Yannick Matthieu Maltha

HOUSEHOLD CAR OWNERSHIP IN THE NETHERLANDS

THE CHANGING INFLUENCE OF FACTORS EXPLAINING HOUSEHOLD CAR OWNERSHIP LEVELS IN THE NETHERLANDS

HOUSEHOLD CAR OWNERSHIP IN THE NETHERLANDS

THE CHANGING INFLUENCE OF FACTORS EXPLAINING HOUSEHOLD CAR OWNERSHIP LEVELS IN THE NETHERLANDS

By

Yannick Matthieu Maltha

in partial fulfilment of the requirements for the degree of

Master of Science
in Systems Engineering, Policy Analysis and Management

at the Delft University of Technology,
to be defended publicly on 29 September 2016

Student number:	4019032	
Supervisor:	Prof. dr. B. van Wee	
Thesis committee:	Dr.ir. M. Kroesen	TU Delft
	Dr.ir. C. van Daalen	TU Delft

This thesis is confidential and cannot be made public until 29 September 2016.

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Summary

To contribute to existing research on the influence of factors on household car ownership, this study addressed the question whether and to what extent the influence of economic, socio-demographic and spatial factors on the number of cars owned by a household, has changed over time. Although this might sound as a forthright question, to the best of the authors' knowledge, there is, however, an absence of studies that investigate the changing influence of factors on car ownership in recent decades. For the purpose of this study, The Netherlands has been chosen as a case study. The Netherlands is considered to have and apply state of the art knowledge about car ownership (modeling) for research and policy purposes. Moreover, The Netherlands provides one of the most resourceful databases regarding car ownership in the world.

The motivation for this study has been twofold. The current observed stabilization of car use in The Netherlands, also referred to as 'peak car', seems to come hand in hand by a decline in car ownership growth. The growth in household car ownership has slowly decreased from 3.1% to 0.6% in 2001-2015. Despite much is unknown on the causes of this trend there is consensus on the claim that the influence of certain factors on car ownership might have changed over time. Secondly, car ownership models that predict household car ownership levels are important to urban planners and decision makers, because car ownership levels are highly correlated with urban sprawl and automobile travel. The consequential question is what policy implications the changing influence of car ownership determinants might bring. For this purpose, special attention will be given to the most dominant car ownership model currently used in The Netherlands: DYNAMO. The model is widely used by Dutch ministries and political parties, in which the effects of general developments and government policy on size, composition and use of the car fleet are modeled.

For the first part of this study, the statistical method of ordered logistic regression has been used on household mobility data on 162,593 households, collected by the national Traffic Survey of The Netherlands in the years 1987, 1991, 1995, 1999, 2003, 2010 and 2014. The results show that the influence of household composition, gender, age, education, working status and urbanization level on car ownership have only moderately changed between 1987 and 2014. However, the influence of household income on car ownership has decreased considerably over time (from 38% in 1987 to 28% in 2014), whereas the influence of household size has increased (from 29% in 1987 to 35% in 2014). Moreover, the relative influence of household income and household size are substantial, contributing to more than 60% of the total influence on household car ownership in all years studied. Finally, we can say with certainty that household composition, education and urbanization levels have increased the average number of household cars for the period studied. Nevertheless, the stabilization in the growth of cars – especially in the last couple of years - cannot necessarily be explained based on the limited outcomes of the analysis.

In the second part, a first comparison has been made between car ownership models that predict household car ownership levels based on constant versus changing influences of car ownership determinants over time. For this part, the statistical method of multinomial logistic regression has been used on household mobility data on 174,393 households, collected by the national Traffic Survey of The Netherlands in the years 1987, 1991, 1995, 1999, 2003 and 2010 to predict household car ownership levels in 2014. The results show that a prediction model that incorporates changing influences of car ownership determinants, at least for the years studied, improves the prediction of household car ownership levels compared to a model that assumes constant coefficient parameters over time.

A comparison with the most important car ownership prediction model in The Netherlands – DYNAMO - shows that the results of this study could serve as the first step to further extend the prediction accuracy of the model. According to one of the main developers of DYNAMO, dr. R. Haaijer, the latest version of DYNAMO (3.0) estimates the main (i.e. constant) effects of household car ownership coefficients from 1990-2010, using national Traffic Survey household data to predict household car ownership levels till 2050. DYNAMO's predictions are, however, of importance to policy makers and urban planners. Short-term and long-term predictions of the number of private cars owned by households are among others used for predicting car mobility and travel behavior, infrastructure building (e.g. roads), noise regulations, spatial planning (e.g. new neighborhoods), and even election programs in The Netherlands.

Based on a qualitative investigation on the influence of household income, short-term (2020) to long-term (2050) policy implications have been formulated. On the basis of the overestimated influence of household income on car ownership levels by DYNAMO, it is expected that prognoses on the total car fleet have been overestimated. Assuming that the influence of household income will continue to decline in the coming decades. Overestimation of the size of the total car fleet in mid-term (2030) and long-term (2050) forecasts for Dutch Ministries (e.g. Ministry of Infrastructure and Environment) could lead to overestimations in car mobility, and will impact on congestion forecasts. Investments in parking facilities, possible congestion charging and planned investments to build new roads could in this way partially lose their effectiveness. Moreover, policies including taxes, noise regulations and car scrapping premiums can be less effective, both financially and in terms of car mobility effects.

An overestimation of car ownership levels on the basis of household income could also impact the effectiveness of tax-based policies of political parties in the short-term (2020). Concerning tax discount on car travel and the kilometer tax, it is supposed that tax revenues from current political programs of Groenlinks and D66 are expected to decline, whereas the tax expenses suggested in the programs of VVD, PVV and CDA are expected to decline till 2020. In addition, fixed taxes on car ownership, including BPM and MRB tariffs, might lose their effectiveness in reducing car ownership levels and car mobility due to a higher price inelasticity of cars than currently is anticipated by DYNAMO.

This study has attempted to improve our understanding on the changing influence of household car ownership factors over time and their potential implication for car ownership prediction models for the sake of policymaking and future research on car ownership modeling. The analysis could be further improved, by including other factors that have been excluded from the national Traffic Survey. For example, attitudinal/psychological factors are considered to become increasingly important factors affecting household car ownership levels. Furthermore, a greater specification of subcategories of factors, including the investigation of the effects of unique categories on car ownership levels, could improve the outcomes of this research. Finally, the effects of some factors can be investigated more thoroughly by using other types of analysis. For example, the influence of the variable age on household car ownership is better described by a quadratic function, rather than a linear function used in this study.

This study has shown that incorporating *interaction effects* of coefficients - from the same datasets and variables used by DYNAMO - could increase the prediction accuracy of future household car ownership levels. With regard to the relevant short-, and long-term policy applications of DYNAMO more research is necessary to further extend our notion of the capabilities of DYNAMO using interaction effects. Especially, the interaction effects of household income and household size have the potential to increase the current prediction accuracy of the DYNAMO model (3.0). Accordingly, if the effects are significant on car fleet forecasts and car mobility, Dutch Ministries and political parties could be informed in the change of effectiveness of preferred policy measures. Furthermore, it would provide policymakers better financial insight in the cost savings or extra investments that are necessary for short-term and long-term policy revisions. This study therefore recommends implementing a module in DYNAMO that allows for interaction effects over time.

Finally, the findings of this study are related to the case of The Netherlands. However, the current debate on the changing influence of factors on car ownership levels applies to multiple industrialized countries. The continuous decline (in the growth) of car ownership levels in Australia, Belgium, Canada, France, Germany, Japan, Sweden, UK and USA could demonstrate that the influence of car ownership determinants has changed over time. It is therefore recommended for these countries to perform an analysis based on longitudinal disaggregate-household data to confirm whether the changing influence of determinants follows a similar pattern as found in The Netherlands. With regard to the policy outcomes of this study, it is recommended to investigate the (sensitivity of) effects of the changing influence of car ownership determinants (if applicable) on car fleet estimations and car mobility forecasts. This is especially important for countries like the UK and USA that formulate short-term and long-term mobility policies based on car ownership models similar to DYNAMO.

" It is the mark of an educated mind to be able to entertain a thought without accepting it "
- Aristotle

Acknowledgements

This thesis is the end result of a five months and a half research endeavour at the Delft University of Technology in The Netherlands. The work is intended for everyone who is interested in car ownership (models) in its broadest sense.

The completion of the work has proven to be a challenge in time and intellect, and I certainly would not have made it without the support of some important individuals - both professionally and personally.

Among my graduation supervisors I would like to specially thank Dr. ir. Maarten Kroesen for his supervision. His approachable attitude, knowledge and dedication have demonstrated to not only be an excellent first supervisor to me, but also a person that has made the research process much more enjoyable.

I thank Dr. ir. Els van Daalen for her dedication in providing me detailed feedback, sharp comments and constructive questions. I specially thank her for her time when I unannounced walked by. I thank the Chairman, Prof. dr. Bert van Wee, for his valuable input and supportive guidance during all meetings, which have elevated the overall quality of this research.

I would like to thank my dear friends and family who supported me during the times I needed it the most.

Yannick Matthieu Maltha
Delft, September 2016

Contents

1

Introduction

Car ownership is an important feature of modern life. The ownership of a car influences our daily travel routines and out-of-doors activities. The increase or decrease of car ownership has an effect on our energy consumption and air-quality locally and globally. Investigating car ownership has therefore been of ample interest among policy makers and researchers in the past decades.

1.1. THE IMPORTANCE OF CAR OWNERSHIP MODELS

For policy makers and urban planners it is of importance to develop models that can identify and explain factors influencing car ownership levels (Fenger, 1999; Ortúzar & Willumsen, 2002; Miller, 2003; Potoglou & Susilo, 2008). Car ownership models are often used in main transport model systems; car ownership is an important determinant for car use, since changes in mobility are the main result of changes in car use for both drivers and passengers (de Jong, Fox, Daly, Pieters, & Smit, 2004; van der Waard et al., 2013; Traa et al., 2014). Furthermore, research has shown that car ownership influences frequency choice of trips (Meurs, 1990), mode choices for work and non-work activities (Uncles 1987; Bhat, 1996) and the destination of non-work activities (Wrigley, 1990). Car ownership is therefore an important determinant for travel behavior and is primarily interlinked with decision-making on motorized trips and residential location (Scott & Axhausen, 2006; Potoglou & Kanaroglou, 2008a).

1.2. PEAK CAR: A NEW TREND IN INDUSTRIALISED COUNTRIES

Given the importance of car ownership and its relation to car use, we should consider a new development: the current stabilisation or even decline of car use faced by different industrialised countries in the world. This trend is often referred to as 'peak car' or 'peak travel' (Goodwin & van Dender, 2013, Kuhnimhof, Zumkeller, & Chlond, 2013; van der Waard, Jorritsma, & Immers, 2013). In Belgium, for example, both private car ownership and driven kilometres per car have been decreasing since 2008 (Janssens, Cools, Miermans, Declerq, & Wets, 2011). In Germany the number of car kilometres under economic growth has not increased since 1995 (Kuhnimhof, Buehler, & Dargay, 2011). The same applies for the UK since the mid-1990s (le Vine & Jones, 2012) and the USA since 2005 (Millard-Ball & Schipper, 2011). Other countries that currently experience this peak car phenomenon are Australia, Canada, France, Japan, Sweden and The Netherlands (Millard-Ball & Schipper, 2011).

The current stabilization of car use in industrialized countries seems to come hand in hand by a decline in car ownership growth (Millard-Ball & Schipper, 2010; Oakil, Manting, & Nijland, 2016a). The decline in the growth of private car ownership in industrialised countries becomes clear when presenting Figure 1, which visualizes the development of vehicle ownership levels for different industrialised countries in the period 1970 till 2008 (years are presented by dots). The USA, Canada, Australia, France, UK, Sweden, Germany and Japan show signs of car ownership growth saturation and even negative growth in the period 1970 till 2008 (Millard-Ball & Schipper, 2010).

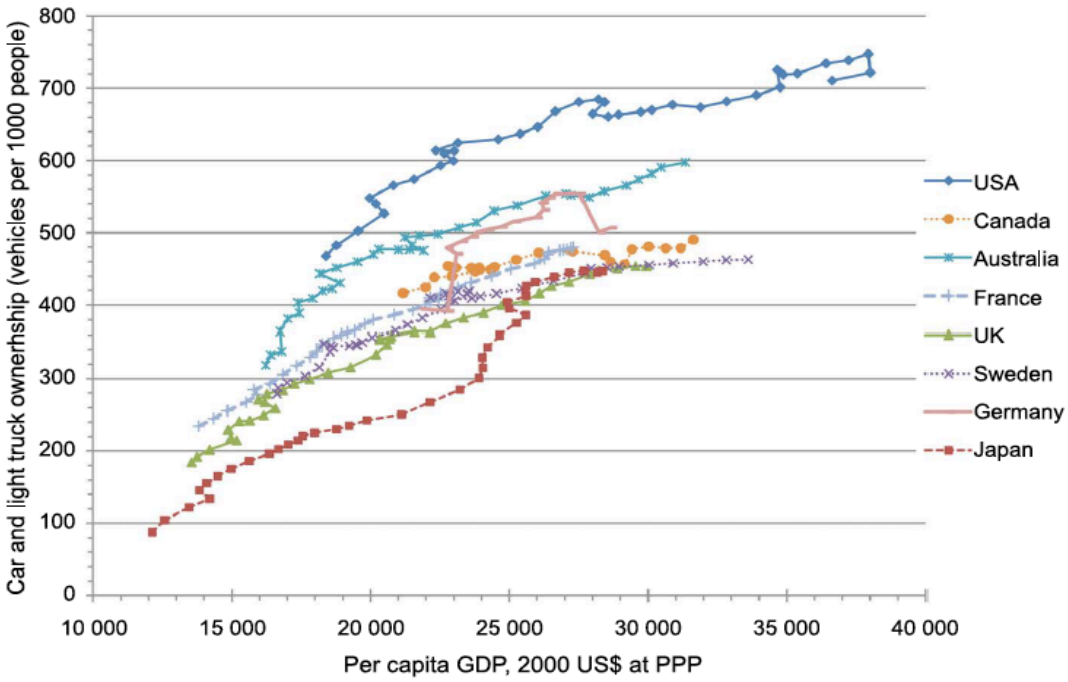


Fig. 1. Signs of saturation of vehicle ownership of different industrialised countries in the period 1970-2008 (Millard-Ball & Schipper, 2010, p. 364)

1.3. THE CASE OF THE NETHERLANDS

Also in The Netherlands the growth of car ownership has been levelling off in the past years (given by Fig 2). Since, 2001, the growth percentage of car ownership has decreased from 3.1% in 2001 to 0.6% in 2015, while the population has steadily increased with a total of 325 thousand persons (Statistics Netherlands, 2016a; Statistics Netherlands, 2016b).

