the data on face value. Not only that, he also made several great visuals of the data, such as Figure 2.7. Vincent Kars has contributed considerably as well, because he stood at the basis of the system SWOV uses to make all these data available. Finally, Marijke Tros took care of the final lay out of the entire thesis, for which I am most grateful.

There are many more to thank. At SWOV no one can excel without the help of all others. Road safety is far too complex to work on one's own. Cooperation is essential. I thank SWOV for letting me join the team, and especially Fred Wegman for his stimulation to write this thesis and his many valuable suggestions to improve the text. The valuable suggestions of Ben Ale and Bert van Wee to merge the separate ingredients into a coherent thesis are without number.

I would like to express my gratitude to the Ministry of Infrastructure and Environment to enable SWOV to perform its scientific and independent research, and thus to have enabled me to write this thesis.

This thesis is a product of the efforts and sacrifices of many colleagues, friends and family members. I am grateful to all, as without their help, I wouldn't have been able to write this thesis. However, in case it contains inaccuracies, they are entirely my responsibility.

## Table of contents

1. Introduction and outline of this thesis ..... 15
1.1. Historical and societal framework ..... 15
1.2. Road safety policy in the Netherlands and globally ..... 16
1.3. Available models ..... 17
1.4. Focus of this thesis ..... 21
1.5. Synopsis of the remaining chapters ..... 23
2. The use of stratification in analysing and modelling road safety data27
2.1. Introduction ..... 27
2.2. Theory ..... 30
2.3. Method ..... 39
2.4. Data ..... 40
2.5. Some examples of stratifications of crash data ..... 45
2.6. Conclusion ..... 55
2.7. Discussion ..... 55
2.8. Some miscellaneous remarks ..... 58
3. Distinguishing traffic modes in analysing road safety development 61
3.1. Introduction ..... 61
3.2. Data (totals) ..... 64
3.3. General risk ..... 67
3.4. A general measure for mobility is not sufficient ..... 69
3.5. Data stratified by traffic mode ..... 70
3.6. Traffic mode risk ..... 74
3.7. The role of the other party ..... 76
3.8. Data stratified by fatality type ..... 77
3.9. Fatality type risk ..... 79
3.10. Conclusion and discussion ..... 83
Acknowledgements ..... 88
References ..... 88
4. Analysing the development of road safety using demographic data ..... 91
4.1. Introduction ..... 91
4.2. Theoretical background ..... 94
4.3. Method ..... 96
4.4. Results ..... 100
4.5. Discussion ..... 105
4.6. Conclusions ..... 109
Appendix. The penalty function for Eq. 6 ..... 110
References ..... 111
5. The expected number of road traffic casualties using stratified data113
5.1. Introduction ..... 113
5.2. Approach ..... 115
5.3. Development of the method, applied to Dutch data ..... 121
5.4. Further stratifications to be applied in the near future ..... 125
5.5. Discussion ..... 133
5.6. Conclusions ..... 140
References ..... 142
6. The effect on road safety of a modal shift from car to bicycle ..... 145
6.1. Introduction ..... 146
6.2. Method ..... 148
6.3. Data ..... 154
6.4. Results ..... 158
6.5. Discussion and conclusion ..... 160
Appendix ..... 164
References ..... 165
7. Towards an extended model for road safety ..... 167
7.1. What have we learned about the stratified approach? ..... 169
7.2. What would a road safety model look like which is both feasible and useful? ..... 173
7.3. Which steps would be necessary next to develop such a model? ..... 178
7.4. Which data are needed for a feasible road safety model? ..... 179
7.5. Which would be the revenues and the limits of such a road safety model? ..... 183
7.6. What makes a good model, and when can policy makers be satisfied with a road safety model? ..... 185
References ..... 189
Summary ..... 197
Samenvatting: De verkeersveiligheid, stukje bij beetje ..... 207
About the author ..... 217
SWOV-Dissertatiereeks ..... 219

## 1. Introduction and outline of this thesis

### 1.1. Historical and societal framework

Traffic and transport are crucial for society, without traffic everything stands still. Since many decades, traffic and transport have become significant parts of everyday life. In the Netherlands, adults travel on average approximately one hour daily (KiM, 2012), and the total fraction of the gross domestic product (GDP) spent on travel and transport amounts to approximately $15 \%$ (KiM, 2012). Unfortunately, traffic has its drawbacks. Air pollution, $\mathrm{CO}_{2}-$ production, noise and road crashes form a continuous threat to public health. These threats ask for effective policies to reduce the side effects of travel to an acceptable level.

Road crashes are a relevant cause of societal loss (O'Reilly et al., 1994), i.e. costs of road fatalities, casualties and material damage. To limit this loss, governments aim their road safety policy at effective and efficient improvement of road safety. To successfully reduce the annual number of casualties, policy makers have to choose effective approaches. Actual causes of road crashes have to be removed, possible health consequences of crashes for road users reduced. Hence, road safety policy gains from knowledge of the relations between road safety and the many relevant factors which influence road safety. In this thesis we shall explore the possibilities to analyse the available data in order to understand road safety developments in terms of changes in distance travelled and other factors.

Between 1950 and 1970 the use of passenger cars faced an exponential growth in many high income countries. In the Netherlands the number of cars and the total car distance travelled increased by almost $15 \%$ annually (see Figure 3.2). Hence, every year the increased amount of distance travelled was approximately $15 \%$ larger than the year before as well. Similar developments were seen in other high income countries (Oppe, 1991a). The number of road fatalities grew likewise. In parallel, road safety became an increasing societal problem. The annual number of fatalities increased by approximately 5\% annually (see Figure 3.1). This asked for an effective road safety policy.

### 1.2. Road safety policy in the Netherlands and globally

Both car industries and governments started to improve safety. Research institutes like SWOV were founded, and knowledge based measures were taken (SWOV, 2012) in the fields of rules and regulations, passive and active vehicle safety devices, improvements of road design etcetera. Data collection of crash data and data on distance travelled was started. After the introduction of computer technology and internet technology, data was made more and more easily available for research. In the Netherlands, especially much was done to supply road safety researchers with sufficient data. This is likely to have been one of the ingredients for a successful, knowledge based road safety policy. A long list of road safety measures have been taken since 1970, of which the introduction of "Sustainable Safety" (Duurzaam Veilig) was an important milestone (Weijermars and Wegman, 2010). The number of fatalities has been decreasing from more than 3264 police reported fatalities in 1972 to 546 in 2011. The actual number of fatalities, known since 1996, is higher, as the police reported data suffer from underreporting, which was $7 \%$ in 1996 and amounted to $17 \%$ in 2011. It is therefore reasonable to assume that the actual number of fatalities in 1972 could have been as high as 3500 . In many other high income countries a likewise increase and subsequent decrease in the number of fatalities was seen, although in countries with a low motorization, the number of fatalities has shown different patterns, e.g. show a summit in a much later year than 1972 (Yannis et al., 2011).

Road safety researchers aim to find out how road safety is related to the relevant factors which influence safety. These are factors like speeding, impaired driving (i.e. driving under the influence of drugs, medicines or alcohol), inexperienced drivers, road design, use of safety belts etcetera. The results of scientific research can be used by policy makers and vehicle manufacturers to implement an effective and efficient approach to improve safety. Hence, the focus of road safety research is to understand the relations between road safety, the factors which enhance or decrease road safety risk, and the effects of road safety measures. In the hypothetical case that all relations were perfectly understood, it would be possible to build a model containing all relevant variables which together determine the number of fatalities under any circumstances described by the variables.

Such a model would be of great help for road safety policy makers. It would enable them to determine what would occur in various scenarios of road
safety policies and developments of travelling behaviour, in such a way that an optimum choice could be made. Of course, this would require that the future development of the state of all relevant variables could be foreseen, as well as the expected effect of new policies. If this were possible, the effect of these developments on road safety, and the policy effects, could be calculated.

To build a model which describes these relations, however, is far beyond our possibilities. And even if such a model could be built, the use of it to forecast the number of road fatalities in a future year would require the forecast of the values of all relevant factors, such as travelling behaviour or safety measures applied in vehicles. Further, policy responsibility is distributed over many institutions and organisations like the government, provinces, municipalities, police, etcetera. This makes it difficult to get a complete picture of the expected future state of traffic. For example, there is currently no central information available describing the state of the road network in the Netherlands. There is no database which contains data on a national level describing the speed limits (and changes in them) of the Dutch roads, their width, curvatures, whether there is road lighting, physical separation of opposite lanes etcetera. Hence, if we want to build such a model, we have to take this into account and focus on small and feasible steps.

### 1.3. Available models

Most models which are currently in use consider traffic as a source of risk. Persons who make a trip are exposed to this risk. This exposure is expressed by means of a single quantity: distance travelled. Usually this is total distance travelled (Bijleveld, 2008), sometimes replaced by distance travelled by motorized vehicles and fuel consumption (Fournier and Simard, 2002). Models without distance travelled are also developed, in case no data on distance travelled are available (Van den Bossche et al., 2007). In many models, the number of fatalities is usually calculated as a product of risk, and the corresponding distance travelled. This risk is assumed to decrease over time, e.g. according to a negative exponential function, as a general result of a general learning process of society, resulting in improved safety factors due to implemented measures, improved driver experience etcetera (Oppe, 1991). Extrapolation of this function to a future year, together with an assumed value of the amount of travel in that future year, provide an estimate of the expected number of fatalities in that year.

The data of the expected future distance travelled is obtained from an external source such as forecasts based on economical and societal models, because it is not considered to be a part of the road safety analysis. Road safety analysis focuses on the development of risk, and on ways to improve risk.

Models with few or no explaining factors and a simple measure for distance travelled have some important advantages. They require little knowledge of the future circumstances (only the future distance travelled is required). Further, the forecasts of risk follow from a straightforward projection. This approach works to a certain extent: the average risk, i.e. the total number of fatalities divided by total distance travelled, was shown to decrease since 1950 (Oppe, 1991b), as is also explained in Chapter 3 of this thesis. Several authors have developed models to describe or forecast road safety fatality time series (Oppe, 1991b, Broughton et al., 2000, Maycock, 2001, Bijleveld, 2008, Eksler, 2009).

Some authors have added stratification to this approach. Broughton (1991) analysed the development of average risk, using a single general safety measure, namely the introduction of the compulsory seat belt in 1983. In 2000 he went a step further and developed a model with 10 subgroups for fatalities, namely 5 types of road users (car occupants, pedestrians, cyclists, motorcyclists and other) and two road types (rural and urban roads). However, data on corresponding distance travelled was not available, hence instead Broughton used total motorcycle distance travelled for motorcycle fatalities, total car distance travelled for car fatalities and total distance travelled for the other modes. Measures were supposed to have a proportional effect to the total number of fatalities. Maycock (2001) made a model for the expected number of $60^{+}$-year old drivers, based on demographic data and a prediction of the number of drivers in future years. In this model, distance travelled was not taken into account. In Gaudry's DRAG-models (Gaudry, 1984, Fournier and Simard, 2002) the total number of fatalities is modelled with a long list of variables, to model distance travelled and fatalities simultaneously. Similar approaches where chosen by researchers in Sweden, France and Belgium (Fridstrøm et al., 1995, Jaeger and Lassarre, 2002, Van den Bossche et al., 2007).

In these models care has been given to solve the problems arising from dependencies between the numbers of fatalities in consecutive years. Improvement of the statistical methods received considerable attention,
resulting in a new state of the art approach, based on state space techniques (Van den Bossche et al., 2007, Commandeur and Koopman, 2007, Bijleveld, 2008). A conceptual difficulty associated with these models is that the influence of specific factors that are included in these models (e.g. motorway speed limits) influence a subset of fatalities only (e.g. motorway crashes only). An approach based on modelling the total number of fatalities does not use the information of variations in the subsets of these totals. Hence it is difficult to correlate these variations to external factors through the fluctuations in the aggregate data.

At the same time, models were developed to relate the probability of a fatality of a specific subtype, such as alcohol crashes, to the development of a specific risk enhancing factor (such as the Blood Alcohol Concentration in the blood of a road user). The general method for this approach was described by Sigrist (2010). However accurate such risk factor models may be, in order to understand the development of the number of fatal crashes of that specific subtype, one needs to know the development of the risk enhancing factor. In the above mentioned example, one needs sufficient information of the distance travelled by impaired drivers, to explain the development of the number of fatal crashes with alcohol impaired drivers.

Most literature on road safety models concerns curve fitting, or forecasting by risk extrapolation, using a relatively simple descriptive general expression for the development of risk as a function of time. Models which explain the road safety developments in the past are extremely difficult, if not impossible to build. Such models should contain all relevant factors as variables, together with the parameters which describe the importance of each factor with respect to the number of fatalities. They should present the exact mathematical form according to which the number of fatalities depends on each specific variable. This is virtually impossible. One of the most important sets of factors is the amount of distance travelled (i.e. mobility, travel). If this factor is not incorporated in the model correctly, and known with sufficient accuracy, an eventual effect of a road safety measure can remain unnoticed because of a compensating effect in distance travelled. The other way around is also possible: that an ineffective road safety measure may seem effective because of an unnoticed decrease in distance travelled. A simple example of the latter is illustrated in Figure 1.1. The figure shows the number of fatalities among moped riders since 1950 in the Netherlands, together with the moped fleet size. The introduction of the mandatory moped helmet in February 1974 coincides with a strong drop in the number of fatalities. A superficial
observer might associate this drop with the safety effect of the helmet itself. However, the contemporary drop in the moped fleet size suggests that a part of the fatality decrease can be associated with the decreased distance travelled by moped.


Figure 1.1. Number of police registered moped fatalities (open circles), actual number of moped fatalities (light grey circles), and moped fleet size (dark grey circles, measured on the first of January each year) in the Netherlands. Helmet use was made obligatory in February 1974.

Significant changes in distance travelled for specific travel modes can occur within a short period. This may have important consequences for the number of fatalities in that period. Hence, any model describing past road fatality numbers should preferably contain the amount of travel for all relevant travel modes in a proper form. This is easier said than done. The travel data required are not always available with the necessary accuracy. For example, the Dutch mobility data are collected by means of an annual household survey. The number of motorcycle trips and car trips in the survey is proportional to the number of people in the survey who report to have made a motorcycle trip or car trip. As this number (and the distance travelled) is approximately 100 times less for motorcycles than for passenger cars, the accuracy of the resulting survey sample is also much less. Consequently, stratification of such data (e.g. by rider age) is hardly feasible for motorcycle mobility data. Furthermore, in many countries data on distance travelled by
different travel modes is not available at all (Lejeune et al., 2007). It is therefore not surprising that analysts of road fatality data confined to the more widely available data on total distance travelled and total mobility, and their quotient, the general risk.

### 1.4. Focus of this thesis

In this thesis the traditional separation of risk and exposure to risk is replaced by a more elaborate approach. Traffic is not only considered as exposure to risk, but also as a source of risk. Hence, in this thesis we will not speak of exposure when we denote distance travelled. A distinction is made between single-vehicle crashes and crashes in which two travellers are involved who crash into one another. For single-vehicle crashes of any travel mode such as moped or passenger car, the distance travelled of mopeds or passenger cars can indeed be seen as an exposure to the risk of a singlevehicle (moped or passenger car) crash. However, for two vehicle crashes such as between a moped and a passenger car, the distance travelled of the passenger cars contributes to the risk of the moped riders, and vice versa. For passenger car drivers, the probability to be involved in a crash with a moped, is a function of moped distance travelled as well.

For road safety policy makers, it is important to know how the number of current or future road casualties relates to external factors such as distance travelled or safety measures. When the number of fatalities is increasing they want to know why. If it is decreasing, they want to know if this decrease will last. Essentially, the answer to these questions requires knowledge of the homogeneity of the collection of crashes. Are all types of crashes increasing or decreasing at the same pace, or are there specific groups that are increasing or decreasing faster than others, and can these changes be explained?

This thesis explores the possibilities to understand the development of the number of road fatalities by bringing Broughton's approach a step further. The approach (Broughton, 1991) is to research the possibilities of a stratified model in which the relevant mobility of the strata (subgroups) is taken into account. This can be the mobility of a single travel mode for single-vehicle crashes, or the mobility of two travel modes for two-vehicle crashes. We will not consider crashes with more than two travel modes involved. The various chapters explore these possibilities in different directions. In Chapters 3 and 6 only a subset of all crashes is taken into account, whereas in Chapters 4 and

5 attempts are made to describe the development of the total number of fatalities or serious injuries by means of well-chosen strata. Whenever feasible, relevant data on distance travelled are taken into account. However, for some subgroups this is not possible, for example because data on distance travelled are not available for all travel modes.

This thesis aims to better understand the relations between road safety and distance travelled, and the properties of distance travelled. The ultimate goal is to explain the number of fatalities in terms of the relevant factors, instead of describing the number of fatalities as a time series, i.e. as a function of time. Probably this goal is only theoretically achievable. Hence, the number of fatalities is no longer described as a function of time, but as a function of relevant factors, i.e. the properties of distance travelled that determine the probability of a fatal crash. Consequently, these factors have to be known as a function of time to determine the expected number of fatalities for past or future years. If it were possible to build such a model, estimation of the expected number of fatalities in a future year would still require the expected values of the external factors in that future year.

The annual number of road casualties is an addition of many different types of crashes and casualties. The development of subsets of the total may show different trends. These differences are associated with different developments of distance travelled or safety measures which apply to these subgroups. To understand and predict the development of the total number, the analysis of the developments of subgroups is useful. The approach in this thesis is based on stratification of data, both of fatalities and on distance travelled. In two vehicle crashes distance travelled of both modes involved is taken into account when possible. The approach has its limits:

1. Stratification is only possible under the condition that both crash data and data on distance travelled can be stratified simultaneously. Hence, there is no stratification by road type, because road type is not available for the data on distance travelled.
2. Stratification implies smaller numbers of fatalities in each subgroup as compared to the total. Therefore there is relatively more statistical noise in the number of fatal crashes.
3. Treatment of two vehicle crashes by taking distance travelled of both modes into account, quickly complicates the mathematics involved. Especially if further stratification by driver age is applied, as there are two drivers involved, the model becomes mathematically complex to a degree that is hardly manageable.

This thesis is meant to explore the possibilities, notwithstanding these limits. We shall try to find a way to improve the understanding of the development in the number of fatalities by making use of the possibilities available in the data sources. By doing so, we encounter a fundamental dilemma in this field. More stratification enables more understanding, but more stratification introduces more statistical uncertainty as well. On the other hand, aggregation simplifies statistical analysis, but lessens the distinctions we can make in the analysis. The challenge in this field of research is to find a balance between the two. One could say that we have to choose between estimation of a sophisticated model with inaccurate parameters and an inappropriate model with accurate parameters. In this thesis, we shall try to find a more sophisticated model, while we care somewhat less about the accuracy of the model. Eventually, the optimum might be found.

Stratification is also useful for policy makers. Road safety policies are usually distributed over different subjects like vulnerable road users, impaired driving, enforcement, vehicle safety etcetera. The impact of measures in one or more of these fields is usually limited to just a specific subset of all fatalities. If a policy maker wants to know the effect of some planned safety measure in a future year, it is desired to have a stratified forecast available in which the specific age group, travel mode or other subset is forecast separately. These subsets may have mutual interaction. Therefore a stratified forecast with multiple simultaneous stratifications may be of value. For example elderly car occupants may be killed in a crash with a younger car driver in the other car. Hence, measures that improve the safe driving behaviour of young car drivers also improves the crash risk of elderly car drivers and other road users.

### 1.5. Synopsis of the remaining chapters

Chapter 2 describes the approach used in this research. Theory, method and data that are used to analyse the development of the annual number of road fatalities in this thesis, are described. Some examples of stratifications are given to illustrate differences in the development of the annual number of fatalities over time.

Chapter 3 shows how the development of aggregated risk differs from the development of risk for six specific subgroups. It is shown that the quantity of "aggregated" risk, i.e. the total annual number of fatalities divided by the total annual distance travelled, is a continuously decreasing function of time
for the entire period between 1950 and 2010. Stratification by traffic mode shows a different development of risk for different strata.

Chapter 4 explores the possibilities to use demographic data to improve the analyses or forecasts of the annual number of road traffic casualties. The possible added value of demographic data for such analyses is shown. Analyses of distance travelled, with and without the use of population data, are compared. In addition, the consequences for the evaluation of risk, i.e. casualties per distance travelled, with and without the use of population data, are explored. Dutch data are used to illustrate the model concept.

Chapter 5 describes an approach to estimate the annual number of fatalities in a future year, using stratified data. We calculate casualty rates from casualty data and data on distance travelled, which is extrapolated and subsequently multiplied by an expected future distance travelled. After correction for separately assessed effects of additional safety measures, the number of casualties is estimated. In this chapter, data on both killed and seriously injured (KSI) crash victims are used.

Chapter 6 describes and applies a method to assess the effect on road safety of a modal shift from car to bicycle. $10 \%$ of all car trips shorter than 7.5 km are assumed to be replaced by bicycle trips. The safety of car occupants and cyclists are taken into account as well as the safety of other road users involved in crashes with either cars or bicycles. The computations are carried out by age and gender of the bicycle rider or car driver. The results are expressed as the expected increase or decrease in the number of casualties, stratified by crash severity (fatalities/serious injuries), age and gender.

Chapter 7 concludes this thesis by describing the findings in this thesis, and it contains suggestions for a further development of road safety models, for both understanding past developments and road safety forecasts. Suggestions for data collections are given, as well as the possibilities and impossibilities of the development of an ideal model. The balance between the desire to stratify the data for conceptual and policy reasons on the one hand, and the practical (im)possibilities on the other hand, is pointed out. To be precise, we will answer the following five questions:

1. What have we learned about the stratified approach?
2. What would a road safety model look like which is both feasible and useful?
3. Which steps would be necessary next to develop such a model?
4. Which data are needed for a feasible road safety model?
5. Which would be the revenues of such a model?

The chapter concludes with a vision on a feasible road safety model.

## 2. The use of stratification in analysing and modelling road safety data


#### Abstract

This chapter describes the reasoning behind the approach in this thesis, of how to understand the development of the number of road casualties. We introduce this approach, based on stratification of crash data and other relevant data. The theory behind the approach is described briefly. The core is that the probability of a crash is related to the presence of a traveller on a public road; hence, the annual number of fatalities is related to the annual distance travelled on public roads. The circumstances during the many different trips vary, and these influence the probability of a crash. The methods used in this thesis to understand the development of road safety data is described. Essentially two methods are used: crash data are related to distance travelled and stratification of both crash data and data on distance travelled is applied. Stratification is carried out by driver age and travel mode mostly. By doing this, the additional information for individual crashes, namely the travel modes involved in a crash, and the age of the drivers involved, is used in the analysis. The data, used in this approach, are described. As in this area of research, data cannot be derived from experiments, the possibilities of research on the national annual number of fatalities is to a high extent depending on availability of data. Data collection quality of crash data and data on distance travelled dominate the maximum achievable accuracy of any analysis. In the Netherlands, both crash data and data on distance travelled have experienced several changes in accuracy over the past 60 years. These developments are described. Further, some examples are given of data stratifications, illustrating the relevance of the approach.


### 2.1. Introduction

Road safety observation starts by counting the number of fatalities, or the number of major or minor injuries over a certain period of time. Here, a fatality is counted if the crash occurred on a public road, and the casualty died within 30 days as a consequence of the crash. In the past, fatalities were considered the more important than serious injuries. As a consequence data on fatalities are usually registered more accurately than data on serious injuries (Derriks and Mak, 2007). In addition, fatalities are uniformly defined internationally. For (serious) injuries this is not yet the case, although the
international interest in injuries is increasing (OECD/ITF, 2012). In the Netherlands, reducing road fatalities has been the main focus of road safety policy makers. Since 2009 a second focus has become more important, i.e. reducing seriously injured road crash casualties, which is justified by its increasing number since 2006. A serious injury is counted if the person is treated in hospital as inpatient, with an injury level of at least an Abbreviated Injury Score of 2 . The burden of disease of seriously injured is comparable to that of the fatalities in the Netherlands (Polinder et al., 2012). However, in this thesis we shall focus on road fatalities.


Figure 2.1. Police registered annual number of fatalities in the Netherlands between 1950 and 2010, depicted against the Matterhorn (SWOV/VOR/BRON, 2012).

In many high income countries, a graph of the development of the annual number of fatalities between 1950 and 2010 appears as a mountain of fatalities, having its summit near 1972. This is illustrated in Figure 2.1 (Netherlands) and Figure 2.2 (several high income countries), based on data available on the SWOV website (SWOV/IRTAD, 2012), and in a database provided by Gaudry (Agora Jules Dupuit, 1999). One might ask how this development can be explained from the development of relevant factors in society, such as distance travelled, modal split of travel, changing circumstances, safer vehicles and infrastructure, and other safety measures. About 2003 and 2004, some developed countries such as Denmark, the Netherlands, Switzerland and France benefited from a sudden temporarily stronger decrease in the annual number of fatalities (Figure 2.2), which again asks for an explanation based on external factors. This is especially relevant
because policy makers should found their decision as to which approach to choose in improving safety on facts about the effects of their decisions. Hence, it is desired that in such assessments, these effects can be successfully separated from other developments.


Figure 2.2. Annual number of fatalities in a number of countries between 1965 and 2010 (SWOV/IRTAD, 2012, Agora Jules Dupuit, 1999).

But how to research these phenomena once they have occurred? Common research methods to try to analyse the available data on the road safety development of the past are unavailable. It is not possible to perform a casecontrol study, to simulate the entire traffic situation in an entire country, or to repeat past years while varying external factors. Instead, we may try to dig into the data and see what can be done to correlate external factors with the number of crashes or casualties. Such data are not always easy to get by, especially data from years before the introduction of the personal computer. And even today it is not easy to find out all relevant developments in the road traffic system. In this chapter we describe several methods to find out more about the relation between road safety and external factors, based on Dutch data that are readily available (SWOV, 2012).

The analyses in this thesis are based on the presumption that the frequency of road safety fatalities in a country is the outcome of a statistical process, as a consequence of intrinsic unsafety in the traffic process. This intrinsic
unsafety, and its relation with the annual number of fatalities in a country, is described in Section 2.2, where we present the probability theory behind road safety. In Section 2.3 the methods used in this thesis to understand the development of the number of fatalities are described. Section 2.4 treats the relevant data used. Section 2.5 presents some examples of time series, stratified in several ways, to illustrate how different subgroups of fatalities show a different development.

### 2.2. Theory

When a person (a traveller) travels, at every moment during the trip there is a small, unknown probability of a fatal crash. This probability depends on the size of the location, the duration of the moment and on many factors that describe the characteristics of the trip, the location, etcetera. These are factors such as road design, weather, vehicle properties, specific human factors and the presence of other road users. Specific human factors are factors such as fitness to drive, which in turn is influenced by e.g. age, gender, experience, sleepiness, blood alcohol concentration (BAC), and characteristics of the person such as emotions (Mesken, 2006). Not only is the driver himself under the influence of these factors, but also his fellow road users are. As a fatal crash is defined on public roads only, the domain where accidents can occur is essentially one dimensional, namely the total length of all public roads.


Figure 2.3. Left panel: Picture of an elementary traffic cell, or ETC: a situation on a length of infrastructure located between $R$ and $R+\mathrm{d} R$, during a time interval between $t$ and $t+\mathrm{d} t$; middle panel: the arrow represents the presence of a trajectory in this ETC; right panel: an ETC with two trajectories.

Figure 2.3, left panel, represents one of the possible moments on a length of road at a location between $R$ and $R+\mathrm{d} R$, and time between $t$ and $t+\mathrm{d} t$. We will call this time-space interval an elementary traffic cell or ETC. A fatal crash is possible in this ETC only when there is traffic present. This presence depends
on the decisions of people to make trips. Now, when a person (a traveller) makes a trip, he follows a trajectory, i.e. he has a time-dependent location $r(t)$. If an ETC contains only one trajectory (Figure 2.3, middle panel) it is able to contribute to the number of single-vehicle fatalities. For a crash in which more than one vehicle is involved, we need an ETC with two trajectories (Figure 2.3, right panel).

For each ETC which contains a length of trajectory, we can define a probability of a crash. This probability depends on the size of the ETC, and of its properties (road width, time, alertness, etcetera etcetera). To make sure that these properties are constant inside the ETC, we choose the ETC to be infinitely small, i.e. $\Delta r \rightarrow \mathrm{~d} r$ and $\Delta t \rightarrow \mathrm{~d} t$. As for small ETC's, the probability of a fatality is proportional to the size of the ETC, the probability of a crash in the ETC will become infinitely small as well. Therefore, for sufficiently small ETC's the probability of a fatality inside an ETC transforms into a fatality density frequency, i.e. an expected number of fatalities per unit of time $\mathrm{d} t$ and per unit of road length $\mathrm{d} r$, denoted as $\Lambda(r, t)$. The expected number of fatalities inside the ETC then becomes $\Lambda(r, t) \mathrm{d} r \mathrm{~d} t$. Here, $\Lambda(r, t)$ depends on the external factors, human factors and properties of the vehicle chosen by the traveller. These factors, such as travel mode or road type are numerous, and correlated in a complex way. Figure 2.4 symbolizes the different factors, and their interconnectivity (Figure 2.4, upper panel). At a closer look many of the symbolized factors mentioned here turn out to be groups of factors actually (Figure 2.4, middle panel): the factor vehicle, for example, not only denotes vehicle type, but also further characteristics such as mass or maximum speed, which all contribute to the fatality density frequency $\Lambda$.

The factors depend on each other in many ways. Not all vehicles can drive on all roads, drivers have to be of a certain minimum age and licensed to drive a car, enforcement is not permanently present everywhere etcetera. Mark that for each ETC, all factors have a single value for that specific ETC, from a collection of many possible values. For example: the factor vehicle type might take the value moped and thus contribute more to $\Lambda$ than if this factor would have had the value passenger car. Further, $\Lambda$ of a moped rider is influenced by its speed, helmet wearing etcetera. These values each contribute to $\Lambda$ inside this ETC, in such a way that it is influenced by all factors. This is illustrated in the lower panel of Figure 2.4. Thus, the outcome of the contributions to $\Lambda$ depends on the values of all relevant factors, represented as the shaded area in the figure. This area can be interpreted as being proportional to $\Lambda$.


Figure 2.4. Symbolic illustration of the different factors involved in the probability of a crash, and their interactions (upper panel), the contributions of each factor to the probability of a fatal crash, depending on the actual value of the factor (middle panel) and the composition of the resulting probability of a crash (the grey area), depending on the contributions of all factors (lower panel).

Each ETC with a trajectory can be characterized by such a composite set of factors, together constructing the probability of a single-vehicle fatality in that ETC (Figure 2.5, left panel). For two vehicle crashes, the situation is even more complex: the probability of a fatality in vehicle 1 does in that case not only depend on the factors defined by traveller 1, but also by the characteristics of traveller 2 and vice versa. E.g. in a crash between two cars, the a priori probability to die is highest in the car with the lower mass. And any car driver is more likely to get killed in a crash with a young driver than in a crash with an elderly driver. Hence, the actual $\Lambda$ of a fatality not only depends on the properties of the traveller and his vehicle, location etcetera, but also on the probability of the presence of any other traveller and the properties he carries with him. This is symbolized in Figure 2.5, right panel.


Figure 2.5. Symbolic illustration of the different factors involved in the probability of a crash in an ETC. The left panel represents an ETC with one trajectory and the probability of a single-vehicle fatal crash, related to the values of the relevant factors. The right panel represents the probability of a two vehicle fatality.

The illustration of Figures 2.4 and 2.5 should by no means be interpreted as an accurate description of the probability of a fatal crash. If there are two travellers present in an ETC, these travellers mutually influence each other's probability of a crash. Further, a truck driver may have a low probability to be involved in a fatal crash as a victim, but on the other hand the probability that someone else is killed in a crash with him involved as a truck driver is much higher. The figure is meant as a conceptual illustration of the contributions to crash risk.

In each ETC, $\Lambda$ is small, and moreover it cannot be observed directly. Only the fact that crashes occur, or do not occur, can reveal some information about this probability. We know that crashes are more likely to occur when factors that enhance $\Lambda$ are present, such as a young inexperienced male
driver, a high BAC, an ill designed road intersection or a powered twowheeler. We know that crashes are less likely, but still possible when no such factors are available. Hence it is likely that the probability of a crash in the majority of the ETC's where a crash occurs is higher than in the majority of the ETC's where no crash occurred. However, to make a scientific statement about that probability, we need a preferably large number of fatalities. This implies we will have to collect many ETC's, e.g. by observing a longer stretch of road or considering a longer interval of time. This way, we accept that, as each ETC has its unique set of circumstances, the final set of ETC's in a certain country and in a certain year is very heterogeneous. Hence, expressing the unsafety in a country by its annual number of fatalities is a gross simplification. This number itself disregards the variety in types of fatalities. On the other hand, outlining national road safety policy requires some sort of aggregation of fatality data, if only to express the level of safety in a country and its progress over time. This progress over time, presented as a time series of the annual number of fatalities is what we want to understand: why the number of fatalities increases in one period, and decreases in another. This requires that we understand how the annual number of fatalities is a result of the crash probabilities during all trips, and how these probabilities are influenced by the relevant factors.

The probability of a crash in the ETC is infinitesimally low, because an ETC corresponds to an infinitesimally small time interval and road length interval. We need both to wait a while, e.g. a second or a minute, and consider a longer stretch of road, e.g. a road segment, to find a finite probability of a fatality. Then, still, this is a very small probability. If we want to apply statistics to sufficiently large numbers of fatalities, we need to wait much longer, e.g. a year, and we must look at a much longer stretch of road, e.g. all roads of a country. Figure 2.6, upper left panel represents all ETC's of all roads in a country, in a period of an entire year. Naturally, each separate cell has its own set of properties; hence the probabilities of a crash in all cells are, in principle, very different. At any time and location, these probabilities may materialize in a fatality (or more than one). Fortunately, they usually don't. Whether or not a fatal crash occurs is a matter of chance. All probabilities in all ETC's together give rise to a number of fatal crashes. Figure 2.6, upper right panel, pictures the realization of these fatalities in an entire country, as a collection of the realizations of all crashes on all roads in an entire year. Here, we imagine all roads of a country as a collection with a total length equal to the sum of the lengths of all roads.


Figure 2.6. Representation of all ETC's on all roads in an entire year (upper left panel), and the realizations of fatal crashes in this road-time space (upper right panel). The collection of all ETC's with their individual probabilities give rise to a probability distribution of the expected number of fatalities (lower left panel), whereas the actual number of fatalities is just one of the theoretically possible outcomes (lower right panel).

Road fatalities present themselves as a statistical process. The expected total number of fatalities in a country in a certain year, which we denote as the frequency $v$, is in a very complex and unknown way related to the probabilities $\Lambda$ of a road fatality in each ETC during that year of all locations on public roads in that country. The total number of fatalities probably is a collection of which some are the result of very rare and very dangerous situations (few ETC with high $\Lambda$ ), whereas others are the result of very common and not so dangerous situations (many ETC with low $\Lambda$ ).

All possible observed realizations of these $\Lambda(r, t)$ with aggregate frequency $v$, are statistically distributed around the most likely value of $v$ (Figure 2.6, lower left panel). This distribution is, in principle, also unknown; it depends on the properties of the entire traffic system. In fact, it depends on the individual probabilities $\Lambda$ of a fatal crash in each ETC. Thus, the actually
observed number of fatalities $N$ in that country and in that year is just one of those possible realizations. We observe $N$, but this value is the outcome of statistical noise around an underlying expected number $v$. Figure 2.6, lower right panel illustrates how the realization of $N$ fatalities, which in reality has one value only (the actual value), could statistically have been some other value, if chance would have let other ETC's materialize their probability into an actual fatality.

The probability of a fatality in an ETC can only be nonzero if there is a trajectory $r(t)$ present. Empty ETC's do not contribute to the number of crashes. In fact, even for an ETC containing a trajectory $r(t)$, outside $r(t)$ the ETC does not contribute to the number of crashes. Hence, the fatality density frequency $\Lambda(r, t)$, can be interpreted as either a fatality density or a fatality frequency along the trajectory, i.e. the probability of a fatality can be defined either per unit of time or per unit of trip length. The relation between $r$ and $t$ of the trajectory $r(t)$ reduces the dimensionality of $\Lambda(r, t)$ to either $\lambda(t)$ or $\lambda(r(t))$. Hence, the probability of a fatality inside an ETC, $\Lambda(r, t) \mathrm{d} r \mathrm{~d} t$, simplifies to either $\lambda(t) \mathrm{d} t$ or $\lambda(r(t)) \mathrm{d} r$. We will choose the latter, for reasons of convenience: $\lambda=\lambda(r(t))$. Fatalities per unit of time would require information about the speed along each trajectory. This information is not available, neither for crash data nor for data on distance travelled.
$\lambda(r(t))$ depends on the values of many external factors. We will denote this set of factors as $\Phi$, and thus $\lambda=\lambda(r(t), \Phi)$. This $\lambda$ can be interpreted as the fatality density present at $(r, t)$, given that there is a trajectory $r(t)$, the properties of which influencing the value of $\lambda$. This implies that risk and distance travelled are inextricably linked: the risk factors are determined by the properties of distance travelled. Thus, if it were possible to know $\lambda$ for all trajectories in a year, it would be a matter of integrating $\lambda$ over the total distance travelled, i.e. the length of all trajectories to find the expected number of fatalities. This is true only if all crashes are independent, and if $\lambda$ is sufficiently low to neglect the fact that people cannot be a fatality more than once.

The fatality density $\lambda$, or (fatality) risk, is calculated as the number of fatalities divided by distance travelled, and it has the dimension of fatalities $/ \mathrm{km}$. An estimation of the average fatality density for all trips in a country during an entire year can be calculated as the quotient of the annual number of fatalities and the annual distance travelled. This average is in itself not necessarily a very meaningful quantity, as the local and temporal
differences of $\lambda$ may be very important. If, for example, a small fraction of all distance travelled is responsible for the majority of all crashes, while the majority of all distance travelled is perfectly safe, the denominator and the numerator would have no relation whatsoever, in the quotient of the number of fatalities and the total distance travelled. For example, suppose there are 50 single-vehicle moped fatalities, corresponding to a negligible moped distance travelled, and 0 car fatalities, corresponding to $100 \times 10^{9}$ (car) km . In that case the total number of fatalities would be $50+0=50$, and the total distance travelled would be $\left(100 \times 10^{9}+0\right) \mathrm{km}=100 \times 10^{9} \mathrm{~km}$. Numerator (50 moped fatalities) and denominator ( $100 \times 10^{9} \mathrm{car} \mathrm{km}$ ) of the average risk calculation have no physical relation to one another. A stratified calculation for moped riders and cars would reveal an extremely high moped risk, and a negligible car risk. The average risk is meaningless.

The concept of a stratified $\lambda$, the variation of which depending on relevant factors, fundamentally offers the possibility of improved insight in the unsafety of the traffic process.

In this thesis we use a straightforward, although complex, approximation of the influence of distance travelled on the number of casualties. We use the following reasoning:

- Travelling is done by persons, who make trips, with a certain purpose, at a certain time.
- A trip is considered to be described as a trajectory, fixing the location $r$ as a function of time $t$. A trip is a sequence of infinitely many ETC's. It has relevant properties in time and space, such as the time of day, date, origin and destination, the choice of the route (with its types of roads, intersections). These properties vary along the trip. They are supposed to be constant for a single ETC.
- The route is a sequence of roads, the existence and quality of which influences the route choice. Road design might or might not be optimized to add to the safety for the road users.
- Road design follows land use planning. As roads last for tens of years, the amount of attention to road safety in land use planning considerably influences road safety.
- The persons who make the trips have several relevant properties, such as age, gender, skills and personal characteristics such as impatience, driving experience or fatigue.
- The trip relates to environmental circumstances such as weather (precipitation, temperature) and traffic (traffic volume, speeds) circumstances.
- The trip is made with a travel mode (including walking), with its own properties such as top speed, stability, active safety properties and mass.
- The use of the vehicle is subject to legislation regarding the vehicle (safety belts, helmets), and its use by the person, on the road of the trajectory. The legislation is supported by enforcement.
- Persons, roads, vehicles, circumstances and legislation form a complex dynamic system of interrelations, together making up the safety of the travel, expressed as $\lambda=\lambda(r(t), \Phi)$, where $\lambda$ is a fatality density: a probability of a fatality per unit of trajectory length. $\lambda$ depends on the trajectory $r(t)$, and on all possible external factors $\Phi$. Here, $\Phi$ depends on the trajectory $r(t)$ as well, and includes the properties of the road (at that time), the persons involved (present at that $r(t))$ etcetera. Some of these factors $\Phi$ are also known as Strategic Performance Indicators (SPI's), such as alcohol use by drivers, road quality, enforcement etcetera.
- The fatality density $\lambda(r(t))$ is defined alongside the trajectory and during the trip only. Away from the trajectories, the probability of a fatality $=0$. It is interpreted as a number of fatalities per unit of trip length. Then, the expected number of fatalities on that road and in that time interval would equal $\lambda(r(t)) \mathrm{d} r(\mathrm{t})$. The expected annual number of crashes in a country $v$ is an aggregation over all trips $r(t)$ within a road domain (a country) and a time domain (e.g. a year) of $\lambda(r(t))$ :

$$
v(\text { year, country })=\int_{r} \lambda(r(t)) \mathrm{d} r(t)
$$

Here, all ETC's are supposed to represent independent possibilities of a fatality.

- The expected annual number of fatalities $v$ is a complex function of all factors mentioned. Hence, the crash probabilities are very different for different ETC's, However, for a set of $s$ ETC's with equal crash probability $\lambda$, and if we can assume that there can occur at most one fatal crash in an ETC, i.e. the possibility of a crash doesn't depend on the occurrence of other crashes, we expect the number of fatal crashes $N$ to be Poisson distributed. This is explained in standard textbooks, see e.g. chapter 15 of Taha (1987). The Poisson distribution is of the following form:

$$
P(N \mid s \lambda)=(s \lambda)^{N} e^{-s \lambda} / N!(N=0,1,2, \ldots) .
$$

- It can be shown that the total number of fatal crashes, resulting from all ETC's with different values of $\lambda$ also results in a Poisson distribution, with parameter $v$ :

$$
P(N \mid v)=v^{N} \mathrm{e}^{-v} / N!(N=0,1,2, \ldots) .
$$

Nicholson and Wong (1993) showed that in fact, the actual statistical distribution of the number of road crashes does not contradict this assumption.

- For single-vehicle crashes, the factors that influence $\lambda(r(t))$ depend on external circumstances, and properties of one vehicle and its driver, and perhaps on the properties of passengers. For two (or more) vehicle crashes, the relation between $v$ and $\lambda$ becomes even more complicated. The probability of a fatal bicycle-car crash (in which the cyclists dies) may depend on car distance travelled, age and experience of the car driver, etcetera. Therefore, it may make sense to define a risk $\lambda^{\prime}(r(t))$, as the probability to be involved in a fatal crash in which someone else is the fatality. Fatalities among passengers form a special case. In any case, $\lambda(r(t))$ depends on the probability of the presence at $r(t)$ of any other road user, together with the factors (age, experience etcetera) relevant for his chance to be involved in a crash. This contribution to risk is an essential new feature of the research described in this thesis.


### 2.3. Method

Essentially, we will use two techniques to further develop a model for $v$ :

1. We shall relate the development of the number of fatalities to distance travelled, also called travel or mobility. As the distance travelled of crash opponents in two vehicle crashes contributes to the risk of such a crash, we will explore the possibilities to incorporate distance travelled of two travel modes simultaneously in the analysis. To do so, we must at least distinguish between single-vehicle crashes and two vehicle crashes.
2. We shall use some of the information available of individual crashes: the crash database contains much more information than just the number of fatal crashes. Most fatal crashes are documented in much more detail than just the fact that it is a fatality. By using information such as the driver age, traffic mode, time and location of the crash etcetera, we enhance the information available to develop a model. As far as these factors influence $\lambda$, we can attribute a mathematical role to them in the model. As a very simple example: suppose there are two traffic modes $\alpha$ and $\beta$, one with constant risk $r_{\alpha}$ and the other with
constant risk $r_{\beta}$. Now, if the distance travelled for modes $\alpha$ and $\beta$ together were constant, the only relevant factor to explain the number of fatalities would be the proportion of distance travelled by modes $\alpha$ and $\beta$. Thus: a stratification of the crashes by traffic mode would result in a deterministic model for the number of fatalities. Of course, traffic safety is more complicated than that.