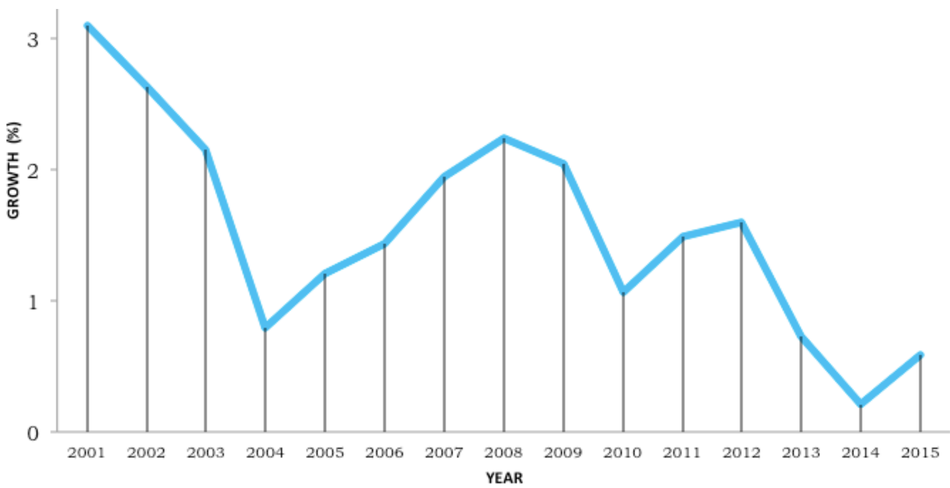


Fig. 2. Growth of car ownership in The Netherlands – own depiction (Source: Statistics Netherlands, 2016a)

There is much unknown on the causes of the current stabilization in car ownership growth in The Netherlands and other countries in the world. Factors considered include a decline of car ownership by young adults, increased urbanization or economic trends. The economic crisis in 2008/2009 might have had an impact on car ownership levels in The Netherlands and other industrialised countries. With regard to the youth, we observe a decline of car ownership levels among young people between 18 and 25 years in The Netherlands (Statistics Netherlands, 2013a). Increased urbanization levels in The Netherlands are suggested to contribute to this development (Statistics Netherlands, 2015a). However, the economic crisis and trends in car ownership among youth do not merely explain why trends of declining car ownership growth were already visible before 2008/2009 (Millard-Ball & Schipper, 2010; Statistics Netherlands, 2016a). If we look at both Figures 1 & 2 we already notice a decline in the growth of car ownership levels in The Netherlands and other different industrialised countries after the turn of the millennium.

Despite the debate about the influence of factors on car ownership levels; ostensibly there is consensus on the claim that the influence of certain factors might have changed over time. In particular, the role of traditional factors (such as income) is suggested to have weakened over the years (Goodwin & van Dender, 2013). To the best of the authors' knowledge, there is, however, an absence of studies that investigate the (changing) influence of factors on car ownership over time. This study will focus on the case of The Netherlands for two reasons:

- The Netherlands is considered to have and apply state of the art knowledge about car ownership (modeling) for research and policy purposes (e.g. Oakil et al., 2016a);
- The Netherlands provides one of the most resourceful databases regarding car ownership in the world (e.g. Potoglou & Kanaroglou, 2008a).

The knowledge gaps that will be pointed out in the next section should be considered as generic and apply to multiple industrialized countries in the world.

1.4. RESEARCH GAP IN CAR OWNERSHIP RESEARCH

Firstly, transport models and urban planning models often assume that the influence of factors (i.e. parameter trends) on car ownership **does not change over time**. The influence of important factors, such as income, gender, and household composition may, however, have changed in the course of time, which could partially explain the current trend in car ownership growth in The Netherlands. For that reason we should investigate whether the influence of important determinants on car ownership has changed over time, and question ourselves whether the assumption of static influence of determinants still holds.

Secondly, transport model systems and urban planning models that make use of car ownership models often **exogenously** determine the number of private cars owned by households (Potoglou & Kanaroglou, 2008a). This contradicts with the view that car ownership modeling plays a vital role in determining motorized trips by households. There's a growing need in developing car ownership models, in which car ownership is endogenously determined (Potoglou & Kanaroglou, 2008a). Understanding how households choose the number of private cars to own is therefore of importance for policy makers and urban planners for establishing mobility scenarios and integrated urban planning models (Potoglou & Kanaroglou, 2008a; Oakil et al., 2016a).

Thirdly, the **type of data** that has been used among research studies to explain car ownership levels in The Netherlands is limited. A considerable proportion of studies have made use of aggregate modelling techniques to obtain insights in the determinants of car ownership in The Netherlands (e.g. Berri, 2009). These techniques have used aggregate economic and demographic data (zonal, regional or national level) to explain car ownership levels. The main reason to use aggregate data is that it is often less time-consuming and costly to obtain. However, studies that use disaggregated data on household level are considered to be more useful, because disaggregate models are statistically more powerful models. Furthermore, aggregated models that use aggregated data, which is highly averaged data, face high levels of co-linearity between explanatory variables meaning that the individual influence of determinants on car ownership levels will be less effectively captured.

Besides, there is a lack of studies that uses longitudinal disaggregate data to date in explaining car ownership levels among households. The use of aggregate cross-sectional data among the majority of recent studies has given us less insight how the influence of car ownership determinants is changing over time. Studies that have made use of longitudinal disaggregate data have primarily been conducted in the 1980s and 1990s (e.g. de Jong, 1990; Golob; 1990; Meurs, 1993). Also the majority of recent studies seem to use relatively out-dated data. Kitamura (2009), for example, explains car ownership as major determinant for mechanized trip generation and modal split by using Dutch National Mobility Panel data. The study is inspired on the work of Ben-Akiva & Lerman (1974) that investigates the dynamics between automobile ownership and mode to work choices. Nevertheless, the data that has been used for this study has been gathered in the spring of 1984, 1985, 1986 and 1987.

Fourthly, the **scope** of current research efforts has often been limited to certain regions (e.g. cities), specific determinants or certain aspects of car ownership related to other fields of study (e.g. de Jong, Kouwenhoven, Geurs, Bucci, & Tuinenga, 2009; van der Waard et al. 2013; Belgiawan et al., 2014; Rubin, Mulder, & Bertolini, 2014; de Groote, van Ommeren, & Koster, 2015; Böcker, Van Amen, & Helbich, 2016; Oakil et al., 2016a). For example, the work of Oakil et al. (2014) researches Dutch households' decisions in changing their car ownership levels using longitudinal data. Nevertheless, the research is specific in the sense that it only uses critical household events (e.g. child birth) for explaining mobility decisions, rather than explaining car ownership from a set of relatively stationary explanatory variables (e.g. income and household size). Most research efforts are mainly driven by themes that are currently enjoying much attention, such as youth, e-communication, urbanization, or peak car (e.g. van der Waard et al., 2013). However, recent studies that look into an extensive set of determinants explaining car ownership levels are difficult to find.

1.5. PROBLEM STATEMENT & RESEARCH OBJECTIVES

To this end, and to the best of the authors' knowledge, there exists no recent study that has investigated the dynamics of the influence of car ownership determinants in The Netherlands based on disaggregate-household data for a longer period of time in recent decades. This gap in current research can be translated into the following problem statement:

> Currently, there is insufficient knowledge on the changing influence of determinants on the choice of the number of private cars to own by households in The Netherlands in recent decades.

To contribute to existing approaches that study the influence of factors on household car ownership in The Netherlands, the aim of this thesis is to provide whether and to what extent the influence of such factors have changed over time. Understanding how such factors have contributed to the households' choice of the number of private cars to own is of importance to urban planners and decision makers, because car ownership levels are highly correlated with urban sprawl and automobile travel (Handy, Cao, & Mokhtarian, 2005). Especially in case car ownership factors are considered to be constant in current car ownership models, it becomes of importance to reveal whether such factors might have changed over the course of time. If this is the case, one should question what implications this might bring for the application of car ownership models that currently predict future car ownership levels in The Netherlands.

Therefore, as an extension to the first research objective, this study aims to provide some policy implications of predicting household car ownership levels based on models that use a constant versus changeable influence of determinants over time. For this purpose, special attention will be given to the most dominant car ownership model currently used in The Netherlands: DYNAMO. The model is widely used by Dutch ministries and political parties, in which the effects of general developments and government policy on size, composition and use of the car fleet are modeled (MuConsult, 2016a).

Insofar, the aim of this study is not to *explain* what factors have contributed to the decline of household car ownership levels in The Netherlands. In contrast, the decline in car ownership levels has led to the awareness that factors influencing car ownership levels might have changed over time – and being the basis of current observed phenomena such as car ownership saturation and peak car. This research tries to bridge this gap, by combining disaggregate household data in recent decades with suitable research methods. The research outcomes might improve our understanding on the influence of car ownership factors over time and their potential role in car ownership prediction models - on account of policymaking, urban planning and future research on car ownership modeling.

1.6. RESEARCH QUESTIONS

Based on the problem statement, the research question of this study is formulated as follows:

To what extent has the influence of determinants on households' choice on the number of private cars to own changed in the Netherlands in recent decades?

The research question should give an answer whether the influence of car ownership determinants on car ownership levels among Dutch households has been subject to change in recent decades. It is important to specify the definition of "influence" to avoid confusion. In this research, "influence" means the micro-level influence of car ownership determinants, which affects the *choice* of a household to obtain a certain amount of cars. The change in micro-level influence of car ownership determinants will be measured by a change in the *parameter trend* over time. The study of trends in car ownership *variables* (e.g. increase of average household income) due to exogenous events (e.g. economic growth), or the changing *relationship* between certain variables and car ownership level (e.g. positive or negative relationship) extend beyond the scope of the research objectives of this study.

To be able to answer the main research question, sub questions have been formulated for this research (Table 1).

Table 1. Overview of research questions and associated research methods

#	Research Questions	Research Methods
1.	What are important determinants that influence a households' choice on the number of private cars to own in The Netherlands?	• Literature review on factors that influence car ownership levels among households.
2.	Which methodology is most suitable to analyse the changing influence of determinants on a households' choice on the number of private cars to own?	• Investigation of literature on methodologies suitable to analyse the influence of determinants on car ownership levels.
3.	What is the influence of determinants over time on a households' choice on the number of private cars to own?	• Descriptive statistics of the determinants of interest. • Perform statistical analysis (modelling). • Perform internal model validation.
4.	What are the policy implications of predicting car ownership levels based on constant versus changing influences of car ownership determinants over time?	• Trend extrapolation of previous research results. • Analysis of outcomes of prediction models for household car ownership. • Formulation of policy implications.

Firstly, the research will start with an investigation of current literature on factors that determine car ownership levels among households (*sub question 1*). Subsequently, suitable research methods will be presented (*sub question 2*). Once the determinants and research methods have been formulated, trends in the determinants of interest and their influence on car ownership levels will be investigated (*sub question 3*). From these observations conclusions will be formulated with regard to what extent the influence of determinants on households' choices has changed in the past decades. Finally, a first comparison between models that predict car ownership levels based on constant parameters versus parameters that are subject to change over time, will be made. This is important to reveal whether the models differentiate in terms of prediction power for household car ownership levels. The outcomes of this analysis will be compared to the DYNAMO model, which is predominantly used by Dutch ministries and political parties. Especially, in case the influence of factors is considered to be constant in the DYNAMO model, we should question ourselves what policy implications this might bring for the future (*sub question 4*).

1.7. THESIS OUTLINE

This study is divided into five parts. The first part comprises the **exploring** phase, a theory-oriented analysis, which has the goal to find an adequate overview of determinants affecting car ownership levels among households. Subsequently, suitable statistical research methods to analyze the changing influence of such determinants will be discussed. The exploring phase will be covered by Chapter 2 and will answer sub questions 1 and 2.

The second part forms the **pre-modeling** part of this research. Chapter 3 will be dedicated to the research methods, data and explanatory variables that will be used for this research. The data and explanatory variables will be selected from the Dutch national mobility surveys OVG/MON/OViN, which have been annually conducted from 1985 till 2014. The datasets record information about trips and background information of respondents, including age, gender, level of education, town of residence, composition of the household, and the possession of vehicles and driving license. In contrast to previous studies, the type of data that will be used for this study is disaggregate and longitudinal, and above all not restricted to certain regions (i.e. nation-wide) or specific determinants. In addition, the main assumptions behind the statistical models that will be used for this study will be discussed.

The third part of this thesis focuses on the **interpretation** of the results of the modeling part to answer sub question 3. Chapter 4 will start with an overview of the descriptive statistics of the determinants of interest. Afterwards, the model results will be discussed with regard to the (changing) influence of determinants on household car ownership levels will be discussed. Chapter 5 will, subsequently, internally validate the statistical models that have been used for this study.

In the fourth part the results from the modeling part will be used to **extrapolate** trends to construct a statistical model that allows for alternative effects of car ownership determinants on the level of household car ownership. The model will be used to explore the prediction capabilities compared to a statistical model that uses a constant (i.e. not changeable) influence of car ownership determinants on the level of household car ownership over time. Furthermore, this part will give an overview of the policy implications, when the influence of factors is considered to be more constant in current car ownership model(s) to predict future car ownership levels in The Netherlands. The goal of this part is to answer sub question 4. Chapter 6 will cover this part.

Finally, the **conclusion**, **recommendations** and **reflection** of this work will be the components of the concluding part of this research. Chapter 7 will answer the main research question and will give recommendations for future research. Finally, Chapter 8 provides a reflection on the research goals that have been adapted during the execution of this study. Furthermore, the limitations of this study, which have impacted the outcomes and depth of this research, will be discussed.

A visualization of the different parts of the thesis is given in Figure 3. By means of a flow diagram (Böhm & Jacopini, 1966) the different research parts and their relation to the end product are shown. Also the relations between every research part and sub questions are given. The sub questions are abbreviated by "SQ" following with the number of the sub question in the same order as presented by Table 1 in Section 1.6.

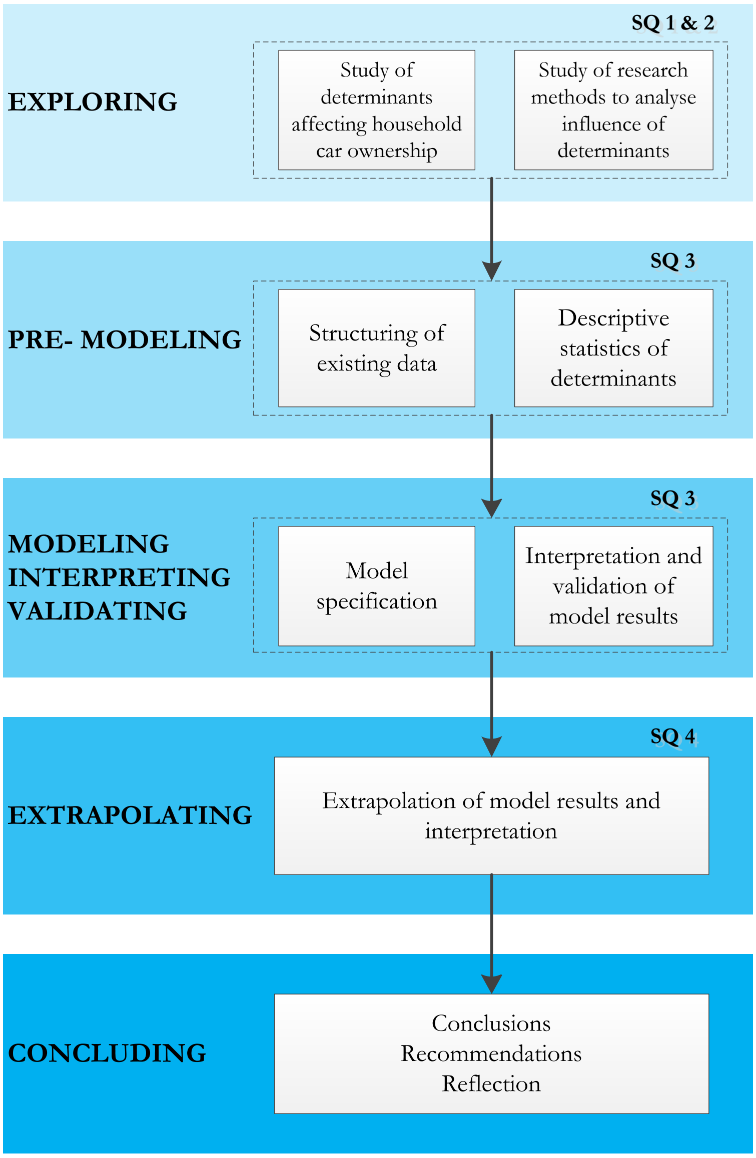


Fig. 3. Visual overview of research approach

2

Literature Overview

This chapter comprises the exploring phase of this research and can be divided into two parts. The chapter will start by giving a literature overview of the most important **determinants** affecting car ownership levels among households (see Section 2.2). The second part will give an overview of statistical **models** that are suitable to analyze the changing influence of car ownership determinants and predict new outcomes of household car ownership levels (see Sections 2.3-2.5).

2.1. RESEARCH APPROACH

To obtain a thorough analysis of car ownership determinants and car ownership models *Scopus, Web of Science* and *Google Scholar* databases have been used. Combinations of the following keywords have led to the articles used in this section: (household) (private) car/vehicle ownership, (car ownership) determinants/factors, car mobility, transport modeling and car ownership modeling.

The research findings of de Jong et al. (2004), Potoglou & Kanaroglou (2008a), Johnstone, Serret, & Bureau (2009), Oakil et al. (2016a) and Wu, Zhao, & Zhang (2016) form main part of the literature overview relating to car ownership determinants in section 2.2. From the references of these articles other articles have been found, often relating to a specific (category of) determinant(s) (e.g. Dargay, 2001; Belgiawan Schmöcker & Fujii, 2010; Oakil et al., 2016b).

Secondly, the publications of Oi & Shuldiner (1963), Bhat & Pulugurta (1998), de Jong et al. (2004), Potoglou & Susilo (2008), Potoglou & Kanaroglou (2008a) and de Jong et al. (2009) are the basis of Sections 2.3-2.5, wherein statistical car ownership modeling methods are described. Especially, the article of Potolgou & Susilo (2008) has proved itself to be valuable, since it provided an extensive comparison of different statistical car ownership models.

2.2. HOUSEHOLD CAR OWNERSHIP DETERMINANTS

Considering literature on car ownership determinants we can distinguish five important categories of determinants that positively or negatively influence car ownership levels among households (de Jong et al., 2004; Johnstone et al., 2009; Oakil et al., 2016a; Wu et al., 2016):

- **Economic** factors (e.g. income, employment, vehicle prices and fuel costs);
- **Socio-demographic** factors (e.g. age, gender, household size and household composition);
- **Spatial/land use** factors (e.g. urbanization level & transport infrastructure);
- **Transportation** factors (e.g. number of passengers per trip and travel distances);
- **Attitudinal** factors (e.g. concern for environment).

Firstly, several economic factors play a considerable role in car ownership levels. The first main variable is the **income**, which can be a disposable or standardized disposable income (Dargay, 2002; Nolan, 2010; Oakil et al., 2014; Potoglou & Kanaroglou, 2008a; Oakil et al., 2016a). Higher income levels have a positive effect on the level of car ownership, whereas lower levels of income have to a lesser extent an effect on car ownership levels. Dargay (2001) tested the asymmetrical relation between car ownership and income using a dynamic econometric model relating household car ownership to income. The study revealed that lower income levels do not directly lead to lower car levels, because of asymmetrical income elasticity. Once a car has been bought, individuals become accustomed to car use, so that car ownership becomes a necessity rather than a luxury. As a consequence, the disposal of car is much more difficult when income falls for individuals who already own a private car (Dargay, 2001; Johansson-Stenman, 2002).

Besides income **vehicle prices** and **fuel costs** also affect car ownership levels (Johnstone et al., 2009). Vehicle costs can be categorized into car purchasing costs and running costs, including maintenance, parts, road tax, tolls, parking and insurance (Dargay, 2002). The vehicle costs and fuel costs elasticity determines the extent to which households are sensitive for price fluctuations. The vehicle and fuel price elasticity could be different for various household groups. To illustrate, the greater 'car dependency' in rural regions results in a lower price elasticity. As a result, the effects of fluctuations in vehicle prices and fuel costs on car ownership levels are often smaller compared to urban regions where the possibility to alter transport mode is greater (Dargay, 2002).

A second category that affects car ownership levels among households relates to socio-demographic factors. Firstly, the **gender** of a person has an effect on the type of household activities and responsibilities, which could in turn affect the ownership of private car(s). Furthermore, research shows that males and females have different attitudes regarding car use and car ownership (e.g. Oakil, Nijland, & Dijst, 2016b). Also the **age** of the households' reference person (i.e. head of the family) is often used as explanatory variable (e.g. Oakil et al., 2016b). Current statistics of Statistics Netherlands on car ownership levels suggest that car ownership among young people is declining while the rate of car ownership among older people is growing (Statistics Netherlands, 2013a; Oakil et al., 2016b). Some studies suggest that young singles and couples will start later with a family so that the possession of a car becomes less stringent (Manting, 2014). The age of a person influences the necessity of car (e.g. work-related, family-related) and the possibility to have a car (income-related).

Socio-demographic factors also include household **size** (relating to the number of members of one household) and household **composition** (relating to the type of members of one household) as explanatory variables for car ownership levels (Potoglou & Kanaroglou, 2008a). Both size and composition (i.e. greater variety) of the household have a positive influence on the number of cars owned by the household. For example, the greater the number of family members the more likely it is that a household will obtain a car for practical mobility purposes. The same is true relating to the composition of a household; e.g. the presence of young children that need to be taken to school will increase the likelihood that a household will posses one or more cars (Nolan, 2010).

In addition, the **educational** level of a households' reference person is considered to be an important explanatory variable for car ownership levels (Potoglou & Kanaroglou, 2008a; Nolan, 2010). The educational level of a person (e.g. primary, secondary or tertiary level education) has a positive influence on the number of cars that is obtained by the household. Higher levels of education have two indirect positive effects: a high level increases the individuals' likelihood move into the labor market and it will increase the probability of higher salaries (Eakins, 2013). Further, higher education levels are positively associated with environmental concern and change the attitude of towards vehicle ownership (Flamm, 2009).

Furthermore, the **working status** (e.g. fulltime, part time or unemployment) of the household's reference person is also found to be a significant determinant of variations in car ownership (Nolan, 2010). Households with a full-time working member(s) are associated with higher levels of car ownership (Potoglou & Kanaroglou, 2008a; Eakins, 2013). Working status or employment status will indirectly positively influence the number of kilometers that need to be travelled weekly by the person. The more kilometers a person needs to travel on weekly basis the more likely it becomes that a person will own one ore more private cars. Besides the working status of a person has also an indirect positive effect on the amount of salary that is earned.

Next to economic and socio-demographic characteristics spatial/land-use characteristics play an important role in determining the number of private cars among households (de Jong et al., 2004; Potoglou & Kanaroglou, 2008a; Nolan, 2010; Oakil et al., 2016a). The most important variable in this category is the **level of urbanization**. Research has shown that households that live in high-density areas are less likely to own a car compared to low-density areas where the necessity of owning a car becomes higher (Oakil et al., 2016a). Other urban variables include the type of housing and attributes related to the place of residence (urban vs. suburban locations) (Bunch, 2000; Potoglou & Kanaroglou, 2008b). Also access to **public transport** has an effect on car ownership levels among households. Especially, public infrastructure that is easily accessible for households (< 5 min) seems to have a significant effect on car ownership levels. Better access leads to households that are less likely to obtain a car (Johnstone et al., 2009; Eakins, 2013).

Psychological factors, including perceptions, attitudes and habits have gained attention among researchers to investigate car use and car ownership levels (e.g. Fujii & Kitamura, 2003). Recent studies on psychological determinants on car ownership reveal interesting insights in the psychology of (non)-car owners. The study of Belgiawan et al. (2011) revealed that primarily **convenience** but also **prestige** and **social orderliness** are significant determinants. Especially, factors as prestige and social orderliness are considered to be factors that might create "anti-car" trends among new generations (Belgiawan et al., 2010). Also **concern for the environment** is a determinant that gains increasing attention from researchers (Johnstone et al., 2009). Environmental concerns relating to air pollution by conventional cars has already led to switching behavior of consumers towards electric cars (e.g. Lieven, Mühlmeier, Henkel, & Waller, 2011) or other modes of transport such as public transport (Beirão & Cabral, 2007). However, "soft" variables as attitude and perception are not easily measured, which makes data on psychological factors difficult to obtain.

Finally, explanatory variables related to transportation, including the average **number of passengers** per private car per trip and **travel distances** for work-related and other activities, are sometimes used to determine car ownership levels (e.g. Wu et al., 2016). However, such variables are often endogenous, meaning that they are dependent on other variables that are related to car ownership. Therefore such variables are considered to be less suitable for analyzing car ownership levels compared to economic/demographic variables and urbanization levels. An overview of the discussed explanatory variables is given in Table 2. Based on the data that will be available for this research the next chapter will make a selection of important explanatory variables.

#	Explanatory Variable	Categorization
1. 2. 3.	Household income Vehicle prices Fuel costs	• **Economic** explanatory variables
4. 5. 6. 7. 8. 9.	Gender of households' reference person Age of households' reference person Household size Household composition Educational level of households' reference person Working status of households' reference person	• **Socio-demographic** explanatory variables
10. 11.	Level of urbanization Access to public infrastructure	• **Spatial/land-use** explanatory variables
12. 13. 14. 15	Convenience Prestige Social Orderliness Concern for environment	• **Attitudinal** explanatory variables
16. 17.	Number of passengers per private car per trip Travel distances for work-related and other activities	• **Transportation** explanatory variables

Table 2. Overview of explanatory variables for household car ownership determination

2.3. STATISTICAL RESEARCH METHODS

In general, statistical car ownership models can be obtained by using aggregate statistical models, which use aggregate data (zonal, regional or national level), or can be attained by disaggregate models, which use the household as unit of analysis (Bhat & Pulugurta, 1998). As discussed in Section 1.4 disaggregate models on household level are more suitable to be used, because disaggregate models are statistically more powerful models, while aggregate models also face high levels of co-linearity between explanatory variables (Oi & Shuldiner; 1963; Bhat & Pulugurta, 1998; Potoglou, & Susilo, 2008; Potoglou & Kanaroglou, 2008a). Therefore, disaggregate models have become the preferred model of modeling discrete car ownership choice among households.

Given the fact that disaggregate-type models are statistically more powerful methods we will focus on two types of car ownership models: static disaggregate models and (pseudo)-dynamic models (de Jong et al., 2004). Both static and (pseudo)-dynamic disaggregate are discrete choice models that use discrete choices for households (e.g. zero, one, two, three or more cars) as explanatory variable (de Jong et al., 2004). Firstly, static disaggregate models using cross-sectional data have predominantly been used to research the relation between determinants and the number of cars owned by households. Applications that have been used in The Netherlands are the Dutch national model system (LMS) for transport (Hague Consulting Group, 1989) and disaggregate models used by Bhat and Pulugurta (1998). To this day, static disaggregate models are still often used for explaining car ownership levels (e.g. Potoglou & Kanaroglou, 2008a). An important example is the DYNAMO model that is being used by Dutch ministries and car companies in which the effects of general developments and government policy on size, composition and use of the car fleet are modeled (MuConsult, 2016a).

On the other hand (pseudo)-dynamic models are increasingly being used among researchers and policy makers to explain car ownership levels among households. The main difference between static and dynamic discrete choice models is that static models do not assume that decisions by households are affected by past or future decisions and events, while dynamic models presume this to be true. Discrete dynamic choice models can be categorized into "holding" models and "transaction" models. Holding models describe the likelihood that a household will hold a set of private cars at a single point in time, while transaction models describe the chance that a transaction of cars occurs in a household (i.e. disposal, replacement or acquiring) (de Jong et al., 2009).

Though dynamic models assume a realistic decision making process by households, the use of static disaggregate model is considered to be more suitable for this research. First, this thesis will not estimate the likelihood whether households have or transact one or more cars. Instead, this research tries to obtain insights whether underlying factors that influence such a decision of time will change. Secondly, dynamic models are still in their infancy and face a number of teething troubles. Dynamic models assume, based on utility theory, that for every point in time the household chooses the best set of private cars with the "highest utility". However, in practice, households do not transact cars so often – among other things due to transaction costs. As a result, dynamic models often suffer from including the right household decision processes, whereby the process of household car ownership is accurately formulated (de Jong et al., 2009). Finally, dynamic models are considered to be sophisticated forecasting models for longer periods of time, especially in case radical changes are expected in the future (de Jong & Kitamura, 2009; Cirillo & Xu, 2011). However, this research is interested in past events, wherein decisions have already been made, and will use a limited time horizon to forecast new outcomes. Accordingly, this thesis proposes to use a static disaggregate model using longitudinal data to gain accurate estimates of households' decisions with regard to car ownership levels.

2.4. STATIC DISAGREGGATE CAR OWNERSHIP MODELS

Static disaggregate car ownership models can be subdivided into two general decision mechanisms: ordered-response and unordered-response mechanisms (Bhat & Pulugurta, 1998; Potoglou & Susilo, 2008). Ordered response models make the assumption that a household's choice of the number of private cars to own is dependent on a one-dimensional latent variable. The latent variable reflects the inclination of the household to own private cars. Such models are referred to as Order Logit Models (ORL) or Order Probit Models (ORP). Unordered response models, based on the principle of random utility maximization, assume that a household chooses the number of cars based on their individual utility and chooses the one with the maximum utility (Potoglou & Susilo, 2008). Multinomial Logit modeling (MNL) is the main representative type of modeling for models that use this unordered response mechanism.