Essentially, we shall apply both methods simultaneously. We shall relate $v$ to distance travelled, and we shall stratify the data with respect to relevant factors. Because of this choice, the possibilities for stratification are limited to those variables which are available both in the crash data and in the data on distance travelled. Thus, the availability of variables in the data sources define our possibilities for stratification and analysis.

We study the development of risk, i.e. the ratio between fatalities and distance travelled for a specific group (age group, travel mode etcetera). Thus, we stratify the annual number of fatalities for a series of years into subgroups using one or more variables (e.g. travel mode and/or driver age) and relate these numbers to a relevant measure of distance travelled, stratified in the same way.

### 2.4. Data

The quality of the data is of crucial importance for road safety modelling. Sufficient completeness and correctness of crash data, and representativeness of survey data make the difference between a useful model and mathematical nonsense. Availability of data describing Safety Performance Indicators (SPI-data) such as blood alcohol concentration survey data, bicycle light use data, enforcement data, data on travelled speeds etcetera, decides for the possibility to estimate their effect on the annual number of fatalities. Here, we describe in some detail the quality of the available Dutch data used in this thesis.

## Crash data

Since 1976 Dutch police registered crash data are digitally available on a level of individual crash records. Police records were coded and made available by the ministry of infrastructure and the environment (SWOV/VOR/BRON, 2012). These data enable stratification with respect to a large number of variables, such as driver or victim factors (age, gender), crash factors (date, time of day, weather conditions), location factors (road/intersection type),
etcetera. Since 1976 these data are gathered and coded by the Ministry of infrastructure and the environment (VOR). In 2004, the data structure was somewhat simplified and retrenched, and since then the database is called BRON. In BRON, crash data since 1985 have been made available. A detailed description is available on the SWOV website (SWOV, 2012). For the years between 1971 and 1975, there are files with digital crash data, but these data are not made available for analysis. For the years still further back, we have tables at our disposal, put together by Statistics Netherlands, the Dutch Central Bureau of Statistics (CBS, 1950 and later years). A detailed description of the available crash data can be found on the SWOV website (SWOV, 2012). The possibilities for stratification of these data are limited to the available tables. Accurate data on the number of fatalities, corrected for police underreporting, are available since 1996. The correction is carried out by Statistics Netherlands. However, this institution doesn't publish all available information on individual crash records. Only specific tables are revealed (CBS, 2012).

## Quality of crash data in the Netherlands

A correct stratification of data by specific variables, e.g. travel mode and road speed limit, requires that the travel mode of each vehicle involved in a registered crash is registered correctly. If errors are made, e.g. mopeds are registered as motorcycles or vice versa, this would influence the correctness of the strata. Further, if these errors depend on the values of other variables (e.g. the registered road speed limit depends on the vehicles involved in the crash), other errors may occur. From the data themselves, it is usually impossible to know that these errors exist. However, a thorough analysis of the data may reveal clues as to such errors. For example: in the Netherlands, until December 1999 mopeds were allowed a maximum speed of $30 \mathrm{~km} / \mathrm{h}$ inside urban areas, and $40 \mathrm{~km} / \mathrm{h}$ outside. In accidents involving mopeds in urban areas on $50 \mathrm{~km} / \mathrm{h}$-roads, policemen sometimes registered a speed limit of $30 \mathrm{~km} / \mathrm{h}$. As a consequence, an analysis of crashes on $30 \mathrm{~km} / \mathrm{h}$ roads (of which there hardly were any before 1995 in the Netherlands), reveals an improbably large amount of crashes before 1995, practically always with mopeds involved. This not only shows for crashes with mopedists as victims, but also for crashes with mopedists as the second party involved. When these errors are suspected, they can sometimes be corrected. In the case mentioned above this could have been done by checking the location of the crash.

Underreporting is also a serious registration problem, especially if it is not uniform over all crashes. If e.g. crashes with elderly, or single-vehicle crashes, are excluded from the crash database more often than on average, this would distort the resulting data for some strata. If, further, this underreporting is not a constant factor in time, interpretation of time series of crash data, specifically of strata, need to be made with care. In this thesis we take care to correct for such registration errors, if they are known or can be suspected. However, the possibility of bias by registration errors cannot be excluded.

The number of fatalities in the Netherlands is accurately known since 1996, when Statistics Netherlands started to check the police registration and compare these data with the statistics of unnatural deaths, and with court files. It appeared that, in 1996 and the first 10 years after that, the number of road fatalities in the Netherlands was approximately $6 \%$ higher than the police registered number. In the years before 1996, only the number of fatalities reported by the police is known.

Unfortunately, there is far less information available about these last 6\% than there is in the police recorded crashes. Moreover, Statistics Netherlands does not make the data on these last $6 \%$ available for research by record, but only in tables only. Since 1996 researchers therefore generally have to choose between either the incomplete police database with its many variables by record, or the information presented by Statistics Netherlands with tables for a few variables only, for an analysis of Dutch fatality data. Until 2008, police registered data for fatalities were accepted as reasonably complete (less than $10 \%$ missing). Unfortunately since 2009 underreporting became a more and more serious problem. In 2011 less than $83 \%$ of the fatalities were registered by the police. Hence police registered data are no longer useful for analyses of data of recent years.

## Mobility data

To analyse risk, we need both crash data and data on distance travelled. A stratified analysis thus asks for stratification of both crash data and data on distance travelled. This limits the possibilities of our stratifications. If all relevant factors were collected in the crash data records, then the possibilities for stratified analysis would entirely depend on the quality of the data on distance travelled. A way to collect data on distance travelled is by means of a (preferably large and accurate) survey: people are asked to describe the trips they made during a day or a week, including trip time, origin and destination, motive, travel-partners, etcetera. Together with data describing
their personal situation, time of day, weather etcetera, a good picture of all distance travelled of a certain population can be made. If such data are unavailable, fleet size or sales of vehicles, or even fuel consumption data (Fournier and Simard, 2002) are used. Sometimes model results can be used, when a model is available that estimates distance travelled, such as the FEATHERS framework for Flanders (Janssens et al., 2007, Bellemans et al., 2010). Then, the possibilities to stratify depend on the model used.

## Quality of mobility data in the Netherlands

Between 1950 and 2000 a statistic on vehicle distance travelled (for passenger car, motorcycle, truck) was also collected. These data cannot be stratified any further. Data on other road travel modes are not available, although estimates were made for distance travelled by e.g. bicycle, based on a registered or estimated bicycle fleet size (Ploeger and Van der Waard, 1997). Recently, the statistic on vehicle distance travelled was resumed. Stratified data on person distance travelled are available since 1984 in the Netherlands, collected in a National Travel Survey (NTS). Until 2004 it was known by the name of OVG, between 2004 and 2009 it was called MON and since 2010 it is named OViN (SWOV/NTS, 2012). These data allow for stratification by age group, gender, travel mode, time and date, purpose of the trip, both for drivers and passengers. The specifications of this survey have changed during the years between 1984 and 2011. More information can be found in e.g. (Moritz and Brög, 1999). On the SWOV website (SWOV, 2012) an overview of the most important changes in the properties can be found. Until 1994, the stratification by age is possible into specific age groups only, as the micro data (i.e. the individual records of the travel survey) of Statistics Netherlands could not be recovered. From 1995 onward, data are available on a level of individual age years. In 1994, distance travelled for 0-11 year old became available. Travel modes in the NTS are limited to walking, bicycle, moped, motorcycle and passenger car (and public transport).

The total person distance travelled and data on vehicle distance travelled were collected using different methods. The National Travel Survey was carried out by post, by telephone and, since 2009 also by internet, among Dutch inhabitants, while motor vehicle traffic volume was collected by multiple traffic volume counts. Hence, these sets differ in many ways, e.g. because the travel survey excludes foreign traffic, business related trips, goods transport and holiday-related traffic. Motor vehicle counts also exclude pedestrians and cyclists. A detailed description of traffic data can be found on the SWOV-website. In this thesis, we consider the person mobility
survey the most valuable one because of the possibility of stratification by several variables.

The quality of the person mobility data has changed significantly since 1985. The number of travellers in the sample increased from approximately 20000 persons between 1985 and 1993 to nearly 150000 persons in 1995 and since then diminished to less than 40000 persons in 2009. This is not without consequence for the applicability of these data for risk analysis, especially in the case of stratified risk assessment. For example, in the calculation of the risk for $75+$ year old moped riders, the small sample size (in recent years less than 20 persons annually) is beyond the limits of meaningful calculations.

For some travel modes, such as trucks or light vans, travel data are not available. In those cases an alternative is to use fleet size or sales data. These data are inferior to travel data, because the use of vehicles in the fleet is unknown. Therefore, using these data asks for great care.

## Demographic data

Demographic data (by year, year of birth and gender), can be a welcome expansion of the available crash data or travel data. Demographic data allow for the calculation of mortality (fatalities per inhabitant) or morbidity (injured per inhabitant).

## Quality of demographic data in the Netherlands.

Demographic data in the Netherlands are gathered by Statistics Netherlands, and based on a supposedly complete count of the Dutch inhabitants, with stratification by year of birth, gender and region. These data have been readily available since 1950 (SWOV/CBS, 2012). The data are assumed to be very accurate, although it cannot be ruled out that the data contain errors. Data are available per January $1^{\text {st }}$ of each year. When the data are to be used to estimate the number of people of a certain age in a specific year, it is advisable to use the mean of the number of people of that age on the $1^{\text {st }}$ of January of both that specific year and the next year, in order to correct for small year to year demographic changes in subsequent years.

## Other data, such as data on SPI's

Several other data sources are relevant for Dutch road safety analyses, such as data on road length, average speed, enforcement activities, blood alcohol concentration in road surveys, vehicle fleet size and sales (for motorized vehicles and bicycles) or road length. Some of these data, and their
descriptions, are available on Statline (CBS, 2012) and on the SWOV-website (SWOV, 2012).

The usefulness of these data sources depends on their quality and on the purpose of the analysis. When data such as fleet size and sales (or even fuel consumption) are used as an approximation of data on distance travelled, the possibilities to stratify distance travelled e.g. by age, gender, road type, type of vehicle are very limited. In addition, some assumptions are necessary to translate the approximation like fuel use or fleet size into distance travelled such as mean distance travelled per vehicle, or fuel consumption per vehicle. In this thesis we prefer not to use these sources to replace data on distance travelled. Sometimes, however, this is inevitable e.g. for analyses of road safety before 1984.

A list of Dutch safety measures is available on the SWOV website. Here, safety measures on a national level, known at SWOV, are listed and described, together with their date of introduction.

Much data on SPI's, relevant for development of a well-defined road safety model for $v$ to fit the past development of road safety, is not available. The precise design and construction of the roads, especially at crash locations, is not known except for what is available in the current databases. Roads characteristics such as curvature of the roads, widths, distance to obstacles near the roads, the presence of traffic lights, the presence of traffic signs, bicycle paths, actual vehicle speeds, etcetera, are not being registered in a central database even now. Speed limits are not yet made available centrally for all roads.

Thus, modelling traffic safety using all relevant factors is a bridge too far right now. Even if we manage to find out how the many factors influence safety, we cannot link that information to the number of fatalities in the Netherlands because we do not know the actual values of the factors. Consequently, we limit ourselves to mobility as the first and most important factor. In this thesis we explore the possibilities to develop a model containing mobility, stratified by age, gender and travel mode.

### 2.5. Some examples of stratifications of crash data

The use of stratification and the incorporation of distance travelled in the analysis of road safety data may be relevant if the development of the total





## 電



|  | 1975 | 1985 | 1995 | 2005 |
| :--- | :--- | :--- | :--- | :--- |
| 700 |  |  |  |  | Figure 2.7. Annual number of police registered fatalities by fatality type: left column: traffic mode of fatality; top row: traffic mode of other traffic mode involved. Fatalities with a pedestrian or two-wheeler (bicycle, light moped, moped or motorcycle) as the other mode involved were added and presented in a single column. Bottom row and right column represent the horizontal and vertical sums by travel mode. Colours represent the fraction of the number of fatalities in the period between 2000 and 2009. The bottom right panel represents the total number of road fatalities.

number of fatalities is indeed inhomogeneous. This fact can be easily demonstrated in many ways. This section presents a number of different stratifications that either show different risk values for different strata, or different developments in the number of fatalities.

A straightforward way to stratify the number of fatalities is by fatality type, i.e. by the combination of the traffic modes involved. For a single-vehicle crash the fatality type is the traffic mode of the victim. For a crash with two (or more) traffic modes, it is the traffic mode of the victim and the vehicle (sometimes a pedestrian) of the most relevant party involved.

In the Netherlands, crash data by fatality type have been available since 1976 for police registered fatalities. Figure 2.7 shows the annual number of fatalities by fatality type. The figure suggests that the development of the number of fatalities in each group does not follow the same trend for all fatality types. The introduction of the light moped (i.e. a moped with a vehicle speed limit of $25 \mathrm{~km} / \mathrm{h}$ that can be ridden without a helmet) about 1990, and the increasing popularity of the small van since about 1980 seem to leave recognizable patterns in the data. Further, the figure illustrates that single-vehicle crashes are important group for almost all traffic modes, except for pedestrians and cyclists.

Other remarkable observations are that the number of fatalities among van and truck occupants is clearly less than the number of fatalities among other road users in crashes where the van or truck is the other vehicle involved. Figure 2.7 also illustrates that the decrease in the number of fatalities in the Netherlands is mainly materialized in just a few important fatality types: pedestrian-car, bicycle-car, moped (in general), car-car, car-truck and carother, where a substantial amount of fatalities decreased by a factor of 4 or more. Many other fatality types showed a less promising decrease since 1976.

Another way to stratify crash data is by a combination of the age of the fatality, and the traffic mode of the fatality. In the Netherlands these data have been available since 1950. Figure 2.8 a shows the number of fatalities as a function of time, for age groups of $0-4,5-9 \ldots$ years of age. For all age groups the scale of the fatality axis is approximately the same. This figure shows how the dominating traffic mode among fatalities strongly depends on the age group. While young children up to four years old almost only died as pedestrians, for 10-14 year old children cycling dominates the casuistic, for 15-19 year old it is moped riding, and for adults older than 20
car occupants die most frequently. For elderly people walking and cycling as an unsafe means of travelling is noticed. Motorcycle riding never dominates the picture, although the figures show that the development of motorcycle fatalities has been different for different age groups: among people of 30 years and older, there were hardly any motorcycle fatalities between 1970 and 1985, while for younger riders, a number of fatalities indeed was registered in these years. This Figure illustrates that the rise and fall of the number of fatalities between 1950 and 2010 in the Netherlands is not uniform over the age groups. The number of fatalities among very young children was relatively high in the fifties to drop enormously afterwards until an almost negligible level in 2010. For some age groups we see just one sharp peak, with an equally sharp dip in it (15-24), whereas for other age groups the development is much less outspoken, sometimes showing several peaks (25-34). Further it is clear that the decrease after 1974 is much less progressive for elderly than for people younger than 25.

It can be assumed that these developments are the consequence of vehicle use. Unfortunately data of vehicle use are not available for the entire period between 1950 and 2010, but we do have demographic data at our disposal. These data can be used to calculate mortality, i.e. the number of fatalities per inhabitant, for all age groups. This is done by using the average number of inhabitants as known on January $1^{\text {st }}$ of the specific year and of the following year. Figure 2.8 b shows the result, again with a uniform scale for all age groups. The figure illustrates how mortality quickly increases with age for adults, while it is also high for 15-24 year old, but actually quite low for younger persons. Only in the first years after 1950, mortality was approximately equal for all ages below 60. Currently mortality has decreased, especially for children.

Comparison between Figure 2.8 a and 2.8 b shows that demography has a relevant influence on the number of fatalities by age group. Whereas in Figure 2.8a the number of fatalities for $75+$ people shows a relatively small decline since 1970, we can see in Figure 2.8b that mortality did decrease considerably stronger, indicating that there is an increasing number of elderly which markedly influences the fatality data.

The development of the number of fatalities is likely to be influenced by distance travelled. Therefore, it is informative to analyse the number of fatalities as a function of distance travelled. Figure 2.9 illustrates how distance travelled relates to fatalities, for four travel modes: powered two-


Figure 2.8a. The number of fatalities as a function of time, for age groups of between $0-4,5-9$, ... years of age.


Figure 2.8b. Road safety mortality as a function of time, for age groups of between 0-4,5-9, ... years of age.
wheelers (i.e. motorcycle, moped and light moped), pedestrians, cyclists and car occupants. Here, data for a period of nine years are presented.

In Figure 2.9, distance travelled is shown along the horizontal axis, and the annual number of fatalities is shown on the vertical axis. As both distance travelled and fatalities vary by several factors of 10 between different combinations of travel mode and age group, logarithmic scales on both axes were chosen. As a consequence, lines of constant risk (i.e. quotient of fatalities and distance travelled) appear as parallel diagonal lines.

Figure 2.9 shows that for any age group, powered two-wheeler risk is approximately a factor 50 higher than car occupant risk. Further, pedestrian risk is almost always higher than bicycle risk, which in turn is almost always higher than car risk, except for 18-29 year old car drivers. Child safety for car passengers turns out to bear by far the least risk of all these groups: only slightly more than $0.5 \cdot 10^{-9}$ fatality $/ \mathrm{km}$. The Figure doesn't show the risk of e.g. truck drivers, for whom the risk may be even less. This is because their distance travelled is not measured in the Dutch National Travel Survey.


Figure 2.9. Fatalities as a function of distance travelled, by traffic mode (diamonds: powered two-wheeler, circles: pedestrian, squares: bicycle and triangles: car) and age group (indicated by colours). Dotted light grey lines denote constant risk, of which the corresponding values are plotted near the horizontal and vertical axes.

The above mentioned examples illustrate that travel mode and age of fatality are certainly relevant variables that discriminate between high and low risk. To illustrate that driver age of crash involved drivers and fatality age are not the same thing, we plotted the number of fatalities for car-car crashes between 1976 and 2010 against two variables: the age of the driver of the vehicle with the fatality, and the age of the driver of the other vehicle (Figure 2.10). In case two drivers died, the crash appears twice in the figure.


Figure 2.10. Average annual number of fatalities in Dutch car-car crashes between 1976 and 2011 by age year of both drivers. Top panel: by age of driver of the fatality on the x-axis, and with different lines by age of the other driver. Bottom panel: the other way around. The average for the $75+$ groups is calculated as if there are 10 age years in this group.

Figure 2.10 shows that drivers of the car with the fatality as well as drivers of the other car involved in the fatal crash are more often young drivers, as the annual number of fatalities per age year of both drivers is highest for car-car crashes where both drivers are between 18 and 24 years of age. Mark that crash involvement is not identical for the different drivers involved. There are more elderly drivers involved in a crash with the victim in their car, than when the victim is in the other car. This can be understood from the fact that in a random car-car crash, the elderly driver is more likely to die (see also Figure 2.11). One should not conclude from Figure 2.10 that young drivers tend to attract each other. The data presented in Figure 2.10 can be well understood from the assumption that driver age of both cars are independent, while in both cars the probability of a fatal crash is highest for young drivers. Hence the distribution of fatalities over driver age is not symmetrical: the driver of the car with the fatality is much more often an older driver. This can be understood as a consequence of the fact that, if there is a crash between two drivers, one of which is an older driver, it is more likely to be fatal because of the probability that the older driver dies.


Figure 2.11. Mortality by age (fraction of people in the Netherlands who died at that age by whatever cause) for male and female inhabitants of the Netherlands in 2008 (SWOV/CBS, 2012).

Older people have a higher a priori mortality, i.e. a higher probability to die (considering all possible causes of death), as is illustrated in Figure 2.11. It is clear that male mortality is higher for virtually every age. Between 15 and 30 years of age, the difference in mortality between males and females is
highest, which is related to the high number of male road crash victims. Between 35 and 95 years of age, mortality increases exponentially with increasing age for both males and females with approximately $10 \%$ annually. This indicates what everyone knows: elderly people die more often than young persons. Thus, it is not surprising that in traffic, the number of fatalities increases with age as well. However, this is no reason not to try and find ways to improve safety, also for elderly people.

### 2.6. Conclusion

Understanding and forecasting the development of road safety in general requires stratification of crash data and data on distance travelled for those variables for which risk (fatalities/km) is different and distance travelled is changing over time. In this thesis the influence of stratified distance travelled is analysed. Both the distance travelled of the fatalities, and the distance travelled of the other travellers involved in a possible crash are taken into account. Hence, whereas for single-vehicle crashes the distance travelled of one travel mode is sufficient, for two vehicle crashes the distance travelled for both traffic modes is needed.

### 2.7. Discussion

In this thesis risk is defined as the fatality density, a probability of a fatal crash per unit of distance travelled. Although this is a common concept to incorporate distance travelled in the analysis of the number of fatalities, here we introduced another reason to do so. By doing so, the two dimensional time vs road length space of the ETC collapses into a one dimensional space, either time or route length. Risk is defined along the trajectories of trips. This is in accordance with literature, as risk and distance travelled are often chosen as the two components which together determine the expected number of casualties (Oppe, 1991b, Bijleveld, 2008, Broughton, 2012). This choice does not imply that the expected number of road fatalities must be assumed to be proportional to distance travelled by definition, nor that risk must be assumed independent of distance travelled. There are several arguments to propose other relations. For two-sided crashes, e.g. bicycle-car, both bicycle distance travelled and car distance travelled contribute to the likeliness of these crashes (Chapter 6). Consequently, if all other factors are constant, the number of car-car crashes may be proportional to the square of car mobility (Chapter 3). On the other hand it is sometimes reasoned that
crash rate is less than proportional to distance travelled, e.g. when roads or intersections with different traffic volume are compared, it is often seen that the number of crashes increases less than linear with traffic volume (Elvik, 2009b). This effect can be a consequence of safer road design of roads with more traffic, or even safer behaviour on roads with more traffic. On the other hand, it is reasonable to believe that the number of crashes on a road is proportional to the length of that road (ceterus paribus), as it is reasonable to expect twice as many fatalities in a country which is twice as large (ceterus paribus). If the latter were untrue, the migration of the separate EU-countries into a single European country would, by itself, improve the safety of all inhabitants in these countries.

For different situations (different roads, years etcetera) with different amounts of distance travelled, the crash rate very much depends on the properties or circumstances of these situations. For example when comparing the crash rate on a road with a high traffic volume to the crash rate of a road with low traffic volume, it is necessary to correct for possible different road designs: roads with high annual traffic volume, such as motorways, are usually designed to have a much lower risk or crash rate than roads with a low annual traffic volume such as urban or rural roads (Martin, 2002). Further, when crash rates for different times of day are compared, low traffic volumes at night are supposed to be related to higher crash rate than high traffic volumes at day and good light conditions (Martin, 2002; Jacket and Frith, 2013). And finally: when comparing current crash rates with those of some years ago, a perhaps lower crash rate in recent years may be the consequence of improved safety factors (safer cars, more experienced drivers), see e.g. (Broughton, 2003).

Road crashes are the result of a stochastic process. To capture the role of chance in this process, statistical analyses of a number of crashes have to be carried out. Analysis of the development of annual road safety crash data invariably means that data are aggregated. Except when performing in depth analyses of individual crashes, aggregation is inevitable: annual numbers of road fatalities are aggregated of equally many different crashes, with different combinations of driver ages, road types, travel modes etcetera. Aggregation of different crashes into large numbers has the advantage that statistical analysis becomes possible. Such an analysis of a large number of different types of crashes is statistically more powerful than separate analyses of the development of small numbers of data of different groups of more similar crashes. There is however also an important disadvantage:
when comparing aggregated crash data of subsequent years, the mutual ratios of the number of crashes in different subgroups might differ from year to year, causing changes in the annual number of fatalities which go unnoticed when looking at aggregated data only. For example: a strong decrease in the number of fatalities in New Zealand between 1988 and 1996 coincided with the import of small second hand cars and a contemporaneous strong decrease in the number of motorcycle crashes (OECD, 2007). This suggests that a modal shift from motorcycle to passenger car was perhaps the major cause of the decrease in the number of fatalities in New Zealand in this period. If such a modal shift remains unnoticed, the observed road safety improvement might unjustly be attributed to contemporary safety policies. Especially when forecasts of the number of fatalities are carried out, it may be misleading to extrapolate aggregated crash data or crash rate data without noticing such a temporary decrease related to a perhaps finite modal shift. Another example is the influence of demographic developments. In many developed countries, the number of births decreased about 1970 when the birth control pill was widely introduced in these countries. This influenced the number of young, inexperienced car drivers some 18 years later, and thus this demographic development can be identified with a lower distance travelled and, consequently, a lower number of crashes with young drivers from about 1990 onward. This also has an impact on the risk of other road users, as there are less young drivers to be involved with in a crash.

There are many examples of aggregations in road safety analyses.

1. Increased blood alcohol concentration (BAC) exponentially enhances risk by a factor or 5 to 10 per g alcohol/ $\ell$ blood (Elvik, 2006, figure 8). In research regarding BAC-effects on risk, differences for different BAC's are made, but usually for other variables such as driver age the data are aggregated.
2. Young, inexperienced drivers show more risk than more experienced middle aged drivers, and elderly drivers are more vulnerable and also have a higher crash rate than middle aged drivers. But in the Netherlands both young and elderly drivers drive in cars with a lower mass than average (Berends, 2009), which also enhances their risk (Evans and Frick, 1993, Evans, 1994).
3. Virtually every analysis is biased because of confounding factors. In countries with a large number of fatalities, more aggregations can be carried out simultaneously before the remaining numbers of fatalities in each group becomes too small to handle statistically. However, in large countries with many fatalities, more stratification may also be necessary
because the inhomogeneity of the circumstances of these crashes. In the USA, there are currently still more than 30000 road fatalities annually (SWOV/IRTAD, 2012). An analysis of these data should take into account that the data were collected in 52 states, each with their own legislations, rules, weather conditions, etcetera.

### 2.8. $\quad$ Some miscellaneous remarks

Road safety data analysis always asks for a compromise between the urge to understand the processes, and the wish to carry out tests and show statistical relations. The former asks for stratification, the latter asks for aggregation. In this thesis we show various ways to stratify crash data, data on distance travelled and risk data, in an attempt to unravel road safety data time series.

The question arises when stratification meets its limits. There are several possible reasons not to further stratify data. One is when there are so few data available, that these data can no longer be interpreted because of the small size of the strata. Another reason not to stratify any further is when the resulting risks are likely to be uniform over the data chosen.

A perfect road safety model would describe how the number of crashes depends on any relevant external factor and measure. To design such a model is, at least at this moment, well beyond the technical possibilities. In this thesis we shall limit ourselves to some first steps only: incorporation of annual distance travelled, stratified by a few relevant factors. Distance travelled, or mobility, is an essential factor, as without distance travelled there are no road crashes. The effect of road safety measures on the road casualty frequency cannot be identified accurately if the frequency of the associated distance travelled is not accurately taken into account. An important reason to incorporate distance travelled is that changes in stratified annual distance travelled can occur relatively abruptly. This may well affect the annual number of fatalities equally abruptly. A sudden modal shift from car to motorcycle, e.g. to avoid motorway taxes (TfL, 2006), was feared to lead to an increase of the number of fatalities in Greater London. Another example is the introduction of the moped helmet in the Netherlands in February 1974 which immediately led to a large decrease of moped use. The positive effect of the use of helmets on the number of moped fatalities is likely to be far outnumbered by the also positive effect of the decrease of moped use. This is illustrated in Figure 1.1, where annual moped fleet size and annual moped fatalities are plotted for the period between 1950 and
2010. The data suggest that the decrease in the number of fatalities coincides with the decrease in fleet size. Although data on distance travelled are not available, this illustrates that changes in distance travelled cannot be ruled out when evaluating safety measures.

A remark about the dimension of annual distance travelled: just as the annual number of fatalities actually is a number per unit of time, and thus a frequency, the annual distance travelled is a distance per unit of time, and thus a speed! Intuitively, this is nonsense: if we say that the Dutch people travel 150 • $10^{9} \mathrm{~km} /$ year, this is not felt as a speed of $1.7 \cdot 10^{6} \mathrm{~km} / \mathrm{h}$ or $470 \mathrm{~m} / \mathrm{s}$ of the population. In this thesis, when we denote the distance travelled per unit of time, we shall just speak of (annual) distance travelled, expressed in km. The number of fatalities is expressed as an annual number. Their ratio, however, is defined as risk or fatalities/distance travelled, expressed in fatalities $/ \mathrm{km}$. In this ratio, the dimension of time disappears. What remains is a fatality density. Actually, this could give rise to some confusion, as in traffic safety literature, the fatality density is sometimes used to express the annual number of fatalities per unit of road length. Mark that although this quantity is described as a fatality density, it is in fact a fatality density frequency (c.f. $\Lambda$, introduced in Section 2.2), expressed as fatalities/km/year.

# 3. Distinguishing traffic modes in analysing road safety development 

This chapter was first published in "Accident Analysis and Prevention" (Stipdonk, H., Berends, E., 2008. Distinguishing traffic modes in analysing road safety development. Accident Analysis and Prevention 40 (4), 1383-1393).


#### Abstract

Changes in mobility influence road safety. The effects of safety measures may even be overshadowed by the effects of temporary mobility fluctuations. Usually mobility is corrected for by defining risk as the ratio between fatalities and mobility. Due to lack of sufficient data, mobility is often approximated by car mobility. In this chapter we will show that the resulting "general" risk is misleading. Stratification by traffic mode shows that the risk for car drivers, motorcyclists and truck drivers is roughly constant between 1950 and 1970 and that it has been decreasing exponentially afterwards. This contradicts the development of general risk, which has been decreasing the whole period between 1950 and today. Further stratification shows that stratification by traffic mode of the fatality alone is still insufficient. Changes in mobility of the other party in a crash are also important. The development of risk of car-car accidents differs significantly from that of single-vehicle car accidents. A comparison between single-vehicle motorcycle accidents and motorcycle-car accidents shows a similar discrepancy. Stratification of mobility by traffic mode, and of fatality data by the relevant traffic modes involved, can enhance the understanding of the influence of mobility on safety.


### 3.1. Introduction

An accurate understanding of the influence of mobility on road safety is required to model the influence of safety measures on the number of fatalities. This is because the effects of each separate safety measure usually only gradually influence a small fraction of all fatalities, whereas mobility changes may be sudden and substantial.

Several authors, e.g. (Oppe, 1991b, a) have studied the total number of traffic fatalities in combination with mobility. Many countries have known a period with a strong increase in the total number of traffic fatalities, whereas afterwards a decrease suddenly set in. This often resulted in a cusp-like
development of the number of fatalities (OECD, 2007). Researchers tried to explain this phenomenon by a simultaneous increase in mobility and a decrease in risk. Oppe calculated risk as the ratio between all fatalities and traffic volume (vehicle kilometres). Other authors have used a similar approach, using a generic measure of mobility, either car (i.e. passenger car) distance travelled or total (motorized) traffic (Broughton et al., 2000), or any approximation for this such as fuel consumption (Fournier and Simard, 2002). Oppe found that a simultaneously exponentially decaying risk and an S-shaped increasing mobility explain the cusp-like trend in the number of fatalities in six countries reasonably well. This approach is acceptable, but not accurate. His approach explains the rough form of the cusp in the number of fatalities, but smaller variations and the recent slowdown of the risk decay were not explained. A thorough understanding of this phenomenon is still lacking, as Gaudry mentioned (Gaudry, 2005).

The disadvantage of the above methods is that a general measure (e.g. car mobility or total motor vehicle mobility) for mobility was used to take into account the influence of mobility on road safety. Mobility of each modality should be taken into account when analysing or modelling road safety. This is because some modalities bear a higher risk than others (e.g. motorized two wheelers). The purpose of this study is to analyse mobility and its influence on road safety. The authors show that a general measure for mobility is not sufficient. We searched for better ways to take into account the development of mobility. Not only car mobility or total mobility was included, but mobility of each modality. This method is more complex because it involves a distinction between different traffic modes, as well as between singlevehicle and multivehicle accidents.

The focus of the chapter will be on understanding rather than on modelling road safety data. Such understanding of the influence of mobility on safety is a first necessary step in modelling road safety development, before models can be made for safety measures. To develop an accurate functional form for any model that describes the influence of mobility on road safety, we consider it necessary to gain insight in their relation. Without this insight, models may be inaccurate, such that changes in the number of fatalities due to changes in mobility may obscure the effect of safety measures. Therefore, the chapter will not discuss calculations of parameters of models. Furthermore, we will not provide a detailed description of the effects of safety measures. Instead, we will demonstrate the importance of mobility changes by stratification of fatalities and mobility by traffic mode. Not only is
the mobility of the traffic mode of the victims important, also the mobility of the traffic mode of the other parties has a large influence on the number of accidents.

As we wanted to include a period with both changes in the increase/decrease of fatalities and mobility in our analysis, we decided to study a 50 year period from 1950 to 2000. Following Gaudry's suggestion, we think that the cusp in the 1970s offers an important opportunity to understand road safety development. In this chapter, Dutch data were analysed because of its accessibility for the authors. However, these analyses can be applied to data of any country if available. The choice of data and its stratification level are discussed in Sections 3.2, 3.5 and 3.8.

We start in Section 3.3 with looking at the total number of fatalities divided by one measure of mobility, in this case car mobility. We will call this ratio the general risk. This is an accepted method to take into account the effects of mobility. In Section 3.4 we will show the shortcomings of such overall approach, in which all fatalities are explained by just one mobility measure. We demonstrate that an understanding of the influence of mobility on the development of fatalities since 1950 cannot be based on motor vehicle mobility only.

In Section 3.6 we will first distinguish between fatalities of different traffic modes. We will limit our analysis to the three traffic modes for which we have mobility data since 1950: car, motorcycle and truck. We will show how the risk of a fatal accident for drivers of these traffic modes, developed over time. The analysis reveals that the influence of changes in traffic volume is very strong, and contradictory to the common belief that the risk decreased over the entire period between 1950 and today.

In Section 3.7 we will show it is necessary to make a further step towards traffic mode separation. The mobility of the other party also plays an important role. We have to distinguish between different fatality types such as single-vehicle accidents (no other party involved), and accidents between two traffic modes. In Section 3.8 we will further stratify the accident data on fatality type. We will separate accidents with car or motorcycle only (singlevehicle accidents), accidents with a pedestrian and either a car or a motorcycle, accidents with two cars, or with both a car and a motorcycle. In Section 3.9 it is shown that the risk for these six fatality types developed differently.

In Section 3.10 we will discuss the results. We will provide possible explanations, and indicate the direction into which we want to continue our research.

### 3.2. Data (totals)

Two types of data are necessary for analysing road safety development, namely accident data and mobility data. To begin with the total number of fatalities and the passenger car mobility were used.

## The total number of fatalities

For all data about fatalities, we used the digitally available accident record data of the Ministry of Transport (SWOV/VOR/BRON, 2012) since 1976, and for the period between 1950 and 1976 we used Statistics Netherlands' tables, (CBS, 1950 and later years).

The data are based on the police road crash registration. Since 1993, police registration data are compared with municipal administration of causes of unnatural death and with court records. This has revealed that underreporting of police fatality data is about $7 \%$. In this analysis we use the police records, and not the higher factual numbers. It is not possible to use the factual numbers in analyses, as there is no additional information available about the accidents that are missing from the police records. Furthermore, these actual numbers are not known before 1993.

In the Netherlands, road fatalities are defined as those involved in a road crash who die due to their injuries within 30 days after the crash. A road crash is defined as an accident on a public road, involving at least one vehicle (not necessarily a motorized vehicle). Accidents involving pedestrians only are not considered a traffic accident, but a single-vehicle bicycle accident is.

The reported numbers of fatalities in the Netherlands since 1950 are shown in Figure 3.1. This figure shows that the numbers of fatalities between 1950 and 1970 have increased strongly. This increase appears to be approximately exponential. This is visualized in the right panel of the figure, where a logarithmic vertical axis is used. Therefore, each exponential increase and decrease is shown as a straight line. About 1970, this exponential increase stops, and from 1972 onward, the number of fatalities decreases almost as strongly as it increased before 1970. After 1985, the decrease lessens. We fitted simple, unweighted exponential lines of constant annual change to
estimate the exponential increases and decreases in the three periods. The exponential slopes $\beta$ of these lines of constant annual change are used to quantitatively compare the developments in the three periods. Each $\beta$ represents the slope of the log-linear function, as in $y(t)=\alpha \cdot e^{\beta \cdot t}$. For small values, this $\beta$ is almost equal to the fractional annual change. Throughout this chapter we will use the value of $\beta$ as a practical indication of the annual change in percentages.


Figure 3.1. Traffic fatalities in the Netherlands since 1950 with a linear (left) and a logarithmic (right) vertical scale. The solid lines are exponential lines of constant annual change for the periods 1950-1970, 1972-1985 and 1985-2005, with the annual change as a percentage.

Figure 3.1 shows an annual increase of approximately $5.4 \%$ before 1970, and a decrease of $-5.7 \%$ afterwards (the minus sign indicates a negative slope). The decrease slows down to a present value of about $-3 \%$ annually. In the following paragraphs we will focus on the cusp near 1970, the strong change in the slope of the log-linear trends before and after 1970. This cusp in the number of fatalities amounts to a difference of $11 \%$ (the difference between $+5.4 \%$ and $-5.7 \%$ ) in both annual change parameters. In the next paragraphs we will point out that this cusp is also found in the development of traffic data.

## Mobility data

In all sections of this chapter, we used historical mobility data of Statistics Netherlands (CBS) with annual data of the distance travelled for cars, motorcycles and trucks. In this section, two other measures for mobility are
presented, namely car fleet size and the length of motorway road (CBS, 1985 and onward, 2000, 2012).

Between 1950 and 1970, the car became very popular in the Netherlands, which is shown in Figure 3.2. This figure shows the development of three traffic quantities in the Netherlands over time: mobility (distance travelled) by car, the car fleet size and the length of motorway road. The figure shows that mobility, fleet size and motorway road length all increased strongly between 1950 and 1970, and that the increase then diminished considerably.

The right panel in Figure 3.2 has a logarithmic vertical axis, such that exponentially increasing or decreasing trends appear as a straight line. The figure shows that all three traffic quantities follow an exponential increase between, roughly, 1950 and 1970 (for road length until 1975) of more than $10 \%$ per year, whereas afterwards this increase is still exponential, but only by a few per cent per year.


Figure 3.2. Car distance travelled, motorway road length and car fleet size, with linear (left) and logarithmic (right) vertical scale. The development in the Netherlands, since 1950. Solid lines are exponential lines of constant annual change, with the annual change as a percentage.

Mobility growth turns out to resemble exponential growth. Exponential growth appears when growth of some quantity is proportional to this quantity (a constant annual growth). We do not know why mobility growth resembles exponential growth. Of course, exponential growth is quite common in nature, both in biological, physical and economical processes. We might speculate that the exponential growth in passenger car mobility as
seen in The Netherlands (and perhaps in many other countries) is a consequence of economical laws of maximum growth. In road safety analyses, it is very common to use an exponential approach. Oppe (Oppe, 1991b, a) and many others use the exponential to describe risk decrease. Its advantage above simple linear behaviour, of course, is that an exponential decrease will never go below zero, whereas linear decrease will. But although we think that exponential models are attractive, that is not the important reason we used exponential trends. That was mostly because the mobility trend (Figure 3.2) shows such a strong resemblance with exponential behaviour.

The cusp like behaviour of Dutch mobility about 1970 is not unique. In the UK a similar cusp-like development of car mobility has taken place (Pooley and Turnbull, 2000). Perhaps the most evident explanation for this phenomenon is that the availability and importance of transport boomed after World War II, and collapsed at the time of the global oil crisis in 1974. In this thesis we accept the development of car mobility as a given fact, and we concentrate on the consequences of this development for road safety.

The yearly change values for all three traffic quantities show a cusp, a sudden change over just a few years' period. This change equals about $13 \%$ for the car fleet, $11 \%$ for car mobility and $9 \%$ for motorway road length. In Figure 3.1 we already saw that the cusp in the number of fatalities is about $11 \%$, the same order of magnitude as the cusp in the traffic data. This suggests that the cusp in the number of fatalities is a direct consequence of the change in traffic growth. The only difference between the cusp in the number of fatalities and the cusp in the traffic quantities is that the latter shows a continuous increase, whereas the former first shows a rise, and then a decline. In the log-linear figures, this means the plots are tilted with respect to one another.

### 3.3. General risk

To analyse the influence of mobility on safety, it is a natural first step to divide fatalities by car mobility (or perhaps fleet size), and to define the result as the traffic risk (fatalities per km driven). Figure 3.3 shows the result for the Netherlands for the years 1950-2000. This approach is the same as Oppe's analysis (1991a) for Japan, the United States, West Germany, Great Britain, Israel and the Netherlands for the period between about 1947-1987. We use a logarithmic vertical axis such that exponential behaviour shows as
a straight line. The almost constant annual decrease in risk of $-6.5 \%$ clearly shows in the case where motor vehicle mobility is used. When car mobility is used, this leads to a slightly higher risk decrease of $-8.4 \%$ until 1985, and a lower decrease afterwards. As both traffic fatalities and car mobility show a cusp of about $11 \%$ about 1970, the slope of the resulting calculated risk does not show any discontinuity near 1972. This suggests that the cusp in the number of fatalities is merely a consequence of a continuously decreasing risk and strong changes in mobility growth. The question of why the resulting general risk in Figure 3.3 has been decreasing, was sufficiently discussed in Oppe's paper. In the last 20 years, the decrease has diminished to about $-4.3 \%$ annually. This is not visible in Oppe's analysis because it is more than 15 years ago that he performed his analysis. In the following sections we will show that the decreasing risk can partly be attributed to the fact that general risk is too inaccurate to understand road safety data.


Figure 3.3. General risk. The total number of fatalities, divided by car distance travelled. The solid lines are exponential lines of constant annual change, with the annual change $\beta$.

Next to total motorized vehicle mobility other general measures for mobility have been used (i.e. population and fuel sales (Fournier and Simard, 2002)). Often, population data are used as an approximation to mobility data (Kopitz and Cropper, 2005), simply because mobility data are unavailable. And indeed, mobility changes may be the result of demographic changes. However, population size alone cannot explain changes in mobility.

### 3.4. A general measure for mobility is not sufficient

The first step in our reasoning is to show that car mobility is not a selfevident approximation of mobility, especially in earlier years (1950-1960) in the Netherlands. Initially, other traffic modes were much more important, while the car was still at the beginning of its life cycle (Filarski, 2004), not only in the Netherlands but also, for example, in Great Britain (Pooley and Turnbull, 2000). In 1950 there were about 1000 fatalities annually in the Netherlands, which is more than today. At that time, accidents with trucks or motorcycles were as serious a problem as accidents with cars. This section shows that car vehicle mobility is insufficient to explain road safety changes.


Figure 3.4. Involvement of cars in fatal accidents. Fraction of accidents where no car was involved (neither as the traffic mode of the victim, nor as the other traffic mode).

It was shown that the general risk is a monotonously descending curve (Figure 3.3). However, this presentation of the facts is somewhat misleading. Indeed, nowadays the car is an important factor in traffic accidents, but this was not always the case. In Figure 3.4 we see the fraction of fatal accidents in which no car was involved, since 1950 (see Section 3.8 for the data source). In the early fifties, many accidents occurred with motorcycles (100 singlevehicle fatalities yearly) or with pedestrians and cyclists, being hit by a cargo truck, a tram or even a horse carriage. Thus, in those days, a car was involved in only one third of the accidents. Nowadays, it is the other way around: in one third of the fatal accidents no car is involved. Typical fatality
types in which no cars are involved are single-vehicle accidents with moped or motorcycle, and accidents between trucks, vans, pedestrians and cyclists.

Therefore, car mobility alone is an insufficient measure of mobility to understand the development of road safety. Dividing fatalities by car mobility gives us a ratio of which two thirds of the numerator is not related to the denominator. Thus, two thirds of the decrease in the resulting risk is a consequence of a division of fatalities by an increasing mobility which is unrelated to these fatalities. Knowing this, an analysis in which different traffic modes are distinguished, is desirable. It is not sufficient to replace car mobility by motorized traffic mobility or total mobility, as different traffic modes have different risks (number of fatalities per km driven). For example, motorcycle risk is much higher than car risk, so an exchange between motorcycle mobility and car mobility (while the sum remains constant) would increase or decrease the number of fatalities. The influence of a measure affecting the total mobility or the mobility of motorized traffic on road safety is susceptible to changes in modal split. When risk is calculated as the number of fatalities per km driven by motor vehicles, changes in modal split cause changes in risk. These changes are difficult to interpret without looking at the modal split. Essentially this means that stratification by traffic mode of both fatalities and mobility is necessary.

Thus, both accident data and mobility data need to be stratified by traffic mode.

### 3.5. Data stratified by traffic mode

Stratifying by traffic mode is not done easily. The number of fatalities (see Section 3.2) can be stratified by the traffic mode of the victim. However, traffic data (mobility) is not generally collected for all modes. Furthermore, only vehicle mobility (and not traveller mobility) is available over a sufficiently long period. The distinction between vehicle mobility and traveller mobility is related to travelling as a passenger or as a driver (for convenience, in this chapter the notion "driver" also includes motorcycle rider).
The restricted availability of mobility data fixes the choice of stratification of accident data. Therefore, firstly, the available mobility data are described. Secondly, the choice of accident data is explained.

## Mobility data

Between 1950 and 2000, random sample traffic data for car, motorcycle and truck was gathered in the Netherlands, see also Section 3.2. Statistics Netherlands (CBS) used these data to estimate the annual vehicle distance travelled. For other traffic modes (pedestrians, cyclists, mopedists) no data are available until 1985. From that year on, Statistics Netherlands and Transport Research Centre AVV started a large National Travel Survey (SWOV/NTS, 2012) among households to collect data on individual mobility (distance travelled by individuals), for all relevant traffic modes except vans and trucks. These data differ slightly from the vehicle mobility data (CBS, 2000) between 1950 and 2000. For trucks and vans, no mobility data have been available since 2000. Section 3.6 will look into the traffic mode risk of cars, motorcycles and trucks, between 1950 and 2000, because the traffic data for the traffic modes is available. This way, only a small part of the total number of fatalities is analysed. In 1950, this is $25 \%$, nowadays this is $60 \%$ of the total.

Figure 3.5 shows the data available of the mobility of cars, motorcycles and trucks between the years 1950 and 2000. The figure shows very different patterns: whereas car mobility grew explosively ( $14 \%$ annually) from 1950, and showed a much smaller growth after 1970 ( $3 \%$ annually), truck mobility has increased more or less continuously. Since 1985, truck mobility (including small vans) has increased by $6 \%$ annually. The use of motorcycles has known several periods of great popularity: about 1955, about 1975 and about 1990.

## The number of fatalities among drivers for various traffic modes

When using mobility data to calculate fatality risk per traffic mode, we must bear in mind that changes in mobility may not be the best measure to explain changes in the number of fatalities. The reason being that vehicle mobility is different from person mobility, as is well known and explained, for example by the ETSC (1999). When a car (or other vehicle) is occupied by more than one person, the probability of a fatality is increased when that vehicle is involved in a serious accident. So when the average number of vehicle occupants changes over time, this may also cause a change in the ratio between mobility and the number of fatalities.

Unfortunately, the traveller mobility of before 1985 is not known in the Netherlands. From that year onward, annual surveys among households were carried out in which people's mobility data was collected as described
above. Before that year, only vehicle traffic data was gathered. This leaves us with two options. One is to relate vehicle mobility to fatal accidents. The other is to relate vehicle mobility to fatalities among drivers. We have chosen the latter. This way we can analyse passenger fatalities separately, which, more so than just analysing fatal accidents, enables a better understanding of the numbers of fatalities. In the next section we will only refer to fatalities among drivers, and compare these to vehicle mobility data. Fatalities among passengers are left out of the analysis for the moment.


Figure 3.5. Mobility (vehicle distance travelled) for cars, trucks \& vans and motorcycles in the Netherlands, between 1950 and 2000. The data on car mobility are identical to those in Figure 3.2

All necessary data on fatal accidents for this stratification were available. The digitally available accident record data of the Ministry of Infrastructure and the Environment (SWOV/VOR/BRON, 2012) since 1976 allow for any stratification of interest. For the period between 1950 and 1976 the CBS' tables (CBS, 1950 and later years) contain the number of driver fatalities per traffic mode.

In Figure 3.6, the numbers of driver fatalities are given for cars, motorcycles and trucks. We use a logarithmic axis to visualize the strong exponential character of the increase in the number of car fatalities. The figure shows several characteristics of the number of driver fatalities. There are two periods of different exponential increase for car driver fatalities. Such
exponential behaviour cannot be observed in the development of motorcycle rider and truck driver fatalities. These data each show their own periods of increasing, decreasing and again increasing numbers. The figure also shows that before 1960, more motorcycle riders were killed than car drivers. Finally, the number of killed truck drivers is very low compared to the other two modes.


Figure 3.6. Fatalities among drivers of car, motorcycle and truck. The solid lines are exponential lines of constant annual change for car fatalities between 1950 and 1970, and between 1970 and 2003, with the annual change.

## Total road safety

In order to analyse road safety in its entirety, all fatalities must be included. Thus, all traffic modes need to be analysed. Furthermore, fatalities among passengers must be taken into account. This can be done by using the average number of passengers per driver, which is not constant over the years. Before 1960, more than half of the fatally injured car occupants were passengers. By restricting the analysis to drivers, we leave out more than half the fatalities for cars in that period. This is much less the case for motorcycles, but for trucks we see the same phenomenon. Therefore, passengers certainly must be taken into account for a complete analysis of the road safety development. However, in this chapter we will focus on understanding the influence of mobility on safety. We will exclude passengers, as they are to be considered a separate factor.


Figure 3.7. Traffic mode risk for drivers. The number of driver fatalities divided by vehicle distance travelled for three traffic modes in the Netherlands since 1950. Solid lines show exponential lines of constant annual change, with the annual change as a percentage.

### 3.6. Traffic mode risk

Figure 3.7 shows the numbers of driver fatalities for cars, motorcycles and trucks per distance travelled. By calculating this ratio, we effectively attribute changes in the number of fatalities to changes in mobility (either because of changes in the number of drivers, or because of changes in the distance travelled per driver). This ratio will be called the driver risk. Several remarkable properties can be observed.

- Between 1950 and 1970, the driver risk increases very slightly, or, at best, remains constant for all three traffic modes. Between 1970 and 2000 the driver risk decreases exponentially, for all three modes. The annual change is more negative than $-5 \%$ for motorcyclists and car drivers, but is only $-2.6 \%$ per year for van and truck drivers.
- The risk is 13 times higher for motorcyclists than for car drivers. This factor has practically remained constant for 55 years.
- The driver risk for cars is 2 to 3 times higher than that for truck and van drivers. This factor has decreased over the years.
- The risk for motorcyclists shows a sudden factor 2 increase about 1970.

This chapter will not provide elaborate explanations for all these observations. However, a few possible explanations will be given in the discussion section. Here we will only suggest possible explanations for the difference between the development of the calculated general risk in Figure 3.3, and that of the stratified risks in Figure 3.7. This difference mainly consists of two aspects. The general risk shows no sudden cusp about 1970, whereas the drivers' risks do. The general risk decreases more or less constantly until about 1980. From then on, the annual decrease of the risk becomes ever smaller. After 1970, all driver risks decrease with a constant rate which does therefore not become smaller with time.

Essentially, the difference between the general risk in Figure 3.3 and the stratified risk in Figure 3.7 is a consequence of the over-simplifying assumption that car kilometres can be used as an overall exposure measure. Risk is a ratio, namely between the number of fatalities and mobility. The numerator of general risk is a sum of two parts: fatalities in accidents in which a car is involved, and all other fatalities. The denominator just consists of car mobility. The two parts of the numerator evolve differently in time. Between 1950 and 1970, car occupant mobility and fatalities increase strongly, and the other part shows a much smaller increase, because mobility of other traffic increases less. Dividing all fatalities by (strongly increasing) car mobility means that both car fatalities and all other fatalities are divided by strongly increasing car mobility. For the fatalities in accidents in which no car is involved, this gives a decreasing contribution to the general risk. This contribution is meaningless, as it is the ratio between fatalities that are not influenced by car mobility, and the increasing car mobility.

As we saw, the risk associated with motorcycle mobility is much higher than that for car mobility. (As we will see in the next section, this also holds for fatalities among pedestrians in accidents with either cars or motorcycles). While the risk for both cars and motorcycles is constant, cars travel the largest distance, and did so even in 1950, modal shift from motorcycle to car generates a general risk decrease. The number of fatalities among motorcyclists decreases because of this modal shift. The same occurs with other fatality types that have nothing to do with cars (pedestrian-truck, bicycle-truck etcetera). Their number has decreased over time, whereas this had nothing to do with the car distance travelled. Essentially, in calculating general risk we divide all fatalities, of which a large and decreasing group has nothing to do with cars, by increasing car mobility. This results in a strongly decreasing general risk, which has little meaning.

### 3.7. The role of the other party

We saw that the development of the number of fatalities depends on the mobility of the victims. However, the number of fatalities among road users with a certain modality depends also on the number and type of other traffic (the potential other parties). For example, the number of fatalities among pedestrians can be expected to have increased between 1950 and 1970 because of the strongly increasing car mobility. To understand the development of the number of fatal crashes between a pedestrian and a car in terms of mobility changes, pedestrian mobility and car mobility are equally important. The same holds for crashes of other combinations of traffic modes.

Some modalities bear a higher risk for other road users than others (e.g. accidents with a truck as other party are more often fatal than accidents with a car). It is therefore necessary to analyse crash data not only by taking traffic mode of the fatality into consideration, but also the traffic mode of different other parties involved. The goal of Section 3.9 is to show that the development of risk also depend on the modality of the other party.

In order to achieve this goal, we will now make another stratification, namely by fatality type (see Section 3.8). We adopt the following definition of a fatality type, which can also be found in the OECD report on achieving ambitious road safety targets (OECD, 2007).

A fatality type $F(P-Q)$ is a fatality in a crash between traffic modes $P$ and $Q$, where a user of traffic mode $P$ was a fatality. A fatality type $F(P)$ is a singlevehicle crash with a fatality in mode P. Fatality types of the type motorcyclecar are therefore fatalities among motorcyclists in a crash with a car. Fatality types of the type "car single-vehicle" are fatalities among occupants of a car, in a crash where no other vehicle was involved (i.e. run off road or a collision with an animal or loose object).

Fatal accidents in which more than two road users are involved (approximately 5\%), are treated as two mode accidents. The fatality is in the first mode, and the second mode is the heaviest of the remaining ones.
We also have to define accident type. An accident type $\mathrm{A}(\mathrm{P}-\mathrm{Q})$ is an accident where either a user of traffic mode $P$ or traffic mode $Q$ is killed. For a stratification of fatality data, the fatality type is to be preferred above accident type. This is because two fatality types contribute in an accident
type. Both may have a different development in time, which becomes invisible in accident types. Also, measures to improve vehicle safety for car occupants will not help much in car-pedestrian accident, because the victims are mostly pedestrians.

### 3.8. Data stratified by fatality type

In this section we will choose some possibly interesting fatality types for further analysis. We concentrate on crashes where a car or motorcycle is involved because of the availability of mobility data (see Section 3.5). We analysed single-vehicle accidents, accidents in which the cars and motorcycle occupants were victims of a crash with a (another) car and accidents in which the car or motorcycle is the other party of a crash with a pedestrian.

We again used the Statistics Netherlands accident data as a basis for fatality types. The data are available from 1950, and are digitally available since 1976. The digitally available data supply us with detailed information, which enable us to distinguish between conflicts $\mathrm{P}-\mathrm{Q}$ and $\mathrm{Q}-\mathrm{P}$ (fatality types). The data from the period 1950-1976 (CBS, 1950 and later years) enable us to count the number of fatal accidents of relevant combinations of traffic modes, but we cannot know (for each combination) whether the fatality is in traffic mode P or Q (accident types), or whether it is a passenger or a driver. This complicates a thorough analysis of many fatality types. Therefore, in this chapter, we will concentrate on fatality types for which it is almost certain which traffic mode has the fatality. The following fatality types were analysed:

1. car single-vehicle
2. motorcycle singe vehicle
3. pedestrian-car
4. pedestrian-motorcycle
5. car-car
6. motorcycle-car

By choosing these fatality types we again effectively restrict the share of fatalities we research. As compared to the analysis in Section 3.5 we have added two fatality types concerning pedestrians, but at the same time we have left out fatalities with trucks, and fatality types such as motorcycletruck or car truck. Effectively, the number of fatalities included in this selection of six fatality types is approximately $25 \%$ in 1950 and $50 \%$ in 2005.

We made the following assumptions on the accident type data before 1976 in order to derive fatality type data from accident data:

1. For pedestrian-car and pedestrian-motorcycle accident types the victim is practically always the pedestrian. Between 1976 and 2006, only 4\%o and $5 \%$ of the fatalities in pedestrian-car and pedestrian-motorcycle accidents, respectively, are car or motorcycle occupants.
2. For motorcycle-car accident types the victim is always the motorcycle rider. Between 1976 and 2006 only 5\% of the fatalities in crashes between a motorcycle and a car are car occupants.

By analysing the number of fatal accidents (and discarding the number of fatalities per accident), the difference between passengers and drivers can no longer be made. This means that part of the risk decrease for fatal car singlevehicle crashes and even more so for car-car crashes can be explained by the decreasing car occupation. For motorcycle crashes this effect is much smaller, because there are not many motorcycle passenger fatalities.
First, Figure 3.8 shows the number of fatal accidents for each fatality type. The graphs for the first three fatality types, on the left side of Figure 3.8, show the exponential lines of constant annual change for 1950-1970, and later years. Each of these three graphs has a sharp cusp in the number of fatalities somewhere between 1970 and 1973; this cusp is the sharpest for carcar accidents. The graphs for car single-vehicle and pedestrian-car (top and bottom left panel) are quite alike, except that they are tilted with respect to one another: pedestrian-car fatalities did at first increase more slowly, and later decrease much faster than car single-vehicle accidents.

The sharpness of the cusps can be expressed as the difference between the two corresponding values of annual change. The results show that for car-single-vehicle and pedestrian-car conflicts, the cusp equals $13 \%$ to $14 \%$. This value is comparable with (but slightly larger than) the value for mobility or fleet size (circa $11 \%$ ). For car-car crashes the cusp is $22 \%$, which is almost twice this value.

The fact that these cusps all coincide in time suggests that the cusps in the number of fatalities stem from the cusps in mobility (see Figure 3.5), i.e. from the explosive increase until 1970, and subsequent calming of the growth. The fact that the cusps in the number of fatalities do not exactly equal those of mobility, may indicate that the relation between mobility and fatalities is not just a linear proportionality, for reasons that are yet not well understood.

Accident types involving motorcycles all show the characteristic triple wave. For motorcycle single-vehicle accidents, this triple wave is superimposed on a decreasing trend, which is even stronger for pedestrian-motorcycle fatal crashes. Here, also, the two graphs are somewhat alike, but tilted: pedestrian-motorcycle accidents decrease in number much faster than motorcycle-single-vehicle accidents. Motorcycle car-accidents follow a triple wave without any decreasing trend since 1970. The triple wave is also visible in the motorcycle mobility (see Figure 3.5).

### 3.9. Fatality type risk

Figure 3.9 shows the fatality type risk. For single-vehicle fatalities (top row), this is straightforward, because there is no other traffic mode involved. Therefore, these accidents can only be influenced by the mobility of the traffic mode involved. For accidents involving pedestrians (second row), this is more difficult, as the number of pedestrians, or the pedestrian distance travelled, may have varied (of which there is only little data). In the Netherlands pedestrian distance travelled has been observed since 1985, and this data shows no significant change over the years. Therefore, we calculated the risk for a car driver or motorcycle rider to be involved in a fatal accident with a pedestrian. This is the number of fatal pedestrian accidents per car km (or motorcycle km ). In doing this, we obtain a calculated risk in which a possible influence of changes in pedestrian mobility is implicitly incorporated. In other words: the result may be influenced by changes in pedestrian mobility. Even without knowing the influence of pedestrian mobility changes, the results enable comparison between the risks for motorcycle and car. For the accidents between two cars, or cars and motorcycles (bottom row), we again simply divide the number of fatalities by mobility of the victim (car mobility for car-car accidents and motorcycle mobility for motorcycle-car) accidents.

Three interesting phenomena occur for car accidents. For car single-vehicle we see that from the $13 \%$ cusp in the number of fatal accidents, a small cusp of $3.8 \%$ remains in the risk. Between 1950 and 1970 there is hardly any risk decrease ( $-1.3 \%$ annual change). This small risk decrease may be explained


Figure 3.8. The number of fatal accidents since 1950 in The Netherlands for six fatality types. In the left panels, solid lines represent exponential lines of constant annual change between 1950 and 1970 and after 1970, with the corresponding annual change as a percentage.


Figure 3.9. The development of fatality type risk: fatal accidents in a fatality type divided by distance travelled by car (left panels) or by motorcycle distance travelled (right panels) since 1950. Solid lines are exponential lines of constant annual change between 1950-1970 and 1970-2000, with their annual change as a percentage. The $\Delta$ shown in the right panels equals the factor between the results for the two lines of constant annual change for 1970.
by the decreasing fraction of car passengers killed. For pedestrian-car accidents, the remaining cusp equals $2.8 \%$. We also see that the risk has a much steeper decrease than that for single-vehicle car fatalities. Fatal car-car accident risk has a strong cusp (11\%), which is expected to be a consequence of the effect of mobility of "the other car" as suggested above. It is easily understood that the $11 \%$ cusp in mobility exactly explains the $22 \%$ cusp in car-car fatalities, when we accept a quadratic relation between car mobility and car-car fatalities. In Section 3.10 we give this possibility some further thoughts.

For accidents involving motorcycles, the risk development also shows a cusp between the two exponential slopes before and after 1970. The difference in slopes before and after 1970 amounts to approximately $6 \%$ for single-vehicle and pedestrian accidents; significantly more than that for car accidents. For motorcycle-car accidents, the cusp equals almost $9 \%$, which is slightly less than that for car-car accidents. Besides the change in slope, there is also an upward jump ( $\Delta$ ) of approximately a factor 2 in each of the three calculated risks. This factor was calculated by estimating the values for both lines of constant annual change in 1970. As this upward jump is visible for all three fatality types concerning motorcycles, it seems to indicate a cause that is specific for motorcyclists in all circumstances. In the discussion we shall propose an explanation.

Comparison of single-vehicle and pedestrian accidents shows that the risk of pedestrian accidents decreases much faster than that of single-vehicle accidents. For pedestrian-car and car single-vehicle the difference is a $5 \%$ to $6 \%$ extra decrease for pedestrian accidents, both between 1950 and 1970, and after 1970. For pedestrian-motorcycle and motorcycle single-vehicle accidents, the difference is $2 \%$ to $3 \%$. It is unclear if this difference is significant. In 50 years, for both crash types, safety has improved much more for pedestrians than for car occupants or motorcyclists. Therefore, in addition to safer driver behaviour, passive vehicle safety improvement, speed limits, safe motorways and many other changes and measures that (also) improve driver safety, there must have been other influencing factors which are especially effective in preventing fatal accidents among pedestrians. Section 3.3 already showed that demographic changes are insufficient to explain this phenomenon. A more probable explanation is that the pedestrian mobility has changed strongly. In other words, people may walk less now than they did in the fifties. Unfortunately, this possibility cannot be easily verified as there are no accurate data on pedestrian mobility available up until 1985.

However, the authors assume that improvements of child safety, e.g. by child education, but to a large extent by separating pedestrians and motorized traffic, (safer playground areas, safe school areas, etcetera), has made an important contribution to the improvement of pedestrian safety.

A final interesting notion follows from a comparison between the general risk in Figure 3.3 and the six fatality type risks in Figure 3.9: All six fatality type risks are different, and none is really alike the general risk development. Whereas the general risk decreases strongly before 1970, four of the six fatality risks hardly decrease or even increase. Also the change in general risk about 1985 is not reflected in any of the fatality type risks. There may be two explanations, namely that either the fatalities not included in the six fatality types have become more important, or that the ratio between for instance single-vehicle car fatalities and car-car fatalities is gradually changing. In fact, the latter is certainly the case, as the number of single-vehicle fatalities is decreasing more slowly than the number of car-car fatalities. So if these two were the only fatality types, it would result in the annual general risk change becoming more and more equal to the lesser of the two annual risk changes, in this case the single-vehicle car fatalities. This notion illustrates the main point of this chapter, namely that stratification of road safety data and mobility data enhances understanding of the development on fatalities and risk.

Concluding, fatality type risks for a pedestrian, a car occupant or a motorcycle driver indeed strongly varies between modalities of the other party involved. Therefore, the other party plays an important role. To understand the influence of mobility on the development of road safety it is therefore essential to include the development of mobility of different traffic modes, both for the traffic mode of the victim and the traffic mode of the other party.

### 3.10. Conclusion and discussion

## Conclusions

This chapter shows that the development of road safety cannot be understood by general risk. Instead, the risk per traffic mode should be used to study road safety. The values of traffic mode risks differ widely from the simple calculation of fatalities per car distance travelled (general risk). Whereas, the general risk shows a decrease between 1950 and 2005, the traffic mode risks for cars, trucks and motorcycles have hardly decreased until 1970.

The total number of fatalities does not follow the development of car mobility. Furthermore, the changes in mobility are not the same for cars, trucks and motorcycles. Thus, the general risk is difficult to interpret, and for that reason it is not a good indicator of what has really occurred.

This chapter also shows that the mobility of the other party (if there is any) also plays an important role in the number and severity of the accidents. This role is necessary to understand the development of road safety. The fatality type risks of two fatality types of which the victim is in the same traffic mode (e.g. motorcycle single-vehicle and motorcycle - car accidents) show a different course over time. Therefore, accident data should not only be stratified by the traffic mode of the victim, but also by the traffic mode of the other party (the combination of the two traffic modes defines the fatality type).

## Stratification

This chapter discusses the relation between mobility (distance travelled) and road safety (traffic fatalities). The key to a meaningful analysis of the relationship between mobility and road safety is the notion that overall numbers (of road safety or mobility) do not give a true understanding of the nature of the influence of mobility on road safety. It has been shown that stratification of traffic modes is vital to understanding the development of risk. It is not possible to analyse risk data (fatalities per distance travelled) in a meaningful way without stratification by traffic mode of the victim, and of the other party. However, it might be necessary to stratify by other characteristics of accidents and traffic such as infrastructural differences, time of day, age and gender of the driver (or drivers). These stratifications may all come with different risks. In this chapter, we choose for stratification to traffic mode as a first step for two reasons. First, our hypothesis was that this stratification is very important, which has been demonstrated. Secondly, data are available for this stratification.

## Single-vehicle car accidents

This section goes into the question why the single-vehicle car risk hardly decreased before 1970. This phenomenon also occurs for the single motorcycle accidents. For motorcycle-car and car-car accidents, we even see an increase of risk in this period. These unexpected developments of risks may stem from the same cause. Here, we discuss possible explanations for the risk development of single car accidents as an example of this phenomenon.

The obvious hypothesis for this unexpected development of risk would be that in the Netherlands hardly any road safety measures were taken before 1970, whereas a stream of measures was initiated afterwards. It is true that many effective safety measures, which continuously improved road safety, were taken after 1970. Alcohol legislation (1972), a speed limit outside rural areas (1974), safety belt legislation (front passengers: introduced in 1956, obligatory since June 1975; rear passengers: introduced in 1967, obligatory since April 1992), crash helmets for motorcyclists (1972) and mopedists (1975), truck speed limits ( $80 \mathrm{~km} / \mathrm{h}, 1974$ ) and truck speed delimiters (1995), are credited with (strongly) reducing the risk. Thus, a sudden change in safety policy in the early 1970s, might explain the bend in risk development of single-vehicle car accidents about 1970. Such a change in safety policy in that period must actually have been the beginning of a permanent and constant risk decrease, whereas before this change in road safety policy supposedly nothing has been done to improve safety. The registration of road safety data was not extensive before 1970 (which can be interpreted as a good support for such a hypothesis). What we know can be expressed in the following observations:

- Some important safety measures, which were expected to be effective, were taken in the 1950s. For example the drivers licence was introduced in 1952, and the speed limit in urban areas ( $50 \mathrm{~km} / \mathrm{h}$ ) was introduced in 1957 (SWOV, 2012). In that time, case studies were not carried out. The accident data shows no clear effect of these measures as it does not show any effects of most present measures. For example, data on the number of trucks with or without speed delimiter, on the number of fatalities related to trucks with a speed delimiter, or data on the role of speeding in truck accidents, are hard to get by. Therefore, it is difficult to transpose the effect of these measures to a national effect.
- Gradual improvement of active and passive vehicle safety is thought to have begun well before 1970. Safety belts were introduced in 1956, and the automobile industry started working on energy absorbing zones, collision avoidance systems etcetera well before 1970. Volvo claims to have introduced a safety cage as early as 1944 and the three point safety belt in 1959. These and other measures can be expected to have gradually penetrated the car fleet in the fifties and sixties (Volvo, 2007).
- Many other, not accurately administrated, measures like crossing zones, bicycle infrastructure, walking zones, traffic lights, roundabouts etcetera were introduced, and are supposed to have contributed to safety as well.
- The explosive growth of car mobility between 1950 and 1970 was accompanied by an almost equally strong increase of motorway road length. Although it is unknown which part of the total mobility took place on motorways, rural or urban roads (even today the necessary data are not available in the Netherlands), the increase in motorway road length indicates that an increasing part of the growing mobility must have moved to these motorways. Whereas most of the fatal accidents in the Netherlands occur on rural $80 \mathrm{~km} / \mathrm{h}$ roads and on urban roads, an increasing amount of the growing mobility takes place on motorways. This leads to the expectation that increasing motorway road length should reduce the risk. The fact that the risk starts to decrease when the motorway length ceases to grow, contradicts this expectation.
- The changing number of passengers in car accidents can also not explain the remaining cusp in risk. Both before and after 1970, the fraction of passengers killed in car accidents decreased by a constant pace of approximately $1.4 \%$ annually. This decrease is expected to be responsible for a part of the decrease in risk, but it cannot explain the change of this decrease about 1970.

We reject the hypothesis that the minor decrease in risk of single car accidents before 1970 is caused by a lack of safety measures before 1970. It is unlikely that all the above-mentioned measures and developments of the traffic system have had no safety effects until 1970. Although it is plausible that a change in safety policy indeed caused an extra decrease in risk, this still does not explain the lack of any risk decrease before 1970. Nor is there a straightforward explanation for the risk decreasing by a constant rate after 1970.

It is conspicuous that the bend in the development of the single-vehicle car risk coincides with the cusp in mobility. This observation might be a clue to an explanation. It is possible that the bend in single-vehicle car risk may be caused by the strongly increasing mobility itself. Oppe (1991a) already pointed out that his calculated general risk tends to depend on distance travelled. He stated that the risk also seems to depend on changes of the distance travelled. The simple fact that the cusp in car distance travelled coincides with the cusps in fatalities, suggests a correlation between the two. As the calculated fatality type risk still contains some cusp, this suggests that there is some remaining influence of changing mobility left.

The authors believe that the bend in the development of the single-vehicle car risk may be explained by driver experience. Before 1970, when mobility increased strongly, the mean level of driver experience must have been significantly lower than in later years. Although until 1965 the age of killed drivers in the Netherlands was uniformly distributed between 20 and 55 years of age, in that period these drivers could not have gathered much experience in the years prior to their fatal crash. After the strong exponential increase of car mobility came to an end about 1970, drivers became more and more experienced, while in the period before 1970, this was not the case. For the upward jump in risk for motorcycle drivers about 1970, the authors suggest a similar explanation. This jump appears to coincide with a sudden sharp increase of young (and inexperienced) driver fatalities, whose enhanced risk may explain the risk jump. This follows from data on the age of killed motorcyclists, which are available for the entire period between 1950 and 2005. This indicates that further stratification of mobility and fatality data by driver age (and experience) may be necessary to further understand the data. Thorough analyses of these possible explanations will be necessary.

## The difference in risk between single-vehicle car and car-car accidents

There is a large difference in risk between single-vehicle car and car-car accidents. The cusp (measured as the difference in annual change) in the number of fatal crashes is almost twice as large for the car-car accidents as for the single-vehicle car accidents ( $22 \%$ versus $14 \%$ ). Division by mobility gives a resulting cusp of $-3.8 \%$ for single-vehicle and $-11.1 \%$ for car-car. It is possible that this difference could be due to a roughly quadratic effect of mobility. In single-vehicle car accidents only the mobility of the victim plays a role. However, in car-car accidents there is also the mobility of the other party. Chances of having a car-car accident increase when the number of other parties that a car encounters during its trips, increases. In dense traffic, driver failures like fatigue, impaired driving or other causes of loss of control are more likely to lead to a car-car crash than in quiet traffic. The mobility of the other party therefore influences risk. Actually, not the mobility of the other party should be taken into account, but the traffic volume which is the mobility divided by the road length. It is desirable to find a way to include the quadratic effect of mobility and the effect of road length in calculating the risk for car-car accidents. Stratification by road type is expected to be important there.

The effect of safety measures may act quadratically as well on the number of car-car fatalities. Thus, when it would appear that e.g. the number of car-car
fatalities, corrected for a quadratic effect of mobility decreases faster than the number of car single fatalities, this might be attributed to such a quadratic effect.

## Acknowledgements

The authors are indebted to Vincent Kars and Niels Bos, who both contributed invaluably to making Dutch accident and mobility data available for research. They also thank prof. dr. Ben Ale from Delft University of Technology and Hansje Weijer from SWOV for reading and improving the text. Henk Stipdonk wishes to commemorate Bob Roszbach (1945-2005), who, in a private communication, expressed his doubts about the traditional concept of fatalities being just the product of risk and mobility (exposure). He based this on the insight that the exposure to risk (i.e. mobility) of one driver may be the cause of risk of the other driver. His doubt was a starting point for the research that led to this paper.

## References

Broughton, J., Allsop, R., Lynam, D., McMahon, C., 2000. The numerical context for setting national reduction targets. Transport Research Laboratory, London.
CBS, 1950 and later years. Statistiek van de verkeersongevallen op de openbare weg 1950. Dutch tabulated data on Dutch traffic accidents. In: CBS ed. Centraal Bureau voor de Statistiek, Den Haag.
CBS, 1985 and onward. Statline, motorvoertuigen: Statistiek van het Nederlandse motorvoertuigenpark. Centraal Bureau voor de Statistiek CBS, Voorburg/Heerlen.
CBS, 2000. Statline, statistiek van de wegen. Centraal Bureau voor de Statistiek CBS, Voorburg/Heerlen.
CBS, 2012. Statline. Centraal Bureau voor de Statistiek, Den Haag.
ETSC, 1999. Exposure data for travel risk assessment: Current practice and future needs in the EU. European Transport Safety Council, Brussels.
Filarski, R., 2004. The rise and decline of transport systems, changes in a historical context, Ministry of Transport and Public Works, Rotterdam.
Fournier, F., Simard, R., 2002. The DRAG-2 model for quebeck. In: Gaudry, M., Lassarre, S. eds. Structural road accident models; the international DRAG family. Pergamon, pp. 37-66.
Gaudry, M., 2005. Presentation at 1st FERSI scientific road safety research conference, BASt, Bergische Gladbach.
Kopitz, E., Cropper, M., 2005. Traffic fatalities and economic growth. Accident Analysis and Prevention 37, 169-178.
OECD, 2007. Towards zero, ambitious road safety targets and the safe system approach, OECD/ITF, Paris.
Oppe, S., 1991a. The development of traffic and traffic safety in six developed countries. Accident Analysis and Prevention 23 (5), 401-412.

Oppe, S., 1991b. Development of traffic and traffic safety: Global trends and incidental fluctuations. Accident Analysis and Prevention 23 (5), 413-422.
Pooley, C.G., Turnbull, J., 2000. Modal choice and modal change: The journey to work in Britain since 1890. Journal of Transport Geography 8, 11-24.
SWOV, 2012. Website database. data on crashes, hospital patients, distance travelled, fleet size, demography, safety measures, etcetera. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam.
SWOV/NTS, 2012. Website database. Dutch data on distance travelled. Adviesdienst verkeer en vervoer/Centraal Bureau Statistieken.

SWOV/VOR/BRON, 2012. Website database. Dutch annual road crash data. Dienst Verkeer en Scheepvaart, Ministry of Infrastructure and Environment.
Volvo, 2007. Safety first, a list of safety measures introduced by Volvo. www.volvocars.com.

## 4. Analysing the development of road safety using demographic data

This chapter was originally published in "Accident Analysis and Prevention" (Stipdonk, H., Bijleveld, F., Van Norden, Y., Commandeur, J., 2012. Analysing the development of road safety using demographic data. Accident Analysis and Prevention, in press (http://dx.doi.org/10.1016/j.aap.2012.08.005).


#### Abstract

The purpose of this chapter is to show that time series analyses of road safety and risk can be improved by using demographic data. We demonstrate that the distance travelled by drivers or riders of a certain age reflects the fluctuations over the years of the number of people of that age within the population. We further demonstrate that the change over time of per capita distance travelled, i.e. distance travelled per person, is often less subject to stochastic fluctuations, and therefore more smooth than the total distance travelled for drivers of that age. This smoothness is used to obtain forecasts of distance travelled, or to average out year-to-year fluctuations of data of distance travelled. Analysis of such data stratified by age group, gender or both reveals that, for most travel modes, per capita distance travelled is to a large extent constant or slowly changing over time. The consequences for the evaluation of risk, i.e. casualties per distance travelled, with and without the use of population data, are explored. Dutch data are used to illustrate the model concept.


It is shown that the analyses and forecasts of distance travelled could gain substantially by incorporating demographic data, as compared to an analysis with data of distance travelled alone. The chapter further shows that, for an analysis of risk and therefore for traffic safety forecasts in the absence of any data of distance travelled, stratified analysis of mortality, i. e. casualties per inhabitant, may be a reasonable alternative.

### 4.1. Introduction

The development of traffic safety can be described by the development of distance travelled and the development of traffic risk. Here, the risk is defined as the number of crashes or fatalities per distance travelled (COST, 2004, Broughton, 2012). Risk itself is not observed directly. Instead, time series of the number of casualties and of distance travelled are needed in order to calculate risk and analyse its development.

It is widely known that different travel modes may have different risks (Elvik and Bjørnskau, 2005, Schofield et al., 2008, Stipdonk and Berends, 2008). Other factors such as age and gender of the driver, or time of day of the road trip, are also known to influence the risk (Massie et al., 1997, Kam, 2003). It is well-known, for example, that young inexperienced car drivers are more at risk of being involved in traffic accidents than more experienced car drivers (McCartt et al., 2009), while elderly drivers also have a higher risk in traffic (Davidse, 2007). In general, moreover, male drivers also have a higher risk than female drivers (Laapotti and Keskinen, 2004, Lenroot et al., 2007, Vlakveld, 2011), (Van den Bossche, 2006, Van den Bossche et al., 2007). Hence, when time series of crash data are studied, stratified risk analysis is relevant.

Changes in overall risk, i.e. the number of all casualties divided by the total distance travelled, can be influenced by a shift from unsafe travel modes to safer travel modes, from unsafe roads to safer roads, or vice versa. Such changes may also be due to changes in circumstances or safety measures, in which case these circumstances or safety measures usually only apply to a subset of all crashes. Thus, to understand the effects of these latter changes on overall risk, stratification may be required.

When analysing or forecasting the number of traffic casualties, stratification by age or transport mode has the drawback that it yields both smaller numbers of casualties, and less accurate data of distance travelled, if available at all. Further, data of distance travelled are often subject to random fluctuations and systematic changes in level, especially when stratified by e.g. different transport modes, driver age, etcetera. The purpose of the research described in this chapter is to show that stratified time series analyses of road safety and risk can be improved by using demographic data in order to average the random fluctuations and explain some of the changes in the level of distance travelled.

In theory, there are many ways in which the overall traffic risk can be stratified. However, for a successful stratification both crash data and data of distance travelled must be available for the stratifications of interest. For many countries, the available data of distance travelled does not allow for stratification by many variables (Yannis et al., 2008). In the Netherlands, for example, it is not possible to stratify by speed limit of the road, because such data are not collected. From the few available variables in the Dutch travel survey, we decided to stratify by travel mode, age and gender, for the reasons just mentioned.

The use of demographic data and per capita distance travelled may help to estimate the actual distance travelled, if a constant or a (log)linear time dependent change of per capita distance travelled can be assumed. In many developed countries, demographic data and demographic forecasts are well documented (European Union, 1995-2011). The development of demographics in the past is very accurately known for many countries, and even future demographic developments are well studied and the methods used are still improving (Booth, 2006). Differences of the population distribution over age for different countries may be part of an explanation of differences in the development of distance travelled by age and of the number of casualties of these countries, see Figure 4.1.

Figure 4.1 shows the number of inhabitants of four European countries, by age year, for 2010 (except for Belgium where the data are displayed for 2008). The figure shows that every country has a population with a characteristic, somewhat whimsical demographic structure, showing specific low values (corresponding to a birth year during World War I or II) or high values (e.g. post war birth years in Italy and the Netherlands), and also sharp and sometimes sudden changes in the population size over consecutive age years. Differences in demographic developments between countries may offer a partial explanation for differences in the development of the number of casualties in different countries, such as described in Yannis et al (2011).

In the next section we first describe how the number of casualties, distance travelled, population size, per capita distance travelled, risk and mortality (or morbidity) are related. Section 4.3 describes how we used population data to modify data of distance travelled, either through stratification by age group (Section 3.1) or by individual age year (Section 3.2). In Section 3.3 a model for risk development is described, stratified by age group. In this chapter we do not discuss the risk model stratified by individual age years, as this was done elsewhere (OECD/ITF, 2012). Section 4.4 presents the results of the models, and the chapter ends with a discussion (Section 4.5) and conclusions (Section 4.6).


Figure 4.1. Population by age year, for four countries. For Germany, Italy and the Netherlands, the population size on January 1st of 2010 is given, while for Belgium this is 2008. Data were extracted from Eurostat (European Union, 1995-2011).

### 4.2. Theoretical background

In this section we describe the relation between the relevant quantities, such as the number of casualties $N_{a, g, m}(t)$, distance travelled $T_{a, g, m}(t)$, population size $P_{a, g}(t)$, per capita distance travelled $\tau_{a, g, m}(t)$ and risk $r_{a, g, m}(t)$. All these quantities are given as a function of time $t$, in years, and for a specific age (or age group) $a$, gender $g$ and travel mode $m . T_{a, g, m}(t)$ can be written as

$$
\begin{equation*}
T_{a, g, m}(t)=\tau_{a, g, m}(t) \cdot P_{a, g}(t), \tag{Eq. 1}
\end{equation*}
$$

Thus, the development of distance travelled $T_{a, g, m}(t)$ either corresponds to a change in per capita distance travelled $\tau_{a, g, m}(t)$ or to a change in the population size $P_{a, g}(t)$ (or both).

Eq. 1 may be used to obtain a forecast of $T$, in case there are no dedicated models available to obtain forecasts using travel behaviour parameters (Nicholson and Wong, 1993). Forecasts of $T$, especially if stratified by $a, g$ and $m$, may be complex, as for many countries, $P_{a, g}(t)$ appears to be a somewhat whimsical but known function of time (see Figure 4.1). However, in the case
that $\tau_{a, g, m}(t)$ is a smooth function of time, forecasts of $T_{a, g, m}(t)$ by forecasting $\tau_{a, g, m}(t)$ and applying the known information on $P_{a, g}(t)$ are straightforward.

Survey data of distance travelled are often binned into (age) groups, in order to average statistical fluctuations of small groups. As a consequence, distance travelled differences within the group remain unobserved, suggesting a uniform value which is not necessarily true. An alternative to this binning is smoothing. When $\tau_{a, g, m}$ is smooth with respect to age, i.e. $\tau_{a, g, m}$ is slowly varying with variations of $a$, survey data stratified by age or age groups may be processed by a smoothing method. A smooth behaviour of $\tau_{a, g, m}(t)$ with respect to $t$ and $a$ may result in a more realistic estimation of $\tau_{a, g, m}(t)$. Hence, smoothing over age is a possible alternative to binning.

When forecasts of the number of casualties are to be assessed, a general method is to assume risk $r$ to be a log linear function of time $t$, and for distance travelled to apply a separate prognosis, that may or may not be based on a (log)linear prognosis (Bijleveld et al., 2008). Here, if $r(t)$ is $\log$ linear, this means that $\log (r(t))$ is linear. Now, when the number of casualties $N_{a, g, m}$, is expressed as:

$$
\begin{equation*}
N_{a, g, m}(t)=\tau_{a, g, m}(t) \cdot P_{a, g}(t) \cdot r_{a, g, m}(t), \tag{Eq. 2}
\end{equation*}
$$

then if both $r$ and $\tau$ may be approximated by a log linear function of time, the product of $r$ and $\tau$ also is a log linear function of time:

$$
\begin{equation*}
\text { If } r(t)=e^{a t+b} \text { and } \tau(t)=e^{c t+d} \text { then } r(t) \cdot \tau(t)=e^{(a+c) t+(b+d) .} \tag{Eq. 3}
\end{equation*}
$$

Thus, application of known demographic data (which is usually far from log linear in time, due to irregular developments of birth rates, see e.g. Figure 4.1) simplifies the forecast of $N$. In that case, the rate $N / P$ (i.e. mortality if $N$ equals the number of fatalities, morbidity if $N$ equals the number of injured) will be an exponential function of $t$.
In an aggregated approach, changes in the overall risk $r$ that are related to changes in either $\tau_{a, g, m}$ or $P_{a, g}$ for specific age years or age groups over time, may remain unobserved. Hence, a change in the overall number of casualties $N$ caused by such a change in $r$, as

$$
\begin{equation*}
N(t)=\sum_{a, g, m} N_{a, g, m}(t)=\sum_{a, g, m}\left(r_{a, g, m}(t) \cdot T_{a, g, m}(t)\right)=r(t) \cdot T(t), \tag{Eq. 4}
\end{equation*}
$$

may be interpreted as an increase of the safety of road users. However, the individual road users do not necessarily experience safer traffic.

In the approach described in this chapter, the number of casualties of any type is assumed proportional to the relevant distance travelled. A more formal description of the relation between distance travelled, risk and the number of casualties should not only take into account the distance travelled of the casualties' vehicles, but also the distance travelled of the other vehicle involved (if any). This is relevant as e.g. the decreasing presence of young male drivers not only affects the number of casualties in that population group, but also the number of casualties in other groups involved in crashes with young male drivers. Van Norden (2009) described a model that takes this influence into account.

### 4.3. Method

Basically, we studied both the development of distance travelled itself, stratified by age, and the consequences of this stratification for risk development. Further, stratification by age was done both into groups and into individual age years. Consequently, four models were analysed: a distance travelled model by age group and a distance travelled model by age year, and a risk model by age group and a risk model by age years. Results of the latter were described in Stipdonk et al. (2010).

## Distance travelled model by age group

In the Netherlands, an estimate of the annual distance travelled by road users is obtained by means of an annual survey (NTS-data, 2010). Between 1987 and 2003, the data of distance travelled in this survey were published for the following 13 age groups: 0-5 years, 6-11 years, 12-14 years, 15-17 years, 18-19 years, 20-24 years, 25-29 years, 30-39 years, 40-49 years, 50-59 years, 60-64 years, 65-74 years, and 75+ years. These data are available for different travel modes, e.g. car drivers and passengers, for bicycle, moped and motorcycle riding and for walking. As an example, the available data of car distance travelled, stratified by age, as used in this research, are presented in Figure 4.2.


Figure 4.2. Observed annual distance travelled ( $10^{9} \mathrm{~km}$ ) by car drivers by age group (thirteen classes) and calendar year (1985-2007).

For every age group, the data of distance travelled as available from the distance travelled survey (NTS-data, 2010) can be decomposed in demographic data and data of per capita distance travelled cf. Eq. 1. Statistics Netherlands provides annual data of the total number of inhabitants of the Netherlands including stratified numbers for each age year of the Dutch population. As an example Figure 4.3 presents the survey data of the distance travelled (for all modes and both genders together) of two age groups: 18-19 year old and 30-39 year old. Data of population of the same group is also presented (CBS, 2012), as well as the per capita distance travelled of both groups.

By application of Eq. 1 to different subsets of the data of distance travelled and demographic data, we obtain calculated values for $\tau_{a, g, m}(t)$ or aggregations thereof. By performing a regression analysis for the observed distance travelled $T_{a, g, m}(t)$ in year $t$ in each of the previously mentioned 13 age groups $a$, where $\tau_{a, g, m}(t)$ is modelled as a linear function of $t$, we have:

$$
\begin{equation*}
T_{a, g, m}(t)=\left(\alpha_{a, g, m}+\beta_{a, g, m} t\right) \cdot P_{a, g}(t)+\varepsilon_{a, g, m}(t), \tag{Eq. 5}
\end{equation*}
$$

where $t=1985, \ldots, 2007$. In Eq. 5, $\alpha_{a, g, m}$ and $\beta_{a, g, m}$ are unknown parameters and $\varepsilon_{a, g, m}(t)$ is the residual in year $t$. A weighted version of the regression model of Eq. 5 was applied to each of the 13 age groups, the weights derived from the accuracy of the survey data as expressed by their standard deviations.