The comparative study of Potoglou & Susilo (2008) that evaluated MNL, ORL and ORP models for car ownership on a number of data-fit measures and theoretical consequences showed that MNL models are more suitable to modeling car ownership. This is because a strong theoretical framework of random utility (RUM) supports MNL models (McFadden, 1973). On the other hand, ORL and ORP models aren't based on a solid behavioral framework, but rely on a single latent variable (Potoglou & Susilo, 2009). In addition, unordered-response mechanisms do not place any restrictions on the effects of explanatory factors in contrast to ordered response mechanisms. Because unordered-response mechanisms use unique parameters per estimated variable, potential non-linearity between explanatory variables and each car ownership level can be uncovered (Bhat & Pulugurta, 2008; Potoglou & Kanaroglou, 2008a). Furthermore, empirical analysis shows that MNL models are more flexible compared to ORL or ORP models allowing for alternative effects of determinants across different car ownership levels. ORL and ORP models are, however, constrained to a unique coefficient per determinant, based on the parallel slopes assumption, which increases the likelihood of producing biased results (Potoglou & Susilo, 2009). This makes MNL models, compared to ORL and ORP models more suitable for forecasting purposes.

2.5. CHOICE OF CAR OWNERSHIP MODELS

A review on car ownership models shows that MNL models seem to be the most appropriate type of car ownership modeling. The flexibility of MNL models comes, however, at a cost, as the interpretation of outcomes is relatively time consuming owing to the substantial number of parameters to be estimated (Potoglou & Kanaroglou, 2008a). To analyze the effects of car ownership determinants on household car ownership levels, both an ORL model (commonly used in research to represent the ordered-response mechanism) and MNL model have been estimated. An analysis of the accuracy of the ORL model and MNL models shows, despite the greater inflexibility of the ORL model, that the accuracy of both models is almost identical (Chapter 5). Because ORL models estimate a unique coefficient per determinant, the interpretation of the ORL model outcomes are more effortless and easier to interpret compared to MNL models. Therefore, for the *exploratory* part of this study - to investigate the changing influence of car ownership determinants over time - the ORL model outcomes will be presented. During the validation of the ORL model, the accuracy of the MNL models will also be shown to prove that the ORL model performs at the same level of accuracy as compared to the MNL models.

For the *predictive* part of this study - to compare a model that uses constant versus changing car ownership coefficients over time - the unordered response model (MNL) will be used. MNL models allow for more alternative effects of determinants on car ownership levels, which make their application more suitable to *predict* household car ownership levels compared to ORL models.

A visualization of the steps that have led to the choice of using an ORL model for the explanatory part, and the MNL model for the predictive part of this research is given in Figure 4. The choice to use both ORL and MNL models in this study makes the availability of disaggregate data of crucial importance for this research. The next chapter will discuss which disaggregate data sources are available for this research. On the basis of the data and variables that are available for this research the chapter will further specify the statistical methods that will be used in this study.

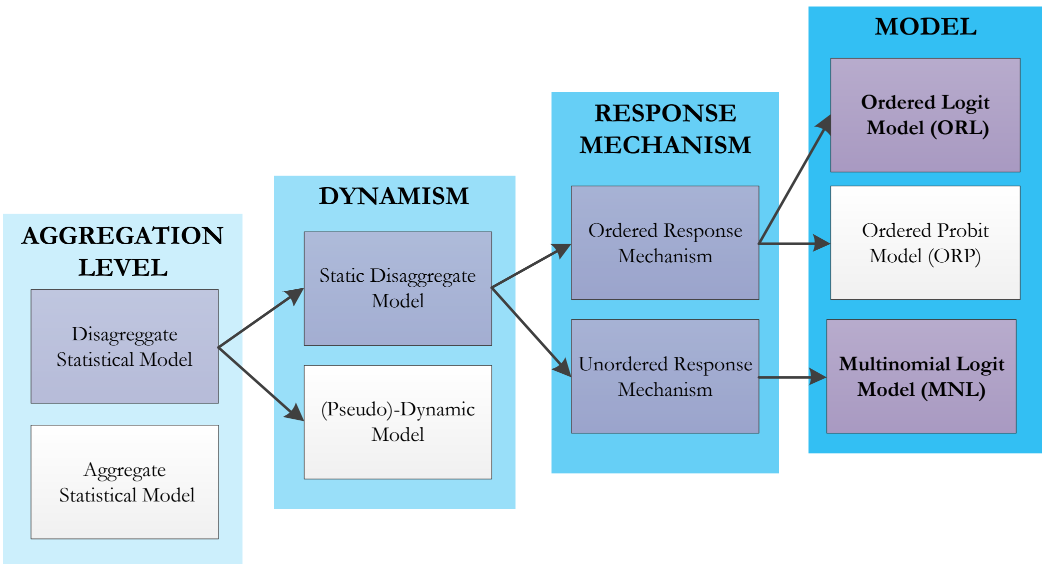


Fig. 4. Decision tree for preferred type of modeling

3

Methods & Model Specification

This chapter forms the pre-modeling part of this research, which describes the research methods and model speciation. The chapter will start by discussing the data that will be used for this research. Subsequently, the chapter will describe which determinants for car ownership levels will be used on the basis of the data that has been made available for this research. Moreover, the most important assumptions that have been made with regard to the model specification will be given. Finally, the chapter will give a specification of the methods that will be used for the rest of this study.

3.1. DATA SOURCES

On the basis of the discussion in Section 2.5, the existence of disaggregate household data is essential. Disaggregate data is, however, costly and time-consuming to obtain (Potoglou & Kanaroglou, 2008a). In The Netherlands, disaggregate data on mobility among households has been collected by means of a national mobility survey among approximately 40,000 individuals on annual basis (SWOV, 2013). Until 2003, the data was collected by Statistics Netherlands and was called the Traffic Survey (*Onderzoek Verplaatsingsgedrag,* OVG). Afterwards, the survey was carried out by the Centre for Transport and Navigation and was called the Dutch Mobility Survey (*Mobiliteitsonderzoek Nederland,* MON). From 2010, the survey has been carried out again by Statistics Netherlands under the name Traffic Survey of The Netherlands (*Onderzoek Verplaatsingen in Nederland,* OViN).

This study has access to the data provided by the three surveys (OVG/MON/OViN) from 1985 till 2014. The datasets record all trips and trip stages for one day among participants and includes demographic as well as economic characteristics of the respondents. The next sections will discuss which issues must be considered when the data is used for research, and which determinants of interest can be researched based on the data that has been made available for this research.

3.2. DATA CHARACTERISTICS

The three types of annual mobility surveys (OVG/MON/OViN) that have been carried out use a significant sample size of the total population of The Netherlands – up to 50,000 unique cases. Data that has been gathered by the three surveys records the trips made by the respondent on a specific day that has been assigned (SWOV, 2016). In addition to trip records, the surveys gather the distance travelled, the transport modes that have been used, the places of departure and arrival, and the reason why the journey has been made. Besides information about the journeys, the respondent is asked to provide further background data, including age, gender, level of education, town of residence, composition of the household, and the possession of vehicles and driving license.

Despite the fact that the OVG/MON/OViN are comprehensive datasets with large sample sizes and a multitude of variables the following limitations need to be taken into account when the data will be used for research (SWOV, 2016):

1. Mobility of Dutch inhabitants outside The Netherlands is not included in the survey;
2. Before 1994 children in the ages 0-11 were not included in the study, resulting in approximately 14 billion kilometers missing from total distance traveled compared to later years;
3. About 80% of the mobility in The Netherlands is estimated to be covered by the OVG/MON/OViN data;
4. Over time changes in research methods among the different surveys have been made (e.g. formulation of questions and sampling frame);

The first two limitations of the data provided will not affect the outcomes of this study; mobility patterns of Dutch inhabitants abroad and the inclusion of children between 0-11 do not change survey results regarding ownership levels among the households. The third and fourth limitation could, however, affect research outcomes and their relevance. Firstly, the total coverage of 80% of mobility in The Netherlands by the surveys excludes lorry traffic (approximately 12%), holiday traffic of individuals (approximately 5%) and distance traveled by individuals who do not reside in The Netherlands (approximately 4%) (SWOV, 2013). The percentage is a measure of the representativeness of the sample size with regard to mobility patterns of the total Dutch population. Systematic biases that occur when respondents forget to fill in their kilometers traveled or make wrong estimates are prevented by checking whether the distance travelled with the reported mode can be made within the timeframe of travel that is given by the respondent. The same is true if car ownership levels from the datasets are compared to the total population of The Netherlands. For example, the survey outcomes of OViN-2010 show differences with regard to car ownership levels found in the Dutch population, as a result of under- or over-representation of certain groups in the OViN survey (Statistics Netherlands, 2016c). Weights have, therefore, been included in the surveys to adjust the survey outcomes to same levels as found in the Dutch population for the year of reference.

Secondly, changes that have been made in the surveys' research methods over the course of time could affect the results of this research. Before 2010, the study was a household survey, wherein all members of the household were asked to fill in the survey. After 2010, the survey has been carried out as an individual survey. Furthermore, in 1985, 1994 and 1999 some major corrections have taken place with regard to the design and execution of the study (SWOV, 2016). The corrections resulted in a divergence of trends and methods, wherefore the data in the period 1985-1999 has been become incomparable to datasets since 1999. However, in response to this issue the data of the period 1985-1998 has been corrected so that the datasets are now comparable to datasets from 1999 onwards. Consequently there is a valuable series of data available from 1985, which makes the period of 1985-2014 suitable to be analyzed.

3.3. EXCLUSION OF *MON* DATASETS

As described in the previous two sections this study can make use of the existing datasets OVG/MON/OViN from 1985 till 2014. Despite the datasets cover a relatively extensive set of determinants, not all determinants of interest that have been found in literature are presented by the datasets (see Section 2.2). Additionally, this research will make use of different types of datasets so that not all variables and measuring scales will be equivalent. Data will thus be an important aspect of this study that will limit the outcomes of this research. Most problematic are the MON datasets, which do not measure household income. Because household income is seen as one of the most important explanatory variables, it would be unrealistic to leave this variable out on the basis of the MON datasets. To measure the influence of household income through time this study will only make use of the comprehensive OVG and OViN datasets, which cover almost 80% of the total years between 1985 and 2014; the MON datasets cover six years of the total time horizon (2004-2009). Hence, possible measurement errors in the datasets can still be detected by using two different types of datasets.

3.4. TIME HORIZON OF ANALYSIS

To obtain accurate insights in the effects of the determinants on car ownership levels a multitude of years from the OVG and OViN datasets has been selected for analysis. The years that will be analyzed are in ascending order 1987, 1991, 1995, 1999, 2003, 2010 and 2014. The selection of years concerns a trade-off between accuracy and time efficiency/computing power. Selecting all years between 1985 and 2014 for analysis purposes would introduce a substantial number of variables, leading to significant amounts of data, which requires substantial computing power. Analyzing all years would therefore lead to computing problems and/or long computing time. Therefore, a procedure has been chosen that selects years based on a four-year time step from 2014 backwards with the exclusion of the years measured by the MON datasets. This led to a considerate time horizon covering 27 years from 2014 till 1987. It is assumed that the seven years used for analysis will give a clear picture of the (changing) influence of determinants on car ownership levels over time.

3.5. INCLUSION OF DETERMINANTS

Table 2 summarizes the five most important categories of household car ownership determinants. These categories contain economic, socio-demographic, spatial-related, transport-related and psychological variables. The datasets of OVG and OViN cover eight out of the seventeen determinants of household car ownership. Table 3 gives an overview which determinants have (not) been included. The datasets cover the important variable of household income (by excluding the MON dataset) and all socio-demographic variables. Furthermore, the datasets cover the most important spatial/land-use explanatory variable, which is the level of urbanization. However, the datasets do not cover vehicle prices, fuel costs, and transportation infrastructure. In addition, transportation and attitudinal explanatory variables are completely absent.

Included determinants	Excluded determinants
I. **Economic** explanatory variables	
1. Household income	2. Vehicle prices 3. Fuel costs
II. **Socio-demographic** explanatory variables	
4. Household size 5. Household composition 6. Gender 7. Age 8. Educational level 9. Working status	
III. **Spatial/land-use** explanatory variables	
10. Urbanization level	11. Transportation infrastructure
IV. **Transportation** explanatory variables	
	12. Number of passengers 13. Travel distances
V. **Attitudinal** explanatory variables	
	14. Convenience 15. Prestige 16. Social orderliness 17. Concern for environment

Table 3. Overview of included and excluded explanatory variables – OViN and OVG datasets

3.6. IMPLICATIONS OF DATA LIMITATIONS

The number of years and the number of variables that will be analyzed are considered to be the main limitations of the data used for this research. The selection of years concerned a trade-off between accuracy and time efficiency/computing power. Seven years (i.e. timestamps) will be analyzed in a time period that comprises roughly three decades. It is assumed that the number of years will give sufficient indication to what extent the influence of car ownership determinants have changed in this period of time. However, no effects will be analyzed of years that have been excluded, which result in a fragmented overview of the actual development of the influence of car ownership determinant over time. Especially, trend extrapolation that will be used in Chapter 6 to predict new model outcomes will suffer from the limited number of time stamps, leading to less accurate model predictions.

Secondly, the total number of variables that is included in the datasets comprises eight out of seventeen car ownership determinants that have been identified affecting household car ownership levels. The variables that have been included are exogenous (compared to endogenous transport-related variables) and considered to be of importance (e.g. household income, size and level of urbanization). The complete absence of especially attitudinal factors is considered to be an important limitation of this study. Variables such as convenience and environmental concern are gaining increasing attention in the debate on the influence of factors on car ownership levels and car use (e.g. Johnstone et al., 2009). The exclusion of certain variables will provide insights in the effects of determinants, but will also lead to less accurate descriptive and predictive models of household car ownership. Accordingly, it is expected that the variance in actual car ownership household data can only be partly explained by ORL and MNL models that will be estimated. The implications of the data restrictions will be further discussed in Chapter 8.

The conceptual model based on the datasets of OVG-1987, OVG-1991, OVG-1995, OVG-1999, OVG-2003, OViN-2010 and OViN-2014 is visualized in Figure 5. The economic, socio-demographic and spatial explanatory variables are visualized in the left column. A further specification of these eight variables will be given in Chapter 4.