Aggregated analyses for both genders and all transport modes were also carried out.


Figure 4.3. Data of distance travelled and population (top panels) and per capita distance travelled (bottom panels), for 18-19 year old (left panels) and 30-39 year old (right panels).

## Distance travelled model by age year

When analysing past risk data, e.g. when assessing the possible effect of a certain safety measure, it may be valuable to have $\tau$ for each age year of the road user population. However, data of distance travelled may not be available for individual age years. In the Netherlands, up until 2003 data of distance travelled were made public for age groups only. From 2004 onward, data for individual age years are made public. Information on data of distance travelled and population size was combined to obtain smooth estimates of the Dutch distance travelled of car drivers stratified by age year for the years between 1985 and 2003. This was done for several travel modes. The method is described here with the use of car driver data.

The distance travelled by car drivers in the four age groups 0-5 years, 6-11 years, $12-14$ years, and $15-17$ year are all zero, for the obvious reason that persons younger than 18 years are not allowed to drive a car in the Netherlands. The annual data of distance travelled of car drivers available for all age groups and the annual data of population size of all age years (as available from Statistics Netherlands) were combined to obtain smooth
estimates of the annual distance travelled by car drivers in 1985-2007 for each age year.

This was achieved by optimizing the agreement between the model and the data by minimizing the so-called loss function $f\left(\alpha_{a}, \beta_{a}\right)$ with respect to all the intercepts $\alpha_{a}$ and slopes $\beta_{a}$.

$$
\begin{equation*}
f\left(\alpha_{a}, \beta_{a}\right)=\sum_{t=1985}^{2007} \sum_{j=1}^{13} w_{j t}\left(T_{j t}-\sum_{a=c_{j, L}}^{c_{j, V}} T_{a}(t)\left(\alpha_{a}+\beta_{a} t\right)\right)^{2}+\text { penalty } \tag{Eq. 6}
\end{equation*}
$$

The observed distance travelled survey data $T_{\mathrm{jt}}$ are available in age groups $j=$ $1, \ldots, 13$, where $j=1$ corresponds to $a=0 . .4$ year, $j=2$ corresponds to $a=6 . .11$ year etcetera. Thus each group is bounded by a lower value $c_{j, L}$ and an upper value $c_{j}, u$.

Weights $w_{j t}$ are used to account for differences in accuracy of the survey data. In order to optimize the model of Eq. 6, information of survey data per individual age year, available for the years 2004-2007 is used. We used survey data per age category up to 2003 to predict distance travelled per individual age year for 2004 and 2005. Smoothing parameters were chosen such that the model of Eq. 6 optimally predicts distance travelled per individual age year for 2004 and 2005.

The practical effect of including the penalty function in the loss function in Eq. 6 is that it is minimized in such a way that the developments over consecutive age years are smooth. In other words, the estimated per capita distance travelled will be similar for consecutive time years as well as for consecutive age years. Details about the penalty function are given in the Appendix.

Finally, using the penalty coefficients obtained from the analysis using survey data per age category up to 2003, the model of Eq. 6 was optimized again, but now on survey data per age category between 1985 and 1993 only. The results were used to predict distance travelled (by estimating $\alpha_{\mathrm{a}}$ and $\beta_{\mathrm{a}}$ ) per individual age year for the years 2004-2007.

## Risk model by age group

A multivariate structural time series model is used to obtain forecasts of distance travelled, the risk, and the numbers of fatalities and inpatients
casualties for the years 2008-2040 for each age group and gender combination. Three different versions of these models are analysed:

- $\quad$ The $d$ model: in this model only data of distance travelled were used
- The $p \mathcal{E} d$ model: in this model data of population size and on distance travelled were used
- The $p$ model: in this model only data of population size were used. In all cases, distance travelled or per capita distance travelled is modelled with a log linear trend, by contrast with the model described in Section 3.1 where the per capita distance travelled was modelled as a linear trend. This was done for practical purposes, as risk too is usually modelled with a $\log$ linear trend.

Log linear multivariate structural time series methods were used. These methods have the important advantage over standard regression techniques that they can handle inherent dependencies between observations in time series data, thus yielding unbiased standard errors for the construction of confidence intervals and for statistical tests. This is achieved by allowing parameters such as the intercept and the slope that are fixed in standard regression models to vary over time in structural time series models. For standard text books discussing these models see Harvey (1989), Durbin and Koopman (2001), and Commandeur and Koopman (2007); for the specific multivariate extension used here, see Bijleveld et al. (2008). Such models are also discussed in Commandeur et al. (1994). Technical details are provided in Appendix A4.3 of Commandeur et al. (2008). The models were fitted with Ox (Doornik, 2001) and SsfPack (Koopman et al., 1999).

Dutch data of the number of fatalities in road crashes for the years 1985-2007 was used, as well as on distance travelled according to the travel survey, disaggregated by age and gender. In addition annual data of the size of the Dutch population disaggregated by age and gender were used.

### 4.4. Results

## Distance travelled model by age group

All results are presented graphically in Commandeur et al (2008). Here we will demonstrate the method by giving a few examples. As a first example, we will present the distance travelled for all modes and both genders for two age groups: 18-19 years and $30-39$ years (cf. Figure 4.3). Figure 4.4 shows the result of the total distance travelled and the per capita distance travelled of these groups.


Figure 4.4. Annual distance travelled by all male and female inhabitants of two age groups, both data and model. Upper panels: 18-19 years of age, lower panels: 30-39 years of age. Left hand panels: total distance travelled ( $10^{9} \mathrm{~km}$, dots, and $95 \%$ confidence intervals for the estimate. The thick lines represent the prediction by the mode of eq. 2). Right hand panels: per capita distance travelled ( $10^{3} \mathrm{~km} / \mathrm{inhabitant}$ ) according to the model of eq. 2 .

The annual per capita distance travelled for 18-19 year old drivers was approximately constant in the observed period. For the group of 30-39 year old, the annual per capita distance travelled increased in the observed period. However, for this age group the total annual distance travelled increased at first but decreased after 2002, which corresponds to the changing population size. The results show that for each age group, the observed annual distance travelled by all road users is described quite well by the product of the actual annual population size and the estimated annual distance travelled stratified by age group.

The model of Eq. 5 was also used to analyse further stratifications of the traffic process (i.e., also to several combinations of age, gender and road user type). The results, given in Commandeur et al (2008) confirm the general idea of this chapter: if the annual distance travelled by a specific age group is decomposed into a product of the annual population size of that age group and the estimated linearized per capita distance travelled of that age group,
the actual distance travelled of that age group as measured in the National Travel Survey is reasonably well represented.


Figure 4.5. Smooth estimates (blue plane) and original data (red histogram) of per capita distance travelled per age year $\left(10^{3} \mathrm{~km}\right)$ for the years 1985-2007.

## Distance travelled model by age year

The results of the optimization of Eq. 6 were verified using survey data per individual age year for 2006 and 2007. These years were not used in the optimization. The resulting smoothed estimate of the linear per capita distance travelled for car drivers per age year is presented in Figure 4.5.

Finally, the distance travelled model was optimized for data between 1985 and 1993, using the penalty coefficients of the former model. The model was extrapolated to 2004-2007, and the results were compared to the survey data for these years. The results of this analysis for car drivers are given in Commandeur et al (2008).

Figure 4.6 indicates that the model recovers the actual survey data more than ten years after the time frame used in the analysis, even though the survey data in 1985-1993 are only available for age groups of car drivers. This suggests two important characteristics of long term development of distance travelled:

- The per capita car driver distance travelled could indeed be assumed to be a smooth function of age between 1985 and 2007.
- The per capita car driver distance travelled could indeed be assumed to be a slowly and predictably changing function of time between 1985 and 2007.


Figure 4.6. Forecasts (line) and original data (dots) of per capita distance travelled by car drivers per age year in 2004 through 2007. The forecasts are a linear extrapolation based on the analysis of the time period 1985-1993.

The availability of population data thus enables us to improve the forecasts of distance travelled, if per capita distance travelled can be assumed to be constant or (log-)linearly changing over time. In the Netherlands, Statistics Netherlands presents actual forecasts for the development of the Dutch population per age year. These forecasts - which extend beyond the year 2020 - together with the parameter estimates of the model of Eq. 5 applied to the entire period 1985-2007 were used to predict the distance travelled by car drivers per age year up until 2020, the results of which are presented in Figure 4.7.

Figure 4.7 shows the (total) estimated annual distance travelled by car drivers, which basically consists of two components:

- Population size
- Changes in per capita distance travelled


Figure 4.7. Estimated annual distance travelled $\left(10^{9} \mathrm{~km}\right)$ by car drivers as a function of their age year (from 0 to 100 years of age, right to left) and time (calendar years, front to back), including forecasts until the year 2020 .

Presently we focus on the demographic aspect. First of all, we see that the development of distance travelled by car drivers over time contains a 'wave' for car drivers of approximately 40 years old in 1985 extending diagonally all the way to car drivers of approximately 70-80 years old in 2020. This wave can be identified as the post-World War-II baby-boom of about 1946, as these car drivers were about 40 years old in the year 1985. Such a change in population size can also be found in other European countries, although the exact form differs from country to country (Figure 4.1). Secondly, in the nineties a diagonal valley starts to appear in the development of the distance travelled of car drivers which had almost continuously been increasing until that year for car drivers older than approximately 25 years old. This could be the effect of certain social factors such as the introduction of the birth control pill for women about the end of the 1960s and the individualization of the society in the 1970s, as this reduction in children born about that time only emerges in the data of distance travelled of car drivers some 20 to 30 years later. Again, such an effect can be found in other countries as well, see Figure 4.1. Finally, it is interesting to note that these waves and valleys in the development of distance travelled run approximately parallel with the progressing age of the car drivers themselves. This cannot be inferred from Figure 4.7, which suggests that there are no major cohort effects in the developments of the distance travelled of car drivers over time.

## Risk model by age group

The number of fatalities, stratified in four age groups based on data for the years 1985-2007, was modelled, using the three models ( $d, p \mathcal{E d}$, and $p$ ). The results for the number of fatalities up to 2040 are compared. The results are presented in Figure 4.8.

The number of male fatalities of 0-14, 15-24 and 25-64 years old and the female fatalities aged 15-24 years display substantial reductions of the number of fatalities in the year 2004, but these reductions do not apply to the other subgroups. Furthermore, we see that the inclusion of population data in the analyses is especially relevant for the forecasts of male fatalities in the 15-24, 25-64 and 65+ age groups, and for the forecasts of the female fatalities in the 25-64 and 65+ age groups, since for these groups the forecasts obtained with the three models diverge over time. The inclusion of population data in the analysis is especially relevant for the forecasts of males in the 25-64 age groups, and for the males and females in the 65+ age groups, since it is for these subgroups that the largest differences in forecasts between the three models are found. Traffic safety forecasts based on fatalities and data of distance travelled can therefore be improved by stratification of this data by age, and inclusion of the expected trend of the composition of the population.

### 4.5. Discussion

This chapter presents an assessment of the use of population data in addition to information on distance travelled in analysing the development of the risk of drivers of a certain age group. In summary, the findings suggest that if distance travelled is to be modelled for an age group, it may be useful to consider the development of the size of that age group.

Distance travelled is considered as a measure of traffic volume. Estimates of distance travelled per age group are often obtained from surveys. Survey results are subject to error almost by definition. Sometimes the development of per capita distance travelled is much more smooth than the development of distance travelled itself. In such cases it is possible that the smoothed per capita distance travelled together with the population size yields a better estimate of distance travelled for that group than the actual measurement from the survey. However, without information regarding the actual distance travelled, such claims cannot be verified.


Figure 4.8. Forecasts from 2007 onwards of fatalities in 4 age classes as obtained with data of distance travelled only (model 1, short dashes), combined population and distance travelled data (model 2, solid line), and population figures only (model 3, long dashes) including intervention variables to capture the significant drop in the total number of fatalities in the Netherlands in 2004.

Although this is not studied extensively in this chapter, it turns out that both linear and log linear approaches to modelling the development of per capita distance travelled appears to work well. However, the fact that it may sometimes be difficult to distinguish between a linear and log linear (or other) development adds to the uncertainty of longer-term forecasts. It is demonstrated that the difference between forecasts based on per capita distance travelled together with population size is not very different from forecasts based on population data alone. In practical applications, forecasts for combined age groups should be compared to the results of more extensive (national) forecasting efforts. In the absence of a more extensive (national) forecast, the current approach might be useful. The differences between the results based on linear and log linear assumptions are not elaborated upon in this study.

It would be interesting to know if the smoothness of the per capita distance travelled, as found in The Netherlands for several relevant strata (time and driver age), can also be found in other countries for relevant transport modes. An affirmative result would further support stratification by age, gender and transport mode in road safety models, allowing for population data as a measure of distance travelled.

The use of demographic data as an approximation of distance travelled can be quite erratic when analysing crash data or making forecasts. Aggregated population data offer little added value to crash data analysis, as changes of travel behaviour such as changes in the modal split or increased travel of specific groups are not represented by population changes. For a similar reason, aggregate mobility data are also of limited use. Incorporating these changes in the model, using a stratified analysis may overcome these difficulties, provided that the necessary data are available. In analysing crash data using population data instead of mobility, some precautions have to be made. In that case, one assumes that per capita distance travelled is constant or nearly constant, i.e. its actual changes over time can be neglected. Forecasts using a model that assumes risk to change negative exponentially (or linearly) over time, will not be affected by a change in per capita distance travelled, provided these changes are also approximately exponential (or linear) over time.

In the case that per capita distance travelled can be expected to have unobserved nonlinear changes, e.g. when it follows an s-curve or a sudden step, demographic data cannot take the place of data of distance travelled. In
the absence of data of distance travelled, information on travel behaviour from other sources such as behavioural or economic studies, may give clues to what strata of transport mode, age group and gender could be a reasonable candidate for a risk analysis based on demographic data.

A potential problem is that of possible cohort effects in the development of the distance travelled of age groups. A similar effect may also exist for risks. For example, 50 year old drivers nowadays are far more experienced than 50 year old drivers were in the 1960s. An approach that takes cohorts into account might be preferable if the differences in risk or in distance travelled behaviour between cohorts are stronger than the differences between age years. In this research, we briefly addressed this issue, and there appeared to be no advantage of an analysis by cohorts over an approach by age years.

The assumed linear relation between the number of crashes and distance travelled as adopted in this chapter, is not trivial. The fact that in many countries the distance travelled increases while the number of crashes decreases since 1972 (Gaudry, 2005), implies that the number of crashes is not proportional to distance travelled. However, by defining risk as the quotient of the number of fatalities and distance travelled, the decreasing number of crashes is interpreted as a decrease in risk. Given a constant risk, this implies that the number of crashes is proportional to distance travelled by definition. Unfortunately, observations of the number of crashes and distance travelled in a specific year are not replicable, so there is no experimental way to verify or falsify such linearity given a constant risk.

In Chapter 3, and in (Stipdonk and Berends, 2008) it was shown that during the strong increase in car use between 1950 and 1970, the number of two vehicle crashes in the Netherlands increased much more progressively than the number of single-vehicle crashes. In that period single-vehicle crash risk for cars or motorcycles was almost constant whereas two vehicle (car-car) crash risk increased with the increasing car distance travelled. This suggests that the relation between distance travelled and the number of crashes is far from uniform for different groups of crashes.

As Elvik et al. (2009) point out, when the influence of traffic volume on the number of crashes is researched, other factors need to be controlled for so that the effects of the amount of travel are not mixed up with the effects of other factors. They adopt AADT (Average Annual Daily Traffic) as the measure of distance travelled, and compare different roads. Thus, traffic and
crashes are stratified by roads and aggregated over time. They show that if the number of fatalities is proportional with $\mathrm{AADT}^{\beta}, \beta$ is found to lie between 0.77 and 0.99 with $95 \%$ certainty. Thus, this relation is less progressive than linear (i.e. $\beta<1$ ) although the results hardly differ from 1 significantly. However, the necessary correction for differences in all safety factors for these different roads is difficult to carry out. Further, we must bear in mind that AADT is very different from the distance travelled as it is used in this chapter. We apply stratifications for different age, gender and travel mode, but by contrast, crashes and distance travelled are aggregated over different road types. In AADT it is exactly the opposite: travel modes and driver ages are aggregated, and different road types with different traffic volumes appear as different values of AADT. Further, AADT does not contain a measure of distance: distance travelled equals AADT times length of road. Thus, even if the number of crashes per length of road may show a less progressive than linear increase with traffic volume locally, more distance travelled in an entire country may also mean more length of road, in which case a linear relation is not contradicted by the results as described by Elvik et al. This is especially the case if increased distance travelled is associated with an increase in population and consequently with increased infrastructure. In this chapter we aggregate over the ensemble of all locations, ranging from very unsafe to very safe roads and intersections. The composition of this ensemble contributes to the relation between the number of crashes and total distance travelled. Thus, over the years there may still be linearity between crashes and distance travelled. In the absence of a firmly established nonlinear relation between the number of crashes and distance travelled, we use a linear relation as a first approximation.

### 4.6. Conclusions

This chapter investigates the relevance and effects of including population data in time series analysis on road safety. The development in time of the population size of a specific age or age group may give rise to changes in distance travelled of drivers of that age or age group. These changes can be captured by calculation of per capita distance travelled per person, and by the population size itself. In turn shifts in distance travelled can give rise to changes in the number of casualties of that specific age or age group.

The analysis has shown that the use of population data can give insight in the development in time of distance travelled by age, gender and travel mode. It appears that such stratified distance travelled is quite well described
by population size and a slowly changing (log)linear trend describing per capita distance travelled.

The results for the fatalities forecasts show a remarkable similarity between models including population data, with or without data of distance travelled. Differences become apparent in the distant future (more than 30 years from now) only. Note that the validity of the models at that range is highly questionable. Forecasts based on data of distance travelled only give different results. These differences can be attributed to non-(log)linear development of population size with respect to birth year.

## Appendix. The penalty function for Eq. 6

The time dependence of the per capita annual distance travelled $\tau_{a}(t)$ is supposed to be linear, with coefficients $\alpha$ and $\beta$ that can be different for each age $a(a=1 . .100)$ :

$$
\tau_{a}(t)=\alpha_{a}+\beta_{a} t \text {, where } a=1, . ., 100 \text { and } t=1987, \ldots, 2007
$$

For different $a, \tau_{a}(t)$ is smoothed. The 'penalty' in the smoothing of the per capita distance travelled $\tau_{a}(t)$ over age $a$ in Eq. 6 consists of two terms, with parameters $L_{1}$ and $L_{2}$ :

$$
\text { penalty }=L_{1} p_{1}+L_{2} p_{2}
$$

In the first term, $p_{1}$ was constructed as to calculate the sum of the numerical second derivatives of $\alpha_{a}$ :

$$
p_{1}=\sum_{k}\left[\left(\alpha_{k+1}-\alpha_{k}\right)-\left(\alpha_{k}-\alpha_{k-1}\right)\right]^{2}, \text { where } k=2, \ldots, 99
$$

In the second term, $p_{2}$ was constructed so as to calculate the sum of the numerical second derivatives of $\beta_{a}$ :

$$
p_{2}=\Sigma_{k}\left[\left(\beta_{k+1}-\beta_{k}\right)-\left(\beta_{k}-\beta_{k-1}\right)\right]^{2}, \text { where } k=2, \ldots, 99
$$

## References

Bijleveld, F.D., Commandeur, J.J.F., Gould, P.G., Koopman, S.J., 2008. Model-based measurement of latent risk in time series with applications. Journal of the Royal Statistical Society Series A 171 (1), 265-277.
Booth, H., 2006. Demographic forecasting: 1980 to 2005 in review International Journal of Forecasting 22 (3), 547-581.
Broughton, J., 2012. Post-2010 casualty forecasting. Road safety Web publication. Transport Research Laboratory, London.
CBS, 2012. Statline. Centraal Bureau voor de Statistiek, Den Haag.
Commandeur, J.J.F., Bijleveld, F.D., Van Norden, Y., Stipdonk, H., Bergel, R., Debbarh, M., Yannis, G., Antoniou, C., Katsochis, D., Papadimitriou, E., 2008. Time series applications on road safety developments in Europe. In: Thomas, P. ed. Building the European Road Safety Observatory. BiVV, INRETS, KfV, NTUA, SWOV, Brussels.
Commandeur, J.J.F., Koopman, S.J., 2007. An introduction to state space time series analysis, Oxford University Press, Oxford.
COST, 2004. Models for traffic and safety development and interventions. European Commission. Directorate general for Transport, Brussels.
Davidse, R.J., 2007. Assisting the older driver, SWOV-dissertatiereeks, SWOV Institute for Road Safety Research Leidschendam.
Doornik, J.A., 2001. Object-oriented matrix programming using Ox 3.0, Timberlake Consultants Press, London.
Durbin, J., Koopman, S.J., 2001. Time series analysis by state space methods, Oxford University Press, Oxford.
Elvik, R., Bjørnskau, T., 2005. How accurately does the public perceive differences in transport risks? An exploratory analysis of scales representing perceived risk. Accident Analysis and Prevention 37 (6), 1005-1011.
European Union, 1995-2011. Eurostat. European Union, Brussels.
Harvey, A.C., 1989. Forecasting, structural time series models and the Kalman filter, Cambridge University Press, Cambridge.
Kam, B.H., 2003. A disaggregate approach to crash rate analysis. Accident Analysis and Prevention 35 (5), 693-709.
Koopman, S.J., Shephard, N., Doornik, J.A., 1999. Statistical algorithms for models in state space using SsfPack 2.2. Econometrics Journal Vol. 2, 113-166.
Laapotti, S., Keskinen, E., 2004. Has the difference in accident patterns between male and female drivers changed between 1984 and 2000. Accident Analysis and Prevention 36 (4), 577-584.
Lenroot, R.K., Gogtay, N., Greenstein, D.K., Wells, E.M., Wallace, G.L., Clasen, L.S., Blumenthal, J.D., Lerch, J., Zijdenbos, A.P., Evans, A.C., Thompson, P.M., Giedd, J.N., 2007. Sexual dimorphism of brain developmental trajectories during childhood and adolescence. NeuroImage 36 (4), 1065-1073.
Massie, D.L., Green, P.E., Campbell, K.L., 1997. Crash involvement rates by driver gender and the role of average annual mileage Accident Analysis and Prevention 29 (5), 675-685.
McCartt, A.T., Mayhew, D.R., Braitman, K.A., Ferguson, S.A., Simpson, H.M., 2009. Effects of age and experience on young driver crashes: Review of recent literature. Traffic Injury Prevention 10, 209-219.

Nicholson, A., Wong, Y.-D., 1993. Are accidents Poisson distributed? A statistical test. Accident Analysis and Prevention 25 (1), 91-97.
NTS-data, 2010. Web-accessible database of Dutch national travel survey. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.
O'Reilly, D., Hopkin, J., Loomes, G., Jones-Lee, M., Philips, P., McMahon, K., Ives, D., Soby, B., Ball, D., Kemp, R., 1994. The value of road safety. UK research on the valuation of preventing non-fatal injuries. Journal of Transport Economics and Policy 28 (1), 4559.

OECD/ITF, 2012. Reporting on serious road traffic casualties; combining and using different data sources to improve understanding of non-fatal road traffic crashes. In: IRTAD ed. Organization for Economic Co-operation and Development OECD/International Transport Forum ITF, Paris.
Schofield, G.M., Gianotti, S., Badland, H.M., Hinckson, E.A., 2008. The incidence of injuries traveling to and from school by travel mode. Preventive Medicine 46 (1), 74-76.
Stipdonk, H., Berends, E., 2008. Distinguishing traffic modes in analysing road safety development. Accident Analysis and Prevention 40 (4), 1383-1393.
Van den Bossche, F.A.M., 2006. Road safety, risk and exposure in Belgium: An econometric approach. Dissertation, Hasselt University, Diepenbeek, Belgium, Hasselt.
Van den Bossche, F.A.M., Wets, G., Brijs, T., 2007. Analysis of road risk by age and gender category. Time series approach. Journal of the Transportation Research Board 2019, 7-14.
Van Norden, Y., 2009. A road safety assessment model for the Netherlands. A feasibility study using disaggregation by traffic mode and age. Young Researcher Seminar 2009. European Conference of Transport Research Institutes (ECTRI), Forum of European National Highway Research Laboratories (FEHRL) and Forum of European Road Safety Research Institutes (FERSI), Torino.
Vlakveld, W.P., 2011. Hazard anticipation of young novice drivers, SWOV Institute for Road Safety Research, Leidschendam.
Yannis, G., Antoniou, C., Eleonora Papadimitriou, Katsohis, D., 2011. When may road fatalities start to decrease? Journal of Safety Research 42 (1), 17-25.
Yannis, G., Papadimitriou, E., Chaziris, A., Duchamp, G., Lejeune, P., Treny, V., Hemdorff, S., Haddak, M., Lenguerrand, E., Holló, P., Gábor, M., Cseffalvay, M., Leitner, T., Angermann, A., Hoeglinger, S., Cardoso, J., Bijleveld, F., Houwing, S., Bjørnskau, T., Rackliff, L., 2008. Risk exposure data common framework. In: Thomas, P. ed. Building the European Road Safety Observatory. European Commission, Brussels.

# 5. The expected number of road traffic casualties using stratified data 

This chapter was first published in a special issue of "Safety Science", dedicated to road safety modelling (Stipdonk, H., Wesemann, P., Ale, B., 2010. The expected number of road traffic casualties using stratified data. Safety Science 48 (9), 1123-1133).


#### Abstract

Road safety policy plans often require robust calculation of the expected number of road casualties in a certain target year. The relevance of such estimations should be measured by their power to influence and support safety policy makers. Thus, techniques to evaluate the safety developments and the estimating methods must be sound, robust, and preferably accepted by both policy makers and the scientific community. In this chapter, we concentrate on choosing an appropriate model used for the calculation, rather than on statistical techniques. We calculate a casualty rate from casualty data and mobility (distance travelled) data, which is extrapolated and subsequently multiplied by an expected future distance travelled. After correction for separately assessed effects of additional safety measures, the number of casualties is estimated. We investigate a method where this is done after both mobility data and casualty data are stratified into properly chosen subsets. Projecting these different trends generally leads to a result that differs from the projection of the aggregated data. Also, stratification enables incorporation in the estimation of explaining factors or additional measures related to a specific subset of the casualties. The principles of stratified projections are illustrated by three Dutch projections which were carried out between 2006 and 2008. Also, some preliminary results of further research on stratification are given. The results imply that the rates of change in casualty rate for different traffic modes or driver age, are not necessarily equal. We propose that these specific decreasing trends are a consequence of external influencing factors.


### 5.1. Introduction

National and international road safety policy plans often need a robust forecast of the number of casualties, fatalities or in-patients or both, in a certain target year. Towards zero, ambitious road safety targets and the safe system approach (OECD, 2007) describes the role of targets to achieve ambitious road safety goals. It illustrates how targets can be set and met by analysing data,
calculating projections, estimating effects of measures and applying costbenefit analyses. In this process, the assessment of a reliable forecast plays an important role.

As traffic crashes are a stochastical process, any forecast will always be liable to fluctuations as a consequence of chance. The best that can be done is to determine the expected number of fatalities. Strictly speaking, a forecast cannot be better than the best possible calculation of the expected number of casualties. In this chapter, we refer to forecasts in this sense.

Road safety development analyses can be used to inspire new road safety strategies (De Ceuster, 2005, Stipdonk and Wesemann, 2007), or to evaluate the consequences of a new road safety strategy or specific safety measures, either ex ante (forecasting the development) or ex post (explaining the development after it has taken place). An example of the first possibility is given by Broughton (2012), who describes a general way of making such forecasts. Another example is the assessment of the expected safety effects of road pricing in the Netherlands (Schermers and Reurings, 2009). This report illustrates how the expected reduction in mobility and, as a consequence, in the number of casualties, plays a role in the political decision making process of Dutch mobility policy. Elvik, (2005) gives an example of an ex post evaluation which shows the difficulty of attributing all past developments to actual changes in safety policy.

Several difficulties are encountered in forecasting road safety casualties. It is to be decided how to choose an appropriate model which contains plausible relations between the relevant factors and the expected number of casualties to be calculated. The most important focus of this chapter is to describe our current intermediate results of our research to develop a possible method to develop such a model. Also, there are important technicalities involved in calculating the best forecasts. Further, it is important to determine the actual effects of specific factors that change over time and influence safety (changes in mobility, ageing population, weather effects, safety measures) on the past developments. When such relations are completely unknown, extrapolation of these trends can be meaningful if it is safe to assume that the relationships are stochastic and these relationships can be considered to be the transfer function of an otherwise black box. It remains unknown if such an assumption is solid.

In this chapter, the safety assessments of either past or future road safety developments has two objectives: to reveal effects of safety measures that were taken in the past, and to substantiate additional safety measures to further improve safety in the future. To be acceptable, we believe that the technique to evaluate the development of safety and the forecasting methods needs to be sound, robust, and, preferably, accepted by policy makers and accepted by the international scientific environment.

### 5.2. Approach

## The basic three step approach

The general purpose of a road safety forecast is to accurately calculate the expected number of casualties in a future year, often in the presence of some strategic road safety policy programme. Broughton (2012) distinguishes three steps to carry out such a forecast. In the first step, casualty rates (casualties per distance travelled) in the target year are estimated, based on a trend analysis of past years. In the second step, the result of the first step is multiplied by the predicted distance travelled in the target year. This results in the baseline forecast (the number of casualties in the target year i.e. the projection of the baseline trend). In the third step, the effects of new road strategy measures are estimated. In the Netherlands, SWOV has used this method in several forecasts (Wesemann, 2007, Aarts et al., 2008, Wesemann, 2008, Schermers and Reurings, 2009). In this chapter, we adopt this three step approach as a general principle.

To apply this method, several problems need to be solved. The first problem is how exactly to extrapolate the data from the past to forecast the expected casualty rate for the target year. Forecasting requires a mathematical expression of the casualty rate as a function of time. A priori, this mathematical expression is unknown, and the goodness of fit is not a sufficient indication of the correctness of any mathematical expression (Elvik, 2005). The second problem is a consequence of the fact that the baseline trend of the casualty rate is assumed to contain implicitly the effects of many measures that are not explicitly included in the model (Broughton, 2012). The problem is how to decide to what extent the future development of the baseline trend, which is supposed to contain a continuation of existing safety policies and measures, requires future road safety improvements which are necessary to realize this development. To put it differently: the problem is to determine the effect of future safety measures that is necessary to maintain
the current baseline trend. The third problem is to estimate the expected additional effect of specific measures in truly new policy.

The first two problems are, in our view, connected. If we ever succeed in correctly expressing the past and expected developments of the casualty rate mathematically, the precise effects of past safety measures must be known as a function of the relevant factors. In that hypothetical and ideal case, time would no longer be a necessary parameter in the model: the casualty rate follows from the factors and the estimated parameters. However, as long as we have not reached this ideal situation, the development of the number of fatalities will contain an unexplained component which represents the unexplained development in time. The exact functional form of this component can only be guessed, and the safety measures that are responsible for the development of this component are, by definition, unknown.

This chapter will discuss a possible approach to estimate this unexplained component, such that the baseline forecast depends on unexplained components as little as possible. Our general approach is based on stratification (i.e. disaggregation) of crash data by relevant factors such as traffic mode, age or road type.

Conceptually, the third problem is the most simple one, although when put into practice it is laborious and difficult to manage. The effect of safety measures, either in the past or in the future, can be determined by thoroughly analysing the nature of the corresponding measures, and their quantitative relation to the expected number of casualties. In principle, any measure affects a subset of the total number of casualties only. For this subset, which we will call the domain of the measure, an analysis should reveal the expected effect of the measure on the number of casualties in that domain. Thus, every measure requires a separate analysis that results in a specific partial model: a quantitative relation between some parameter that quantifies the measure and the number of casualties saved. In traffic safety literature, many safety measures and their effects have been described. For example: Evans (1986) calculates the risk reduction as a result of seat belt use. The outcome can be used to estimate the effect of enhanced seat belt use in relation with a gradual increase in seat belt use due to seat belt reminders (Lie et al., 2008). In this chapter, we will not discuss modelling the effect of specific safety measures. The issue is well described by Sigrist (2010). Instead, we will attempt to deal with the first two problems.

## Possible improvements of the method

Apart from suitable stratifications, the forecasts can be further improved if it is known how to add the effects of explanatory factors (such as safety measures, demographic changes, weather influence) in the past. In fact, this means replacing part of the obscure and unexplained trend by an external factor. This way the baseline becomes more accurate. A possible approach could be to formulate and test hypotheses, corresponding to specific stratifications which show a decreasing trend. Another possibility is to use existing and well established relations between a specific Safety Performance Indicator (SPI) and a specific road safety domain (e.g. the proportion of car occupants wearing seat belts vs the number of casualties among car occupants in crashes where a seatbelt can be effective). This also requires sufficiently accurate data on the SPI.

Stipdonk and Berends (2008) have shown that between 1950 and 2005 the development of the number of any subset of casualties in the Netherlands generally depended strongly on the conflict type, characterised by the traffic mode of the casualty and (if any) the traffic mode of the other vehicle. They showed that both the development of the number of casualties and the casualty rate in single-vehicle car crashes differed considerably from the development of that of, for example, car-car crashes, or pedestrianmotorcycle crashes. Thus, any expression that describes such development should preferably reflect these differences, and contain the parameters that explain these differences. Thus, stratification of crash data by traffic mode, or rather, by casualty type (defined by the combination of traffic modes involved), makes it possible to establish completely different trends. These results indicate that the total number of casualties is a sum of many components, each following substantially different trends. Maier and Ahrens (2009) have followed these lines to estimate the expected number of casualties in Germany for 2015 and 2020.

Even when using simple negative exponential trends to model the different components, there are two arguments in favour of projecting stratifications as compared to just fitting a simple negative exponential trend to the sum directly. Firstly, stratification enables using data on distance travelled, stratified by traffic mode. Secondly, if the resulting casualty rates show very different trends (e.g. some are constant whereas others are decreasing), these differences can be taken into account when calculating any expected number of casualties. A third advantage could occur when different trends coincide with differences in the way these trends are understood (e.g. when
motorcycle casualty rates decrease less than car casualty rates as a consequence of an increased seat belt use affecting car occupants only). This way, stratification allows us to add explanatory factors that are specific for a specific stratification. Our suggestion to solve problems one and two, therefore, is to sufficiently stratify crash data, such that different risk developments are mathematically described accordingly, by different trends.

## An indication of the improved correctness of the model

By applying properly chosen stratifications, we hope to find better forecasts, in the sense that the calculated uncertainty of these forecasts is as little as possible. As an illustration of the effect of which method is chosen for the projection, on the outcome and uncertainty of the projection, we calculated the expected number of fatalities in 2020 in three ways. The results are shown in Figure 5.1. The mobility data used are motor vehicle data extracted from Statistics Netherlands (2000), which are based on a combination of surveys, traffic counts and other sources. The results are obtained with state space modelling as described by Commandeur and Koopman (2007), applying the above approach, and without any stratification. The three different ways to calculate the expected number of fatalities in 2020 are (see Table 5.1):

1. We assume that the expected mobility is a stochastic quantity, its uncertainty being defined by the fluctuations in the mobility data. The model further contains a mathematical intervention, to take care of a sudden decrease in the number of fatalities in The Netherlands in 2004 of more than $20 \%$. The purpose of this intervention is to demonstrate its effect on the calculated forecast.
2. As 1, but here we assume that the projected mobility for 2020 is exact.
3. As 2 , but the mathematical intervention is removed.

|  | $\mathbf{1}^{\text {st }}$ result | 2 $^{\text {nd }}$ result | $3^{\text {rd }}$ result |
| :--- | ---: | ---: | ---: |
| Expected number of fatalities | 480 | 480 | 440 |
| Expected number $+\sigma$ | 930 | 600 | 560 |
| Expected number $-\sigma$ | 250 | 390 | 340 |
| $68 \%$ confidence interval | 680 | 210 | 220 |

Table 5.1. Results of three state space projections for the crash rate and the mobility in 2020, based on crash data between 1950 and 2007, and mobility data from traffic counts, between 1950 and 2000, with its $68 \%$ uncertainty interval. In the first result a mathematical intervention was added to the model to allow for the sudden decrease in 2004. The expected mobility was considered as a stochast. The second result was similar to the first, but the expected mobility was considered exact. The third result was similar to the second, but the mathematical intervention in 2004 was removed.


Figure 5.1. State space projection for the crash rate and the mobility in 2020, based on crash data between 1950 and 2007, and mobility data from traffic counts, between 1950 and 2000, with its $68 \%$ uncertainty interval. Top and middle panel: in 2004 a mathematical intervention was added to the model to allow for the sudden decrease in 2004. Lower panel: no intervention in 2004. In the top panel, the future mobility data were projected as well, in the middle and lower panel the future mobility data were assumed to be exact.

The uncertainty in the expected number of fatalities decreases by more than a factor of 3 when assuming that the expected mobility is an external and exact factor. The results further show how the presence of an intervention in the model influences both the forecast and its uncertainty.

In the ideal case of a model that explains, and correctly predicts the development of the number of fatalities in terms of all relevant external factors, the variance $\sigma^{2}$ of the prediction should be equal to the prediction. Our aim is to find a model that brings down this variance by applying stratifications and subsequently searching for relevant external factors. This does not guarantee that we will eventually find the perfect model and thus solve the first two problems mentioned above. However, the variance of the prediction is an independent indication of model correctness that may help to find improvements of the model. This can be done by adding stratifications and by introducing the influence of relevant external factors.

## How much stratification is necessary and how much is sufficient?

It is difficult to say a priori which stratifications are necessary. One possible strategy is to stratify data with distinct accident causes, as it is likely that different causes indicate that they are sensitive to different measures (Hale, 2001) and therefore show different trends. Unfortunately we know little of the actual causes of road accidents. Thus, other criteria for choosing the proper stratifications are required. There are several ways to achieve this. One consists of looking for stratifications with different trends, where we compare the parameters associated with these trends, and decide if their difference is statistically significant. It must be noted that in this case an arbitrary choice must be made concerning the preferred significance level. Another way is to look for systematic misspecifications in the trend for aggregated data: if a supposedly negative exponential trend is actually a sum of two different trends (e.g. a constant + a negative exponential), the aggregate trend analysis will in principle show a misspecification, which indicates that some (unknown) stratification may be necessary.

At present no theoretical framework is available that helps to choose the factors for which stratification improves the model. To decide on stratification for different crash groups, we consider two indications, and two conditions:

1. Crash groups with different developments of the crash rates over time;
2. Crash groups with different crash rates, combined with different development of the mobility, e.g. in the case of a shift in mobility between different crash groups over time;
3. Availability of stratified mobility data, as a condition for stratification.
4. A sufficient number of crash data, as a condition for stratification.

In The Netherlands, stratified mobility data can be obtained from Statistics Netherlands (OVG/MON, 1985-2009). These data are based on an annual household survey, currently with a response of approximately 30,000 households, corresponding to circa $70 \%$ of the current size of the survey. The response decreased from 60,000 in 1999 to 30,000 in 2005, due to a decrease in the size of the survey. The results of the survey are used by Statistics Netherlands to calculate the mobility for the entire Dutch population. Traffic of non-inhabitant drivers is not included, nor is freight traffic. Under the assumption that the fraction of mobility by non-inhabitants is constant, indeed the accuracy of the data enables stratification by traffic mode, age and gender, at least to some extent. To illustrate this, Figure 5.2 shows the estimated per capita mobility by age and gender of drivers of three traffic modes (passenger car, bicycle and moped), and for three consecutive years. The figure indicates that the estimated data indeed fluctuate, partly due to chance. These fluctuations can be considered as an upper limit of the inaccuracy of these estimates, as these variations may also stem from fluctuations in the actual mobility. The figure also shows a consistent agedependent pattern, and differences between male and female mobility by age and gender, that are significantly more pronounced than the size of the fluctuations. Thus, these age- and gender-dependent patterns are informative about actual differences in male and female age-dependent mobility. The majority of these data are available through the SWOV-website (NTS-data, 2010).

### 5.3. Development of the method, applied to Dutch data

Between 2006 and 2009, Wesemann et al. (2010) carried out three road safety forecasts for The Netherlands. In 2006, Wesemann $(2007,2008)$ forecast the expected number of road fatalities and hospitalized injuries for the year 2020, based on road safety and mobility data up to 2005. Knowing that stratification was important, he experimented with three different stratifications: by traffic mode, by age and by road type. Each method resulted in a different expected number of casualties in 2020. This result is a consequence of the fact that the different subsets of casualty types show
different developments in time, which indicates that the development is indeed not uniform. This result can be seen as an indication of the necessity of stratification.


Figure 5.2. Annual per capita mobility by age, for male and female bicyclists, mopedists and passenger car drivers. The left panel shows the data for 1995 and 2008, the right panel shows averaged three year data (1995-1997 and 2006-2008).

At that time, a recent drop in the number of fatalities in $2004(-20 \%$ as compared to 2003) complicated the forecasts. Thus, Wesemann had to decide how to incorporate this drop in his analysis. Three different scenarios were analysed. In scenario 1 the drop was considered to be an incidental, temporal drop, to be undone during the next year or years. In scenario 2 the drop was considered to be a unique and incidental, but permanent downward step in
an otherwise steadily decreasing trend. In scenario 3 the drop was attributed to chance: it was treated as the outcome of the combination of known and unknown processes, to be considered as a regular part of the traffic process. These different scenarios are illustrated in Figure 5.3, which shows simple negative exponential trends of the annual number of fatalities (hence not corrected for mobility) between 1987 and 2005. For reasons of completeness, data of 2006-2008 are added to the figure. The figure shows that the projection for scenario 2 is the most optimistic for 2020, although scenario 3 can be expected to be the most optimistic in the long run.


Figure 5.3. Negative exponential trends of the yearly number of fatalities (hence not corrected for mobility) between 1987 and 2005, for three different scenarios to incorporate the sudden decrease in 2004. . For scenario 1 the trend was calculated without the use of the 2004 and 2005 data. For scenario 2 the trend of scenario 1 was multiplied by an appropriate factor less than 1, to adjust for the sudden decrease in the number of fatalities in 2004. For scenario 3, no special treatment was given for the data of 2004 and 2005. For reasons of completeness, data of 2006-2008 are added to the figure.

In 2008, two more forecasts were carried out, both using stratification of casualty rate by traffic mode. One forecast was aimed at estimating a new ambitious and realistic target for the number of fatalities in 2010 (Aarts et al., 2008), and the other was an analysis of the road pricing effects (Schermers and Reurings, 2009), aimed at calculating the expected number of fatalities in 2020, with and without road pricing to be introduced in 2012. The ministry considers many different pricing schemes, such as pricing specific routes, or
pricing roads to and from the main cities, either during peak hours only or around the clock, and, for all variants, whether they should only be applied to the main infrastructure or to rural roads also. In all variants, road pricing is expected to decrease the mobility during peak hours. Assuming an invariant casualty rate, the safety increases on the relevant roads.

In these forecasts, it was chosen to stratify by casualty type: for single-vehicle crashes the casualty type is defined by the traffic mode of the casualty; for crashes between more than one vehicle, the casualty type is defined by the traffic mode of the casualty and the most important other vehicle. Thus, the casualty type is determined only by the vehicles involved. Although other stratifications (e.g. by road type or by age of driver) are still considered important, they have as yet not been applied. This is because we have experienced that all relevant stratifications should actually be carried out simultaneously (Wesemann, 2007, 2008), and not separately, as was done earlier by Wesemann. Combined stratifications are indeed technically feasible, as was shown by Stipdonk (2008) and Van Norden (2009).

In Aarts et al. (2008), only fatalities were analysed, stratified into several casualty types or groups of casualty types. Ten casualty types, each contributing significantly to the total number of fatalities, were analysed separately, following the approach described above. The remaining casualty types were grouped with the same dominant traffic mode for which there was a forecast of mobility data available (passenger car, van, motorcycle, truck, moped). For fatalities of casualty types that could not be related to any dominant traffic mode, such as bicycle-bus or pedestrian-bicycle, the casualty rate could not be analysed. Instead, the total number of fatalities was projected for this group. As the number of fatalities, 117 annually on average, has approximately been constant for more than 30 years, this number was kept constant in the forecast. This was the only group for which the number of casualties was analysed instead of casualty rates. Hence, the number of fatalities was considered to be independent of mobility changes. For all other casualty types or groups, a negative exponential trend was estimated. A similar approach was followed by Maier and Ahrens (2009).