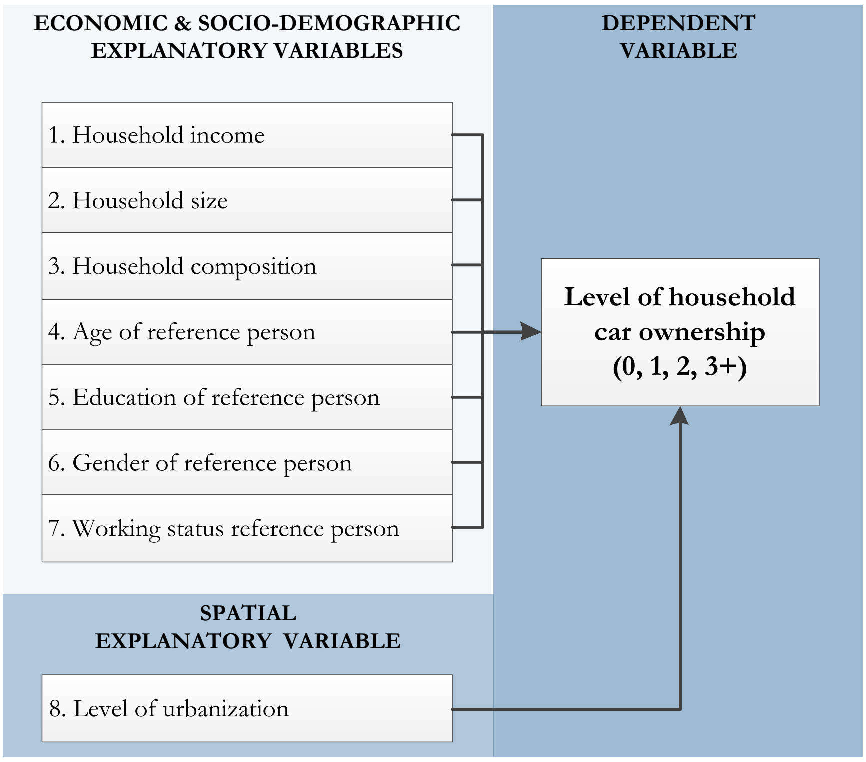


Fig. 5. Conceptual model of household car ownership

3.7. MODEL SPECIFICATION ASSUMPTIONS

Some important demarcations with regard to the model specification need to be explicated. Firstly, for this study four car ownership level alternatives are specified: **zero, one, two, three or more private cars**. It is important to notice that the term "private car", which has not been further specified in the surveys, could also include minivans, vans, light-duty vehicles and pick-up trucks next to passenger cars. Also the data does not distinguish between cars for personal use and company cars.

Secondly, the explanatory variables household income, size and composition (i.e. the presence of children) are all measured on a *household* level. The explanatory variables gender, age, education and working status cannot, however, be measured on household level. In order to obtain useful insight in explanatory variables that cannot be averaged on household level, this study will take **one person** (male/female) as the reference person per single household. Gender, age, education and working status are, therefore, measured on an *individual* level and are so-called proxies. Considering households with two "heads of family" (e.g. married couples) one of the two heads has been randomly selected as reference person. The reason for this random selection is to not discriminate on other aspects that need be analyzed by the model, to secure relevant outcomes. For example, by only selecting males as head of the household we would discriminate on gender and possibly obtain biased results regarding the influence of gender on car ownership levels.

Thirdly, as already explained in the Section 3.2 the survey results can be corrected such that they reflect the Dutch population for the year of reference. In order to achieve this, **weights** have been specified in the surveys (i.e. multiplication factors and correction factors). This study decided to not make use of such weights to analyze the data. Firstly, because a multitude of years will be studied, comprising for different weights would mean that merging datasets couldn't be done, unless all cases are multiplied by their weights accordingly. This would lead to a substantial database (millions of cases) that reflects the entire population of The Netherlands multiplied by seven years. It is expected that such a dataset would be problematic to analyze in terms of computing power. Nevertheless, the already substantial datasets facilitate to investigate relatively small subgroups within the population, such as households owning three or more cars. Previous research using the datasets suggests that especially data on transport modes in the OVG and OViN datasets (e.g. motorcycle, moped and public transport) are under- or overrepresented. For a number of transport-modes and sub-groups the number of trips is extremely small (SWOV, 2016). However, these variables and subgroups will not be used in this study.

Finally, **children** have been excluded from the datasets (0-17 years). Hence, children under eighteen in The Netherlands are legally not permitted to obtain a car. This group is therefore not relevant to be analyzed during the research. All other steps that led to the structuring of the data and re-categorization of the variables are described in Appendix I.

3.8. OPERATIONALIZATION OF DETERMINANTS

Eight explanatory variables will be used to measure the effect on household car ownership, being either zero, one, two or three or more private cars. The independent variable household income is defined by six categories: 0-10,000, 10,000-20,000, 20,000-30,000, 30,000-40,000, 40,000-50,000 and 50,000 or more euros/year. The classification of household income is based on the same classification used by Statistics Netherlands. Furthermore, the household income represents here the *disposable income* of a household, which is the cumulated gross income of the persons living in one household (excluding children) after deducting income transfers, premium income insurance, health insurance premiums, taxes on income and capital and inflation (Statistics Netherlands, 2015e; Statistics Netherlands, 2016f). Although we are quite certain and assume in this study that the disposable income is adjusted for inflation (Figure 6), we could not exclude this has not been the case. Therefore, an additional analysis has been executed to investigate the effects of an adjustment for inflation if we assume that disposable household income has not been adjusted for inflation (Appendix VI). The analysis shows that the discrepancies in terms of effects on the outcomes are relatively modest for all variables. The same trends are visible but to a certain extent weakened for household income and household size (Appendix VI). Nevertheless, for the rest of this study, we will assume that household disposable income has been adjusted for inflation.

Statistics Netherlands uses the disposable income to create the standardized disposable income, which is corrected for the type (e.g. children) and number of persons living in a household. Using equivalence factors all incomes are reduced to the income of a single-person household to make the purchasing power among households comparable (Statistics Netherlands, 2008; Bouhuijs & Engelen, 2015). This research preferred to use the household disposable income instead of standardizing household disposable income by household type and household size. The reason is that this research is interested in the total income of the household, which is not reduced to the income of a single-person household. Furthermore, it is also assumed that car ownership depends more on total household income than on standardized household income (Oakil et al., 2016a).

Household size is categorized based on the number of persons, corresponding to 1, 2, 3, 4, 5, 6 or more persons per household. Household composition includes two types: families with and without children. The age of the reference person has been categorized into 18-19 years, 20-29 years, 30-39 years, 40-49 years, 50-64 years, 65-74 years, and 75 years and older. The educational level considered four categories into primary education (BO, LO), "old" secondary education (LBO, VGLO, LAVO, MAVO, MULO), "new" secondary education (MBO, HAVO, Atheneum, Gymnasium, MMS, HBS) and tertiary education levels (HBO/University). Working status or employment status was either non-employed, "part time" employed (<30 hours per week) or "fulltime" employed (>30 hours per week). Finally, the number of addresses per km^2 defined the urbanization level of the household. Based on the official categorization of Statistics Netherlands, five urbanization levels are recognized: i) very high-density areas (≥ 2,500 addresses per km^2); ii) high density areas (1,500-2,500 addresses per km^2); iii) moderately high-density areas (1,000-1,500 addresses per km^2); iv) low density areas with (500-1,000 addresses per km^2); and v) very low-density areas (<500 addresses per km^2).

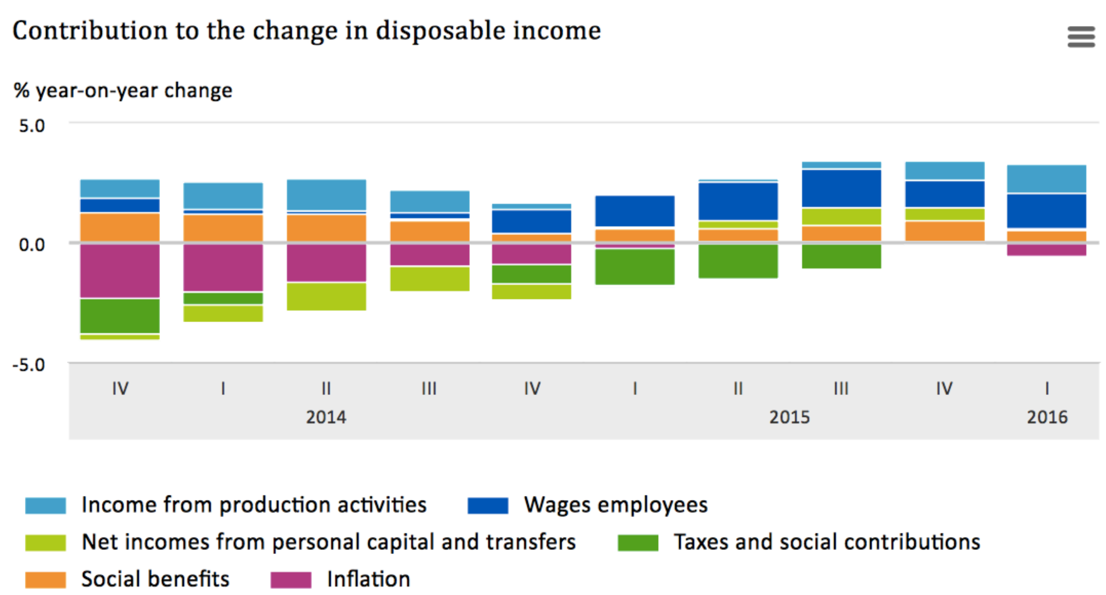


Fig. 6. Example of the elements in disposable income used by Statistics Netherlands (2016f)

3.9. SPECIFICATION OF RESEARCH METHODS

Based on the findings from Chapter 2 and Chapter 3, this conclusive section will give a detailed overview of the methods that will be used in the coming sections in this research.

I. EXPLORATION

In order to assess the effects of the explanatory variables (Fig. 5) on household car ownership levels for the years 1987, 1991, 1995, 1999, 2003, 2010 and 2014, interaction effects have been created by k-1 dummy variables (where k = number of years). By doing so, the year 1987 will be chosen as reference year. Possible interaction between explanatory variables, such as age and the level of urbanization, or income and working status is also controlled for via this method. As discussed in Section 2.5, ordered logistic regression analysis (ORL) is chosen in order to unravel the changing effects of the determinants on household car ownership levels between 1987 and 2014. The ORL model is estimated by means of the statistical program SPSS. Chapter 4 will present the outcomes of the ORL model that has been estimated for this research.

II. VALIDATION

To validate the ORL model outcomes, the model will be internally validated. For this purpose, a holdout validation sample will be used. Randomly 20% of the total number of unique cases for every OVG and OViN dataset in 1987, 1991, 1995, 1999, 2003, 2010 and 2014 has been subtracted. The coefficients estimated by the model will be used to predict the response of the respondents in the holdout sample. After the ORL model outcomes have been internally validated, the accuracy of two MNL models (using constant and changing car ownership coefficients) will be compared to prove that ORL model performs at the same level of accuracy (Chapter 5). The key reason to use the ORL model is that the interpretation of the model outcomes is more effortless compared to MNL model outcomes.

III. PREDICTION

Finally, to compare prediction models that use constant versus changing car ownership coefficients over time, MNL modeling will be used. MNL models are considered to be the most suitable type of models to predict household car ownership levels (Potoglou & Susilo, 2008). In order to achieve a most accurate comparison between the two types of models (based on the data available for this research) the OVG and OViN datasets of 1987, 1991, 1995, 1999, 2003 and 2010 will be used to predict household car ownership levels in 2014. The reason to choose 2014 as year of reference is to select as many as possible years (i.e. timestamps) for the prediction, and compare the predicted outcomes based on actual data that has been made available for 2014. To predict household car ownership levels in 2014 the coefficient values in both models will be extrapolated. With regard to the MNL model using constant coefficients, the averaged coefficient values between 1987 and 2010 have been extrapolated to 2014 resulting in the same average coefficient values. Concerning the MNL model that allows for interaction effects (i.e. changing effects of coefficients), linear trend extrapolation of coefficient values in previous year (1987 – 2010) has been executed to estimate coefficient values for 2014. Both models will be estimated via the statistical program SPSS. Trend extrapolation will be done via Excel.

4

Results

This chapter will discuss the results of the modeling part of this research. The chapter will describe the descriptive statistics of the explanatory variables for household car ownership levels. In the second part of this chapter the results of the ORL model will be presented and discussed.

4.1. DESCRIPTIVE ANALYSIS

This section gives an overview of the measured quantities, expressed in percentages, of the subcategories of the eight independent variables and the dependent variable household car ownership for the years 1987, 1991, 1995, 1999, 2003, 2010 and 2014. The descriptive statistics that are presented by this section do only describe the datasets, but do not give any correlations between variables nor give indications on the influence of specific variables. Nonetheless, the distributions of the measured quantities of the determinants of interest could lead to a better understanding of the outcomes that will be produced by the ORL model. The coming subsections will give a visualization of the measured quantities per category for every explanatory variable. Important trends will be discussed and possible implications of such trends will be briefly mentioned. Other descriptive statistics, including means, standard deviations or the total number of cases per dataset can be found in Appendix II. Additionally, more detailed distributions of the measured quantities of subcategories per variable can be found in Appendix III.

I. DEVELOPMENT OF HOUSEHOLD INCOME

Levels of household income have changed quite profoundly from 1987 till 2014. Figure 7 shows that the average household income has steadily increased; varying from levels of 0-30,000 euros in 1987-1995 to 0-50,000 euros or more in 2014. Through the years the variety of different income levels has also increased. In 1987-1995 only three categories of household income existed, compared to six categories from 1995 onwards. A first explanation of the development of income levels in the datasets is related to the re-categorization of income levels. The OVG datasets have been adjusted to create one uniform categorization of household income (see also Appendix I). The household income categorization is based on the same type of categorization used by Statistics Netherlands (Statistics Netherlands, 2015b). Because the highest income levels in the OVG datasets have been formulated as "amount or more", income levels higher than 30,000 euros have not been specified. Therefore, all highest income levels have been adjusted to average income levels following a similar income increment as the second highest income category (e.g. 20,000+ euros has been adjusted to an average amount of

25,000). By following this procedure information has been lost on income levels that earned significantly more (outliers) compared to the average amount. As a result, Figure 7 shows some significant differences in household income in the years studied. Secondly, household income has gradually grown, due to an increase of household wealth in the past decades (Statistics Netherlands, 2015b; Statistics Netherlands, 2015d). This trend is visible for the years 2010 and 2014, wherein no adjustments have been made regarding the categorization of household income.

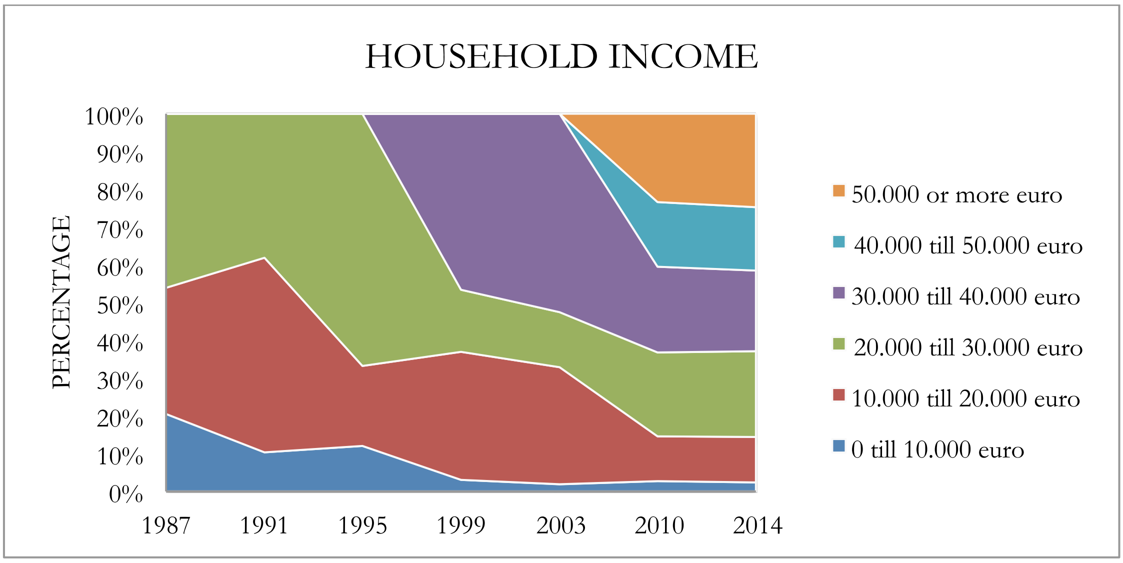


Fig. 7. Development of household income from 1987 till 2014

II. DEVELOPMENT OF HOUSEHOLD SIZE

The average household size (i.e. number of persons living in one household) has changed, to some extent, in the period 1987-2014 (Fig. 8). One can notice an average decline in the percentage of larger households (4+ persons) and an average increase in the percentage of smaller households (especially 2 person households) from the reference year 1987 till 2014. The average increase of the percentage of smaller households in the datasets is plausible, when the data is compared with household data provided by Statistics Netherlands (Statistics Netherlands, 2015c). In 1987 an average household consisted of 2.49 persons, whereas in 2014 an average household consisted of 2.18 persons, which implicates a 13% decline of the average household size in less than three decades. Moreover, the years 1999 and 2003 consistently deviate (also for household composition, gender, education and urbanization) from other years. The reason is that 1999 the research method was changed to the so-called "Neu Kontiv Design" to increase the response percentage from 40% to 70% from 1999 onwards (Kadrouch & Moritz, 1998). From 2004 the study was taken over by the Centre for Transport and Navigation (DVS), which have led to new adaptations in the sample. More surveys are now carried out on workdays and in small provinces (SWOV, 2013). The deviations in the method from 1999 till 2004 has resulted in a series in this period that is being less comparable to the other years studied (SWOV, 2013).

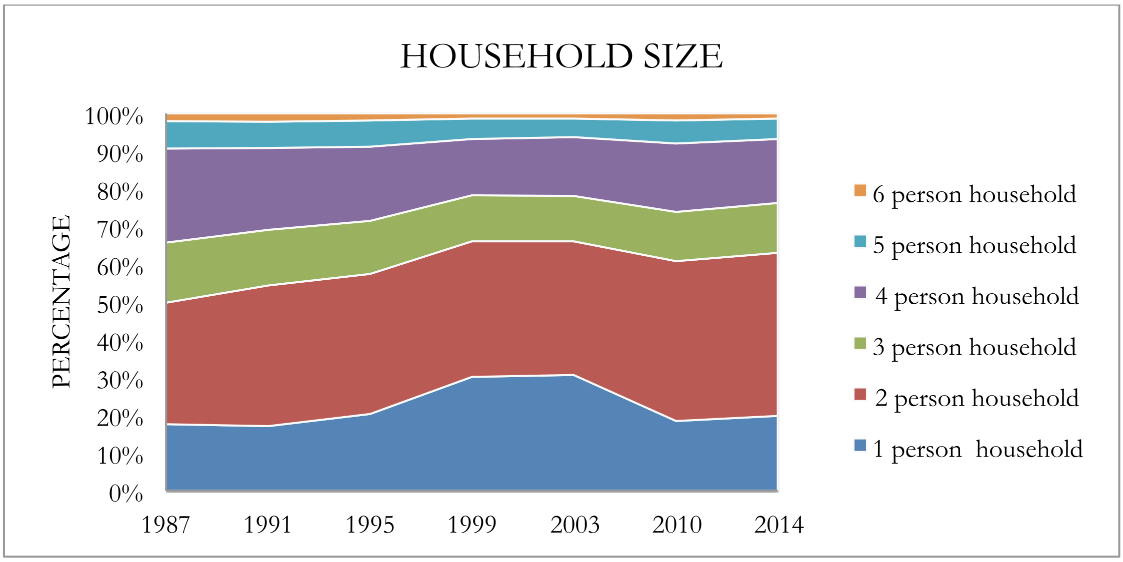


Fig. 8. Development of household size from 1987 till 2014

III. DEVELOPMENT OF HOUSEHOLD COMPOSITION

The household composition (i.e. the existence of children in one household) has changed through time (Fig. 9). One can observe a relatively continuous decrease in the percentage of households with children compared to households without children from 1987-2014. The development of an increasing percentage of households with no children is also found in population data provided by Statistics Netherlands (Statistics Netherlands, 2014a; Statistics Netherlands, 2015c). An explanation for this trend is that younger generations produce on average fewer children compared to older generations. This development also partially explains the decreasing average household size as discussed in the previous subsection.

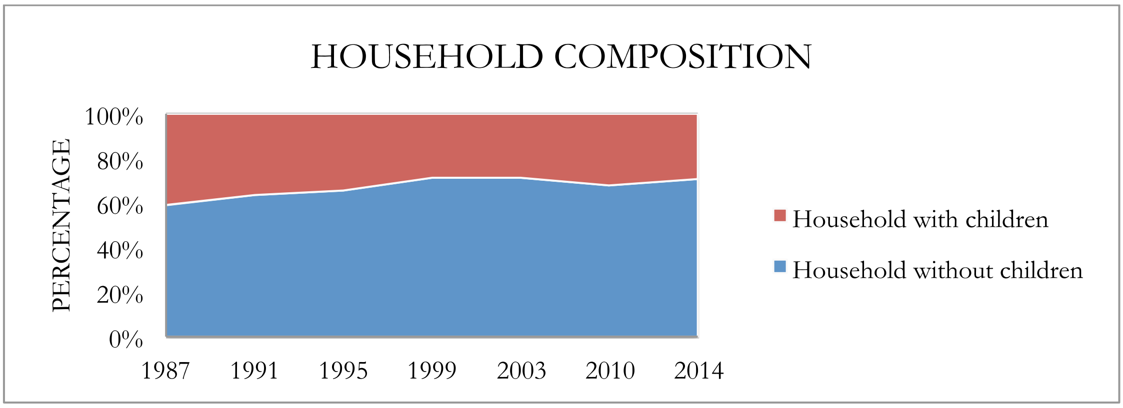


Fig. 9. Development of household size from 1987 till 2014

IV. DEVELOPMENT OF GENDER

On average the ratio male/female has been relatively stable in the period 1987-2014 (Fig. 10). Nonetheless, one can observe some interesting ratio-shifts during the years 1999 and 2003. The ratio male/female shifted from a 50/50 ratio to a 60/40 ratio. The reason for this shift is due to number of males that have been questioned in the OVG datasets 1999 and 2003. Men, assigned as "head of the family" have been predominantly asked to fill in the 1999 and 2003 OVG surveys due to a shift in sample method (Neu Kontiv Design) in these years (SWOV, 2013). This resulted in less representative sample size regarding the gender ratio compared to the Dutch population, which has not nearly changed in the past three decades (Statistics Netherlands, 2016d). To illustrate, the ratio male/female in the Dutch population in 1987 was equal to 49.43/50.57 (14.615.125 inhabitants measured on January 1, 1987) compared to a ratio of 49.52/50.48 in 2014 (16.829.289 inhabitants measured on January 1, 2014) (Statistics Netherlands, 2016d).

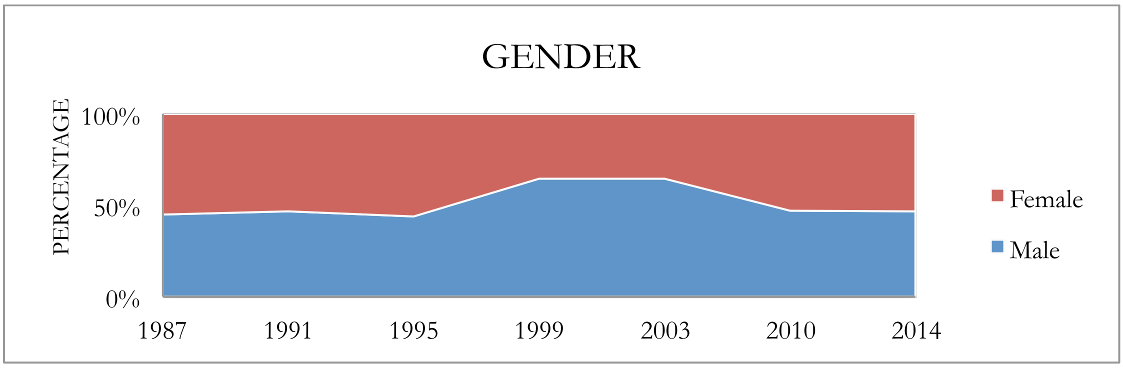


Fig. 10. Development of gender ratio from 1987 till 2014

V. DEVELOPMENT OF AGE

The development of the magnitude of age-classes (from 18-19 years to 75+ years) has solidly changed throughout the years (Fig. 11). The total share of younger age-classes (i.e. 18-39 years) has decreased, while older age-classes (50+ years) have increased in percentage. A likely explanation for this trend is the continuous ageing of the Dutch population due to higher life expectancies. According to the population pyramid provided by Statistics Netherlands the ageing process started already decades ago, which could also partially explain the trends shown by Fig. 10 (Statistics Netherlands, 2016e).

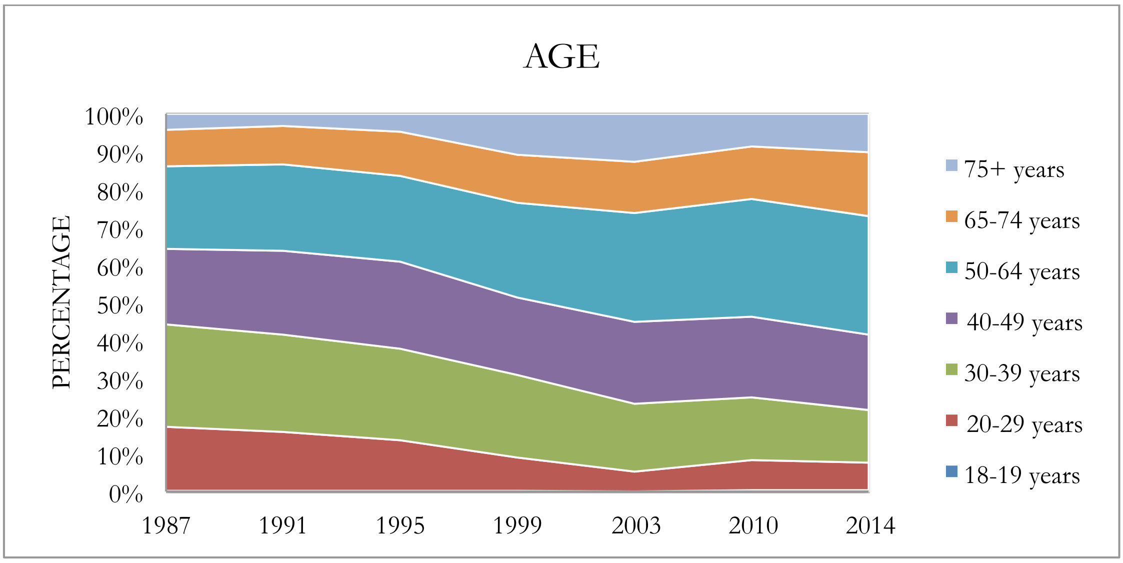


Fig. 11. Development of age-classes from 1987 till 2014

VI. DEVELOPMENT OF EDUCATION

The development of the shares of the types of educational levels shown in Fig. 12 shows two clear trends. "New" educational levels have grown compared to older educational levels. This development seems to be plausible, as the educational system has evolved through years and new educational forms have been introduced. Secondly, one can notice a more noteworthy development: the level of education (secondary and tertiary) has continuously increased from 1987 till 2014. Especially the group of higher professional education (HBO/University) has grown significantly. The increase in education levels is also found in Dutch population in the period 2001-2012, which has been made publicly by Statistics Netherlands (2013b). The average increase in the percentage of higher education levels in The Netherlands has mainly been the result of younger generations, who achieved to obtain diplomas in higher education levels.

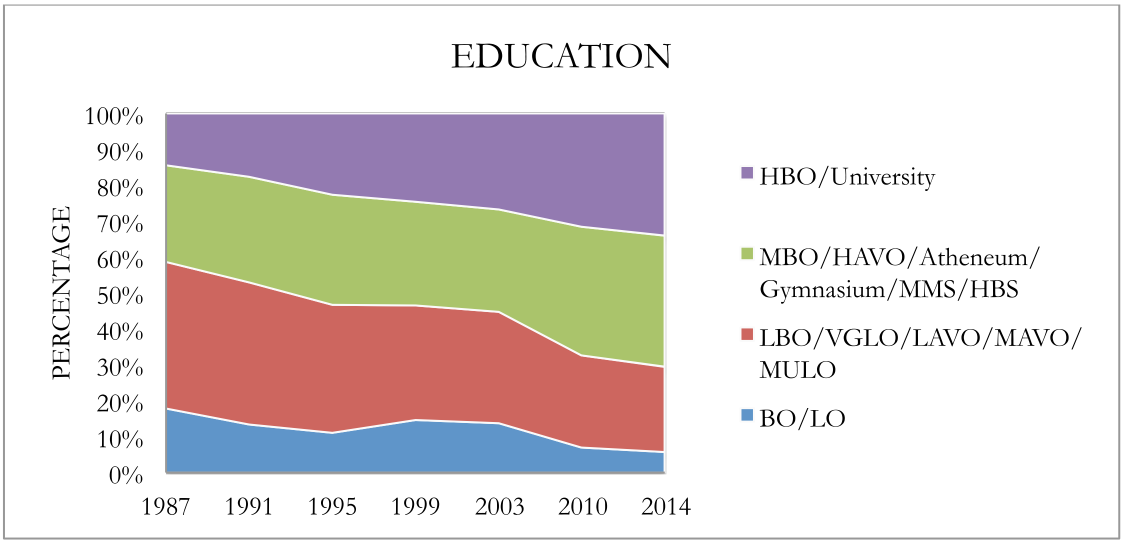


Fig. 12. Development of educational levels from 1987 till 2014

VII. DEVELOPMENT OF WORKING STATUS

The working status or employment status has been categorized into "no employment", "employment till 30 hours per week" and "employment activity of 30 hours or more per week". Figure 13 shows a decrease of the share of "fulltime" employment (>30 hours) from 1999 onwards, whilst "part time" employment (<30 hours) has increased from 1987 till 2014. The percentage of respondents falling into the "no employment" category has declined till 1999, but steadily increased from then on. Both trends of an increase in part-time employment and decrease in full-time employment have also been found in the Dutch population in the period 1992-2013, which has been made publicly by Statistics Netherlands (2014b). An explanation for this trend is that more women in younger generations started to participate in the labor market, whereas younger men started to work fewer hours per week (Smits & de Vries, 2013).