Schermers and Reurings (2009) altered the applied method. This was necessary because the Dutch Ministry of Transport requested a distinction be made between different road types. In general, this would require using mobility data by road type, and such data is not available. For this specific analysis, the authors estimated the mobility development, stratified by road
type and traffic mode. We will not give a detailed description of the methods used here.

The several approaches mentioned above thus show different results for different partially stratified analyses. E.g. partial stratification by age leads to a different expected number of casualties than partial stratification by transport mode. This indicates that some or all of the partial stratifications are in itself suboptimal. We propose to solve this problem by carrying out all relevant stratifications simultaneously. Thus, each separate subset is extrapolated separately. Therefore, there is no longer any discrepancy between different stratification schemes.

### 5.4. Further stratifications to be applied in the near future

## Some preliminary results

Van Norden (2009) analysed several casualty types, for example bicycle-carcrashes, which she further stratified by age of both drivers. Due to the unavailability of mobility data stratified by individual age years at that time, their methods were not yet suitable to be applied to all relevant casualty types taken together (which is necessary when forecasting the total number of fatalities or in-patients). However, mobility data for individual driver age have become available recently, enabling further refinement and application of this method. The method implies that single-vehicle crashes are stratified by driver age, and two-vehicle crashes are stratified by both the driver age of the traffic mode of the casualty, and the driver age of the opponent; crashes with a casualty in each vehicle are counted twice. In the assessment, the mobility, stratified by driver age, is in fact a mobility per capita, thus enabling the usage of past and future development of demographics in the calculations. In case of two-vehicle crashes, this method takes the mobility of both vehicles involved into account. The method involves smoothing techniques, assuming gradual changes in per capita mobility, and in casualty rates over time and for driver age. For two-vehicle crashes, the smoothing is applied to the number of crashes divided by the relevant distance travelled by both traffic modes, i.e. crash rate per km (i.e. crashes per km per km). Thus, the number of casualties is implicitly assumed to be locally proportional to distance travelled by age of driver of both modes

By analysing the per capita mobility, we correct for changes in the mobility over time that are caused by demographic changes. In Stipdonk (2008) it was shown that for many driver ages and traffic modes the per capita mobility is
not very time dependent. Thus, for forecasts of the mobility of pedestrians or mopeds reasonable forecasts could possibly be achieved by just applying the demographic development to a constant or slowly changing age dependent per capita mobility. Furthermore, it is assumed that in general mobility is not very different for similar ages (e.g. 42 year old drivers are supposed to drive approximately equal distances as 43 year old drivers). Thus by smoothing the per capita mobility data, we are able to estimate the actual mobility stratified by age group and calendar year.

Using stratified mobility data makes it possible to take different mobility and risk behaviour into account. By doing so, we make use of available information from the mobility database and the accident database (i.e. information on the age of drivers) that would be otherwise undisclosed and unused. Even if stratification results in small numbers with high variance in the database, this should generally enable an improvement of the model.

In addition, Van Norden (2009) showed that the development in time of the crash rate associated with crashes of drivers of a specific age, depends on driver age. To calculate this crash rate, she smoothed the risk with respect to age, and allowed for a (negative) exponential change in risk over time. In general, for young drivers, the resulting crash rate decreases less than for experienced drivers. This suggests that safety measures are not effective for all drivers. This may be related to the nature of the safety measures involved.

In this analysis it was not possible to give confidence intervals. For an accurate calculation of confidence intervals we would have to analyse data by adding several calendar years and age groups. It would then be possible to present results of risk changes by driver age over time by calculating this risk for periods of 5 years, and for different age groups. This way, it would become possible to show the same results, for discrete age groups and time intervals of several years, together with their variances. In principle, the results would be the same. We have chosen to stress the principles of disaggregation instead of calculating accurate results and confidence intervals.

The results of Van Norden (2009) can be illustrated by the following examples:

1. The risk for young car drivers of being involved in a single-vehicle car crash is high and it hardly decreases, while there still is some decrease of the risk for elderly drivers. The same holds for car-car crashes. The
risk of being involved in a car-car crash is higher for young drivers, and does not decrease much, whereas for elderly drivers, the risk of being involved is decreasing. This effect may be due to the fact that measures such as enhanced enforcement mainly have an effect on experienced drivers. Young drivers first need to be exposed to some of this enforcement, while elderly drivers, based on their experience, simply respond to a general announcement of enhanced enforcement. In addition, the average experience of elderly drivers is likely to have increased over the last decades, whereas young drivers always have had hardly any experience at all. This is illustrated in Figure 5.4, which shows the estimated casualty rate for car drivers (KSI $/ \mathrm{km}$ ) in singlevehicle car crashes, by age of driver, between 1987 and 2007. The figure illustrates that during these 20 years the estimated casualty rate has been decreasing less for young drivers (a risk reduction of less than $20 \%$ for driver ages below 25) as compared to more experienced drivers (a reduction of almost $50 \%$ for driver ages between 40 and 70). The data are based on Van Norden (2009).


Figure 5.4. Estimated casualty rate for car drivers (KSI/km) in single-vehicle car crashes, by age of driver, for 1987 (solid line) and 2007 (dashed line).
2. For cyclists, the risk to be involved in a bicycle-car crash was seen to have decreased most for children (6-12 years), and much less for other cyclist ages. In 1987, the children's risk was much higher than that for other ages, and between 1987 and 2007 it decreased by more than $70 \%$
for cyclists younger than the age of 10 . For 30 -year-old cyclists, this reduction was $13 \%$. Thus, if this decrease was a consequence of safety measures, they were not uniformly effective for all cyclist ages. To achieve further improvement of cyclist safety in bicycle-car crashes, it is likely that new measures that also affect 30-year-old cyclists are to be thought of. Today, the risk for 10-year-old cyclists is about 1.3 times that for 30-year-old cyclists. In 1987 this difference amounted to a factor of 4.3. This may be due to the safety around schools which has been improving during the last 15 years. This is illustrated in Figure 5.5, which shows the estimated casualty rate for cyclists (KSI/km) in bicyclecar crashes, by age of cyclist, between 1987 and 2007. The data are based on Van Norden (2009).


Figure 5.5. Estimated casualty rate for bicycle drivers (KSI/km) in bicycle car crashes, by age of cyclist, for 1987 (solid line) and 2007 (dashed line).

Furthermore, the involvement in serious crashes by age is influenced by different mechanisms, such as vulnerability and experience. This can be illustrated by comparing the risk to be involved in a car-car crash as an inpatient (Figure 5.6, solid lines) with the risk to be involved in a car-car crash as a driver of the opponent car (Figure 5.6, dashed lines), for crashes with at least one injured driver. It shows the involvement of elderly drivers in carcar crashes as the injured driver strongly increases with age, whereas this is much less the case for their involvement in car-car crashes as the opponent
driver. For younger drivers this difference is much less pronounced. The distinction between these two ways a driver can be involved in a car-car crash, supports our approach which intends to apply this distinction in the future models. Mark that for younger drivers, the risk of being involved in a car-car crash as an injured driver is clearly lower than being involved in a single-vehicle car crash. For elderly drivers the car-car risk is higher.

## Our stratification scheme for the near future

It is our intention to improve our forecasts where possible. One approach could be to stratify the data by groups for which better - or more plausible forecasts can be made. A balance has to be found between the advantages of further stratifications and data reliability and model complexity.


Figure 5.6. Estimated casualty rate for car drivers in a car-car crash, either when they are involved as injured driver (solid lines) or as the driver of the opponent car (dashed lines). Curves are given for 1987 and 2007.

There is no theoretical objection against forecasts using casualty rates and mobility data stratified for many variables. However, in practice the level of stratification will, in practice, be determined by the available mobility data, and the available crash data. In The Netherlands, the resolution of the available mobility depends on the traffic mode. For several traffic modes including cycling and car driving, National Traffic Survey (NTS) data are available. For those traffic modes, some further stratification by driver age and gender is also possible. This stratification is also desirable because of the
different casualty rates. For other traffic modes, no mobility data are available. Further stratification, such as by road type, is more difficult to carry out due to lack of mobility data by road type.

We want to apply stratifications that meet the indications given in Section 5.2 , i.e. either significantly different crash rate development or a combination of significantly different crash rates and significantly different mobility development over time. Also, the two conditions, i.e. availability of mobility and crash data, have to be met. If these conditions allow some further stratification, we intent to research if the indications show that stratification is indeed called for. This gives rise to the following scheme, which is derived from the availability of the mobility and crash data. The first part of the scheme goes into the available mobility data, and the second part describes how to cope with small numbers of crashes or small mobility survey samples.

1. The available mobility data enable the level of stratification. This availability divides the traffic modes into three groups, described below. A group 1a for which NTS-mobility data is available, a group 1 b for which no NTS-mobility data is available, but useful substitutes can be used instead, and a group 1c for which no mobility data are available. For each group, the mobility data is incorporated as follows: For casualties in single-vehicle crashes with traffic modes from group 1a or 1 b , the mobility data are used to calculate the casualty rate which is to be forecast. For casualties in twovehicle crashes in which one of the traffic modes involved are from group 1a or 1 b , the mobility data of that modes is incorporated in the analysis, where the number of casualties is divided by the mobility of that modes to arrive at a casualty rate which is to be forecast. For casualties in two-vehicle crashes in which both of the traffic modes involved are from group 1a or 1 b , the mobility data of both modes are incorporated in the analysis, where the number of casualties is divided by the mobility of that modes to arrive at a casualty rate per km, which is to be forecast.

1a. Traffic modes, for which NTS-data are available. These are pedestrian, bicycle, moped (including light moped), motorcycles and passenger cars. For those traffic modes, stratification by age and gender is also possible, although it depends on the traffic mode if the available data allow stratification by individual age year or if perhaps age groups are to be made. The majority of fatal and serious crashes in The Netherlands involve at least
one of these modes, e. g. today in $70 \%$ of all fatal crashes a passenger car is involved (Chapter 3, Stipdonk and Berends, 2008).

1b. Traffic modes for which no mobility data are available in the NTS household survey, but for which an approximation of distance travelled can be used (e.g. road traffic volume statistics, fleet size data or sales data). These are small vans or trucks. For those traffic modes, one of these approximations will be used, assuming that a suitable approximation is proportional to the actual mobility. As a consequence, the resulting casualty rate forecasts concerning these modes will be vulnerable to unknown changes in the relation between the used approximation and mobility itself. For example, if fleet size is used as an approximation of mobility, a change in the mean annual distance travelled per vehicle will show as a change in the casualty rate.

1c. Crashes between traffic modes for which no mobility data or approximation is available. These are crashes such as between buses and miscellaneous vehicles. For those traffic modes the number of casualties is modelled without any mobility. As a consequence any change in mobility of these modes that consequently affects the number of casualties involving that mode, will remain unnoticed. The calculated casualty rates and numbers of casualties cannot be corrected for those changes.
2. For some small groups, it is necessary to add groups of stratified data together, either by aggregating age years, traffic modes, gender, or even calendar years. This is the case for small population groups, with either few crashes, or where the sample size of the mobility data becomes too small and the in-sample variance cannot be trusted to give a reasonable indication of the population variance. This will depend on the sample error variance of the available mobility survey data.

2a. For age groups that are few in population (e. g. for very old persons) and thus have very little distance travelled and are involved in a low number of crashes, stratification by age is not useful, in which case these age groups have to be aggregated.
$2 b$. For many two-vehicle casualty types, the annual number of casualties in the Netherlands does not allow stratification by age and gender for individual age years for each separate casualty type. In those cases, it is usually possible to analyse the driver age dependence for several casualty
types simultaneously. Examples are casualty types with a passenger car as opponent. In The Netherlands, the age-dependence of those casualty types is independent of the traffic mode of the casualty. This is illustrated in Figure 5.7, where the number of seriously injured in a crash with a passenger car as the opponent, is given for the driver age of the opponent. The age dependence seems to a large extent of the same form, except for a different constant between different traffic modes. We have as yet not decided which casualty types will be analysed simultaneously this way.

2c. Not all casualty types will be extrapolated separately. Casualties in crashes where all traffic modes involved are a member of group 1c will be analysed as one group. Crashes between both traffic modes from group 1c and one other traffic mode, will be analysed as one group. Some more casualty types might be analysed together. This still has to be decided.

This scheme describes stratifications that, to our view, are necessary: we know that traffic mode, by age and gender, shows many different mobility developments, and we also know the associated casualty rates may differ strongly. However, other stratifications may be necessary too. Stratification by road type may be a good candidate, if mobility develops differently for different road types.


Figure 5.7. Number of fatalities and serious injuries in The Netherlands, between 1976 and 2005, in a crash with a passenger car as the opponent traffic mode, by age of the driver of the opponent passenger car. Data are shown for nine different casualty transport modes, on a logarithmic scale.

### 5.5. Discussion

## Decreasing trend

When forecasting a decreasing number of casualties by estimating trend lines for the number of casualties or for the casualty rate, the use of the negative exponential trend line has some important consequences. One is that the trend always tends to zero casualties in the (distant) future, although it will never reach zero. To this end, the negative exponential trend line has an important advantage as compared to a linear trend line which indeed will cross the zero casualty rate and become negative at some future date; this is unrealistic, both for extrapolations of casualties or casualty rates. However, a negative exponential trend as such may also be unrealistic. Many examples (OECD, 2007) illustrate that part of the casualties cannot be associated with a decreasing trend; this becomes apparent after stratification. This implies that the common choice of a negative exponential trend line may be considered fundamentally incorrect, as a trend of the form
$Y(t)=A+B \mathrm{e}^{-\mathrm{C} \cdot t}$
would be more appropriate. Here, $\mathrm{Y}(t)$ denotes the casualty rate (or, if mobility is constant, the number of casualties) in year $t$. The constant $A$ denotes some unchanging component in Y, for instance because part of the unsafe conditions seemed immune to safety measures in the past. Stratification of data by whichever relevant dimension (traffic mode, age, road type) helps to identify the candidates for contribution to this constant. However, subsets of casualties, obtained by stratification that still show a clearly decreasing trend, may nevertheless be obscuring constant components that will only be revealed after further stratification. An example: in Aarts et al. (2008) the number of fatalities was shown to contain a constant of approximately 117 fatalities, which had been independent of time for a period of more than 30 years. This constant became apparent after stratification by traffic mode.

## Explaining the overall trend

In principle, stratification results in different time series of the number of casualties by conflict type, road type, age of driver etcetera. These stratified results enable a more straightforward way to assess the effect of entirely different safety measures and other relevant influencing factors, if these measures and factors apply to specific subsets. Different trends in crash rate may be the consequence of different explanatory factors (measure or
development). Thus, the overall decreasing trend in casualties or casualty rates may be the result of a decreasing trend in a limited set of subsets of crashes for which explanatory factors were relevant. For all practical purposes, one may at first restrict oneself to subsets which show the most dominant decreasing trends.

## Unexplained components

Ideally, the development of the casualty rate could be entirely understood from the introduction of safety measures or the development of influencing factors. Decreasing crash rates should as much as possible be identified with explicit external factors and safety measures. In practice, this goal is not achievable yet, either because the relation between the measures or factors and the number of casualties is not sufficiently well known yet, or because the development of the factor itself isn't sufficiently accurately known. Thus, if there is a decreasing trend, it is uncertain if this trend will continue without additional effort. The continuation of the known measures and factors will not necessarily ensure a continuation of the decreasing baseline trend. Therefore it is important to make an explicit inventory of the measures that are supposed to have been responsible for the development of the trends in the past. Thus, at a later stage, it will be possible to correct the assessment for the future; at the same time it prevents that measures are inadvertently double-counted.

Following the same idealistic principles, sudden major changes in the number of casualties demand an explanation in terms of a change in distance travelled or casualty rate as well. In that respect, the 2004 drop in fatalities in The Netherlands should eventually be connected to external factors. At this moment, this drop cannot be explained satisfactorily. Research on Dutch crash data has shown that a considerable part of the drop can be associated with young car drivers. I. e. the drop is only manifest in accidents in which at least one car driver of at most 35 years of age is involved.

As an example, we will look at the average number of fatalities in singlevehicle car accidents for a specific driver age group, and divide this number by the population of this age group. We call the resulting quotient the number of single-vehicle car fatalities per driver, or fpdssc . When we compare the $\mathrm{fpd}_{\text {svc }}$ for drivers of up to 35 years of age, and drivers older than 35 , the results show a striking difference. For the younger drivers, the fpdsve decreased from $4.7 \times 10^{-6}$ to $3.1 \times 10^{-6}$. The largest drop shows between 2003 and 2004. For the older drivers, the fpd svc is about $0.9 \times 10^{-6}$ and it is steadily
decreasing by $3.7 \%$ yearly. There is not a hint of a specific drop in 2004 for this age group. This is illustrated in Figure 5.8, top panel. This difference in $\mathrm{fpd}_{\text {svc }}$ may be related to a change in mobility per car driver, or mpdc. The development of mpdc for the same two age groups shows a steady average increase of about 75 km yearly, whereas the $\mathrm{mpd}_{\mathrm{c}}$ for the younger group indeed has decreased considerably in the last five years. However, this is a gradual decrease, not a strong decrease in a short period. In 2004, hardly any decrease in $\mathrm{mpd}_{\mathrm{c}}$ is seen at all in 2004. So the question is if the mobility data nevertheless can explain the fatality data?

The answer to this question is not known, but it might be found in the effect of driver experience. In Chapter 3 (Stipdonk and Berends, 2008) it was suggested that a change in the number of new car drivers could have a large effect on the safety, but a much smaller effect on the mobility. To verify this hypothesis, we need data on the experience of car drivers, involved in a crash. At this moment, no such data is available. The explanation of the sudden decrease in the number of fatalities in 2004 in the Netherlands can therefore not be answered yet. Still, we think that this example shows that a stratification by conflict type and age (group) may reveal interesting targets for possible new safety policies.

In the past, the lack of quantitative explanations for the overall decreasing trend has given rise to the suggestion that there is an autonomous process that improves the safety level (Oppe, 1991b). This theoretical concept refers to a process of continuous 'learning' by road users, car manufacturers, road authorities, etcetera. In our view it is uncertain if such a process, for which there may well have been evidence in the past, is indeed autonomous, i.e. that it can indeed be held responsible for a permanent and lasting annual decrease. Firstly, many subsets of the number of casualties (traffic modes, crash types, driver age groups) show little or no decreasing trend. Secondly, other factors, such as demographic changes, do provide explanations for some subsets with a decreasing trend. Nevertheless, specific forms of 'learning processes` could explain decreasing trends in some other subsets. For example, look at elderly car drivers who nowadays have far more driving experience than the same age group twenty years ago, whereas 18 year old drivers have no driving experience at all, which is no different from the situation twenty years ago. However, this learning process may come to an end when all car drivers have learned to drive at the same early age of 18 . Thus, application of a theoretical concept of general learning, as an explanation of the decreasing trend in the casualty rate, bears the risk that
policy makers believe that little needs to be done, whereas in reality, the decreasing trend will eventually come to a stop.


Figure 5.8. Comparison of the number of fatalities per driver in a single-vehicle car crash (fpdsve, top panel), and the mobility per driver (mpdc, bottom panel), for two age groups: below and above 35 years of age. The top panel shows the average fpdsvc for the younger group in the two most recent periods of five years, and also the average annual relative decrease of the fpdscc for the elderly group. The bottom panel shows the average annual linear increase in the mpdc for the elderly group. The points in this figure are connected with thin solid lines that serve as a guide to the eye only.

## Limits to stratification

A perpetual drive to stratify data further and further is, from a conceptual perspective, preferable. As an illustration, consider stratification of crashes on junctions into crashes on either roundabouts or intersections. Such a stratification could reveal that $A$. the normalized casualty rates (i.e. the casualty rate per junction) of roundabouts and intersections are both constant, and that B. a decrease in the casualty rate on junctions can be explained by an increase in the number of roundabouts.

However, given a finite number of casualties in a period of N years, there seems to be a limit to the level at which modelling simultaneous stratifications is still feasible. We suggest that this problem is less serious than it seems, according to the following reasoning. When we stratify crash data by traffic mode, age, gender, road type etcetera, we no longer have a single time series. Instead, our aggregated single time series is replaced by multiple time series, each available for a period of N years, and each of which can be modelled with a specific small set of parameters for that time series. Thus, a stratified model not only explains the aggregated number of casualties, but also the separate stratified subsets. Also, by stratifying, extra information from the data is used. This justifies the use of additional parameters.

This reasoning also holds for data on distance travelled. Stratification by traffic mode and age group allows for the use of more parameters, which enables estimating different time series for different traffic modes and age groups. However, subsets in the stratification with a high crash rate, such as moped or motorcycle crashes, or young car drivers, do present a difficulty. For these traffic modes with a relatively small total distance travelled, the sample size of the traffic survey is small in relation to the associated number of crashes, because of the high crash rate. This means that the calculated high crash rates have a relatively low accuracy. This, however, is no reason for not stratifying distance travelled by traffic mode. By not stratifying the data it is a priori not possible to take into account changes in distance travelled by these groups.

Thus, by stratification into subsets, the relative uncertainty of the crash data is higher than the relative uncertainty for the aggregate. For Poisson distributed crash data the standard deviation of the data is as a first approximation proportional to the square root of the observed values. On the other hand, the amount of information (the number of bits in each value) is
proportional to the sum of the logarithms of the stratified numbers. This follows from the length of the binary representation of any number (Shannon, 1948). Thus, if some value $v$ is stratified into $v_{1}$ and $v_{2}$, then $\operatorname{var}\left(v_{1}\right)$ $+\operatorname{var}\left(v_{2}\right)=\operatorname{var}(v)$, but $\log \left(v_{1}\right)+\log \left(v_{2}\right)>\log \left(v_{1}+v_{2}\right)$. The same reasoning holds for mobility data, although the accuracy of these data is a much more complex function of the observations that underlie the estimated mobility. Still, as stratification introduces extra information, this enables development of a better model (given that the extra information can be interpreted to a sufficient extent).

However, care should be taken that the model remains feasible, and does not grow into an unnecessary complex of formulas. Additional factors and parameters should improve the understanding of the safety development, or should else be avoided. In general, only well-established relations between factors and the number of casualties should be taken into account.

## Past vs future

Models that help explain the past traffic safety development, and models that enable estimating the future number of casualties, have a lot in common. It is clear that forecasting models and explanatory models are used differently.

An explanatory model aims to describe the fluctuations in the number of casualties in the past by finding relations between these numbers and external factors (such as mobility, SPIs and other safety measures). These relations, when parameterized, are then optimized to find the maximum likelihood of the known data, given the model. Thus, for an explanatory model, in principle all data is available, and the better these data, the more accurate the model will be.

For a forecasting model, just as for an explaining model, we need a, preferably accurate, mathematical description of the relation between the number of casualties, and the factors this depends on. In principle, the explanatory model should provide us with all we need to forecast road safety. However, there are some drawbacks. For example, to build a perfect explanatory model for forecasts, if this can be done at all, we need to know the values of all relevant factors in the model for the target year. If the explanatory model for instance uses motorcycle mobility, safety belt use and the number of roundabouts as explanatory factors, we need to know or estimate these factors in the target year explicitly.

Also, we need to know what additional measures will be taken, and what their effects will be. Notwithstanding these drawbacks, we know of no arguments for the idea that in its mathematical description, an explanatory model is different from a forecasting model, and the best possible explanatory model should be a good start for a forecasting model. The more these explanatory models contain constant parameters and time dependent measured quantities such as mobility, and SPIs, the more their time dependence will be determined by the time dependent observations of external factors. At this moment, the models still hardly contain constant parameters and measured quantities, but are dominated by time dependent trends. As these trends may come to a halt, such explanatory models should be handled with care when used to estimate a future number of casualties.

## Additional road safety measures.

In this chapter, we did not go into an analysis of methods to assess the effect of additional measures, because such assessments are, to a large extent, dependent of the measure. There is a large number of measures for which such assessments are available (Elvik and Vaa, 2004). As the domain of each measure usually is only a (small) subset of the total number of casualties, stratification helps identify that subset. Hence, it may be that a general problem of assessing the effect of safety measures, namely the assessment of a combination of different safety measures, can be partly overcome by stratification. Elvik (2009a) showed that estimates of the simultaneous effect of measures on the total number of casualties, tends to be less than would be expected according to the separate effects. Once logic attaches the effect of each separate safety measure to the correct domain of the crashes that are affected by the measure, it may turn out that the expected effect of each separate measure becomes smaller, and that a more accurate estimate of a set of simultaneous measures becomes possible.

## Availability of mobility data

Countries that do not have stratified mobility data (by traffic mode, by driver age) at their disposal, cannot easily stratify their analysis. In that case, it may be an idea to hypothesize a constant (or slowly increasing) per capita distance travelled (by traffic mode), and to try to check if this is a reasonable assumption, for instance by analysing driving licence data, fleet ownership data, public transport use data, traffic offence data or any other dataset that may serve as an approximation of per capita mobility. Once it has been established that a constant (or slowly changing) per capita distance travelled is a reasonable assumption, it is possible to analyse stratified casualty trends.

The dependence between mobility, crash rate and the severity of the crash
Analysing road safety data implies careful treatment of crash data, crash severity data and mobility data, as was extensively described by Bijleveld (2008). We give two examples of the subtleties involved.

In our approach we prefer to analyse mobility and crash rate simultaneously, rather than independently. This is because, in fact, mobility and crash rate are not independent factors. This is illustrated by the fact that in two-traffic mode accidents, the crash rate of one traffic mode depends on the mobility of the other mode.

In The Netherlands, underreporting of crashes is known to be a serious problem. Even for fatal crashes only about $92 \%$ is reported by the police, and of the seriously injured (MAIS2+) in-patients only about $30 \%$ are reported. The underreporting of accidents without injuries is assumed to be even more serious. As a consequence, there is a correlation between the registered number of crashes and their severity. This implies that crash severity and crash occurrence are not independent (at least, as far as registered crashes are concerned), and that it is therefore not possible to correctly consider the number of casualties as the outcome of separate processes for the number of crashes and the probability of a serious injury. This is because crashes involving vehicles carrying more than one passenger are more likely to be registered and more likely to have more casualties. To overcome this problem in our approach as described in Section 5, we analyse casualties among drivers only and do not consider casualties among passengers yet.

### 5.6. Conclusions

The three step approach described by Broughton (2012), is a suitable method to carry out road safety forecasts. In principle, by introduction of the intermediate quantity casualty rate, the number of casualties per distance travelled, this approach enables a correct analysis of the effects of mobility and other factors.

As casualty rates differ widely between traffic modes, it is necessary to stratify both casualty data and mobility data by traffic mode. Essentially, the same holds for any other factor for which the casualty rate differs significantly, such as driver age or road type.

As the decreasing trends of casualty rates may also differ widely, trend extrapolations demand stratification as well; even if risks do not differ strongly, but (decreasing) trends do, it is necessary to stratify, to prevent the forecast from being incorrect.

A further reason for stratification of mobility and casualty data is that influencing factors such as SPIs may be effective in a small subset of the total number of casualties only. We have called this subset the domain of the influencing factor. Therefore, stratified models enable incorporation of explanatory factors by offering explicit domains of the explanatory factors (assuming that the stratifications are appropriate). The same holds for implementation of additional safety measures; their future effects, if limited to a specific subset, can conveniently be calculated if this subset is specified by appropriate stratifications, such that the projection of this subset is separately carried out in the model.

Often, it appears that negative casualty trends, when stratified, decompose into trends that decrease quickly and trends that decrease more slowly or not at all. This suggests that the trend for the aggregated set is not, as is often assumed, negative exponential, but is in fact the sum of several negative exponential trends and possibly a constant. The assumption that trend lines generally contain a constant allows model parameter optimizations to generate solutions that do not tend to decrease to zero, in cases where this is in fact incorrect.

The fact that, in practice, so many subsets of casualty rates show very differently decreasing trends, indicates that the sometimes suggested general social developments that cause a general decreasing casualty rate, may not exist. Assuming such developments to be the implicit cause of a decreasing trend, may lead to the incorrect conclusion that less needs to be done to ensure a further improvement of traffic safety. If it were true that the learning society can be seen as a cause for a continually improving safety system, then one would expect this to be effective for different traffic modes, driver ages etcetera. Without actual proof of such an effect, the assumption that there might be some autonomous decreasing casualty rate could be misleading

## Acknowledgement

The authors wish to thank Frits Bijleveld and Jacques Commandeur, for their contribution to the analyses described here, and for sharing their valuable thoughts and knowledge on this subject.

## References

Aarts, L., Weijermars, W., Schoon, C., Wesemann, P., 2008. Maximaal 500 verkeersdoden in 2020: Waarom eigenlijk niet? : Maatregelpakketten en effectschattingen om te komen tot een aangescherpte verkeersveiligheidsdoelstelling. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2008-5.
Bijleveld, F., 2008. Time series analysis in road safety research using state space methods, Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam.
Broughton, J., 2012. Post-2010 casualty forecasting. Road safety Web publication. Transport Research Laboratory, London.
CBS, 2000. Statline, statistiek van de wegen. Centraal Bureau voor de Statistiek CBS, Voorburg/Heerlen.
Commandeur, J.J.F., Bijleveld, F.D., Van Norden, Y., Stipdonk, H., Bergel, R., Debbarh, M., Yannis, G., Antoniou, C., Katsochis, D., Papadimitriou, E. 2008. Time series applications on road safety developments in Europe. In: Thomas, P. ed. Building the European Road Safety Observatory. BiVV, INRETS, KfV, NTUA, SWOV, Brussels.
Commandeur, J.J.F., Koopman, S.J., 2007. An introduction to state space time series analysis, Oxford university press, Oxford.
De Ceuster, G., 2005. Assessment of the contribution of the TEN and other transport policy measures to the mid term implementation of the white paper on the European transport policy for 2010. In: DG TREN, E.C. ed. TML, Leuven.
Elvik, R., 2005. Has progress in improving road safety come to a stop? Institute of Transport Economics TØI, Oslo.
Elvik, R., 2009. An exploratory analysis of models for estimating the combined effects of road safety measures. Accident Analysis and Prevention 41 (4), 876-880.
Elvik, R., Vaa, T., 2004. The handbook of road safety measures, Elsevier Ltd, Oxford, UK.
Evans, L., 1986. Double pair comparison-a new method to determine how occupant characteristics affect fatality risk in traffic crashes. Accident Analysis and Prevention 18 (3), 217-227.
Hale, A.R., 2001. Conditions of occurrence of major and minor accidents. Journal of the Institution of Occupational Safety and Health 5 (1), 7-21.
Lie, A., Krafft, M., Kullgren, A., Tingvall, C., 2008. Intelligent seat belt reminders. Do they change driver seat belt use in Europe? Traffic Injury Prevention 9 (5), 446-449.
Maier, R., Ahrens, G.-A., 2009. Entwicklung der strassenverkehrssicherheit und ihrer rahmenbedingungen bis 2015/2020. Forschungs- und Entwicklungsvorhaben. Technische Universität Dresden, Dresden.
NTS-data, 2010. Web-accessible database of Dutch national travel survey. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.
OECD, 2007. Towards zero, ambitious road safety targets and the safe system approach, OECD/ITF, Paris.

Oppe, S., 1991. Development of traffic and traffic safety: Global trends and incidental fluctuations. Accident Analysis and Prevention 23 (5), 413-422.
OVG/MON, 1985-2009. Onderzoek Verplaatsingsgedrag en Mobiliteitsonderzoek Nederland. Statistics Netherlands, Dutch annual mobility survey.
Schermers, G., Reurings, M., 2009. Verkeersveiligheidseffecten van de invoering van anders betalen voor mobiliteit. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2009-2.
Shannon, C.E., 1948. A mathematical theory of communication. The Bell System Technical Journal Vol. 27 (July, october), 379-423; 623-656.
Sigrist, S., 2010. Towards a method to forecast the effectiveness of national road safety programmes. Safety Science 48 (9), 1106-1110.
Stipdonk, H., Berends, E., 2008. Distinguishing traffic modes in analysing road safety development. Accident Analysis and Prevention 40 (4), 1383-1393.
Stipdonk, H., Wesemann, P., 2007. Long-term forecasts of road traffic fatalities in the European union : General methods and results. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2006-15.
Van Norden, Y., 2009. A road safety assessment model for the Netherlands. A feasibility study using disaggregation by traffic mode and age. Young Researcher Seminar 2009. European Conference of Transport Research Institutes (ECTRI), Forum of European National Highway Research Laboratories (FEHRL) and Forum of European Road Safety Research Institutes (FERSI), Torino.
Wesemann, P., 2007. De verkeersveiligheid in 2020 : Verkenning van ontwikkelingen in mobiliteit, ongevallen en beleid. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2006-27.
Wesemann, P., 2008. De verkeersveiligheid in 2020; een verkenning van ontwikkelingen in mobiliteit, ongevallen en beleid. Tijdschrift Vervoerwetenschap 44 (juni), 36-45.
Wesemann, P., Van Norden, Y., Stipdonk, H., 2010. Road safety in 2020; future developments of exposure, crashes and policy. Safety Science 48 (9), 1098-1105.

# 6. The effect on road safety of a modal shift from car to bicycle 

This chapter was first published in "Traffic Injury Prevention" (Stipdonk, H., Reurings, M., 2012. The effect on road safety of a modal shift from car to bicycle. Traffic Injury Prevention 13 (4) 412-421 ).


#### Abstract

Objective To describe and apply a method to assess the effect on road safety of a modal shift from car to bicycle.


Method
$10 \%$ of all car trips shorter than 7.5 km were assumed to be replaced by bicycle trips. Single-vehicle and multivehicle crashes involving cars and/or bicycles were considered. The safety of car occupants and cyclists was taken into account as well as the safety of other road users involved in such crashes. The computations were carried out by age and gender. Assuming constant risk (casualties per distance travelled), the expected number of accidents is proportional to the mobility shift. Several types of risk were considered: the risk of being injured as a car driver or cyclist and the risk of being involved as a car driver or cyclist in a crash in which another road user is injured.

## Results

The results indicated that the total gain of the modal shift is negative for fatalities, which means there is a net increase in the number of fatalities. The modal shift was advantageous for young drivers and disadvantageous for elderly drivers. In addition, it was more positive for males than for females. The turning point was about the age of 35 . For hospitalized casualties, due to the strong influence of the many hospitalized cyclists in nonmotorized vehicle crashes, there was a strong negative overall effect, and the modal shift resulted in a positive effect for 18- and 19-year-old males only.
Overall, a small increase (up to $1 \%$ ) in the number of cyclist fatalities and a greater increase of 3.5 percent in the number of inpatients are expected. The increase in casualties is mainly due to the proportion of single-vehicle bicycle crashes with serious injuries in relation to the total number of injured cyclists. The effect of the modal shift was shown to depend on age and gender, resulting in fewer casualties for younger drivers and for women.
Conclusions
It is possible to provide a first approximation of the effect on road safety of a mobility shift from cars to bicycles. This approximation indicates that, in
general, road safety does not benefit from this modal shift. The effect differs for gender and age groups. Elderly drivers are safer inside a car than on a bicycle. For the number of hospitalized casualties, the modal shift increases the number of casualties for practically all ages and both genders.

### 6.1. Introduction

The role of motorized transport in $\mathrm{CO}_{2}$ emissions and of interventions to reduce car travel, have been widely researched and described in literature (Graham-Rowe et al., 2011) (and references therein). Often, such interventions involve a modal shift from passenger car to another travel mode, for example, public transport or cycling. Such a modal shift might also influence road safety, because the risk to road users of being injured or even killed in a road crash is not the same for all travel modes. For example, the safety of cyclists or bus passengers differs from the safety of car occupants. Cyclists are easily injured when involved in a crash, particularly when a motorized vehicle is involved. Car occupants are far less vulnerable than cyclists. A modal shift from passenger car to public transport has a completely different impact on road safety, because bus passengers have a very low risk of being killed in a traffic accident. Furthermore, cyclists do not endanger other road users as much as car or bus drivers do due to their much lower speed and mass. A 60-year-old may be a very experienced car driver but a vulnerable cyclist, while an 18-year-old may be accustomed to cycling, but not experienced in driving a car. So whereas cycling produces much less $\mathrm{CO}_{2}$ per distance travelled (approximately 15 times less than a passenger car, (Hermans, 2012) and is therefore to be preferred from an environmental point of view, it is relevant to know what the effects would be on road safety if more people switched to riding a bicycle instead of driving a car.

It is not immediately clear whether road safety will improve or deteriorate if car trips are replaced by bicycle trips, because such a shift has both positive and negative impacts on road safety. An increase in the number of cyclists will lead to more casualties among cyclists; especially in crashes not involving motorized vehicles, i.e., in single-vehicle crashes or crashes between 2 bicycles. There will also be an increase in the number of collisions between bicycles and other road users, but these crashes will hardly lead to casualties among these other road users. A decrease in car use will probably result in fewer crashes with cars. This will lead to fewer casualties among car
occupants as well as to fewer casualties among other road users, such as pedestrians and cyclists.

Crashes between bicycles and cars are an interesting crash type in the light of the above. The increase in bicycle mobility could lead to more of these crashes; the decrease in car mobility, however, could result in fewer of these crashes. A phenomenon called safety by numbers may also play a role in this issue. Several studies suggest that the risk for cyclists of colliding with a motorized vehicle is highly nonlinear with bicycle traffic volume, which means that a motorist is less likely to collide with a cyclist when more people are cycling. In other words, the risk that a cyclist will collide with a motorized vehicle is inversely proportional to the amount of cycling (Jacobsen, 2003). A study by Elvik (2009b) showed that for this reason the total number of these crashes would decrease if a substantial share of trips by motorized transport were instead bicycle trips. However, Wegman, Zhang \& Dijkstra (2012) stated that there is no sound evidence that Jacobsen's conclusion can be explained by numbers alone. They suggest there may be no safety by numbers effect if only the amount of cycling increases without risk reducing measures. In this paper, a possible effect of safety by numbers is ignored for the above reason. Also, a supposedly positive effect that more cycling may have on the risk of cyclists, may partly be counteracted by a similar negative effect that a reduction in car driving may have on the risk of drivers. Furthermore, if these effects indeed exist, they are of unknown size. The calculated results of this paper are to be considered as a first approximation to the solution.

In the 'Discussion' section of his paper Elvik (2009b) states that:

- the registration of crashes involving cyclists or pedestrians in official road crash statistics is not complete.
- a large proportion of these (nonregistered) crashes are single-vehicle bicycle crashes, for which the safety by numbers effect does not apply.

These crashes are left out of his computations, probably leading to an overestimation of the road safety effect of a modal shift from car to bicycle.

In The Netherlands, single-vehicle bicycle crashes are rarely reported by the police (Derriks and Mak, 2007, Bos et al., 2009). A recent study (Reurings, 2010, Reurings and Stipdonk, 2011) shows that in 2008 almost 7800 cyclists were seriously injured in crashes not involving motorized vehicles, of which only $4 \%$ were registered in the official road crash database. According to
emergency room surveys, $70 \%$ of the crashes in which no motor vehicle was involved were single-vehicle bicycle crashes (Van Kampen, 2007). The remaining crashes are mainly crashes between two bicycles.

In the Netherlands, this group of casualties should not be ignored when the effect of a modal shift from car to bicycle is estimated. Therefore a new method was developed, which is the subject of this paper. This method was developed as part of a study in which the health benefits are estimated of substituting short-distance car trips with short-distance cycling trips (Van Kempen et al., 2010, Stipdonk and Reurings, 2010). However, the method was also applied to other modal shifts. For example, Schermers \& Reurings (2009) estimated the safety effect of Different Payment for Mobility (DPM, i.e. a scheme for road pricing). Because motorcycles would be exempt from DPM, it is possible that DPM leads to a shift from cars to motorcycles. Due to a much higher crash rate, even if only a small group of car drivers shifted to motorcycle, the increase in distance travelled on motorcycles could result in decreased road safety. A method similar to the one presented in this paper was used to calculate the safety effect of this shift.

### 6.2. Method

## Principles of the method

The effect of the modal shift from car to bicycle was estimated for all casualties in the following crash types:

1. crashes in which at least one car is involved but no bicycles;
2. crashes in which at least one bicycle is involved but no cars;
3. crashes in which at least one bicycle and one car are involved.

It is reasonable to assume that casualties in other crash types would not be affected by this modal shift. The casualties that are discussed here are not only killed or injured car drivers and cyclists, but also include casualties among the other road users.

Because a car may have one or more passengers, a car trip may be replaced by multiple bicycle trips. These bicycle trips made by car passengers, however, are not taken into account in this paper because, in general, it is not easy to determine how car passenger trips should be replaced by bicycle trips in the computations. Hence, bicycle trips made by passengers are ignored. As a consequence, casualties among car occupants influenced by the modal shift are limited to car drivers. This means that car passengers who would not
have been killed or injured if their car trips had been replaced by bicycle trips were not included in the computations.
This led to the following categories of casualties:

- 1a: casualties among car drivers in single-vehicle car crashes.
- 1b: casualties among car drivers in crashes with parties other than a car or bicycle.
- $\quad 1 \mathrm{c}$ : casualties among road users other than car occupants and cyclists in crashes involving a car.
- $\quad 1 \mathrm{~d}$ : casualties among car occupants, in crashes between two cars
- 2a: casualties in single-vehicle bicycle crashes.
- $\quad 2 \mathrm{~b}$ : casualties among cyclists in crashes with parties other than a car or bicycle.
- $\quad$ 2c: casualties among road users other than car occupants and cyclists in crashes with a bicycle.
- 2 d : casualties among cyclists in crashes between two bicycles.
- 3a: casualties among cyclists in crashes with a car.
- 3b: casualties among car occupants in crashes with a bicycle.

In crashes involving more than two vehicles (which are rare) the existence of the third party is neglected.

For each of these groups the effect of the modal shift was estimated separately. This was done by comparing the expected number of casualties after the modal shift took place to the actual number of casualties in the present situation (before the modal shift). The precise calculations are given in the following subsection.

It is known that the risk of being involved in a road crash differs between males and females and between age groups. In addition, car and bicycle mobility is not uniformly distributed across age and gender. Therefore it is important to take the age and gender of car drivers and cyclists (to whom the modal shift applies) into account. A consequence of replacing only the car trips made by drivers and not those made by passengers is that only the modal shift of persons older than 18 ( 16 in the United States) is included; therefore only casualties among car drivers and bicyclists over age 18 are considered.

Detailed data on the properties of the car trips that will be substituted, and of the bicycle trips that will substitute them are not available in the Netherlands. Therefore, several assumptions need to be made. The main assumption is that car mobility which is substituted by bicycle mobility has
the same safety properties as the remaining car mobility. Subsequently, the added bicycle mobility is also assumed to have exactly the same properties as the bicycle mobility before this addition. The same properties means that for each characteristic (except age and gender), such as time of day and road type, car mobility decreases and bicycle mobility increases by the same factor. For example, if $1 \%$ of car mobility is shifted to bicycle mobility, every road type loses $1 \%$ of car mobility; if the modal shift results in a $10 \%$ increase in bicycle mobility, every road type that can be used by bicycles gains that $10 \%$ of bicycle mobility.

## Mathematical description of the method

Crashes involving cars and no bicycles.
Let $M_{c}(a, g)$ be the distance travelled by car drivers of age $a$ and gender $g$ and let $M_{b}(a, g)$ be the distance travelled by cyclists of age $a$ and gender $g$, both before the car-bicycle shift. The amount of car mobility that is replaced by bicycle mobility is denoted by $\mu(a, g)$. Thus, car and bicycle mobility after the car-bicycle shift, denoted by $M^{\prime}(a, g)$ and $M^{\prime}{ }^{\prime}(a, g)$, depends on the original car and bicycle mobility and the shifted mobility, $\mu(a, g)$ :

$$
\begin{aligned}
& M_{c}^{\prime}(a, g)=M_{c}(a, g)-\mu(a, g) ; \\
& M_{b}^{\prime}(a, g)=M_{b}(a, g)+\mu(a, g) .
\end{aligned}
$$

The relative changes in car and bicycle mobility, $\phi_{c}$ and $\phi_{b}$, are calculated as follows:

$$
\begin{aligned}
\phi_{c}(a, g) & =\mu(a, g) / M_{c}(a, g) ; \\
\phi_{b}(a, g) & =\mu(a, g) / M_{b}(a, g) .
\end{aligned}
$$

Next it is shown how the expected change in the number of casualties resulting from the car-bicycle shift is calculated for each type of casualty.

## Group 1a: casualties among drivers in single-vehicle car crashes

Let $N_{c s}(a, g)$, resp. $N^{\prime}{ }_{c s}(a, g)$, denote the number of casualties among drivers of age $a$ and gender $g$ in single-vehicle car crashes before, resp. after, the carbicycle shift. The risk $R_{c s}(a, g)$ of becoming such a casualty is defined as:

$$
R_{c s}(a, g)=N_{c s}(a, g) / M_{c}(a, g) .
$$

Assuming that $R_{c s}(a, g)$ will not change if $M_{c}(a, g)$ changes, we can calculate $N^{\prime}{ }_{c s}(a, g)$ :

$$
\begin{aligned}
& N_{c s}^{\prime}(a, g)=R_{c s}(a, g) \cdot M_{c}^{\prime}(a, g) \\
& =R_{c s}(a, g) \cdot\left(M_{c}(a, g)-\mu(a, g)\right) \\
& = \\
& =R_{c s}(a, g) \cdot M_{c}(a, g) \cdot\left(1-\phi_{c}(a, g)\right) \\
& =N_{c s}(a, g) \cdot\left(1-\phi_{c}(a, g)\right) .
\end{aligned}
$$

It now follows that the decrease in the number of casualties due to the carbicycle shift, $n_{c s}(a, g)$, is given by:

$$
n_{c s}(a, g)=N_{c s}^{\prime}(a, g)-N_{c s}(a, g)=-\phi_{c}(a, g) \cdot N_{c s}(a, g) .
$$

The total decrease in the number of casualties in single-vehicle car crashes for all ages and both genders, denoted as $n_{c s}$, is equal to:

$$
n_{c s}=-\sum_{\mathrm{a}, \mathrm{~g}} \phi_{c}(a, g) \cdot N_{c s}(a, g)
$$

Group 1b: casualties among car drivers in crashes with parties other than a car or bicycle
The number of casualties in this group is given by $N_{c o}(a, g)$, where $a$ and $g$ denote age and gender of the driver of the involved car. Because the mobility of other vehicles than cars and bicycles does not change, the change in $N_{c o}$ (denoted by $n_{c o}$ ) is again proportional to the change in car mobility. Therefore, analogous to the computations for group 1a:

$$
n_{c o}(a, g)=-\phi_{c}(a, g) \cdot N_{c o}(a, g) \text { and } n_{c o}=-\sum_{\mathrm{a}, \mathrm{~g}} \phi_{c}(a, g) \cdot N_{c o}(a, g) \text {. }
$$

Group 1c: casualties among other road users than car occupants and cyclists in crashes involving a car
The number of casualties in this group is denoted as $\widetilde{N}_{o c}(a, g)$, where $a$ and $g$ denote age and gender of the driver of the involved car (and not of the casualty). The tilde is added above the symbol $N$ to emphasize the fact that the casualty here is not the car driver. The following relations hold:

$$
\tilde{n}_{o c}(a, g)=-\phi_{c}(a, g) \cdot \tilde{N}_{o c}(a, g) \text { and } \tilde{n}_{o c}=-\sum_{\mathrm{a}, \mathrm{~g}} \phi_{c}(a, g) \cdot \tilde{N}_{o c}(a, g)
$$

## Group 1d: casualties among car occupants in crashes between two cars

Now the age and gender of two different car drivers have to be taken into account. One drives car $\mathcal{C}_{1}$ and is the considered casualty, the other drives car $c_{2}$, and is not the considered casualty. After a car-bicycle shift, both drivers
drive less, by an amount $\phi_{c}\left(a_{1}, g_{1}\right)$ and $\phi_{c}\left(a_{2}, g_{2}\right)$ respectively (with $a_{i}$ and $g_{i}$ the age and gender of the driver of car $c_{i}, i=1,2$ ). This provides a double effect on the expected new number of casualties. Both effects are calculated separately. First the decrease in casualties due to a change in mobility of possible casualties is calculated (car $c_{1}$ ). Let $N_{c c}\left(a_{1}, g_{1}\right)$ be the number of casualties among car drivers of age $a_{1}$ and gender $g_{1}$ in crashes with another car. The change in this number due to the car-bicycle shift is denoted by $n_{c c}$ and is calculated as follows:

$$
n_{c c}=-\sum_{\mathrm{a}, \mathrm{~g}} \phi_{c}\left(a_{1}, g_{1}\right) \cdot N_{c c}\left(a_{1}, g_{1}\right)
$$

Next, let $\widetilde{N}_{c c}\left(a_{2}, g_{2}\right)$ denote the number of casualties among car occupants in crashes with another car, where the driver of that other car has age $a_{2}$ and gender $g_{2}$. The effect of the change in car mobility in this number is denoted by $\widetilde{n}_{c c}$ and calculated as follows:

$$
\tilde{n}_{c c}=-\sum_{\mathrm{a}, \mathrm{~g}} \phi_{c}\left(a_{2}, g_{2}\right) \cdot \tilde{N}_{c c}\left(a_{2}, g_{2}\right)
$$

The total effect is the sum of $n_{c c}$ and $\widetilde{n}_{c c}$, equal to

$$
n_{c c}+\tilde{n}_{c c}=-\sum_{\mathrm{a}, \mathrm{~g}} \phi_{c}(a, g) \cdot\left(N_{c c}(a, g)+\tilde{N}_{c c}(a, g)\right)
$$

This is not exactly equal to the decrease of the number of casualties in crashes between two cars resulting from a change in car mobility. Indeed, the decrease in the number of casualties in crashes in which both drivers shift to cycling are counted twice. As long as $\phi_{c}$ is small, this is a negligible error. The error yields a small overestimation of the assessed decrease in the number of car-car casualties.

Crashes involving bicycles and no cars
Group 2a: casualties in single-vehicle bicycle crashes

$$
n_{b s}=\sum_{a, g} \phi_{b}(a, g) \cdot N_{b s}(a, g)
$$

Group 2b: casualties among cyclists in crashes with parties other than a car or bicycle

$$
n_{b o}=\sum_{a, g} \phi_{b}(a, g) \cdot N_{b o}(a, g) .
$$

Group 2c: casualties among road users other than car occupants and cyclists in crashes with a bicycle

$$
\tilde{n}_{o b}=\sum_{a, g} \phi_{b}(a, g) \cdot \tilde{N}_{o b}(a, g) .
$$

Group 2d: casualties among cyclists in crashes between two bicycles

$$
n_{b b}+\tilde{n}_{b b}=\sum_{\mathrm{a}, \mathrm{~g}} \phi_{b}(a, g) \cdot\left(N_{b b}(a, g)+\widetilde{N}_{b b}(a, g)\right) .
$$

Crashes between a bicycle and a car.
For casualties in crashes between bicycles and cars, the effect of the decreasing car mobility as well as the effect of the increasing bicycle mobility needs to be taken into account. If $\phi_{c}$ and $\phi_{b}$ are assumed to be small, it can be assumed that the effects are approximately linear. Therefore, the assumption is that the number of car drivers does not change when the effect is estimated of a car-bicycle shift on the number of casualties among cyclists and vice versa.

## Group 3a: casualties among cyclists in crashes with a car

The increasing bicycle mobility will lead to an increase in the number of these casualties, whereas the decreasing car mobility, on the other hand, will lead to a decrease. Let $N_{b c}(a, g)$ denote the number of casualties among cyclists of age $a$ and gender $g$ in crashes with a car and let $\widetilde{N}_{b c}(a, g)$ denote the number of casualties among cyclists in crashes with a car, where the driver of that car has age $a$ and gender $g$. Then, analogously to the computations for crash groups 1d and 2d, the increase in the total number $n_{b c}$ of casualties among cyclists in crashes with cars is calculated as follows:

$$
n_{b c}=\sum_{a, g}\left(\phi_{b}(a, g) \cdot N_{b c}(a, g)-\phi_{c}(a, g) \cdot \tilde{N}_{b c}(a, g)\right)
$$

Again, a small error occurs, as the extra cyclists of age $a$ and gender $g$ cannot crash with the car drivers of the same age and gender that no longer drive a
car. Due to this error we perforce slightly underrate the decrease in the number of casualties.

Group 3b: casualties among car occupants in crashes with a bicycle.
This is a very small group because usually, in a collision between a bicycle and a car, the cyclist will be the casualty, and not the car occupant. Nevertheless, these numbers are also estimated. In the same way as in group 3a it follows that:

$$
n_{c b}=\sum_{a, g}\left(\phi_{b}(a, g) \cdot N_{c b}(a, g)-\phi_{c}(a, g) \cdot \widetilde{N}_{c b}(a, g)\right)
$$

In this formula, $N_{c b}(a, g)$ denotes the number of casualties among car drivers with age $a$ and gender $g$ and $\tilde{N}_{c b}(a, g)$ denotes the number of casualties among car drivers who crashed with a cyclist of age $a$ and gender $g$.

### 6.3. Data



Figure 6.1. Distance travelled on short and long car trips for males and females in different age groups and for the periods 1999-2000 and 2005-2006; open circles: 19992000; closed circles: 2005-2006; solid line: trip length < 7.5 km ; dashed lines: trip length $>7.5 \mathrm{~km}$ (SWOV/NTS, 2012).

## Mobility data

The mobility data for the computations were obtained from the National Travel Survey (SWOV/NTS, 2012). This survey uses a sample of households,
and each person within these households is requested to record all journeys made on one particular day. Age and gender of the persons are known, where age is binned in age groups. Some other variables recorded are trip length and mode of transport. Road type is not available.


Figure 6.2. Distance travelled on short and long bicycle trips for males and females in different age groups and for the periods 1999-2000 and 2005-2006 open circles: 19992000; closed circles: 2005-2006; solid line: trip length < 7.5 km ; dashed lines: trip length $>7.5 \mathrm{~km}$ (SWOV/NTS, 2012).

To calculate the safety effect of a car-bicycle shift, mobility data were analysed in two periods: 1999-2000 and 2005-2006. Figures 6.1 and 6.2 show the distance travelled for males and females for different age groups. From Figure 6.1 it follows that males travel a far greater distance by car than females, especially in long car trips. Figure 6.2 shows a completely different picture. Females travel a greater distance by bicycle in short trips than males. It follows from Figure 6.2 that most bicycle trips are shorter than 7.5 km .

The relative decrease of car mobility as a function of age and gender was calculated for the situation in which $10 \%$ of the short trips (less than 7.5 km ) by car drivers was substituted by bicycle trips; see Figure 6.3. This figure shows that, as a result, between $1 \%$ and $2 \%$ of the total distance travelled by car was replaced by bicycle mobility. The relative decrease of car mobility was calculated for two periods, namely 1999-2000 and 2005-2006. The mobility exchange fractions are almost the same for both periods and
therefore it can be assumed that this holds for the whole period 1999-2006. Therefore the average over the whole period 1999-2006 is used.


Figure 6.3. The relative reduction of car mobility by age and gender, if $10 \%$ of the short passenger car trips (shorter than 7.5 km ) are substituted with bicycle trips. Open circles: 1999-2000; closed circles: 2005-2006; solid line: average.


Figure 6.4. The relative increase of bicycle mobility by age and gender, if $10 \%$ of short passenger car trips (shorter than 7.5 km ) are substituted with bicycle trips. Open circles: 1999-2000; closed circles: 2005-2006; solid line: average.

The small reduction in car mobility (between $1 \%$ and $2 \%$ ) results in a large increase of bicycle mobility (about 10\%); see Figure 6.4. The relative increase in bicycle mobility was much larger than the relative decrease in car mobility, since in The Netherlands the total mobility of bicycles is roughly one tenth of that of passenger cars.

## The number of casualties

Only fatalities and inpatients were included in the computations. In The Netherlands information on these casualties is generally available from the Registered Road Crash Database BRON (SWOV/VOR/BRON, 2012) of the Dutch Ministry of Transport's Centre for Transport and Navigation. For almost all crashes in BRON, age, gender and traffic mode of the parties involved are registered. Thus, the conflict types are also known. However, as already mentioned in the Introduction, only about $4 \%$ of the injured casualties in crashes not involving motorized vehicles are registered in BRON.

Therefore, information on these casualties has to come from another source: the National Medical Registration (SWOV/LMR, 2012). All inpatients of Dutch hospitals, and hence also hospitalized road casualties are registered in this registry. The inpatients are registered in the LMR by age, gender and several other variables such as traffic mode, the presence of a motorized vehicle in the crash, and injury severity, so it is possible to select the hospitalized cyclists in crashes in which no motorized vehicles were involved in the LMR. According to emergency room surveys, $70 \%$ of the crashes in which no motor vehicles are involved were single-vehicle bicycle crashes (Van Kampen, 2007). The remainder were mostly crashes between two bicycles. Of the latter group, the age and gender of the other cyclists were not known. Therefore, this group cannot be treated as group 2d (the number of casualties among cyclists, resulting from crashes between two bicycles). Instead, all hospitalized cyclists are treated as if they belong to group 2a (casualties in single-vehicle bicycle crashes). This results in the effect of extra cycling being underestimated, because the quadratic effect of mobility on the increase in the number of crashes between two bicycles was ignored.

An exploratory analysis showed that about a quarter of the hospitalized cyclists registered in the LMR as a casualty of a crash in which no motor vehicles were involved actually had a crash with a car (Reurings et al., 2007) . Therefore, only 75\% of LMR registered, hospitalized cyclists in crashes in
which no motor vehicles were involved, were included in this analysis. The remaining $25 \%$ were assumed to be covered by the calculation of group 3a.

The number of hospitalized cyclists in single-vehicle crashes was large. In fact, this number was almost entirely responsible for the results found for hospitalized cyclists. To make sure that all those hospitalized are really seriously injured, only injured persons with an injury severity of at least 2 on the internationally used injury severity scale "Maximum Abbreviated Injury Scale, MAIS" (Gennarelli and Wodzin, 2005) were included. For all other hospitalized casualties in this calculation, we do not make this correction. This overestimates the number of other hospitalized casualties, compared to cyclists, because in those cases also minor injuries (less than 2 on the MAIS) are counted. On the other hand, these data were not corrected for police underreporting though the cyclist data were. The hospital data on cyclists were more accurate (there is less underreporting) than the police recorded data. The effects of overestimation and underreporting were assumed to approximately level out.

As some of the numbers of casualties are quite small, we used police record crash data for the period 1999-2006. For serious road injuries of singlevehicle bicycle crashes we used data for 1997-2005, as these were the only data available. In doing so, we ignore the fact that the development in time might be different for different groups. Still, as a first estimation of the order of magnitude of the effect, this approximation is a useful indication.

The number of fatalities and seriously injured for each group are provided in table A in the Appendix.

### 6.4. Results

For each group of casualties the change in the number of casualties due to the car-bicycle shift was determined. This was done for the time period 1999-2006. The results per age group are shown in Table 6.1 for male casualties and in Table 6.2 for female casualties. For example, the effect of the increase of bicycle mobility on fatalities among male drivers aged 18 to 19 was an increase of 0.4 fatalities, whereas the effect of the decrease of car mobility on the same fatalities was a decrease of 2.4 fatalities. Overall, the mobility shift considered in this paper will lead to a decrease of 2.0 fatalities among male car drivers aged 18-19.

In this first approximate assessment, the total gain was unfavourable for safety, as there was a net increase in the number of fatalities. The car-bicycle shift was advantageous for young drivers, and disadvantageous for elderly drivers. In addition, it was more positive for males than for females. The turning point was approximately about the age of 35 . For hospitalized casualties, due to the strong influence of the many hospitalized cyclists in nonmotorized vehicle crashes, there was a strong negative overall effect, and the car-bicycle shift resulted in a positive effect for 18- and 19-year-old males only.

| Age | Male drivers, fatalities |  |  | Male drivers, hospitalized |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Effect of $\phi$ b | Effect of $\phi$ c | Net effect | Effect of $\phi$ b | Effect of $\phi$ c | Net effect |
| 18-19 | 0.4 | 2.4 | -2.0 | 17.1 | 25.6 | -8.5 |
| 20-24 | 2.5 | 6.6 | -4.2 | 101.7 | 66.6 | 35.1 |
| 25-29 | 2.0 | 3.2 | -1.2 | 158.3 | 40.6 | 117.7 |
| 30-34 | 3.1 | 3.6 | -0.4 | 241.2 | 39.8 | 201.3 |
| 35-39 | 5.3 | 2.5 | 2.9 | 247.5 | 35.3 | 212.2 |
| 40-44 | 4.5 | 2.4 | 2.1 | 278.8 | 31.2 | 247.6 |
| 45-49 | 4.6 | 2.0 | 2.6 | 258.6 | 26.0 | 232.6 |
| 50-54 | 5.5 | 1.9 | 3.6 | 278.5 | 23.0 | 255.5 |
| 55-59 | 5.6 | 1.5 | 4.2 | 251.1 | 20.6 | 230.5 |
| 60-64 | 5.0 | 1.3 | 3.7 | 183.0 | 19.7 | 163.3 |
| 65-69 | 7.5 | 1.4 | 6.1 | 171.5 | 18.1 | 153.4 |
| 70-74 | 8.1 | 1.6 | 6.5 | 184.3 | 17.9 | 166.4 |
| 75-79 | 11.7 | 1.4 | 10.4 | 250.7 | 15.4 | 235.3 |
| 80+ | 15.7 | 2.8 | 12.8 | 250.8 | 17.4 | 233.4 |
| Total | 81.6 | 34.6 | 47.0 | 2872.9 | 397.2 | 2475.8 |

Table 6.1. The expected change in the number of male killed and hospitalized casualties (in eight years) due to an increase in bicycle mobility ( $\phi \mathrm{b}$ ) and decrease in car mobility ( $\phi \mathrm{c}$ ).

The overall result of a modal shift of $10 \%$ of all short car trips ( $<7.5 \mathrm{~km}$ ) to bicycle trips would thus amount to an additional 8 fatalities annually (about $1 \%$ of the total number of fatalities, on average ca 1000 annually between 1999 and 2006). Also, the effect would have been an additional 500 inpatients annually ( $3.5 \%$ of the total number of inpatients, on average ca 14,000 annually between 1999 and 2006).

| Age | Female drivers, fatalities |  |  | Female drivers, hospitalized |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Effect of $\phi$ b | Effect of $\phi$ c | Net effect | Effect of фb | Effect of $\phi$ c | Net effect |
| 18-19 | 0.3 | 0.5 | -0.2 | 13.9 | 9.0 | 4.9 |
| 20-24 | 1.1 | 1.6 | -0.6 | 68.0 | 34.5 | 33.5 |
| 25-29 | 1.6 | 1.6 | -0.1 | 97.9 | 32.4 | 65.5 |
| 30-34 | 2.3 | 1.7 | 0.6 | 140.8 | 42.9 | 97.9 |
| 35-39 | 1.9 | 2.0 | -0.1 | 166.7 | 49.3 | 117.4 |
| 40-44 | 2.4 | 1.7 | 0.7 | 183.5 | 37.0 | 146.5 |
| 45-49 | 2.4 | 1.5 | 0.9 | 163.9 | 27.0 | 136.9 |
| 50-54 | 2.7 | 0.9 | 1.8 | 201.9 | 21.1 | 180.8 |
| 55-59 | 2.8 | 1.1 | 1.7 | 187.1 | 17.4 | 169.7 |
| 60-64 | 2.5 | 0.6 | 1.9 | 131.0 | 11.1 | 119.8 |
| 65-69 | 2.2 | 0.7 | 1.5 | 129.0 | 10.3 | 118.7 |
| 70-74 | 2.3 | 0.6 | 1.8 | 169.4 | 7.5 | 161.9 |
| 75-79 | 3.9 | 0.7 | 3.2 | 186.4 | 8.0 | 178.4 |
| 80+ | 3.4 | 0.7 | 2.8 | 163.6 | 6.4 | 157.2 |
| Total | 31.9 | 16.0 | 16.0 | 2003.1 | 313.8 | 1689.3 |

Table 6.2. The expected change in the number of female killed and hospitalized casualties (in eight years) due to an increase in bicycle mobility ( $\phi \mathrm{b}$ ) and decrease in car mobility ( $\phi \mathrm{c}$ ).

### 6.5. Discussion and conclusion

## Discussion

The assessment of the effect of a modal shift from car to bicycle is a first estimate. The assumptions on which the presented calculations are based, may provide an incomplete picture of what actually might occur as a result of such a modal shift. Several effects have not been taken into account. These will be discussed in this section.

In the calculations the road type on which mobility takes place and where casualties occur have not been taken into account. It is, however, known that a large part of car mobility takes place on motorways, while most of the severe crashes take place on urban roads and streets. It is likely that if the total distance travelled in short car trips decreases with $10 \%$, the distance travelled on motorways decreases much less than the distance travelled on other roads. The latter road types have a higher crash risk than motorways,
thus the actual reduction of the number of casualties due to a decrease in short trips is expected to be significantly higher than average.

Unfortunately, no accurate data are available on mobility by road type, age and gender. A simple estimation of the order of magnitude of this effect can be made on the basis of two assumptions: 1) that about half the car mobility takes place on motorways, and 2) the simplification that the relevant crashes occur on all other roads (actually $90 \%$ of all crashes with motorized vehicles occur on all other roads). Thus the positive effect on car-involved casualties would slightly less than double, and therefore reveal a possibly indifferent effect for fatalities, instead of a negative effect. The negative effect on hospitalized casualties would still be very large, because relatively many casualties occur in crashes in which no motorized traffic is involved (falling off one's bike, bicycle-bicycle crashes etcetera). These crashes may even occur on all roads.

The presented assessment assumes constant risk when people cycle more or drive less. Actually, more cycling might mean several things: it may be that people who do not cycle at all, would start cycling, and consequently have a higher than average risk. Another possibility is that if people who cycle not very often, increase the distance they travel by bicycle, they will gain experience which might decrease the average risk. In The Netherlands, cycling is quite common, so it is not expected that there would be people who do not cycle at all. However, in other countries, enhanced bicycle use might lead to enhanced risk. This is especially true for elderly people without much previous experience. People over 70 will be less able and less disposed to change from car to bicycle.

In a study where single-vehicle crashes in different municipalities were compared to the amount of cycling in these municipalities, Schepers (2011) concludes that cyclists are less likely to be involved in a severe single-bicycle crash in municipalities with a high amount of cycling. Perhaps a car-bicycle shift is more likely where there is sufficient bicycle-friendly infrastructure. In that case, the shift could lead to both shorter bicycle trips and safer bicycle routes. This would lead to fewer casualties after the shift than calculated here.

A Dutch study (AVV, 2005) aimed at a similar estimation of the effect of a car-bicycle shift, suggests that in general bicycle trips are $20 \%$ shorter than the car trips they replace. In this report this possibility was not researched,
but the assumption may be valid, in which case the net result of the mobility exchange would be $20 \%$ more advantageous.

If some short car trips with a driver and a passenger were to be replaced by two or more bicycle trips, the safety effects would be different. The passenger, who does not have to be of the same age and gender as the driver, bears a different risk. Therefore, such a calculation is much more complicated than a calculation involving drivers only. It can be carried out with the data available, if it can be assumed that a driver and passenger car trip would be replaced by two bicycle trips. However, this is not necessarily the case:

- If the passenger were a small child, the car trip should probably be replaced by a single bicycle trip with the child as a passenger on the bicycle, but in other cases the car trip should be replaced by multiple bicycle trips.
- If the purpose of the car trip were to bring the passenger to his destination, i.e., the driver acting as the chauffeur of the passenger, different car trips may have to be replaced by a single bicycle trip.
- If the purpose of the car trip were to take both the driver and the passenger(s) to the same destination, the car trip should be replaced by two (or more) bicycle trips.
- If the car driver takes the passenger to a destination which is on his way to his own destination, this results in even more complex replacing of bicycle trips.

The increased safety due to fewer car trips with passengers could be assumed to be equal to that of only a driver's trip, as far as car occupant safety is concerned. Concerning the safety of other road users, the safety effect of the replacement would be the same. The decrease in safety because of the extra bicycle trips for former car passengers would be the same as that of the replacement of a car trip by a bicycle trip, although care should be taken if the age of the passenger substantially differs from the age of the driver. For an accurate calculation, it is necessary to know the reasons for car trips with passengers, the age of all passengers and what would actually occur if these trips were to be exchanged for bicycle trips. Therefore, passenger safety is left out of the calculations.

The study in this paper originated from environmental issues; it was part of a study in which the health benefits of the substitution of short-distance car trips with short-distance bicycle trips were estimated. That study found that
the disease burden related to physical activity was reduced by a maximum of $1.3 \%$ after one year and that the health benefits due to reduction in road traffic noise levels and traffic-related air pollution are relatively small (Van Kempen et al., 2010). Furthermore, De Hartog et al. (2010) calculated the health benefits of cycling instead of driving a car. They concluded that the overall health benefits of cycling were substantially greater than the negative effect on road safety caused by a car-bicycle shift.

The subject is also interesting from another point of view: more cycling instead of driving will probably decrease traffic jams in rush hours. The effect on road safety of more cycling in rush hours can also be calculated. This could be done analogously to the computations in this report, requiring not only stratification by age and gender, but also by time of day. Again, both the mobility effect and the effect on the number of casualties are expected to depend on time of day. As both crash and mobility data is available, for age groups, gender and time of day, the computations are possible.

## Conclusion

A first approximation of the effect on road safety of a mobility shift from car to bicycle was provided. The approximation indicates that, in general, road safety does not benefit from this mobility shift. When $10 \%$ of the short car trips are exchanged for bicycle trips for all ages, our calculations suggest an annual increase of up to $1 \%$ of the fatalities and an approximately $3.5 \%$ increase of all hospitalized casualties in The Netherlands.

The effect of the mobility shift differs for gender and age groups. For example, the research shows that, after a car-bicycle shift, the number of fatalities decreased for young car drivers ( $<35$ years) but increased for older car drivers. So, especially older drivers contribute more to safety inside a car than on a bicycle. The mobility shift increased the number of casualties requiring hospitalization for practically all ages. It was found that it is only beneficial for 18- and 19- year-old males to switch to cycling. Even if the figures mentioned above are an overestimation by a factor 2 , the annual increase in the number of casualties requiring hospitalization would still be 250 , although in this calculation only $10 \%$ of the short trips, which is just $1 \%$ of car mobility, was exchanged.

## Appendix



Table 6.A. The total number of male and female drivers/riders killed and hospitalized casualties in each group, in 1999-2005 (for hospitalized cyclists: 1997-2006).

## References

AVV, 2005. Effect toename fietsaandeel op de verkeersveiligheid. Adviesdienst Verkeer en Vervoer, Delft, the Netherlands.
Bos, N., Reurings, M., Mak, P., 2009. Correction for underreporting of road traffic casualties in the Netherlands; relevance and method. Road safety data: Collection and analysis for target setting and monitoring performances and progress. In: Oecd ed. proceedings of the 4th IRTAD Conference. Seoul, Korea.
Derriks, H., Mak, P., 2007. Underreporting of road traffic casualties. In: OECD ed. IRTAD special report. Paris.
Elvik, R., 2009. The non-linearity of risk and the promotion of environmentally sustainable transport. Accident Analysis and Prevention 41 (5), 849/855.
Gennarelli, T.A., Wodzin, E., 2005. Abbreviated Injury Scale (AIS). Association for the Advancement of Automotive Medicine AAAM, Barrington, ILL.
Graham-Rowe, E., Skippon, S., Gardner, B., Abraham, C., 2011. Can we reduce car use, and if so, how? A review of available evidence. Transportation Research Part A 45, 401418.

De Hartog, J.J., Boogaard, H., Nijland, H., Hoek, G., 2010. Do the health benefits of cycling outweigh the risks? Environmetal Health Perspectives 118, 1109-1116.
Hermans, J., 2012. The energy survival guide, AUP Leiden University Press, Leiden.
Jacobsen, P.L., 2003. Safety in numbers: More walkers and bicyclists, safer walking and bicycling. Injury Prevention 9, 205-209.
Reurings, M., 2010. Ernstig verkeersgewonden in Nederland in 1993-2008: In het ziekenhuis opgenomen verkeersslachtoffers met een MAIS-score van tenminste 2. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, the Netherlands, R-2010-15.
Reurings, M., Bos, N., Van Kampen, L.T.B., 2007. Berekening van het werkelijk aantal in ziekenhuizen opgenomen verkeersgewonden, 1997-2003: Methode en resultaten van koppeling en ophoging van bestanden. SWOV Institute for Road Safety Research, Leidschendam, the Netherlands, R-2007-2.
Reurings, M., Stipdonk, H., 2011. Estimating the number of serious road injuries in the Netherlands. Annals of Epidemiology 21 (9), 648-653.
Schepers, P., 2011. Does more cycling also reduce the risk of single-bicycle crashes? Injury prevention DOI 10.1136/injuryprev-2011-40097.
Schermers, G., Reurings, M., 2009. Verkeersveiligheidseffecten van de invoering van anders betalen voor mobiliteit. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2009-2.
Stipdonk, H., Reurings, M., 2010. The safety effect of exchanging car mobility for bicycle mobility; substituting a small number of short car trips with bicycle trips. Institute for Road Safety Research SWOV, Leidschendam, R-2010-18.
SWOV/LMR, 2012. Website database. National medical registration SWOV/KIWA Prismant.
SWOV/NTS, 2012. Website database. Dutch data on distance travelled. Adviesdienst Verkeer en Vervoer/Centraal Bureau voor de Statistiek.
SWOV/VOR/BRON, 2012. Website database. Dutch annual road crash data. Dienst Verkeer en Scheepvaart, Ministry of Infrastructure and Environment.
Van Kampen, L.T.B., 2007. Gewonde fietsers in het ziekenhuis; een analyse van ongevallenen letselgegevens. The Netherlands: SWOV Institute for Road Safety Research, Leidschendam, R-2007-9.

Van Kempen, E., Swart, W., Wendel-Vos, W., 2010. Exchanging car trips by cycling in the Netherlands. National Institute for Public Health and the Environment RIVM, Bilthoven.
Wegman, F., Zhang, F., Dijkstra, A., 2012. How to make more cycling good for road safety. Accident Analysis and Prevention 44 (1), 19-29.

## 7. Towards an extended model for road safety

This thesis shows that the number of casualties of any travel mode can be understood as a combination of risk and distance travelled of that mode. The approach in this thesis is based on the definition of risk as a fatality density, i.e. the expected number of fatalities per unit of distance travelled. Further, stratification by travel mode, age and gender are applied and for two vehicle crashes the distance travelled by two travellers is taken into account. Distance travelled is therefore a necessary condition for a road fatality. As a consequence the factors which determine the associated risk for a specific piece of distance travelled can all be associated with that specific piece of distance travelled. So risk and distance travelled are inextricably linked. Risk is a property of distance travelled. Each risk factor is only meaningful along a trip. When someone is inexperienced as a car driver this only plays a role when he is driving a car. Ill designed roads are only dangerous when people are using these roads, etcetera. The probability to encounter other traffic also contributes to this risk; hence distance travelled of the other travel modes is also a relevant risk factor.

Risk is defined as the number of fatalities per distance travelled. The most straightforward way to analyse the development of risk is by calculating risk irrespective of whatever property of distance travelled, such as travel mode, age of driver etcetera. This requires data on total distance travelled (car, moped, bicycle etcetera) only. Mobility data are available since 1950, but not for all travel modes, and mobility time series are not consistently measured between 1950 and 2010. Data on pedestrian, bicycle and moped mobility have not been available until 1985. Hence, researchers have used approximations for total distance travelled, such as motor vehicle distance travelled, or car distance travelled (which has been the prevailing travel mode in the entire period). The resulting risk developments, both for motor vehicle distance travelled and for car distance travelled as the denominator, are illustrated in Figure 7.1 using several data sources (cf. Figure 3.3 where car distance travelled was used as denominator). The figure shows a clear decrease of risk in both cases. This suggests that traffic safety has been improving almost incessantly between 1950 and 2011.


Figure 7.1. Comparison between general risk defined as fatalities/total motor vehicle distance travelled (circles) and defined as fatalities/car distance travelled (diamonds). Open symbols denote registered fatalities divided by road count mobility data (CBS, 2000). Closed symbols denote the actual number of fatalities divided by mobility data based on odometer registrations (Nationale AutoPas, NAP) data (CBS, 2011).

However, this is not the end of the story. It was shown that both risk itself and the development of risk over time strongly depend on travel mode and driver's age. Modal shifts and demographic developments may therefore be associated with changes in the number of road casualties. This implies that if one wants to understand changes in the number of road fatalities, it is necessary to stratify, at least by the main travel modes and by age, and preferably by other characteristics for which both the risk and the development of the distance travelled show strong differences between different strata. Because risk and distance travelled are inextricably linked, both data on distance travelled and crash data need to be stratified.

If the risks, associated with the strata of distance travelled are known and constant, it is sufficient to know the development of distance travelled of the different strata in order to explain or forecast the expected number of fatalities. Unfortunately, this risk is not always accurately known and constant for each relevant subgroup. Even if it is known it is not always uniform for all distance travelled within the subgroup. This is due to the fact that within the chosen subgroup some factors vary in such a way that this implies different risk levels and different developments of the distance
travelled. Further stratification with respect to one or more of these varying factors might reveal this.

Simultaneously, stratified analyses of road crash data involves massive data manipulation and requires both detailed crash data and data on distance travelled. Such data are not always readily available. The question arises to what extend this approach is worth further investigation and what further steps are to be taken.

In this chapter we shall consider the results and try to describe the revenues of the approach of this thesis as described above. We shall answer five questions:

1. What have we learned about the stratified approach?
2. What would a road safety model look like which is both feasible and useful?
3. Which steps would be necessary next to develop such a model?
4. Which data are needed for a feasible road safety model?
5. Which would be the revenues of such a model?

The chapter concludes with a vision on a feasible road safety model.

### 7.1. What have we learned about the stratified approach?

In Chapter 2 of this thesis, it was demonstrated that the development over time of the number of fatalities showed different trends for different subgroups. The development of risk also differs for different subgroups, such as car single-vehicle and motorcycle single-vehicle crash risk. Two vehicle crash risk was shown to be likely to be influenced by the development of the distance travelled by the other travel mode involved. Examples were car-car crashes or pedestrian-motorcycle crashes. These risk trends did not follow a smooth decreasing function of time, as was shown in Chapter 3. Specifically, about 1972, the development of risk showed strong anomalies. In order to understand the development of the number of twomode crashes, the incorporation of distance travelled by the other mode involved, was shown to be relevant. In Chapter 3 we saw that the increasing risk of car-car crashes between 1950 and 1972 could be better understood by taking into account the increasing number of available opponent cars. One might suggest that the sharp cusp in the risk of a car-car crash round 1972 in the Netherlands (Figure 3.9, lower left panel) was caused by the development of external factors and safety measures before and after 1972. This would mean that the unsafety first increased between 1950 and 1972.

The increase in risk between 1950 and 1972 is in contrast with the risk decrease for other types of crashes. Hence, it is unlikely that this risk increase was caused by a decreasing amount of safety measures before 1972 or other risk-increasing hypothetical external factors. If that would have been the case there is no explanation why this pattern was not also found for other crash types with a car involved, such as pedestrian-car, motorcycle-car or singlevehicle car crashes. Although there is no way to know for certain that the distance travelled of the other modes was causal for the risk development found, there is sufficient evidence that in understanding these developments it is helpful to incorporate the distance travelled by all modes involved.

Since 1972 the development of risk was shown to be very different for different subgroups as well. Chapters 4 and 5 demonstrate that some subgroups, defined by travel mode and driver's age, showed a much stronger risk decrease than others. Thus, the risk of some subgroups decreased faster than average, whereas the risk of other subgroups decreased more slowly (or even increased). The projection of the trend for such a subgroup for a future year may yield an expected value for this subgroup which is considerably lower or higher than the value we would expect when all subgroups would have the same average trend. However, the sum of the projections of these strata is not necessarily significantly different from the aggregated projection. This is because an increase in the prediction of one subgroup is often cancelled out by a decrease of another. Further, the actual future number is expected to deviate statistically from the underlying expected value, hence a verification of which model is the better one is practically impossible.

In Chapters 4 and 5 , road safety forecasting methods using time projections of risk are described. Forecasts will always contain uncertainties, as they are based on data influenced by chance. The fact that these forecasts are essentially based on mathematical projections of risk, adds considerably to these uncertainties. This is the more so because the time series of the risk is a function of time, such as $r=\alpha+\beta t$ though there is no proof that $\alpha$ and $\beta$ will remain constant in the future. The statistical uncertainty of the steepness of slope $\beta$, together with the fact that the reason why there is a slope anyway (and whether it will last), adds to the uncertainty of the forecast. The former contribution to the uncertainty cannot be eliminated. However, the latter contribution to the uncertainty would be less if the slope is zero, i.e. if risk is constant over time and there is reason to believe that this is a consequence of the absence of factors which influence risk. In that case, the accuracy of the forecast
would still depend on the accuracy of the forecast of the associated distance travelled and of the statistical accuracy of the determined risk. But the uncertainty which is a consequence of the unknown reason of a risk change would be removed.

The alternative for a model which uses projections of risk would be to develop a model which links the number of casualties to external factors (i.e. distance travelled, stratified by all relevant factors), which in turn requires the corresponding risk values for all strata. If these risk values would be invariant with time, we would get rid of the uncertainty resulting from the projection technique. Such a situation would emerge if for a certain group the risk decrease could be attributed to a shift from relatively unsafe distance travelled to relatively safe distance travelled, while each subset would have a constant risk. An example might be the increased use of safety belts in cars. Since the introduction of safety belts, the (unsafe) distance travelled without safety belt was gradually replaced by (safe) distance travelled with safety belt. Aggregation of car distance travelled regardless of safety belt use would reveal a decreasing risk, whereas stratification of car distance travelled with respect to safety belt use might show two groups of constant risk. The risk decrease in the aggregated group would transform to a gradual shift between two groups of different risk.

A mathematical representation of this example might clarify this reasoning: Suppose total car mobility $M$ can be stratified into two strata $M_{1}$ and $M_{2}$, where 1 stands for "no safety belt" and 2 stands for safety belt use. Further suppose that driving without safety belt bears a high constant risk $\mathrm{r}_{1}$, whereas driving with safety belt bears a low constant risk $\mathrm{r}_{2}$. If in year 2000, $\mathrm{M}_{2}=$ (i.e. all distance travelled is without safety belt), and in year $2010 \mathrm{M}_{1}$ is 0 (everyone wears a safety belt), then in year 2000 the aggregate risk equals $\mathrm{r}_{1}$ and in year 2010 the aggregate risk equals $\mathrm{r}_{2}$. This risk apparently decreased from $r_{1}$ to $r_{2}$ in 10 years, because of the increase of safety belt use. An aggregate analysis shows this risk decrease, a stratified analysis shows a modal shift from car use without belt to car use with belt. The decrease of the aggregate risk can be explained by the stratification into strata with constant risk.

A model in which all strata show a constant risk seems attractive, but this would require knowledge of all relevant variables associated with the strata. We know there are many variables such as travel mode, age, gender, BAC, speed, road type, etcetera. A model which describes the (future) number of
casualties in terms of the expected distance travelled by travel mode, age, gender, BAC, speed, road type etcetera, would be quite complex. Even if it were possible to develop such a model, it would be virtually impossible to know the future distance travelled stratified with respect to all these variables. On the other hand, if such a model were available, it could be used to evaluate road policies (including the expected travel by mode). The uncertainties of the expected number of fatalities in a future year would then stem from uncertainties in the values of the external factors in that future year. This would enable calculation of the expected number of fatalities for different scenarios of the developments of these factors.

This illustrates that an accurate estimation of a future number of fatalities, would require an accurate picture of future travel (supposing this were technically feasible). As long as this picture is unknown, the prediction of future number of fatalities must be uncertain, even when the road safety model were perfect and would not contain uncertainties due to extrapolations of risk. This is a consequence of the uncertain development of (stratified) travel itself. Only if per capita distance travelled, stratified by e.g. age and gender were known to be constant, a large part of this uncertainty would be eliminated (see Chapter 4). However, the model enables us to describe how the relevant factors influence road safety, which is still relevant for outlining road safety policy. The fact that the expected number of fatalities is uncertain because of an unknown future modal split is probably more easily accepted by society than an uncertainty because of the unknown development of risk.

In Chapter 6 we showed that the stratification of crash data by travel mode, gender and age can be used to assess what would have occurred in case of a modal shift between car and bicycle. The results for fatalities and for serious injuries were shown to differ considerably. On the one hand, for a small modal shift the number of fatalities is more or less indifferent to a modal shift from car to bicycle, on the other hand such a shift is most unfavourable for the number of non-fatal serious road injuries. Further, it was shown that the effect of a modal shift depends on both age and gender of the driver/rider. This is a further illustration that age and gender are candidates for stratification. Of course a real life verification of the results in this case is not possible at all. It is impossible to measure the actual number of casualties in two different mobility scenarios. Society is no laboratory. We have to believe the model here, and then decide whether such a modal shift is desirable and
worth aiming at with new policies. The uncertainties which stem from the projection technique do not apply in this type of assessment however.

## Summary:

- Road fatalities can only occur under the condition of distance travelled. As a consequence, the (annual) number of road fatalities can be defined as a fatality density (fatalities per distance travelled) times (annual) distance travelled.
- Anywhere along the trips of any traveller, the circumstances of any "piece of distance travelled" determine the associated fatality density or risk.
- Analyses and forecasts of the annual number of road fatalities benefit from stratification by relevant factors of distance travelled, provided that the resulting strata are sufficiently large groups to allow for statistical analyses.
- Relevant factors are those factors for which both risk is not uniform and the development of distance travelled is not constant for different values of that factor.
- Examples of those factors were shown to be travel mode, driver's age and gender, and factors connected to the distance travelled of the other travel mode involved (in case of two vehicle crashes).
- Forecasts of the number of road fatalities suffer from uncertainties in the expected risk and from uncertainties in the expected distance travelled. The latter uncertainty can partly be eliminated if it were likely that stratified per capita distance travelled is approximately constant.
- The uncertainty in risk forecasts based on projections of (negative) exponential time dependent risk models stem from statistical uncertainties, and from the unknown cause of the development of (decreasing) risk. Risk forecasts of strata for which the risk is likely to be constant, suffer less from the latter uncertainty.


### 7.2. What would a road safety model look like which is both feasible and useful?

The most important application of a road safety model would be to support policy making. This means it would enable policy makers to better understand which the past and future effects of their policies were or would be. The feasibility of a road safety model would depend on the possibility to build it. Building it would only be possible if the essential parameters could
be estimated, based on the available data or data which would be made available.

First, there is a distinction to be made between a forecasting and a descriptive road safety model. Estimating the number of road fatalities in the future requires a model which has been developed by analysing and describing the developments of the past sufficiently well. Forecasting models therefore are not essentially different from models which describe past developments. In both cases, the model contains a list of parameters, estimated to fit reality. The model describes the (expected) number of associated fatalities as a function of distance travelled stratified by all relevant factors. However, a forecasting model also requires that we know the values of all relevant factors (to be precise: the amount of distance travelled, stratified by all relevant factors) in the future. Hence, the usefulness of a descriptive model for a forecast is limited to the forecasts of these external factors, even if the model perfectly describes the past. An explanatory model "only" requires that we know the values of these factors in the past, i.e. the distance travelled associated with all combinations of these factors.

Essentially, crashes occur because distance travelled bears a certain risk. The number of fatalities in a specific subgroup equals the risk $\lambda(r(t), \Phi)$ times the travelled distance by that subgroup. Here, $\lambda$ symbolizes the fatality density, i.e. the probability of a fatality per unit of distance travelled. This fatality density is a function of the location $r$ which in turn is a function of time $t$. Here, $r(t)$ denotes a trajectory of a traveller and $\Phi$ denotes the specific values of all relevant factors which together define the risk. For example, the number of single-vehicle moped crashes with 16 year old male riders on modified mopeds depends on the risk associated with the modification, with the young driver's inexperience, with their collective distance travelled and many other factors present along their routes, which we shall not discuss now for the sake of clarity). Now suppose that police enforcement to prevent modification of mopeds was increased in year $x$, in such a way that the fraction of modified mopeds decreased. Further suppose that this fraction is known and the risk associated with modified and non-modified mopeds is known. Then, if distance travelled for both groups is known, we may estimate the changes in the number of fatalities on mopeds. If we also know which mopeds in moped crashes were modified, we are able to check if our reasoning is correct, by analysing the number of crashes with and without modified mopeds.