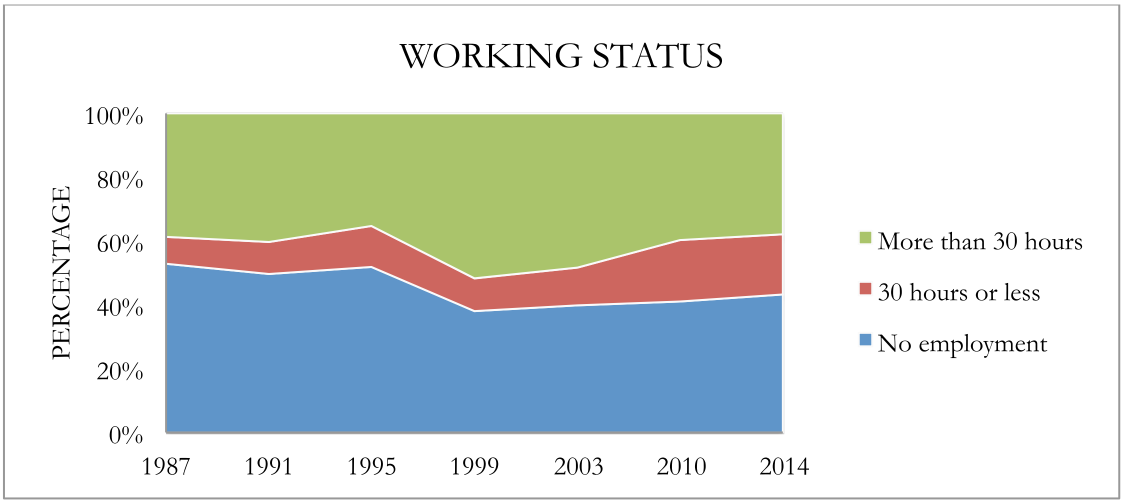


Fig. 13. Development of working status from 1987 till 2014

VIII. DEVELOPMENT OF URBANIZATION

The percentages for different levels of urbanization (wherein households reside) have changed from 1987 till 2014 (Fig. 14). Especially, the years 1987 and 1991 show differences compared to the years from 1995 and ahead. Hence, the variable urbanization for the years 1987 and 1991 has been re-categorized. Both datasets use an outdated system to measure urbanization (i.e. rural communities, urbanized rural communities and urbanized municipalities). The old categories have been re-organized into five new categories based on documents on urbanization rates provided by Statistics Netherlands (1983). Yet, the results of 1987 and 1991 give a distorted picture of the actual degree of urbanization during those years. For the years from 1995 onwards, one can observe that the share of households living in different urbanized regions has been relatively constant between 1995 and 2014 (highly to very-high urbanized regions accounted for 40% in this period). The trend of gradual-continuous urbanization in The Netherlands is therefore less visible in the datasets (PBL, 2013). An explanation can be found in the relatively homogenously distribution of the surveys in different urban areas to include relatively small subgroups within the population. For example, from 2004 more surveys are now carried out in small provinces, and consequently, less in larger (more urbanized) provinces (SWOV, 2013).

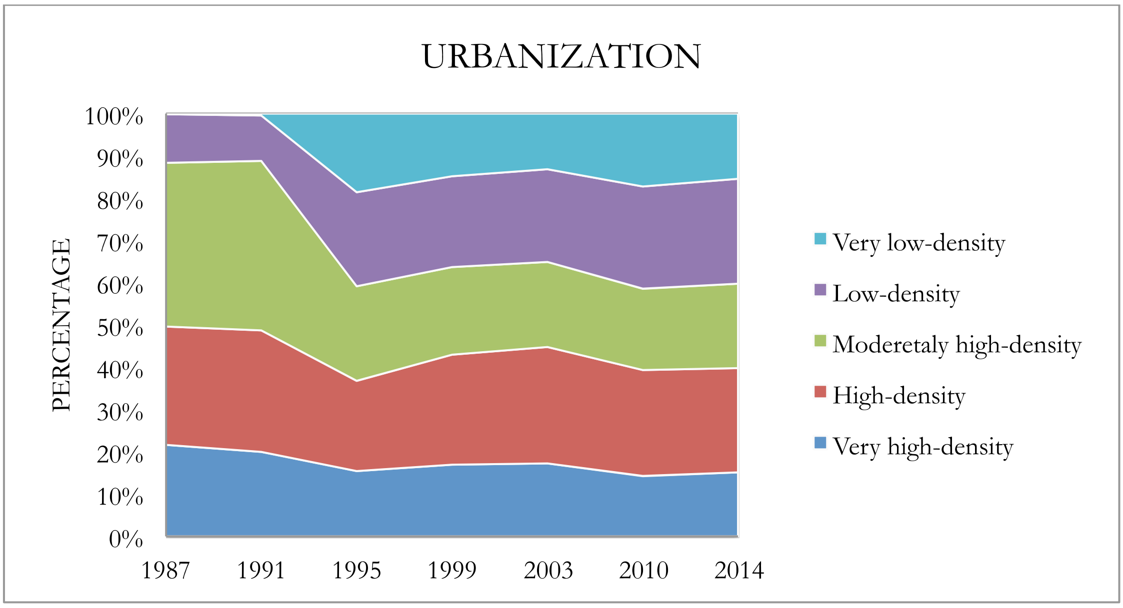


Fig. 14. Development of urbanization levels from 1987 till 2014

IX. DEVELOPMENT OF THE NUMBER OF PRIVATE CARS

Results on the measured frequencies in the datasets of 1987-2014 on the number of private cars owned by households show a percentage-growth of households obtaining two or more private cars (Fig. 15). Especially, households obtaining two private cars have grown significantly. The increase in the percentage of households obtained two private cars is also observed in the total Dutch population in the period 2001-2013 (RIVM, 2013). As shown by Figure 14 the total number of private cars per household has grown significantly in 2003, but shows stabilization from 2010 onwards. As discussed in the introduction of this thesis, the stabilization of the growth in car ownership has also been observed in the total Dutch car fleet (Statistics Netherlands, 2016a; Statistics Netherlands, 2016b).

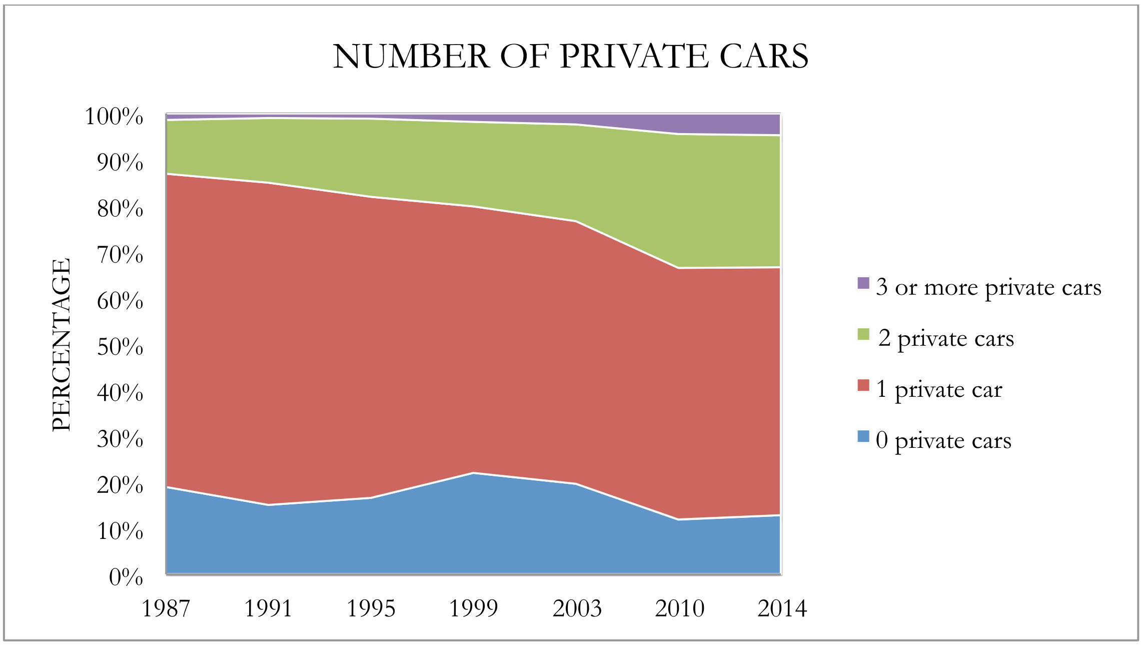


Fig. 15. Development of the number of private cars levels from 1987 till 2014

4.2. ORDERED LOGISTIC REGRESSION ANALYSIS

Following a descriptive analysis of the datasets, an ORL model has been estimated to investigate the changing effects of determinants on household car ownership. Figure 16 presents the results of the ORL estimated logit coefficients of the explanatory variables (income, size, etc.) on household car ownership. In general, the model performed well, as indicated by the relatively high pseudo R-squares (Appendix IV). For example, the Nagelkerke pseudo R-square of 0.423 is considered to be a relatively high. The model fitting information of the estimated ORL model also shows there is a statistical improvement of the model by including the explanatory variables (household income, size, etc.) in the model. Furthermore, the variance in the model outcomes is explained proportionately by the explanatory variables.

The majority of estimated parameters were statistically significant (Table 4). The logit coefficients of the explanatory variables indicate that the effects of the explanatory variables on household car ownership have changed between 1987 and 2014. To provide a means for comparing the effects of the explanatory variables, which are measured in different metrics, relative effects of the logit coefficients have been estimated (Fig. 17). Relative importances or relative (maximum) effects are obtained by first multiplying the logit coefficient value with the minimum and maximum category of the explanatory variable. For example, the logit coefficient value of 0.9 in 1987 for household income has been multiplied with 1 and 6 referring to the lowest category of household income and the maximum category of household income. Subsequently, the difference of both products is then calculated: (1*0.9) - (6*0.9) = 4.5. This calculation has been applied to all explanatory variables. The relative importance of household income in 1987 is then calculated by dividing 4.5 with the total sum of all product differences of the explanatory variables (Vermunt & Magidson, 2005). For household income in 1987 this would result in: 4.5/(total sum) = 4.5/12.2 = 0.38. The relative importance of household income in 1987 related to the other variables is therefore 38% (see Fig. 17).

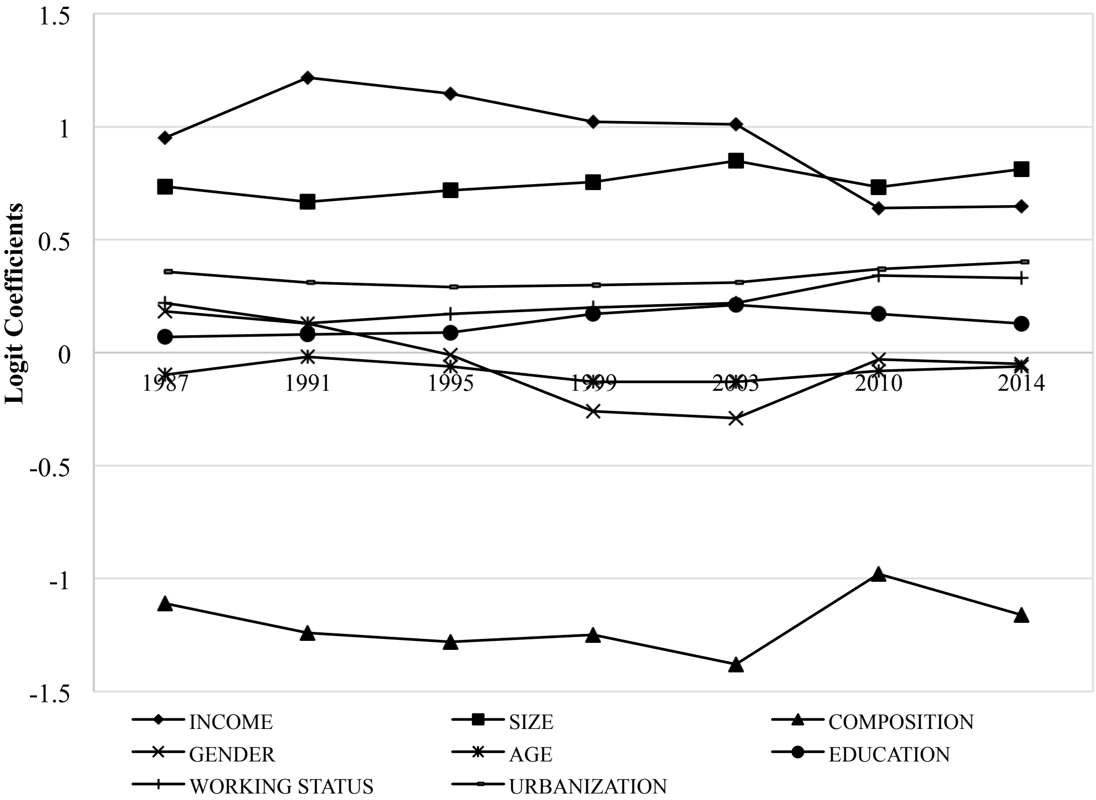


Fig. 16. Ordered Logistic Regression Model – Estimation results (coefficients) for explanatory variables