If we were able to make a model which correctly describes the variation of the probability density $\lambda(r(t))$, it would contain a tremendous amount of variables for all relevant factors to start with. In order to be useful for outlining policy, a model would also have to give insight in the effect of safety measures. As the effects of these safety measures are expressed as a change in the values of one or more of these factors in the model, we would also want to know how this influence works exactly. For example the calculation of the effect of the introduction of a safety helmet for bicyclists involves the estimation of the effect of the helmet itself (i.e. which proportion of the fatal bicycle crashes actually becomes non-fatal because of the helmet having been worn), and the estimation of the penetration of the measure into the distance travelled by bicycle (how many cyclists decide to actually wear a helmet). This is equivalent to knowing the distance travelled with and without helmet. Compliance can be expected to be one of the influencing factors, which in turn will be influenced by enforcement. We further need to know whether the distance travelled by cyclists remains the same as before the introduction of the measure (c.f. the decreased moped use after the introduction of the moped helmet in the Netherlands in 1972, described in the introduction of this thesis). However, in the road safety models described in this thesis distance travelled (including route choice, choice of mode etcetera) is considered an external variable. Finally, we need to know the size of the target group (the number of fatalities among cyclists not wearing helmets (Sigrist, 2010). This requires that crash data contain information on helmet wearing in case of bicyclists involved. Such a model can only be developed if data on these quantities (target groups, penetration, and specific effect) are available for all relevant factors and measures. In 2012, this is beyond the practical possibilities as such data are currently unavailable both in the Netherlands and elsewhere on a scale of national road safety.

Let's assume for a moment that these data are indeed available. Then it would in theory be possible to develop a model in which every imaginable subgroup with respect to age, travel mode, road type, and all other attributes such as bicycle helmet wearing, weather conditions or modified moped use, would be identifiable. There are virtually infinitely many subgroups imaginable. A feasible model cannot contain all these possible subgroups separately. On the other hand, future road safety policies may focus on any of those subgroups. Policy making requires a reasonably accurate estimate of the effects of all kinds of road safety measures on the expected number of fatalities in a future year. The estimates of the effect of future safety measures require some knowledge of the expected number of casualties in the relevant
subgroup if no measures are taken, and of distance travelled in that subgroup. We know now that development of demographics, risk development and expected distance travelled may strongly differ for different subgroups, and therefore it is not preferable to perform rough aggregated forecasts only. Such forecasts cannot be used to estimate policy effects. However, these measures may concern a subgroup which was not specified in the model. This does not necessarily yield a serious problem, as long as two conditions are met:

1. The subgroup involved consists of one or more subsets of any of the subgroups defined in the model.
2. The subsets of interest can be considered a fixed fraction within the subgroups which are specified in the model. This assumption includes the assumption that the distance travelled associated with this subgroup is also a fixed fraction of the distance travelled associated with the subgroups in the model.

One could for instance think of measures regarding road infrastructure. Suppose that a new policy aims at reducing motorway fatalities (or highway intersection fatalities). Further assume that both the proportion of motorway crashes is fixed and the proportion of distance travelled on motorways is fixed for each subgroup in the model. The effect of this measure in a future year can then be expressed as a decrease of this proportion. Of course, this can only be done if the effect of the measure is known for each subgroup.

The results of this thesis suggest that a practical model should make use of demographic data to enable the incorporation of mortality and per capita mobility data, because the development of these quantities in time are more smooth than the time series of mobility and the number of fatalities. Further, an estimate of distance travelled by the most relevant travel modes and age groups is a necessary ingredient. The age groups don't have to be as small as single age years (as in the research described in Chapter 5), because the model doesn't have to be smooth over age years in order to be meaningful. The use of age groups instead of individual age years simplifies the model estimation and therefore increases the feasibility of the model. The age groups may be somewhat larger as long as both the average distance travelled per capita and risk are representative of the group. This would require an approximately linear relation between both distance travelled per capita and risk as a function of age year within the age group. However, this simplification is not entirely without consequences. For example, if demographic developments cause a change in the distribution of age years
within the age group, this may give rise to temporary shifts in the average risk of that group. This is true unless the risk is constant within the age group, which is almost never the case.

Further, for two vehicle crashes it is important that distance travelled of both travel modes is taken into account. We can assume that the distribution over driver age of the probability of being involved in a crash is independent of either the travel mode or the age of the driver of the opponent. Hence, the estimates of risk by travel mode and driver's age can be done separately for both travel modes involved, instead of stratifying by driver's age and travel mode of both vehicles involved simultaneously. This simplifies the analysis. This way, it is not necessary to end up with very small subgroups of combinations of both a specific age group of the driver or rider of each travel mode. This enhances the amount of further stratifications in the model.

## Summary

- Define two types of models: a descriptive model which enables analysis of the development of the number of road fatalities in the past, and a forecasting model, which in some way enables projections of (stratified) risk into a future year.
- In practice the estimation of the effect of safety measures not only requires the size of the effect of the measure in terms of individual crashes, but also of the amount of relevant distance travelled which is influenced by the measure.
- Both the descriptive and the forecasting model should contain as many stratifications as is practically feasible, as far as both the risks associated with the strata and the trends in mobility of the strata are sufficiently different.
- Stratification into subgroups is desirable when these subgroups both bear or are expected to bear a different risk and showed or are expected to show a different development of distance travelled in the past or the future.
- Effects of safety measures of subsets which are smaller than the modelled subgroups can be estimated if both the fraction of crashes and the fraction of distance travelled of this subset of the modelled subgroup are and will remain constant.
- Simplifications of the modelling process are possible if the influence of different variables such as driver's age of the two vehicles involved, can be assumed independent. These simplifications enhance the amount of stratification.


### 7.3. Which steps would be necessary next to develop such a model?

The modelling described in this thesis can be seen as a set of experiments to explore the modelling potential based on the information available in the Netherlands, and to experience the limitations of available data. There are lessons to be learned from the results. The stratification by age of driver, especially if applied simultaneously on the ages of two drivers involved, involves laborious and complex calculations, A feasible model would for example not necessarily need to contain strata with individual age years. Taking the method described in Chapter 5 as a starting point, some other techniques to incorporate the effect of age in the model are worth further exploration.

I propose to develop the following improvements of the model regarding the influence of driver age:

1. The approach described in Chapter 5 can be simplified by choosing driver's age groups, as used in Chapter 4, instead of smoothing data for individual age years. This simplification gives room for other stratifications such as rider's age stratification for the group of crashes with motorized two-wheelers.
2. Crash involvement of car drivers depends on driver's age. This dependence is different for drivers of the car with the casualty and drivers of the other car involved. This difference is approximately an increasing function of age. Research of the development of this function over time might reveal that this function can be easily estimated, and used to simplify the driver's age dependency.
3. Research of the characteristics of the influence of age (and gender) dependent driver behaviour, driver's experience and vulnerability might enable parameterization of the influence of driver's age on risk. A successful parameterization would greatly enhance the possibilities of using road safety models to understand and predict the number of fatalities as a function of distance travelled. Therefore it is worthwhile to carry out such research.

Other improvements could be the incorporation of stratifications that correspond to other relevant factors. These would have to be stratifications of which it is known that the different strata correspond to sufficiently different risks, while the distance travelled of these strata developed (or is expected to develop) with different trends. Further, the number of fatalities in these
strata should be sufficient to allow for a statistical analysis. Incorporation of these factors in the model is desirable if these conditions are known to be met. I propose to research if this is indeed the case for factors such as BAC, speed limit and road type, vehicle safety devices, helmet use, etcetera. Such research would require additional data, as described in the next section.

### 7.4. Which data are needed for a feasible road safety model?

It is clear that a model describing the number of road fatalities (or a model describing the number of serious injuries) requires data on road safety crashes and on distance travelled. Regarding the statistical uncertainties arising from interpreting small numbers of fatalities or of serious injuries, this means that it would be preferable to register every fatality (and serious injury). Further, in case some crashes are missed in the registration it is important to know whether the missed fatalities are a random selection from the total or not. In the latter case a problem arises. The same holds for data on distance travelled. Here, we would like to have information about each trip. As there are billions of trips annually in the Netherlands a subset will suffice as long as the measured subset is sufficiently large and unbiased (Commandeur, 2012). We have to be reasonably certain that the nonregistered trips (and hence also the registered trips) are a random selection of the entire set. If this is not the case, the mobility associated with the nonregistered trips cannot be accurately estimated. Further, the registered subsets need to be sufficiently large regarding the statistical uncertainties arising from the size of the survey on distance travelled. All data requirements mentioned above hold for both descriptive and forecasting models. Forecasting models also need estimations about future distance travelled.

The next question is: which information is needed for each registered fatality and for each road trip in the travel survey? For a researcher crash data and mobility data are generally needed at the most stratified level possible for as many relevant variables possible. However, a practical choice has to be made. The reasoning to decide about this choice is the same as the one behind the desirable stratifications into subgroups as described in the former section. Hence, we are looking for factors for which the subgroups have different risk and for which the development of distance travelled is different. For a stratification by a factor to be of practical use, we must further require that the relevant subgroups associated with the factor are of a sufficient size,
i.e. the corresponding number of fatalities in each subgroup justifies a separate analysis from a statistical point of view, and probably also from a policy making point of view.

Some factors for which stratification is necessary have already been described well in the former chapters of this thesis. These factors are travel mode, driver's age and gender. They satisfy the two conditions: some of the different travel modes had clearly very different risks, and also the use of these modes showed significant different developments over time (e.g. motorcycle and passenger car as was shown in Chapter 3). Further, risk is different for different driver's age and gender as was shown for passenger car drivers and bicycle riders in Chapter 6. The development of travel behaviour is also different for different age groups and gender as was described in Chapter 4. These differences stem from two effects. Firstly, they are the consequence of demographic developments. Secondly, for some travel modes the development of per capita distance travelled is different for different age and gender. A problem with the data on distance travelled is that the risks are high for some travel modes, while the proportion of corresponding distance travelled is low. Hence, while the number of fatalities associated with these travel modes (such as moped and motorcycle) is substantial, the accuracy of the measurement of distance travelled is much lower for these subgroups as compared to other groups such as passenger car drivers. The distance travelled of high risk trips should preferably be measured with at least the same accuracy as that of low risk trips.

There are several other factors for which the two conditions are met. Some of these will be described below.
The first factor known to have a strong influence of risk is the blood alcohol concentration (BAC, measured in mass of alcohol per volume of blood). The risk of a crash increases progressively with increasing BAC. The proportion of impaired distance travelled has changed during the past years. Policies aiming at a reduction of impaired driving will hopefully be successful and lead to a further reduction. To incorporate this effect in the development of the number of fatalities accurately, it is necessary to measure distance travelled by impaired drivers as a function of the BAC. These measurements are currently carried out by the police in regular road check-ups of road users. A problem with the data which are collected in this way is that the probability to encounter a driver with a very high BAC in a regular alcohol check-up is very low, while on the other hand the crash risk of such a driver is extremely high, both the risk to harm himself as the risk to harm another
road user. Hence, the accuracy with which the distance travelled by those impaired drivers with the highest risk is measured, is the least accurate because of the low number of drivers in the sample of checked drivers. For the crash data, it is necessary to know the BAC of all drivers involved in a crash. Unfortunately, it is assumed that the majority of drunken drivers are not identifiable in the crash database (Assum and Sørensen, 2010). The same requirements hold for drugs, though the relation between drug concentration and driving capability is more complicated (Houwing et al., 2013).

Another factor which is important for road safety is speed. It is generally understood that high speed induces high risk, where risk increases progressively with excess speed (Aarts and Van Schagen, 2006). Speed measures (such as speed humps in $30 \mathrm{~km} / \mathrm{h}$ roads) are meant to reduce speed. For other roads speed limits have sometimes changed upward in the past. Hence, we can assume that speed behaviour has changed for some roads. However, at this moment neither for distance travelled nor for crashes, information about actual speed is registered systematically. The road speed limit of the roads used along the trips is not registered in the NTS either. Fortunately a road database containing the speed limits of most of the roads is available since a few years. With this database, combined with a route choice algorithm, the available origin-destination pairs of each trip in the NTS-database could be used to estimate the distance travelled on the most important different road types.

A third possibly important factor is driving experience. Many researchers showed that driving experience has a strong influence on crash involvement (Vlakveld, 2011). Risk differences up to a factor of 10 are mentioned between novice drivers and very experienced drivers. Further, it can be safely assumed that the development of the proportion of distance travelled has changed over time. In the sixties, there were hardly any car drivers with more than 10 years of driving experience, while probably each year more than $10 \%$ of the drivers were inexperienced. Nowadays, the number of novice drivers is approximately $2 \%$ of the population, while a high proportion of the $40+$ year old drivers can be assumed to be very experienced. Unfortunately, the proportion of distance travelled, stratified by experience is not collected at this moment, and the experience of crashinvolved drivers is also not registered.

There are many more factors, which we shall not discuss here at great length, but which are worth mentioning. First there is driving offence behaviour. It is
known that a small subset of distance travelled is associated with a large amount of offences and a much higher risk (Goldenbeld et al., 2012). Then, there are numerous features of distance travelled such as modification of mopeds and light mopeds, safety belt wearing, helmet wearing (by riders of mopeds, light mopeds and bicycles), proper lighting of bicycles, phoning and even texting (by all road users, including cyclists). The risk increase caused by these factors is known to be substantial. For most of these factors the proportion of distance travelled is not systematically measured. Crash data do not contain information on these factors either.

Finally, there are factors worth measuring which weren't mentioned here or which we do not even know yet. It is necessary to pay attention to this question regularly, in order to identify the relevant crash factors of today. Then hopefully it will be possible to implement both a crash registration method and an NTS which enables reporting all relevant factors systematically, for those crashes and for those trips for which these factors are relevant. Perhaps electronic devices can be used which contain software which makes it easy for registering officers to register these factors. Perhaps the NTS can be designed to include new features regularly, in order to meet the requirements of road safety research.

## Summary

- Road safety modelling as described in this thesis requires data on road crashes and data on distance travelled. Crash data should preferably be compete or virtually complete, at least for fatalities and seriously injured. Missing records should not be biased with respect to specific properties.
- Each crash record and each record of a trip in the NTS should contain information on relevant factors.
- Some of these factors are already mentioned in this thesis. These are travel mode, driver's age and gender.
- Other factors worth registering are BAC and drug's use in road users, driver experience, speeding and offence behaviour.
- An inventory should be made of factors worth registering on a regular basis. The crash registration and the travel survey should preferably be organized in such a way that emerging relevant factors can be registered if necessary.
- Subgroups with high risk and low proportion of distance travelled need to be registered in the NTS, with the accuracy at least equal to the accuracy of other subgroups.


### 7.5. Which would be the revenues and the limits of such a road safety model?

If a road safety model together with the data as described above, were available, this would have several important advantages for outlining road safety policy and scientific research. This is the more true if such a model could be developed not only for road fatalities but also for serious road injuries, since the number of serious injuries has been steadily increasing since 2005 (Reurings et al., 2012). Because of the differences between fatal crashes and crashes that cause serious injuries, these models would have different parameters for risk values.

A model as described above (or a simpler version) would have its merits in the possibilities for policy making. It would enable estimating the development of the number of fatal crashes for the important subgroups, in terms of age, travel mode and perhaps some factors such as alcohol or drug's use or any other relevant road safety policy issue. Such a model could serve policy makers by enabling a readily available assessment of the possible effects of their proposed road safety policies. It would enable the calculation of the number of fatalities under different policy scenarios, for all kinds of relevant subgroups. Further, the estimation of the consequences of changes in travel like a modal shift for the number of fatalities, would become feasible. Finally, with such a model the risk differences for different subgroups and the consequences of these differences for the development of road safety would become apparent.

A stratified model describing the development of fatalities and serious injuries for different strata would help policy makers to focus on future road safety problems. In the Netherlands, risk decrease is stronger for fatalities than for serious injuries (Reurings et al., 2012), and stronger for older drivers than for young drivers (Chapter 5). In addition, the crash risk for motorized two-wheelers decreases less than that for the majority of the car occupants (SWOV, 2007). If these trends will last, road safety fatalities will be an issue of many serious injuries and few fatalities within 10 to 20 years, except for young drivers, motorized two wheelers and elderly road users. This means that new policies are necessary, which makes it worthwhile to know the expected development for different subgroups. Policy making will always require new inspirations, because past results do not guarantee future successes.

With a better model, the effect of the "autonomous" risk decrease (the annual decrease as used in the models used in Chapters 4 and 5) might preferably be replaced by a risk decrease which can be shown to be the revenue of specific safety measures and other changing factors. When these factors are not explicitly brought into the model, it is unknown to which extent the decrease is a consequence of actual safety measures. Other factors such as increasing driver's experience, may also have contributed to the decreasing risk in the past. In fact, this is a serious candidate to explain the cusps in risks as described in Chapter 3. It was concluded that it is likely that the collective experience of the driver's population hardly improved between 1950 and 1972. Every subsequent year $15 \%$ of all drivers were likely to be inexperienced. After 1972 the number of new drivers increased much more slowly thereby allowing the driver's population to gain experience. In those years there were probably not many very old drivers yet. This process can be expected to have contributed to the decreasing risk trend since 1972 in the Netherlands. Hence, the decreasing risk since 1972 may be partly (or even largely) attributed to the increased collective driver's experience.

A model which takes the effect of increasing collective driver's experience into account, could help to differentiate between effects of either Dutch or European safety policies and "autonomous" effects, thereby preventing policy making from leaning back, trusting the risk decrease to continue indefinitely. This is not likely to occur.

There are also some limitations to the use of stratified models. Generally, models that provide an improved description of the road safety reality do not necessarily enable better estimations of the actual future number of fatalities. Even with a "perfect" model (i.e. a model which consists of true and exact relations between relevant factors and their influence on the probability of fatalities in traffic), there are so many unknown but necessary future parameters, that a calculation of the expected future number of fatalities is a bridge too far. A model which is capable of estimating the number of non-fatal serious injuries would be even more difficult to develop. Such a model would require a clear definition of a serious injury and sufficient information about the non-fatal crashes involving serious injuries. Such a definition has been adopted in the Netherlands in 2009 (Reurings and Bos, 2009). Unfortunately, the quality of the crash reports are known to be rather poor (Reurings and Bos, 2012).

## Summary

- A model (and the necessary data to build such a model) would enable policy makers to distinguish between different risk groups with a different rate of expected risk decrease and to anticipate on the main future problems in road safety.
- The model would enable estimation of the effect of different road safety policy scenarios and different travel behaviour scenarios.
- An accurate model would enable to distinguish between effects of Dutch safety policies, European safety policies and "autonomous" development (such as the overall increase of driver's experience since 1972).
- The most important drawback of the model proposed here is that its accuracy to predict the number of casualties in a future year will still strongly depend on the expected distance travelled, stratified by the relevant variables. The more accurate the model, the more accurate data on future distance travelled will be needed for predictions.


### 7.6. What makes a good model, and when can policy makers be satisfied with a road safety model?

Modelling road safety is too complex to aim for perfection. On the other hand, very simple models are of little assistance to policy making. So the question arises to what extent road safety models might be enhanced with stratifications and additional factors and still be feasible? And consequently, which efforts are we prepared to make to improve the data collection?

The traditional approach, as was also explored in this thesis, is to find models which fit the data best. The approach assumes that the model follows from the data. The model contains parameters, such as values of the risk or the annual decrease of risk. These values are estimated by minimizing the difference between model values and actual time series of values of risk of specific subgroups. This approach limits the number of subgroups because of the decreasing number of fatalities in each subgroup as the number of subgroups increases.

However, the amount of stratification which is necessary to incorporate all essential risk differences and differences in risk development in the model, probably far outnumbers the practically feasible number of subgroups. A further improvement of road safety models requires a solution to this problem.

It is worth investigating whether improvement could be achieved by incorporating externally derived risk values in the model. For example assume that the risk decrease related to roundabouts as compared to intersections would be known with sufficient certainty. (Actually, the risk decrease due to the replacement of an intersection by a roundabout is known to some extent, but it is not very accurate (Churchill et al., 2010). Then, the ratio of the number of roundabouts and intersections, (if it were known as a time series), could be inserted in the model as a risk decreasing factor. In that case the value of this factor is not estimated by the crash data and the data on distance travelled on roundabouts and intersections. Instead, it is estimated separately and not during the process of optimizing the model. The challenge here lies in obtaining the data on the number of roundabouts and intersections, and their ratio, weighted with traffic volume. Similarly, if it were possible to express the difference in risk between male and female drivers as a constant factor, it would no longer be necessary to distinguish driver's gender. However, such an approach requires that all kinds of causes of these differences in risk between males and females wouldn't change over time. An example for such a cause is the systematic difference between car mass of cars driven by males or females (Evans and Frick, 1993, Evans, 1994, Evans and Gerrish, 2001, Berends, 2009). A third example could be the influence of enhanced enforcement, decreasing the fraction of extreme speeding. If the relation between enforcement or driving offences (Goldenbeld et al., 2012), and driving (mis)behaviour were known, changes in risk associated with data on changes in enforcement or offences can be estimated and incorporated in the model.

Are these improvements worth the effort? I am convinced this is the case. A model which enables calculation of the effect of a roundabout in terms of decrease in the number of fatalities or injuries, is cheap compared to the cost of even a single roundabout. There are about 4000 roundabouts in the Netherlands (Churchill et al., 2010). It is impossible to say if the locations of these roundabouts were the best choice in terms of road safety improvement. Nevertheless, it seems very unlikely that this is the case. A similar reasoning holds for enforcement. The current efforts to enforce traffic rules, using cameras and road side checks are not the outcome of a strategy which optimizes road safety. It is unlikely that the current balance between road side checks and camera's is the best possible way to invest the available budget in enforcement. Hence a complete lack of models is certainly unsatisfactory, although models can always be criticized because of their uncertainties.

A model which can be used to support stratified policy making, requires stratification, incorporation of stratified distance travelled and incorporation of safety measures. The more complex such a model, the more complex the statistics, and consequently the less statistical certainty of the correctness of the model. That is the choice which has to be made: either we leave the data in an aggregated form, thereby allowing greater statistical accuracy, or we accept the statistical noise associated with stratification. The rich possibilities of exploring the latter make it worth to choose this path of research further.

## References

Aarts, L., Weijermars, W., Schoon, C., Wesemann, P., 2008. Maximaal 500 verkeersdoden in 2020: Waarom eigenlijk niet? : Maatregelpakketten en effectschattingen om te komen tot een aangescherpte verkeersveiligheidsdoelstelling. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2008-5.
Aarts, L.T., Van Schagen, I.N.L.G., 2006. Driving speed and the risk of road crashes: A review. Accident Analysis and Prevention 38 (2), 215-224.
Agora Jules Dupuit, 1999, www.e-ajd.net, Safety page, Excel files AJD-7.
Assum, T., Sørensen, M., 2010. Safety performance indicator for alcohol in road accidentsinternational comparison, validity and data quality. Accident Analysis and Prevention 42 (2), 595-603.
AVV, 2005. Effect toename fietsaandeel op de verkeersveiligheid. Adviesdienst Verkeer en Vervoer, Delft, the Netherlands.
Bellemans, T., Kochan, B., Janssens, D., Wets, G., Arentze, T., Timmermans, H., 2010. Implementation framework and development trajectory of feathers activity-based simulation platform. Transportation Research Record: Journal of the Transportation Research Board 2175, 111-119.
Berends, E., 2009. De invloed van automassa op het letselrisico bij botsingen tussen twee personenauto's; een kwantitatieve analyse. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2009-5.
Bijleveld, F., 2008. Time series analysis in road safety research using state space methods, SWOV dissertatiereeks, SWOV Institute for Road Safety Research, Leidschendam.
Bijleveld, F.D., Commandeur, J.J.F., Gould, P.G., Koopman, S.J., 2008. Model-based measurement of latent risk in time series with applications. Journal of the Royal Statistical Society Series A 171 (1), 265-277.
Booth, H., 2006. Demographic forecasting: 1980 to 2005 in review International Journal of Forecasting 22 (3), 547-581.
Bos, N., Reurings, M., Mak, P., 2009. Correction for underreporting of road traffic casualties in the Netherlands; relevance and method. Road safety data: Collection and analysis for target setting and monitoring performances and progress. In: OECD ed., Proceedings of the 4th IRTAD Conference. Seoul, Korea.
Broughton, J., 1991. Forecasting road accident casualties in Great Britain. Accident Analysis and Prevention 23 (5), 353-363.
Broughton, J., 2003. The benefits of improved car secondary safety. Accident Analysis and Prevention 35 (4), 527-535.
Broughton, J., 2012. Post-2010 casualty forecasting. Road safety Web publication. Transport Research Laboratory, London.
Broughton, J., Allsop, R., Lynam, D., McMahon, C., 2000. The numerical context for setting national reduction targets. Transport Research Laboratory, London.
CBS, 1950 and later years. Statistiek van de verkeersongevallen op de openbare weg 1950. Dutch tabulated data on Dutch traffic accidents. In: CBS ed. Centraal Bureau voor de Statistiek, Den Haag.
CBS, 1985 and onward. Statline, motorvoertuigen: Statistiek van het Nederlandse motorvoertuigenpark. Centraal Bureau voor de Statistiek CBS, Voorburg/Heerlen.

CBS, 2000. Statline, statistiek van de wegen. Centraal Bureau voor de Statistiek CBS, Voorburg/Heerlen.
CBS, 2011. Statline, nationale autopas mobiliteitsdata. Centraal Bureau voor de Statistiek CBS, Voorburg/Heerlen.
CBS, 2012. Statline. Centraal Bureau voor de Statistiek, Den Haag.
Churchill, T., Stipdonk, H., Bijleveld, F., 2010. Effects of roundabouts on road casualties in the Netherlands. SWOV, Leidschendam, R-2010-21.
Commandeur, J.J.F., 2012. Sampling techniques and naturalistic driving study designs. Deliverable 6.2 B of the EC FP7 project DaCoTA.
Commandeur, J.J.F., Bijleveld, F.D., Van Norden, Y., Stipdonk, H., Bergel, R., Debbarh, M., Yannis, G., Antoniou, C., Katsochis, D., Papadimitriou, E., 2008. Time series applications on road safety developments in Europe. In: Thomas, P. ed. Building the European Road Safety Observatory. BiVV, INRETS, KfV, NTUA, SWOV, Brussels.
Commandeur, J.J.F., Koopman, S.J., 2007. An introduction to state space time series analysis, Oxford University Press, Oxford.
COST, 2004. Models for traffic and safety development and interventions. European Commission. Directorate general for Transport, Brussels.
Davidse, R.J., 2007. Assisting the older driver, SWOV-dissertatiereeks, SWOV Institute for Road Safety Research Leidschendam.
De Ceuster, G., 2005. Assessment of the contribution of the TEN and other transport policy measures to the mid term implementation of the white paper on the European transport policy for 2010. In: DG TREN, E.C. ed. TML, Leuven.
Derriks, H., Mak, P., 2007. Underreporting of road traffic casualties. In: OECD ed. IRTAD special report. Paris.
Doornik, J.A., 2001. Object-oriented matrix programming using Ox 3.0, Timberlake Consultants Press, London.
Durbin, J., Koopman, S.J., 2001. Time series analysis by state space methods, Oxford University Press, Oxford.
Eksler, V., 2009. Road mortality in Europe: A regional approach, Université de Versailles, Saint-Quentin-en-Yvelines.
Elvik, R., 2005. Has progress in improving road safety come to a stop? Institute of Transport Economics TØI, Oslo.
Elvik, R., 2006. Laws of accident causation. Accident Analysis and Prevention 38 (4), 742-747.
Elvik, R., 2009a. An exploratory analysis of models for estimating the combined effects of road safety measures. Accident Analysis and Prevention 41 (4), 876-880.
Elvik, R., 2009b. The non-linearity of risk and the promotion of environmentally sustainable transport. Accident Analysis and Prevention 41 (5), 849/855.
Elvik, R., Bjørnskau, T., 2005. How accurately does the public perceive differences in transport risks?: An exploratory analysis of scales representing perceived risk. Accident Analysis and Prevention 37 (6), 1005-1011.
Elvik, R., Vaa, T., 2004. The handbook of road safety measures, Elsevier Ltd, Oxford, UK.
ETSC, 1999. Exposure data for travel risk assessment: Current practice and future needs in the eu. European Transport Safety Council, Brussels.
European Union, 1995-2011. Eurostat. European Union, Brussels.
Evans, L., 1986. Double pair comparison-a new method to determine how occupant characteristics affect fatality risk in traffic crashes. Accident Analysis and Prevention 18 (3), 217-227.

Evans, L., 1994. Driver injury and fatality risk in two-car crashes versus mass ratio inferred using newtonian mechanics. Accident Analysis and Prevention 26 (5), 609-616.
Evans, L., Frick, M.C., 1993. Mass ratio and relative driver fatality risk in two-vehicle crashes. Accident Analysis and Prevention 25 (2), 213-224.
Evans, L., Gerrish, P.H., 2001. Gender and age influence on fatality risk from the same physical impact determined using two-car crashes. SAE 2001 World Congress. PA: Society of Automotive Engineers, Detroit.
Filarski, R., 2004. The rise and decline of transport systems, changes in a historical context, Ministry of Transport and Public Works, Rotterdam.
Fournier, F., Simard, R., 2002. The DRAG-2 model for quebeck. In: Gaudry, M., Lassarre, S. eds. Structural road accident models; the international DRAG family. Pergamon, pp. 37-66.
Fridstrøm, L., Ifver, J., Ingebrigsten, S., Kulmala, R., Thomsen, L., 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.
Gaudry, M., 1984. DRAG, model of the demand for road use accidents and their severity, applied in Quebec from 1956 to 1982., Université de Montréal, Montréal.
Gaudry, M., 2005. Presentation at 1st FERSI scientific road safety research conference, BASt, Bergische Gladbach.
Gennarelli, T.A., Wodzin, E., 2005. Abbreviated Injury Scale (AIS). Association for the Advancement of Automotive Medicine AAAM, Barrington, ILL.
Goldenbeld, C., Stipdonk, H., Reurings, M., Van Norden, Y., 2012. Crash involvement of motor vehicles in relationship to the number and severity of traffic offences. An exploratory analysis of Dutch traffic offences and crash data. Traffic Injury Prevention DOI:10.1080/15389588.2012.743125.
Graham-Rowe, E., Skippon, S., Gardner, B., Abraham, C., 2011. Can we reduce car use, and if so, how? A review of available evidence. Transportation Research Part A 45, 401418.

Hale, A.R., 2001. Conditions of occurrence of major and minor accidents. Journal of the Institution of Occupational Safety and Health 5 (1), 7-21.
De Hartog, J.J., Boogaard, H., Nijland, H., Hoek, G., 2010. Do the health benefits of cycling outweigh the risks? Environmetal Health Perspectives 118, 1109-1116.
Harvey, A.C., 1989. Forecasting, structural time series models and the Kalman filter, Cambridge University Press, Cambridge.
Hermans, J., 2012. The energy survival guide, AUP Leiden University Press, Leiden.
Houwing, S., Hagenzieker, M., Mathijssen, R.P.M., Legrand, S.-A., Verstraete, A.G., Hels, T., Bernhoft, I.M., Simonsen, K.W., Lillsunde, P., Fravretto, D., Ferrara, S.D., Caplinskiene, M., Movig, K.L.L., Brookhuis, K.A., 2013. Random and systematic errors in case control studies calculating the injury risk of driving under the influence of psychoactive substances. Accident Analysis and Prevention 52 (1) 144153.

Jacobsen, P.L., 2003. Safety in numbers: More walkers and bicyclists, safer walking and bicycling. Injury Prevention 9, 205-209.
Jacket, M., Frith, W., 2013, Quantifying the impact of road lighting on road safety - a New Zealand study, International Association of Traffic and Safety Sciences Research, 36, 135-145.
Jaeger, L., Lassarre, S., 2002. The TAG-1 model for France. In: Proceedings of the Structural road accident models; The international DRAG family, pp. 157-184.

Janssens, D., Wets, G., Timmermans, H.J.P., Arentze, T.A., 2007. Modeling short-term dynamics in activity-travel patterns: Conceptual framework of the feathers model. In: Proceedings of the 11th World Conference on Transport Research, Berkeley CA, USA.
Kam, B.H., 2003. A disaggregate approach to crash rate analysis. Accident Analysis and Prevention 35 (5), 693-709.
KiM, 2012. Mobiliteitsbalans 2012. Kennisinstituut voor Mobiliteitsbeleid, Ministerie van Infrastructuur en Milieu, Den Haag.
Koopman, S.J., Shephard, N., Doornik, J.A., 1999. Statistical algorithms for models in state space using SsfPack 2.2. Econometrics Journal Vol. 2, 113-166.
Kopitz, E., Cropper, M., 2005. Traffic fatalities and economic growth. Accident Analysis and Prevention 37, 169-178.
Laapotti, S., Keskinen, E., 2004. Has the difference in accident patterns between male and female drivers changed between 1984 and 2000. Accident Analysis and Prevention 36 (4), 577-584.
Lejeune, P., Treny, V., Duchamp, G., Hemdorff, S., Haddak, M., Hollo, P., Cardoso, J., Arsenio, E., Yannis, G., Papadimitriou, E., Bijleveld, F., Houwing, S., Bjørnskau, T., Rackliff, L., Angermann, A., 2007. First classification of the EU member states on risk and exposure data. In: Thomas, P. ed. SafetyNet. European Union, Brussels.
Lenroot, R.K., Gogtay, N., Greenstein, D.K., Wells, E.M., Wallace, G.L., Clasen, L.S., Blumenthal, J.D., Lerch, J., Zijdenbos, A.P., Evans, A.C., Thompson, P.M., Giedd, J.N., 2007. Sexual dimorphism of brain developmental trajectories during childhood and adolescence. NeuroImage 36 (4), 1065-1073.
Lie, A., Krafft, M., Kullgren, A., Tingvall, C., 2008. Intelligent seat belt reminders. Do they change driver seat belt use in Europe? Traffic Injury Prevention 9 (5), 446-449.
Maier, R., Ahrens, G.-A., 2009. Entwicklung der strassenverkehrssicherheit und ihrer rahmenbedingungen bis 2015/2020. Forschungs- und Entwicklungsvorhaben. Technische Universität Dresden, Dresden.
Martin, J.-L., 2002, Relationship between crash rate and hourly traffic flow on interurban motorways, Accident Analysis and Prevention 34 (5), 619-629.
Massie, D.L., Green, P.E., Campbell, K.L., 1997. Crash involvement rates by driver gender and the role of average annual mileage Accident Analysis and Prevention 29 (5), 675-685.
Maycock, G., 2001. Forecasting older driver accidents and casualties. Department of the Environment, Transport and the Regions.
Mesken, J., 2006. Determinants and consequences of drivers'emotions, SWOV dissertatiereeks, SWOV Institute for Road Safety Research, Leidschendam.
McCartt, A.T., Mayhew, D.R., Braitman, K.A., Ferguson, S.A., Simpson, H.M., 2009. Effects of age and experience on young driver crashes: Review of recent literature. Traffic Injury Prevention 10, 209-219.
Moritz, G., Brög, W., 1999. Redesign of the Dutch travel survey, paper presented at the TRB conference on personal travel, Washington D.C. Statistics Netherlands, Heerlen.
Nicholson, A., Wong, Y.-D., 1993. Are accidents Poisson distributed? A statistical test. Accident Analysis and Prevention 25 (1), 91-97.
NTS-data, 2010. Web-accessible database of Dutch national travel survey. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.
O'Reilly, D., Hopkin, J., Loomes, G., Jones-Lee, M., Philips, P., McMahon, K., Ives, D., Soby, B., Ball, D., Kemp, R., 1994. The value of road safety. UK research on the valuation of
preventing non-fatal injuries. Journal of Transport Economics and Policy 28 (1), 4559.

OECD, 2007. Towards zero, ambitious road safety targets and the safe system approach, OECD/ITF, Paris.
OECD/ITF, 2012. Reporting on serious road traffic casualties; combining and using different data sources to improve understanding of non-fatal road traffic crashes. In: IRTAD ed. Organization for Economic Co-operation and Development OECD/International Transport Forum ITF, Paris.
Oppe, S., 1991a. The development of traffic and traffic safety in six developed countries. Accident Analysis and Prevention 23 (5), 401-412.
Oppe, S., 1991b. Development of traffic and traffic safety: Global trends and incidental fluctuations. Accident Analysis and Prevention 23 (5), 413-422.
OVG/MON, 1985-2009. Onderzoek Verplaatsingsgedrag en Mobiliteitsonderzoek Nederland. Statistics Netherlands, Dutch annual mobility survey.
Ploeger, J., Van der Waard, J., 1997. Waar komt de groei vandaan? In: Infrastructurele ontwikkelingen, Adviesdienst Verkeer en Vervoer, Rotterdam.
Polinder, S., Haagsma, J.A., Toet, H., Beeck, E.F.V., 2012. Epidemiological burden of minor, major and fatal trauma in a national injury pyramid. British Journal of Surgeon Society Ltd 99(suppl1) (114-121).
Pooley, C.G., Turnbull, J., 2000. Modal choice and modal change: The journey to work in Britain since 1890. Journal of Transport Geography 8, 11-24.
Reurings, M., 2010. Ernstig verkeersgewonden in Nederland in 1993-2008: In het ziekenhuis opgenomen verkeersslachtoffers met een MAIS-score van ten minste 2. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2010-15.
Reurings, M., Bos, N., 2009. Ernstig gewonde verkeersslachtoffers in Nederland in 1993-2008; het werkelijk aantal in ziekenhuizen opgenomen verkeersslachtoffers met een MAIS van ten minste 2 . Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2009-12.
Reurings, M., Bos, N., 2012. Ernstig verkeersgewonden in de jaren 2009 en 2010. SWOV, Leidschendam, R-2012-7.
Reurings, M., Bos, N., Van Kampen, L.T.B., 2007. Berekening van het werkelijk aantal in ziekenhuizen opgenomen verkeersgewonden, 1997-2003: Methode en resultaten van koppeling en ophoging van bestanden. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, The Netherlands, R-2007-2
Reurings, M., Commandeur, J., 2006. International orientation on methodologies for modelling developments in road safety. SWOV, Leidschendam, R-2006-34.
Reurings, M., Stipdonk, H., 2011. Estimating the number of serious road injuries in the Netherlands. Annals of Epidemiology 21 (9), 648-653.
Reurings, M., Stipdonk, H., Minnaard, F., Eenink, R., 2012. Waarom is de ontwikkeling van het aantal ernstig verkeersgewonden anders dan die van het aantal verkeersdoden? SWOV, Leidschendam, R-2012-9
Schepers, P., 2011. Does more cycling also reduce the risk of single-bicycle crashes? Injury prevention DOI 10.1136/injuryprev-2011-40097.
Schermers, G., Reurings, M., 2009. Verkeersveiligheidseffecten van de invoering van anders betalen voor mobiliteit. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2009-2.
Schofield, G.M., Gianotti, S., Badland, H.M., Hinckson, E.A., 2008. The incidence of injuries traveling to and from school by travel mode. Preventive Medicine 46 (1), 74-76.

Shannon, C.E., 1948. A mathematical theory of communication. The Bell System Technical Journal Vol. 27 (July, october), 379-423; 623-656.
Sigrist, S., 2010. Towards a method to forecast the effectiveness of national road safety programmes. Safety Science 48 (9), 1106-1110.
Stipdonk, H., Berends, E., 2008. Distinguishing traffic modes in analysing road safety development. Accident Analysis and Prevention 40 (4), 1383-1393.
Stipdonk, H., Reurings, M., 2010. The safety effect of exchanging car mobility for bicycle mobility; substituting a small number of short car trips with bicycle trips. SWOV Institute for Road Safety Research, Leidschendam, R-2010-18.
Stipdonk, H., Wesemann, P., 2007. Long-term forecasts of road traffic fatalities in the European union : General methods and results. SWOV Institute for Road Safety Research, Leidschendam, R-2006-15.
SWOV, 2007. De top bedwongen. Balans van de verkeersonveiligheid in Nederland 19502005, Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam.
SWOV, 2012. Website database. data on crashes, hospital patients, distance travelled, fleet size, demography, safety measures, etcetera. SWOV Institute for Road Safety Research, Leidschendam.
SWOV/CBS, 2012. Website database. Dutch demographic data. Centraal Bureau Statistieken, Statistics Netherlands.
SWOV/IRTAD, 2012. Website database. Annual road crash data of IRTAD countries. Dienst Verkeer en Scheepvaart, Ministry of Infrastructure and Environment.
SWOV/LMR, 2012. Website database. National medical registration SWOV/KIWA Prismant.
SWOV/NTS, 2012. Website database. Dutch data on distance travelled. Adviesdienst verkeer en vervoer/Centraal Bureau Statistieken.
SWOV/VOR/BRON, 2012. Website database. Dutch annual road crash data. Dienst Verkeer en Scheepvaart, Ministry of Infrastructure and Environment
Taha, H.A., 1987. Operations research, Macmillan publishing company, New York.
TfL, 2006. Central London congestion charging 4th annual report. London.
Van den Bossche, F.A.M., 2006. Road safety, risk and exposure in Belgium: An econometric approach. Dissertation, Hasselt University, Diepenbeek, Belgium, Hasselt.
Van den Bossche, F.A.M., Wets, G., Brijs, T., 2007. Analysis of road risk by age and gender category. Time series approach. Journal of the Transportation Research Board 2019, 7-14.
Van Norden, Y., 2009. A road safety assessment model for the Netherlands. A feasibility study using disaggregation by traffic mode and age. Young Researcher Seminar 2009. European Conference of Transport Research Institutes (ECTRI), Forum of European National Highway Research Laboratories (FEHRL) and Forum of European Road Safety Research Institutes (FERSI), Torino.
Van Kampen, L.T.B., 2007. Gewonde fietsers in het ziekenhuis; een analyse van ongevallenen letselgegevens. The Netherlands: SWOV Institute for Road Safety Research, Leidschendam, R-2007-9.
Van Kempen, E., Swart, W., Wendel-Vos, W., 2010. Exchanging car trips by cycling in the Netherlands. National Institute for Public Health and the Environment RIVM, Bilthoven.
Vlakveld, W., 2011. Hazard anticipation of young novice drivers. SWOV-dissertatiereeks, SWOV Institute for Road Safety Research, Leidschendam.
Volvo, 2007. Safety first, a list of safety measures introduced by Volvo. www.volvocars.com.

Wegman, F., Zhang, F., Dijkstra, A., 2012. How to make more cycling good for road safety. Accident Analysis and Prevention 44 (1), 19-29.
Weijermars, W., Wegman, F., 2010. The balance struck; evaluation of ten years sustainable safety. Transportation Research Record: Journal of the Transportation Research Board.
Wesemann, P., 2007. De verkeersveiligheid in 2020 : Verkenning van ontwikkelingen in mobiliteit, ongevallen en beleid. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam, R-2006-27.
Wesemann, P., 2008. De verkeersveiligheid in 2020; een verkenning van ontwikkelingen in mobiliteit, ongevallen en beleid. Tijdschrift Vervoerwetenschap 44 (juni), 36-45.
Wesemann, P., Van Norden, Y., Stipdonk, H., 2010. Road safety in 2020; future developments of exposure, crashes and policy. Safety Science 48 (9), 1098-1105.
Yannis, G., Antoniou, C., Eleonora Papadimitriou, Katsohis, D., 2011. When may road fatalities start to decrease? Journal of Safety Research 42 (1), 17-25.
Yannis, G., Papadimitriou, E., Chaziris, A., Duchamp, G., Lejeune, P., Treny, V., Hemdorff, S., Haddak, M., Lenguerrand, E., Holló, P., Gábor, M., Cseffalvay, M., Leitner, T., Angermann, A., Hoeglinger, S., Cardoso, J., Bijleveld, F., Houwing, S., Bjørnskau, T., Rackliff, L., 2008. Risk exposure data common framework. In: Thomas, P. ed. Building the European Road Safety Observatory. European Commission, Brussels.

## Summary

This thesis describes a method in order to get a better understanding of the development of traffic safety. The unsafety of traffic, expressed as the annual number of road fatalities, has shown a remarkable development in the Netherlands over the past years. In about 1950 there were approximately a thousand fatalities annually. Between 1950 and about 1972 this number increased approximately exponentially to almost 3500. Afterward, a decrease followed to about 650 fatalities in 2012. This means that the annual number of fatalities decreased during most of the years, but not always and not at the same pace. In some years there was even an increase (as in 1976 and 1977, when there was an increase of more than 100 fatalities). And also in 2011 and 2012 the number of fatalities increased in comparison to the previous year. In many other high income countries a similar development was seen: an increasing number of fatalities until 1972 and a decrease thereafter, with irregular jumps up and down in the annual number of fatalities.

The goal of this thesis is to research how these developments can be related to factors such as the amount of traffic, road safety measures implemented by the government, etcetera. The thesis is a search for the way to define how these factors determine road safety. The emphasis is on fatalities, because the data for those are the most reliable ones, both in the Netherlands as in many other countries. Here we focus on the analysis of Dutch data. However, the final goal is to develop a general method which can also be used to analyse road crash data for other countries.

In the past, a lot of research was done with the same purpose. We now know that the development can be explained roughly by a combination of increased car distance travelled and increased road safety. The initial increase was attributed to the explosive increase of car distance travelled by more than $13 \%$ annually. This increase was so strong that the measures to increase road safety were insufficient to compensate the increase. After 1972 this increase diminished to 3\% annually, and in 2011 the increase had almost come to a standstill. Because of this much smaller increase the numerous safety measures could notably influence the annual number of fatalities.

This explanation immediately shows that the amount of traffic is one of the determining quantities in traffic safety. In this thesis this quantity is called mobility or distance travelled and it is expressed in km . The other
determining quantity is the number of fatalities per distance travelled. In this thesis this quantity is called risk and it is expressed in fatalities $/ \mathrm{km}$. Risk and distance travelled are both important to understand the development of road safety, but there is also an important difference. Risk is the core subject of road safety policy and -research, whereas mobility is seen as an autonomous quantity by Dutch road safety researchers and policy makers. Hence, Dutch road safety policy making does not make use of the possibility to improve road safety by limiting distance travelled. Policy making exclusively aims at reducing the number of road casualties, given the amount of distance travelled. In the terminology of this thesis this means we want to decrease risk.

Road safety policy makers often make use of a long term political target: a maximum expected number of road fatalities in a certain future year. To know which road safety policy is necessary to achieve a chosen target, it is necessary to understand how the development of road safety depends on past policies. For that purpose the rough explanation of the development of the number of fatalities since 1950 mentioned above (at first mostly an explosive mobility increase, and afterwards mostly successful safety measures) is of little help. Those who want to know e.g. which measures contributed to which extent to the road safety improvements, have to analyse further. Knowledge of such relations is important to know to what extent the effects of past policies will continue in the future.

Research of the development of road safety, also described as road safety assessment and forecasting, aims at three challenges:

1. Understanding the developments of past road safety
2. Construction of models with which these developments can be understood and predicted
3. Estimating the expected developments in the future

This thesis describes an approach which consists of stratified (i.e. subdivided into groups) analysis of the development. Mobility is not always and everywhere equally dangerous. A 10 km ride on a modified moped bears much more risk for the rider than a 10 km trip by a truck driver. Every part of a trip has a matching risk, determined by numerous factors. The driver, the vehicle, the road design, traffic rules, weather conditions, enforcement... all are factors which matter.

Stratification of data is also a logical step for the estimation of the effects of road safety measures. Road safety measures almost always relate to only one
subset of the mobility, hence to only one subset of crashes. Think of the safety belt, the moped helmet, the blind spot mirror for trucks, alcohol laws etcetera.

Approximately $30 \%$ of all road fatalities occurs in a single-vehicle crash. A single-vehicle crash is a crash in which the victim does not collide with another road user, but falls or collides with a tree or pole. In the registration of such crashes there is no other road user involved. This is different for twosided crashes. For those crashes, the vehicle type of the other road user, age and experience of the driver and his blood alcohol concentration (BAC) are also relevant. In this thesis we investigate how the influence of distance travelled of other road users can be taken into account in the calculation of risk.

Subgroups can be formed in many different ways: casualties among moped riders or car occupants, among children or elderly, casualties with or without a certain BAC, etcetera. But one can also consider fatalities in a crash where a car was involved as the other party, or where a younger person drove the other car, etcetera. The developments of the number of fatalities were not uniform for all those different subgroups. Risk and distance travelled sometimes developed quite differently for different kinds of mobility. In Chapter 2 the applied method of stratification is further explained and founded. The chapter describes the way in which different kinds of factors determine the developments and which data are needed to carry out the analyses. Crash data (such as the VOR-database for years between 1976 and 2003, and the BRON database since 1985) and data on distance travelled (among others the OVG/MON/OViN database) are very important.

VOR/BRON and OVG/MON/OViN distinguish several variables such as travel mode, age and gender of the traveller. Hence it is possible to carry out separate analyses for specific travel modes or age groups. However, VOR/BRON has been available from 1976 onward only, and OVG/MON/OViN was available since 1985. As a consequence it is much more complicated to perform stratified analyses of mobility and risk before 1985. The only stratified mobility data in that period can be found in the motor vehicle mobility statistic of Statistics Netherlands (CBS). Statistics Netherlands has estimated the amount of distance travelled for passenger cars, trucks and motorcycles, based on random roadside traffic counts. For data of road fatalities we must make use of tables, also edited by Statistics Netherlands. In Chapter 3 these data are used to show that there is a
considerable difference between analysis of the aggregated number of fatalities and stratified analyses. First, the development of aggregated risk is calculated, i.e. the total number of fatalities is divided by the total distance travelled by motorized vehicles. The result shows that aggregated risk decreased by $6 \%$ to $7 \%$ annually between 1950 and 2000. Next, the development of the number of fatalities is studied for three subgroups, i.e. for car occupants, truck occupants and motorcycle riders. The result shows that the number of fatalities in these subgroups developed differently. Mobility appears to have been an important factor. The development of the number of fatalities among motorcycle riders showed several highs and lows that coincided with similar highs and lows in motorcycle use. The risks for these three travel modes appear to have developed in approximately identical ways: risk slightly increased between 1950 and 1972 and decreased afterwards. For motorcyclists we also see a particular risk increase about 1970, when risk doubled within one or two years. Hence we find a remarkable difference between the development of these risks and aggregated risk, especially between 1950 and 1972.

We can further zoom in by taking into account the travel mode of the other party (if any) as well. As an example six groups, or fatality types are analysed: car single-vehicle, car-car, pedestrian-car, motorcycle-car, motorcycle-single-vehicle and pedestrian-motorcycle. Here, too, the influence of mobility appears to be important. The developments of the number of fatalities in the first three groups (in which a car was involved) all show a strong increase between 1950 and 1972, and a decrease afterwards. The strong increase until 1972 is characteristic for the also exponential car mobility increase in that period. The last three groups show the sequence of highs and lows mentioned before which are characteristic for the development of motorcycle use. When we look at risk the developments are different for each of these six groups, especially in the period until 1972. In motor-car crashes and car-car crashes we even observe an increasing risk of approximately $4 \%$ annually. Chapter 3 shows that for these different groups both risk and mobility developed differently, in such a way that the contributions of these groups to the development of the total number of fatalities differed.

Since 1985 we have both detailed OVG-MON-OViN mobility data and BRON crash data at our disposal in the Netherlands. This enables further stratification of road safety development. Both mobility data and crash data can be stratified by travel mode, age and gender. Mobility data are in
principle available for age groups and since 2004 also for individual age years. As a consequence of further stratification, the groups of crashes and groups for which the distance travelled needs to be known, become smaller and smaller. This may lead to such small groups that a meaningful statistical analysis of the development of risk is no longer possible. Chapter 4 describes a partial solution of this problem which is found in the incorporation of demographic data in the analysis. The analysis of the development in time, stratified by age and gender can be simplified in that way. This is because the demographic developments resound in the development of mobility of specific age groups. For example, right after World War II the birth rate was considerably higher than in the years before or after. Therefore, this baby boom cohort is larger than the cohorts in the years before or after. As a consequence, their mobility, and their involvement in road crashes during their lives is also larger than that of other cohorts. However, the mobility and crash involvement per person of this cohort is not necessarily exceptional. Something similar goes for the effect of the decreasing birth rate from 280000 in 1969 to 180000 in 1975. Related to this decrease the number of 16 year old mopedist started to decrease in 1985 and the number of 18 year old car drivers started to decrease in 1987.

By taking the demographic data into account in the analysis of mobility data, the development of per capita distance travelled can be obtained, by age and gender. It appears to be relatively constant for many age groups, although there are some exceptions. For example, per capita car distance travelled of people born after 1970 has been decreasing since at least 15 years, which implies that for this group not only the size of the group but also the mobility per person has been decreasing. On the other hand, for elderly people especially car distance travelled by females born before 1970, is still increasing. These developments have their effect on crash involvement.

Profound stratifications are possible by expressing and analysing the development of both distance travelled and crash involvement per capita, and by assuming that these quantities vary only slowly in time and age. The analyses presented in Chapter 4 are based on mobility data stratified by age groups. At the time these analyses were carried out, there were no mobility data available for individual age years for the years before 2004. Nevertheless it appeared to be possible to model data on distance travelled with a model which was both smooth in age year and in calendar year. Comparison of the results of the analysis with the available mobility data per
individual age year for 2004 and later years showed that the mobility model obtained for mobility by age groups seems to describe reality fairly well.
Afterwards, at the request of SWOV, mobility data stratified by individual age year were made available by Statistics Netherlands. The possibilities to use these data in the analyses and forecasts of the development of road safety are further researched in Chapter 5 for the period between 1995 and 2007. For single-vehicle car crashes, car-car crashes and bicycle-car crashes analyses were carried out in which stratification was applied by individual age year for both the age of the driver of the vehicle with the casualty and the age of the driver of the other vehicle. Because of the small number of fatalities for each individual age year, this analysis was extended to police registered hospitalized casualties. For the modelling by age of both distance travelled and risk a smooth function was assumed, which means that the difference in the estimated value (for distance travelled and risk), varying with age could not change too much. For the modelling of mobility as a function of time a smooth function was assumed as well. For the modelling of risk as a function of time, a (negative) exponential function appeared to describe the usually decreasing risk reasonably well. With such a model it was possible to extrapolate these developments to a year before 1995 or after 2007. The results show that the risk of a crash with a fatality or hospitalized injury for young car drivers (up to approximately 35 years of age) hardly decreased between 1987 and 2007, in contrast to the risk for older car drivers. This is true both for single-vehicle crashes and car-car crashes, and in the latter crash type for the age of both drivers. Further, the risk to be involved in a car-car crash as the driver of the car with the casualty was compared to the risk to be involved as the other driver. For older drivers approximately 35 and up), the risk to be involved as the driver of the casualty decreased faster than the risk to be involved as the other driver. Finally, between 1987 and 2007 the risk for cyclists to be involved in a crash with a car decreased most for children and $40+$ riders. These analyses show it is plausible that road safety improvements are not equally profitable for all road users.

The analysis of the number of single-vehicle crashes requires that we know the distance travelled of the relevant mode. Then we can check whether risk of single-vehicle crashes has developed in a positive or negative way. That way we can see whether an increase in the number of fatalities is a consequence of an increase in traffic volume or of safety level. For two vehicle crashes this is different. In that case it is also necessary to know the development of distance travelled by the other party involved. Indeed, the safety of cyclists is also influenced by the presence of cars. In a situation
where there are no cars, just bicycles, cyclists would no longer be killed by a car. This notion is used to answer a much posed question about the safety of cycling in relation to car driving: will the number of casualties increase or decrease when some of the short trips currently made by car would be made by bicycle instead. This is not the same as a comparison between the risk of car driving and cycling. When car driving decreases, this also leads to a decrease in the number of killed pedestrians, cyclists etcetera. This question is researched in Chapter 6. A differentiation is made between fatalities and serious injuries, between age groups and gender. Further, crashes between bicycles and cars were taken into account, but also crashes between bicycles or cars and other traffic modes, and single-vehicle crashes for both bicycles and cars. For the calculation it was assumed that $10 \%$ of all car trips with a trip length of up to 7.5 km in the period between 1999 and 2005 would have been replaced by bicycle trips. This corresponds to approximately $1 \%$ to $2 \%$ of all car distance travelled and to approximately $10 \%$ of all bicycle distance travelled. We assume that all factors which determine risk would not be influenced by the mobility shift and hence that all new bicycle trips occur on the same roads, times of day etcetera as the other bicycle trips (and vice versa for the car trips which expire). It is found that as a consequence of this modal shift from car to bicycle, the number of fatalities would have increased by approximately 8 annually. This increase is mostly due to a negative effect of the modal shift for elderly. When young car drivers change to cycling, traffic would become safer. For injured, the result is quite different. The net result would amount to 400 to 500 more injured. Only for 18 and 19 year old men the modal shift would be beneficiary to safety. This negative effect is a consequence of the many bicycle crashes in which there is no motorized vehicle involved, but in which nevertheless the cyclist gets seriously injured.

It is desirable to use all available road safety knowledge to develop a model which describes the development of road safety Such a model should enable researchers to unravel the development of the number of fatalities into different components. These can be effects of specific developments in distance travelled, road safety measures and effects of other influencing factors such as the weather. With such a model the researcher can estimate the future effect of specific road safety measures in the past. If the model enables the researcher to estimate the number of fatalities expected in the future, he can advise policy makers about further steps to take. What is their expected effect, what will occur if nothing is done, what is necessary to achieve a certain future target? Chapter 7 explains what the findings of this
thesis mean for research of the development of road safety, modelling and forecasts of road safety.

If different groups show substantially different risks, while the development of distance travelled also differs, it is worthwhile to distinguish these groups and to determine the development of the number of fatalities in these groups separately. This is advantageous with respect to aggregation of both groups. The advantage follows from the fact that the average risk changes because the composition of the aggregation changes. Even if the risk for both subgroups is constant, average risk will change as a consequence of the different developments of distance travelled in both subgroups. Without stratification this change cannot be understood and will be interpreted as an "autonomous" decrease. This impedes a good estimation of the future number of fatalities, because it is unknown how and for how long the risk decrease will last when the origin of the risk decrease is unknown. The more we have to count on a continuation of the steady risk decrease of the past for which we do not know the origin, the more uncertain the quality of the forecast will be.

However, it is impossible to consider all possible relevant stratifications at the same time in the analyses. An alternative may be that stratifications are replaced by a simpler approach. An example would be an approach in which the age dependency of risk is modelled by a formula, instead of an approach using stratification by age. Another possibility is the use of risk reduction values which are known from literature on Safety Performance Indicators (SPI's) like the use of safety belts, blood alcohol concentration of intersections replaced by roundabouts. This requires data collection of these safety performance indicators. In the meantime it remains important to achieve sound data about road crashes and distance travelled. This is the more true for the registration of crash data. The standard of registration in the Netherlands has diminished to a level which is unworthy of the excellent position of the Netherlands as a country with a high traffic safety standard worldwide. It would credit the Netherlands if we would design and introduce a crash registration system which would enable policy makers and scientists to give a new impulse to the development of road safety. Such a system should preferably be flexible, in order to enable new attributes of crashes in the registration. Examples of such attributes are the driving experience of crash involved drivers, the presence of active and passive vehicle safety means, phoning and texting behaviour right before the crash etcetera. There will always be new issues which appear to be important in
road safety, and the road crash register should enable researchers and policy makers to gather the relevant facts. For the same reason it is necessary that the collection of data on distance travelled gives room to the registration of attributes which are relevant for road safety.

This thesis shows that a good understanding of the development of the number of road fatalities at least requires that the total number of fatalities is stratified by different groups. The distinction between these groups becomes essential if both the development of risk and the development of distance travelled are different for these groups. Moreover, it is important in the analysis to take the influence of other road users into account in the determination of the development of risk. This approach will not guarantee that all fluctuations in the number of fatalities can be related to external factors like safety measures, but without stratification it is unlikely that such a relation can be stated.

## Samenvatting:

## De verkeersveiligheid, stukje bij beetje

Dit proefschrift beschrijft een methode om de ontwikkeling van de verkeersonveiligheid beter te begrijpen. De verkeersonveiligheid, uitgedrukt in het jaarlijks aantal verkeersdoden, heeft in Nederland in het verleden een markante ontwikkeling laten zien. Omstreeks 1950 vielen er circa 1000 verkeersdoden per jaar. Tussen 1950 en (omstreeks) 1972 steeg dit aantal ongeveer exponentieel tot bijna 3500. Daarna volgde een daling tot circa 650 doden in 2012. Dat wil zeggen dat het aantal doden in de meeste jaren afnam, maar niet elk jaar en niet in elk jaar in gelijke mate. In sommige jaren was er zelfs sprake van een stijging, zoals in de jaren 1976 en 1977 (met ruim honderd doden meer dan het jaar daarvóór). En ook in 2011 (en 2012) steeg het aantal verkeersdoden met enkele tientallen ten opzichte van het jaar ervoor. In veel andere geïndustrialiseerde landen zag men een soortgelijke ontwikkeling: een stijgend aantal doden tussen 1950 en 1972, een daling daarna, en onregelmatige sprongen in het aantal doden omhoog en omlaag.

Het doel van dit proefschrift is om na te gaan hoe die ontwikkelingen kunnen worden gerelateerd aan factoren zoals de hoeveelheid verkeer, verkeersveiligheidsmaatregelen van de overheid, enzovoort. Het proefschrift is een zoektocht naar de manier waarop we kunnen vaststellen hoe die factoren de verkeersveiligheid bepalen. Daarbij ligt de nadruk op doden, omdat hiervan de gegevens het meest betrouwbaar geregistreerd zijn, zowel in Nederland als in veel andere landen. Hier richten we ons op analyse van Nederlandse gegevens. De bedoeling is echter om een algemeen geldende aanpak te ontwikkelen waarmee ook de verkeersonveiligheidsgegevens van andere landen kunnen worden geanalyseerd.

In het verleden is al veel onderzoek verricht aan dit vraagstuk. We weten inmiddels dat de ontwikkeling ruwweg wordt verklaard door een combinatie van de gestegen mobiliteit en de toegenomen verkeersveiligheid. De aanvankelijke stijging wordt toegeschreven aan een explosieve groei van het autoverkeer met jaarlijks 13,5\%. Deze groei was zo sterk dat de maatregelen om de verkeersveiligheid te vergroten ontoereikend waren om deze groei te compenseren. Na 1972 nam die groei af tot jaarlijks 3\% en in 2011 was deze groei zo goed als tot stilstand gekomen. Door deze veel geringere groei na 1972 konden de talrijke maatregelen om de
verkeersveiligheid te verbeteren merkbaar een netto positief effect sorteren op het jaarlijks aantal verkeersdoden.
Deze verklaring maakt meteen duidelijk dat de hoeveelheid verkeer een van de bepalende grootheid in de verkeersonveiligheid is. Deze grootheid wordt in dit proefschrift de mobiliteit of de afgelegde afstand genoemd, en uitgedrukt in km. De andere bepalende grootheid is het aantal verkeersdoden per afgelegde afstand. Deze grootheid wordt in dit proefschrift het risico genoemd, en uitgedrukt in verkeersdoden $/ \mathrm{km}$. Risico en afgelegde afstand zijn allebei belangrijk om de ontwikkeling van de verkeersonveiligheid te begrijpen, maar er is ook een onderscheid. Het risico is namelijk het eigenlijke onderwerp van verkeersveiligheidsbeleid en -onderzoek. Nederlandse verkeersveiligheidsexperts en -beleidsmakers zien de mobiliteit daarentegen als een autonome grootheid met zijn eigen dynamiek. Verkeersveiligheidsbeleid bedient zich in Nederland dan ook niet van de mogelijkheid om de mobiliteit te beperken. Het beleid is er uitsluitend op gericht om het aantal slachtoffers te verminderen, gegeven de mobiliteit. In de terminologie van dit proefschrift betekent dit dat we het risico van de mobiliteit willen verminderen.

Verkeersveiligheidsbeleidsmakers bedienen zich voor hun keuze van maatregelen vaak van een politieke beleidsdoelstelling: een maximum aantal te verwachten verkeersdoden in een zeker toekomstig jaar. Om te weten welk verkeersveiligheidsbeleid nodig is om een gekozen doelstelling te behalen, is het nodig om te begrijpen hoe het aantal verkeersdoden afhangt van het beleid in het verleden. De eerder genoemde globale verklaring van de ontwikkelingen van het aantal verkeersdoden sinds 1950 (eerst vooral een explosieve mobiliteitsstijging en daarna vooral succesvol veiligheidsbeleid) is daartoe weinig informatief. Wie bijvoorbeeld wil weten welke maatregelen in welke mate bijdroegen aan de verbetering van de verkeersveiligheid, moet verder analyseren. Kennis van dergelijke relaties is belangrijk om te weten in hoeverre het effect van het beleid uit het verleden nog zal voortduren in de toekomst.

Het onderzoek naar de ontwikkeling van de verkeersveiligheid, in Nederland ook wel de planbureaufunctie van verkeersveiligheidsonderzoek genoemd, is gericht op drie uitdagingen.

1. De ontwikkelingen van de verkeersveiligheid in het verleden begrijpen.
2. De bouw van modellen waarmee de ontwikkelingen kunnen worden begrepen en voorspeld.
3. De ontwikkelingen in de toekomst voorspellen.

Dit proefschrift beschrijft een aanpak waarbij de ontwikkeling van het risico gestratificeerd, dat wil zeggen in groepen verdeeld, wordt bestudeerd. Mobiliteit is immers niet altijd, overal en voor iedereen even gevaarlijk. Zo is een ritje van 10 km op een opgevoerde bromfiets voor de berijder veel riskanter dan een ritje van 10 km van een vrachtautochauffeur met veiligheidsgordel om. Elk stukje afgelegde afstand heeft een bijbehorend risico dat wordt bepaald door talloze factoren. De bestuurder, het voertuig, de weginrichting, regelgeving, weersomstandigheden, handhaving...... allemaal factoren die ertoe doen.

Ook voor het bepalen van de effecten van verkeersveiligheidsmaatregelen is stratificatie van de gegevens van verkeersongevallen een logische stap. Verkeersveiligheidsmaatregelen hebben immers altijd betrekking op slechts een deel van de mobiliteit en dus ook op een deel van de verkeersongevallen. Denk aan de autogordel, de bromfietshelm, de dodehoekspiegel voor vrachtwagens, alcoholwetgeving, enzovoort.

Ongeveer $30 \%$ van alle verkeersdoden valt bij een enkelvoudig ongeval. Dat is een ongeval waarbij het slachtoffer niet in botsing komt met een andere weggebruiker, maar valt of tegen bijvoorbeeld een boom of paal rijdt. Voor dergelijke ongevallen is er dus volgens de beschikbare gegevens geen ander verkeer bij het ongeval betrokken. Voor tweezijdige ongevallen is dat anders. Daar is de vervoerwijze van het andere voertuig, de leeftijd en ervaring van de bestuurder en of deze onder invloed was van drugs of alcohol ook van belang. In dit proefschrift wordt nagegaan hoe de invloed van de mobiliteit van anderen in rekening kan worden gebracht in het risico.

Er kunnen op allerlei manieren subgroepen worden gevormd: slachtoffers onder bromfietsers of automobilisten, onder kinderen of ouderen, slachtoffers die wel of geen alcohol in het bloed hadden, enzovoort. Maar ook verkeersdoden door een ongeval waarbij een auto als andere partij betrokken was, of waarbij een jonge bestuurder in de andere auto reed, enzovoort. De ontwikkelingen van het aantal verkeersdoden sinds 1950 waren niet uniform voor al die verschillende subgroepen. Het risico en de afgelegde afstand hebben zich voor verschillende soorten mobiliteit soms zeer verschillend ontwikkeld. In Hoofdstuk 2 wordt de toegepaste methode van stratificatie van mobiliteit en risico verder uitgelegd en onderbouwd. Daarbij komt aan bod hoe de verschillende factoren bepalend zijn voor de ontwikkeling van de verkeersveiligheid en welke gegevens nodig zijn om de analyses uit te voeren. Gegevens over de verkeersongevallen (tussen 1976 en

2003 het VOR-bestand en vanaf 1985 het BRON-bestand) en over de mobiliteit (onder andere het OVG/MON/OViN-bestand) zijn daarbij zeer belangrijk.

VOR/BRON en OVG/MON/OViN onderscheiden diverse variabelen zoals vervoerswijze, leeftijd en geslacht van de reiziger. Daardoor is het mogelijk om voor bepaalde leeftijdsgroepen of vervoerswijzen apart analyses uit te voeren. VOR/BRON is echter pas vanaf 1976 beschikbaar, en OVG/MON/OViN vanaf 1985. Daardoor is het veel lastiger om voor de periode vóór 1985 gestratificeerde analyses uit te voeren van mobiliteit en risico van groepen verkeersdoden. De enige gestratificeerde gegevens van de mobiliteit in die periode zijn te vinden in de motorrijtuigenstatistiek. Het CBS heeft op basis van verkeerstellingen de jaarlijkse mobiliteit van onder meer personenauto's, vrachtauto's en motoren geschat. Deze gegevens zijn vanaf 1950 beschikbaar. Voor gegevens over verkeersdoden in de periode tussen 1950 en 1976 moeten we ons baseren op tabellenboeken, eveneens van het CBS. In Hoofdstuk 3 wordt op basis van die gegevens aangetoond dat het sterk uitmaakt of we de ontwikkeling van het aantal verkeersdoden geaggregeerd of gestratificeerd analyseren. Eerst wordt de ontwikkeling van een geaggregeerd risico berekend. Dat wil zeggen dat het totaal aantal verkeersdoden wordt gedeeld door de afgelegde afstand door gemotoriseerd verkeer. Dan blijkt dat dit geaggregeerde risico sinds 1950 vrijwel steeds is gedaald met $6 \%$ á $7 \%$ per jaar. Vervolgens wordt het aantal verkeersdoden voor drie subgroepen apart bestudeerd, namelijk voor auto-inzittenden, vrachtauto-inzittenden en motorrijders. Het blijkt dat het aantal verkeersdoden voor die drie groepen zich verschillend heeft ontwikkeld, waarbij de mobiliteit steeds een belangrijke factor blijkt te zijn geweest. De ontwikkeling van het aantal verkeersdoden onder motorrijders kende bijvoorbeeld achtereenvolgens pieken en dalen die samenvielen met soortgelijke pieken en dalen in het gebruik van de motor. De risico's voor deze drie groepen blijken zich op ongeveer vergelijkbare wijze te hebben ontwikkeld: het risico was licht stijgend tussen 1950 en 1972 en daalde daarna. Bij de motorrijders zien we bovendien een extra piekje omstreeks 1970 toen het risico in een of twee jaar tijd verdubbelde. We vinden dus een opvallend verschil met het geaggregeerde risico, vooral tussen 1950 en 1985.

We kunnen nog verder inzoomen door ook de vervoerswijze van de (eventuele) tegenpartij in de stratificatie te betrekken. Als voorbeeld zijn zes groepen, of typen verkeersdoden, onderzocht: auto-enkelvoudig, auto-auto, voetganger-auto, motor-auto, motor-enkelvoudig, en voetganger-motor. Ook
hier blijkt de invloed van de mobiliteit van belang. De ontwikkeling van het aantal verkeersdoden in de eerste drie groepen (waarbij een auto betrokken was) vertoont een scherpe stijging tussen 1950 en 1972 en daarna een daling. De scherpe stijging tot 1972 is karakteristiek voor de eveneens exponentiële stijging van het autogebruik in die periode. De laatste drie groepen ongevallen (waarbij een motor betrokken was) vertonen de eerder genoemde opeenvolgende pieken en dalen die karakteristiek zijn voor de ontwikkeling van de motormobiliteit. Kijken we naar het risico, dan blijkt dat de ontwikkelingen bij elk van de zes genoemde groepen anders verlopen, vooral in de periode tot omstreeks 1972. Bij motor-auto ongevallen en autoauto ongevallen zien we tussen 1950 en 1972 zelfs een stijgend risico met zo'n $4 \%$ per jaar. Hoofdstuk 3 laat zien dat voor deze verschillende groepen zowel het risico als de mobiliteit zich verschillend ontwikkelden, waardoor de bijdragen van deze groepen aan de ontwikkeling van het totaal aantal verkeersdoden verschilden.

Sinds 1985 beschikken we in Nederland zowel over mobiliteitsgegevens van het OVG/MON/OViN, als over gedetailleerde ongevalsgegevens van BRON. Daarmee is het mogelijk om de analyse van de ontwikkeling van de verkeersonveiligheid nog verder te stratificeren. Zo kan zowel de mobiliteit als het aantal verkeersdoden of verkeersgewonden worden gestratificeerd naar vervoerswijze, leeftijd en geslacht. Leeftijd is daarbij voor de mobiliteitsgegevens in beginsel beschikbaar in leeftijdsklassen en pas vanaf 2004 ook in individuele leeftijdsjaren. Bij verdere stratificatie worden de groepen verkeersongevallen en ook de groepen waarvoor de afgelegde afstand apart bekend moet zijn, wel steeds kleiner. Dit leidt soms tot zulke kleine groepen dat een zinvolle statistische analyse van de ontwikkeling van het risico niet meer mogelijk is. In Hoofdstuk 4 wordt voor dit probleem een gedeeltelijke oplossing beschreven waarbij demografische gegevens in de analyse worden betrokken. De analyses van de ontwikkeling in de tijd naar leeftijd en geslacht worden daarmee eenvoudiger. Dat komt omdat in de ontwikkeling van de mobiliteit van bepaalde leeftijdsgroepen de demografische ontwikkeling "doorklinkt". Zo werden er vlak na de Tweede Wereldoorlog ineens veel meer kinderen geboren dan daarvoor en ook dan daarna. Dit babyboomcohort is dus groter dan dat van andere geboortejaren. De mobiliteit, en ook de betrokkenheid bij ongevallen van dat cohort is daardoor hun leven lang ook groter dan dat van cohorten van de jaren vlak ervoor en erna. Echter, per hoofd van de bevolking is de mobiliteit en de ongevalsbetrokkenheid van dit cohort niet noodzakelijkerwijs uitzonderlijk. Iets soortgelijks geldt voor het effect van de daling van het geboortecijfer van
circa 280000 in 1969 tot circa 180000 in 1975. In samenhang met deze daling zette omstreeks 1985 een daling in het aantal jonge bromfietsers in en in 1987 een daling van het aantal jonge automobilisten.

Door de demografische gegevens in rekening te brengen in de analyse van de mobiliteitsgegevens kan de ontwikkeling van de mobiliteit per hoofd van de bevolking worden verkregen, naar leeftijd en naar geslacht. Deze blijkt voor veel leeftijdsgroepen vrij constant te zijn, hoewel er uiteraard uitzonderingen zijn. Zo neemt de automobiliteit per hoofd van de bevolking van mensen die na 1970 geboren zijn al zeker 15 jaar af, zodat voor deze groep dus niet alleen het aantal mensen, maar ook de mobiliteit per persoon steeds verder afneemt. Voor ouderen daarentegen blijkt vooral de mobiliteit van vrouwen die vóór 1970 geboren zijn nog wat toe te nemen. Deze ontwikkelingen hebben hun weerslag op de betrokkenheid bij ongevallen.

Door zowel de mobiliteitsgegevens als de ongevalsgegevens per hoofd van de bevolking uit te drukken en te analyseren en deze gegevens te beschouwen als langzaam variërend in de tijd en over de leeftijd, zijn verregaande stratificaties mogelijk. De in hoofdstuk 4 gepresenteerde analyses zijn gebaseerd op mobiliteitsgegevens in leeftijdsgroepen. Er waren voor die analyses nog geen mobiliteitsgegevens op basis van individuele leeftijdsjaren beschikbaar van vóór 2004. Niettemin bleek het mogelijk om de mobiliteitsgegevens te modelleren met een vloeiend model dat zowel voor leeftijdsjaar als voor kalenderjaar een geleidelijke verandering veronderstelde. Door de resultaten van de analyse te vergelijken met de gegevens per individueel leeftijdsjaar vanaf 2004, bleek dat het op basis van de data in leeftijdsgroepen verkregen mobiliteitsmodel de werkelijkheid behoorlijk beschrijft.

Nadien heeft het CBS op verzoek van de SWOV mobiliteitsgegevens vanaf 1995 per individueel leeftijdsjaar beschikbaar gesteld. De mogelijkheden om deze gegevens te gebruiken voor analyses en voorspellingen van de ontwikkeling van de verkeersonveiligheid zijn in Hoofdstuk 5 nader onderzocht voor de periode 1995-2007. Voor enkelvoudige auto-ongevallen, auto-auto ongevallen en fiets-auto-ongevallen zijn analyses uitgevoerd waarbij de leeftijd van de bestuurder van het voertuig van het slachtoffer én die van de tegenpartij (indien van toepassing) naar individueel leeftijdsjaar zijn gestratificeerd. Vanwege het geringe aantal verkeersdoden per individueel leeftijdsjaar is in deze analyse ook het aantal door de politie geregistreerde in een ziekenhuis opgenomen gewonden betrokken. Voor de
modellering naar leeftijd van zowel mobiliteit als risico is een gladde functie verondersteld, dat wil zeggen dat het verschil in geschatte waarde (voor mobiliteit en risico), variërend met de leeftijd niet te sterk mocht veranderen. Voor de modellering van de mobiliteit naar kalenderjaar werd eveneens een glad model verondersteld. Voor de modellering van het risico naar kalenderjaar bleek een (negatief-) exponentiële ontwikkeling het veelal dalende risico goed te beschrijven. Met een dergelijk model was het mogelijk om deze ontwikkeling eenvoudig naar een jaartal vóór 1995 of na 2007 te extrapoleren. De resultaten wijzen uit dat het risico op een ongeval met een verkeersdode of (geregistreerde) ziekenhuisopname voor jonge autobestuurders (tot circa 35 jaar) tussen 1987 en 2007 nauwelijks is afgenomen, in tegenstelling tot het risico van oudere autobestuurders. Dit geldt zowel voor enkelvoudige ongevallen als voor auto-auto ongevallen, en in het laatste geval voor de leeftijd van beide bestuurders. Verder is een vergelijking gemaakt tussen het risico dat een bestuurder van de auto met het slachtoffer bij een ernstig ongeval tussen twee auto's betrokken raakt, met het risico om als andere bestuurder betrokken te raken. Het risico om als bestuurder van het slachtoffer betrokken te raken nam voor bestuurders vanaf circa 35 sneller af dan het risico om als andere bestuurder betrokken te raken.' Ten slotte blijkt dat het risico van fietsers om dood of gewond te raken in een ongeval met een auto vooral bij kinderen en 40-plussers, te zijn afgenomen tussen 1995 en 2007. Dergelijke analyses maken aannemelijk dat niet alle weggebruikers in dezelfde mate profiteren van de verbeteringen van de verkeersveiligheid.

De analyse van de aantallen enkelvoudige ongevallen vergt dat we de mobiliteit van de betreffende vervoerswijze kennen. Daarmee kunnen we nagaan of het risico op enkelvoudige ongevallen zich positief of negatief heeft ontwikkeld. Zo kunnen we zien of een stijgend aantal verkeersdoden het gevolg is van een verkeerstoename of van een verandering in het veiligheidsniveau. Voor tweezijdige ongevallen ligt dat dus ingewikkelder. Daarbij is het ook nodig om de vervoerswijze van de tegenpartij bij de analyse te betrekken. Immers: de veiligheid van fietsers wordt ook beïnvloed door de aanwezigheid van auto's. Als er alleen fietsers waren en geen auto's, dan zouden er geen fietsers meer overlijden door een aanrijding met een auto. Dit is gebruikt bij de bepaling van het antwoord op een veel gestelde vraag over de veiligheid van fietsen in vergelijking met autorijden: neemt het aantal slachtoffers toe of af indien men voor korte ritjes vaker de fiets zou nemen in plaats van de auto? Dit is niet hetzelfde is als het vergelijken van het risico van autorijden en fietsen. Immers, wanneer mensen minder
autorijden heeft dat ook tot gevolg dat er minder voetgangers, fietsers enzovoort door een auto zullen worden aangereden. Dit vraagstuk wordt onderzocht in Hoofdstuk 6. Daarbij is onderscheid gemaakt naar letselernst (doden en gewonden), naar leeftijd en naar geslacht. Verder is gekeken naar ongevallen tussen fietsers en auto's, maar ook tussen fietsers of auto's en ander verkeer en naar enkelvoudige ongevallen met fietsers en met auto's. Bij de berekening is aangenomen dat $10 \%$ van de autoritten met een lengte tot $7,5 \mathrm{~km}$ in de periode 1999-2005 zou zijn vervangen door fietsritten. Dat correspondeert met $1 \%$ à $2 \%$ van de totale automobiliteit en circa $10 \%$ van de fietsmobiliteit. We hebben aangenomen dat de factoren die het risico bepalen bij de mobiliteitsverschuiving gelijk blijven en dus dat de nieuwe fietsritten op dezelfde wegen plaatsvinden, op dezelfde tijdstippen etcetera, als alle andere ritten (en vice versa voor de autoritten die vervallen). Dan blijkt dat het aantal doden als gevolg van een dergelijke verschuiving van autogebruik naar fietsgebruik licht zou zijn gestegen met circa 8 per jaar. Deze stijging is vooral te wijten aan een negatief effect van de verschuiving bij ouderen. Wanneer jonge automobilisten gaan fietsen, wordt het verkeer veiliger. Voor gewonden is het beeld geheel anders. Daarbij blijkt dat het aantal gewonden zou zijn gestegen met 400 à 500 per jaar. Alleen voor 18-19 jarige mannen is er dan nog een positief effect. Dit nadelige effect is het gevolg van het grote aantal ongevallen met fietsers waarbij géén motorvoertuig betrokken is, maar waarbij de fietser wel ernstig gewond raakt.

Het is gewenst om de beschikbare kennis over verkeersveiligheid te gebruiken voor de ontwikkeling van een model dat de verkeersveiligheid van de verkeersveiligheid beschrijft. Een dergelijk model moet de onderzoeker in staat stellen om de ontwikkelingen in het aantal slachtoffers te ontleden in verschillende componenten. Dit kunnen effecten van specifieke mobiliteitsontwikkelingen, verkeersveiligheidsmaatregelen en effecten van andere invloeden zoals het weer zijn. Met een dergelijk model kan de onderzoeker een schatting maken van het toekomstig effect van bepaalde veiligheidsmaatregelen uit het verleden. Als het model hem in staat stelt om een uitspraak te doen over het verwachte aantal slachtoffers in een toekomstig jaar, dan kan hij beleidsmakers adviseren over de te nemen maatregelen. Wat is hun effectiviteit, wat staat er te wachten indien we niets doen, wat is nodig om een toekomstige doelstelling te halen? In Hoofdstuk 7 wordt uiteengezet wat de bevindingen uit dit proefschrift voor het onderzoek naar de ontwikkelingen van de verkeersveiligheid, de modellering, en verkenningen van de verkeersveiligheid betekenen.

Als het risico voor bepaalde subgroepen onderling sterk verschilt, terwijl ook de ontwikkeling van de mobiliteit verschilt, is het raadzaam on deze groepen onderling te onderscheiden en de ontwikkeling van het aantal slachtoffers in beide groepen apart te bepalen. Dit heeft een voordeel ten opzichte van het aggregeren van de beide groepen. Dit voordeel bestaat eruit dat het gemiddelde risico van de aggregatie verandert doordat de samenstelling van de aggregatie verandert. Zelfs als het risico van de beide subgroepen constant is, dan zal het gemiddelde risico veranderen door de verschillende ontwikkelingen van de mobiliteit van beide groepen. Die verandering kan zonder stratificatie niet worden begrepen en moet in dat geval worden geïnterpreteerd als een "autonome" ontwikkeling. Dit bemoeilijkt een goede schatting van het toekomstig aantal slachtoffers, omdat het onbekend is hoe en hoelang de risicodaling zich zal voortzetten indien de oorsprong ervan onbekend is. Hoe meer we voor prognoses moeten vertrouwen op een continuering van de gestage risicodaling uit het verleden zonder te weten waaraan die daling zijn oorsprong ontleent, des te onzekerder is daarmee de kwaliteit van de prognose.

Het is echter onmogelijk om alle denkbare stratificaties die relevant zijn, tegelijk in de analyses te betrekken. Als alternatief kunnen stratificaties wellicht worden vervangen door een eenvoudiger aanpak. Een voorbeeld hiervan is een aanpak waarbij de leeftijdsafhankelijkheid van ongevalsbetrokkenheid in een formule wordt gevat, in plaats van door stratificatie toe te passen. Een andere mogelijkheid is het gebruik van uit de literatuur bekende risicoreductiewaarden voor veiligheidsindicatoren zoals gordeldracht, alcoholgebruik of door rotondes vervangen kruispunten. Dit vereist dat er gegevens over deze veiligheidsindicatoren worden verzameld. Intussen blijft het belangrijk om deugdelijke gegevens over ongevallen en mobiliteit te blijven verzamelen. Dit geldt temeer voor de registratie van de ongevalsgegevens, waarvan de kwaliteit inmiddels is gedaald tot een niveau dat de voorbeeldfunctie die Nederland als verkeersveilig land in de wereld heeft, onwaardig is. Het zou Nederland passen om een ongevalsregistratiesystematiek te ontwerpen en in te voeren, die de beleidsmakers en wetenschappers in staat stelt om een nieuwe impuls te geven aan de ontwikkelingen van de verkeersveiligheid. Een dergelijk systeem zou bij voorkeur flexibel moeten zijn, zodat het mogelijk wordt om nieuwe kenmerken van ongevallen in de registratie op te nemen. Voorbeelden van zulke nieuwe kenmerken zijn de opgedane rijervaring van bestuurders die bij ongevallen betrokken zijn of de aanwezigheid en het gebruik van veiligheidsmiddelen, telefoneergedrag voorafgaand aan het
ongeval enzovoort. Er zullen altijd nieuwe ongevalskenmerken belangrijk blijken en de registratie van ongevallen behoort onderzoekers en beleidsmakers in staat te stellen hierover kennis te vergaren. Om dezelfde reden is het nodig dat de inwinning van mobiliteitsgegevens ruimte geeft aan de registratie van relevante kenmerken.

Dit proefschrift wijst uit dat een goed begrip van de ontwikkeling van het aantal verkeersdoden in elk geval vraagt dat het totaal aantal verkeersdoden wordt onderscheiden naar verschillende groepen. Het onderscheid tussen verschillende groepen is vooral van belang als zowel de ontwikkeling van het risico als dat van de afgelegde afstand voor die groepen verschilt. Daarnaast is het in de analyse van belang om rekening te houden met de invloed van de mobiliteit van andere weggebruikers bij de bepaling van de ontwikkeling van het risico. Deze aanpak garandeert niet dat alle fluctuaties in het aantal doden kunnen worden gerelateerd aan externe factoren zoals maatregelen, maar zonder stratificatie is het onwaarschijnlijk dat een dergelijke relatie kan worden gelegd.

## About the author

Henk Stipdonk was born in Leiden on February $20^{\text {th }}$ in 1957, as the son of a garage owner and a nurse. He grew up in Leiden and went to school in Wassenaar, where he obtained his athenaeum diploma from Rijnlands Lyceum in 1975. He studied experimental physics at the RijksUniversiteit Leiden, in the Kamerlingh Onnes Laboratory. He graduated in 1981 and did scientific research on the magnetic phase diagram of solid ${ }^{3} \mathrm{He}$ until 1985.

Between 1985 and 2004 he worked at Rijkswaterstaat, the Dutch motorway authority. In 1985 he started as a researcher at the traffic and transport engineering division (DVK/AVV/VWL) in the field of traffic theory. Subsequently he was manager of inland waterway safety research projects and from 1993 he was manager of several departments which were concerned with for example road safety research and development of traffic management engineering systems.

In 2004 he started at SWOV, as head of the road safety assessment and forecasting department, where he carried out the research on which this thesis is based.

## SWOV-Dissertatiereeks

In deze reeks zijn eerder verschenen:

Jolieke Mesken (2006). Determinants and consequences of drivers' emotions.

Ragnhild Davidse (2007). Assisting the older driver: Intersection design and in-car devices to improve the safety of the older driver.

Maura Houtenbos (2008). Expecting the unexpected: A study of interaction behaviour at intersections.

Frits Bijleveld (2008). Time series analysis in road safety research using state space methods.

Saskia de Craen (2010). The X-factor; A longitudinal study of calibration in young novice drivers.

Atze Dijkstra (2011). En route to safer roads; How road structure and road classification can affect road safety.

Charlotte Bax (2011). Processes and patterns; The utilisation of knowledge in Dutch road safety policy.

Willem Vlakveld (2011). Hazard anticipation of young novice drivers; Assessing and enhancing the capabilities of young novice drivers to anticipate latent hazards in road and traffic situations.

Sjoerd Houwing (2013). Estimating the risk of driving under the influence of psychoactive substances.