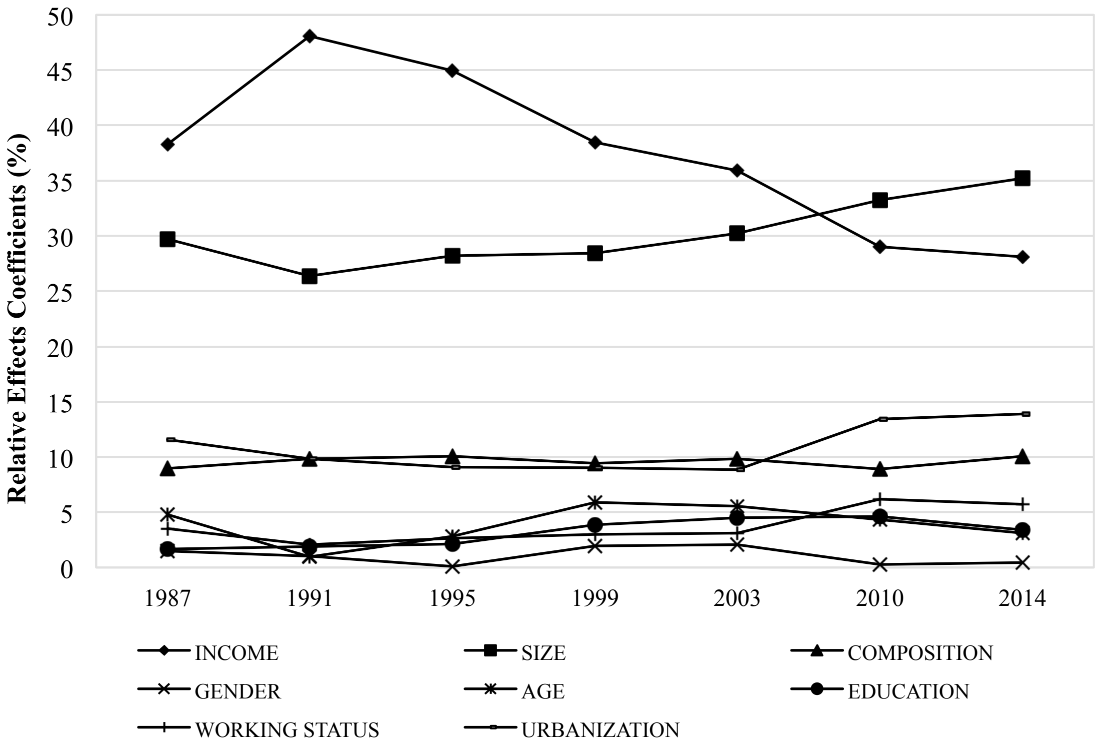


Fig. 17. Ordered Logistic Regression Model – Relative importance of estimated logit coefficients

Table 4. Ordered Logistic Regression Model Estimates

						95% Confidence Interval	
	β*	Std. Error	Wald	df	Sig.	Lower	Upper
Treshold values							
[Car = 0]	2.152	0.051	1775.056	1	0.000	2.052	2.252
[Car = 1]	6.215	0.054	13187.917	1	0.000	6.109	6.321
[Car = 2]	9.189	0.058	25481.145	1	0.000	9.076	9.302
Household income							
Income 1987 (ref*)	0.950	0.039	580.216	1	0.000	0.873	1.027
Income* 1991	0.266	0.063	17.680	1	0.000	0.142	0.390
Income* 1995	0.196	0.043	20.929	1	0.000	0.112	0.279
Income* 1999	0.071	0.042	2.848	1	0.091	-0.011	0.153
Income* 2003	0.059	0.045	1.725	1	0.189	-0.029	0.146
Income* 2010	-0.310	0.041	57.159	1	0.000	-0.390	-0.230
Income* 2014	-0.303	0.041	54.571	1	0.000	-0.384	-0.223
Household size							
Size 1987 (ref*)	0.736	0.034	467.205	1	0.000	0.669	0.803
Size* 1991	-0.069	0.050	1.926	1	0.165	-0.167	0.029
Size* 1995	-0.018	0.036	0.247	1	0.619	-0.090	0.053
Size* 1999	0.019	0.037	0.269	1	0.604	-0.054	0.093
Size* 2003	0.113	0.041	7.738	1	0.005	0.033	0.193
Size* 2010	-0.003	0.039	0.007	1	0.935	-0.079	0.073
Size* 2014	0.075	0.039	3.690	1	0.055	-0.002	0.151
Household composition							
Composition 1987 (ref*)	-1.108	0.084	175.001	1	0.000	-1.272	-0.944
Composition* 1991	-0.128	0.123	1.081	1	0.298	-0.369	0.113
Composition* 1995	-0.171	0.090	3.625	1	0.057	-0.348	0.005
Composition* 1999	-0.138	0.092	2.257	1	0.133	-0.318	0.042
Composition* 2003	-0.275	0.100	7.510	1	0.006	-0.471	-0.078
Composition* 2010	0.130	0.095	1.865	1	0.172	-0.057	0.317
Composition* 2014	-0.053	0.095	0.312	1	0.577	-0.240	0.134
Gender							
Gender 1987 (ref*)	0.183	0.050	13.471	1	0.000	0.085	0.281
Gender* 1991	-0.055	0.072	0.583	1	0.445	-0.197	0.087
Gender* 1995	-0.195	0.053	13.575	1	0.000	-0.298	-0.091
Gender* 1999	-0.446	0.054	67.984	1	0.000	-0.552	-0.340
Gender* 2003	-0.475	0.059	65.525	1	0.000	-0.590	-0.360
Gender* 2010	-0.213	0.055	14.778	1	0.000	-0.322	-0.104
Gender* 2014	-0.233	0.056	17.544	1	0.000	-0.341	-0.124
Age							
Age 1987 (ref*)	-0.099	0.017	34.566	1	0.000	-0.131	-0.066
Age* 1991	0.076	0.024	9.553	1	0.002	0.028	0.123
Age* 1995	0.042	0.018	5.708	1	0.017	0.008	0.076
Age* 1999	-0.032	0.018	3.321	1	0.068	-0.067	0.002
Age* 2003	-0.027	0.019	2.018	1	0.155	-0.065	0.010
Age* 2010	0.024	0.018	1.748	1	0.186	-0.012	0.061
Age* 2014	0.043	0.019	5.509	1	0.019	0.007	0.080
Educational level							
Education 1987 (ref*)	0.069	0.029	5.586	1	0.018	0.012	0.126
Education* 1991	0.010	0.043	0.060	1	0.807	-0.073	0.094
Education* 1995	0.022	0.031	0.511	1	0.475	-0.039	0.083
Education* 1999	0.099	0.031	9.979	1	0.002	0.038	0.161
Education* 2003	0.141	0.034	17.636	1	0.000	0.075	0.207
Education* 2010	0.096	0.033	8.404	1	0.004	0.031	0.160
Education* 2014	0.058	0.033	3.072	1	0.080	-0.007	0.123

Working status							
Working status 1987 (ref*)	0.219	0.030	51.566	1	0.000	0.159	0.279
Working status* 1991	-0.085	0.045	3.624	1	0.057	-0.172	0.003
Working status* 1995	-0.051	0.033	2.485	1	0.115	-0.115	0.013
Working status*1999	-0.015	0.033	0.212	1	0.645	-0.081	0.050
Working status* 2003	0.004	0.037	0.011	1	0.915	-0.068	0.076
Working status* 2010	0.123	0.035	12.455	1	0.000	0.055	0.191
Working status* 2014	0.113	0.035	10.461	1	0.001	0.045	0.182
Urbanization							
Urbanization 1987 (ref*)	0.359	0.028	168.349	1	0.000	0.305	0.413
Urbanization* 1991	-0.050	0.041	1.530	1	0.216	-0.130	0.029
Urbanization* 1995	-0.065	0.029	5.131	1	0.023	-0.121	-0.009
Urbanization* 1999	-0.056	0.029	3.776	1	0.052	-0.112	0.000
Urbanization* 2003	-0.045	0.030	2.244	1	0.134	-0.104	0.014
Urbanization* 2010	0.006	0.029	0.037	1	0.848	-0.052	0.063
Urbanization* 2014	0.039	0.030	0.030	1	0.188	-0.019	0.097

* The coefficients of 1991-2014 can be calculated by summing the main effect (β_{1987}) with the interaction effect of the year of reference. To illustrate, the coefficient value of household income in 1991 can be calculated as follows: *Income 1987 + Income 1991 = 0.950 + 0.266 = 1.216.*

4.3. ORDERED LOGISTIC REGRESSION RESULTS

This section will discuss the model outcomes of the estimated ORL model. The goal is to investigate the changing effects of car ownership determinants on different household car ownership levels (zero, one, two, three or more private cars). The next subsections will discuss the most important trends derived from the logit coefficients that are estimated by the ORL model. The most important conclusions of the ORL model outcomes will be discussed in Section 4.4. Section 4.5 will combine the results of the descriptive analysis and ordered logistic regression analysis.

I. HOUSEHOLD INCOME

Household income shows significant changes regarding its effect on car ownership and car ownership levels from 1987 till 2014 (Fig. 16). Considering the ORL model outcomes in Figure 16, the estimated coefficients are positive, but continuously declining from 1991 onwards. Therefore, a unit-increase of household income (i.e. 10K) in 1991 would lead to a 1.22 logit increase of car ownership, whereas a unit-increase of household income in 2014 would only lead to a 0.65 logit increase of car ownership. As a result, the logit value has on average decreased with almost 50% from 1991 till 2014. Only the logit coefficient of 1987 seems to be a relative "outlier" compared to the other logit coefficients. An explanation for the logit coefficient value in 1987 cannot be directly derived from the results. Furthermore, Figure 17 shows that the relative effects are high for household income compared to the other explanatory variables, but have decreased from 1991 onwards. For example, the relative importance of the logit coefficient in 1991 accounted for almost 50% of the total magnitude of effects compared to 30% in 2014.

II. HOUSEHOLD SIZE

In contrast to household income, the effects of household size on car ownership levels have slightly, but gradually increased from 1987 till 2014 (Fig. 16). Considering the ORL model outcomes in Figure 16, a unit-increase of household size (i.e. 1 person) in 1987 would lead to a 0.7 logit increase in car ownership level, whereas a unit-increase of household size in 2014 would lead to a 0.8 logit increase in car ownership level. As a result, the logit value increased on average with 10% from 1987 till 2014. The relative importance of the logit coefficients is significant compared to other explanatory variables and gradually increased from 30% in 1987 to 35% in 2014 (Figure 17).

III. HOUSEHOLD COMPOSITION

Household composition shows a gradual increase of its negative logit coefficients from 1987 till 2003 (Fig. 16). The ORL model outcome shows that a unit-increase in household composition in 1987 (i.e. the transition of a household of having children) decreased the logit value of car ownership with 1.1 compared to a decrease of 1.4 in 2003. The logit coefficient in 2010 and 2014 (-1.0 and -1.1) are smaller compared to other years, which could be the result of the different surveys that have been used (OVG versus OViN). Compared to household income and size the relative importance of the logit coefficients is considerably lower (around 10% for all years).

IV. GENDER

The logit coefficient of the gender of a person on car ownership levels has changed unpredictably from 1987 till 2014 (Fig. 16). Compared to household income, size and composition the magnitude of the change of the effect is moderate, but has developed much less gradually in the period 1987-2014. To illustrate, a shift from male to female, would result in a logit value increase of 0.2 in 1987, compared to a logit value decrease of -0.3 in 2003 and -0.05 in 2014. Therefore, being either male or female in 2014 has been on average of less importance for the level of car ownership in households compared to 1987. Nevertheless, the relative effect of the logit coefficients of gender compared to household income and size is extremely low (between 0% and 2%).

V. AGE

The magnitude of the logit coefficients of age on car ownership levels has hardly increased from 1991 till 2014, and is significantly small compared to the size of the effects of household income, size and composition (Fig. 16). The relative effect compared to other explanatory variables is also small; a relative effect of 5% in 1987 and 3.4% in 2014 (Fig. 17). An explanation for the small relative effects of age is because only linear effects have been calculated instead of quadratic effects. However, the likelihood to possess a car is relatively small for young ages (e.g. <25), but increases with age till a certain extent (e.g. 40s-60s) and decreases for older ages (e.g. 65+). Linear effects do not take into account a parabolic course, whereas quadratic effects do.

VI. EDUCATION, WORKING STATUS & URBANIZATION

The change in the magnitude of the logit coefficients of education, working status and urbanization are compared to all other explanatory variables the smallest (Fig. 16). The size of the increase or decrease of the logit coefficients is negligible. All logit coefficients of the three variables are positive. To illustrate, an increase of 1 unit in education (i.e. one higher education level) would result in a 0.1 increase of the logit value of car ownership in 2014. The relative effects of education and working status are compared to other explanatory small, and vary between 2% and 6% in the years 1987-2014 (Fig 17.). The relative effect of urbanization is, however, quite proportionate with an average relative importance of 11% from 1987 till 2014.

4.4. CONCLUSIONS LOGISTIC REGRESSION ANALYSIS

Based on the findings in the previous sections some important conclusions can be drawn. Statistical diagnosis shows that including the explanatory variables improves the ORL model outcomes significantly. In general, the model performed well, as specified by the relatively high R-squares (Appendix IV). Furthermore, the majority of estimated parameters were statistically significant (Table 4).

The ORL logit coefficients of the explanatory variables indicate that the effects of the explanatory variables on household car ownership **have changed between 1987 and 2014.** Especially, the relative influence of **household income** and **household size** are substantial, contributing to more than 60% of the total influence on household car ownership in all years studied. Whereas the influence of household income on car ownership decreased over time (from 38% in 1987 to 28% in 2014), the influence of household size has increased (from 29% in 1987 to 35% in 2014). With regard to household income, an explanation for this trend might be that the elasticity of rising income has declined over years compared to the elasticity of falling income, which is also referred to as hysteresis (Dargay, 2001). Both household income and household size show a positive relationship with the number of cars owned by a household.

The same positive relationship is observed for education, working status and urbanization with household car ownership. The relative influence of these variables on household car ownership has gradually grown with an average total of 2% between 1987 and 2014. With regard to household composition a relatively strong negative relationship with car ownership is observed, which has remained relatively stable between 1987 and 2014 (around 10%). Also age and gender show predominantly a negative relationship with the number of cars owned by a household. Their influence on car ownership has decreased over 1% between 1987 and 2014, and was relatively small in 2014.

4.5. QUALITATIVE ANALYSIS OF MODELING RESULTS

This section combines the results from the descriptive analysis with the ORL model outcomes. Table 5 shows the average increase or decrease of the car ownership variables over time (column: 'trend in variable'). The third column ('influence of parameter') shows the positive or negative influence of the parameter on household car ownership. The fourth column ('parameter trend') shows the change in influence of the parameter on household car ownership as estimated by the ORL model. The combined results of the columns give an indication of the increase or decrease of household car ownership levels over time. For example, the average household income has increased over time, and its influence is positive on household car ownership levels. However, the influence of the parameter on household car ownership has substantially decreased over time. Therefore, based on qualitative analysis we cannot conclude whether the average number of cars owned by a household would have increased or decreased over time on the basis of household income. The same applies to household size, gender, age and working status.

Parameter	Trend in variable	Influence of parameter	Parameter trend	Household car levels
Household income	+	+	–	*x*
Household size	–	+	+	*x*
Household composition	–	–	0	**+**
Gender	0	+/–	–	*x*
Age	+	–	0	*x*
Education	+	+	+	**+**
Working status	–	+	+	*x*
Urbanization	0	+	+	**+**
Number of private cars	+	n.a.	n.a.	**+**

Table 5. Qualitative analysis of descriptive statistics and ORL model results

According to the outcomes of Table 5, we expect that household composition, education and urbanization have increased the average number of household cars. Comparing the outcomes with the actual development of the number of cars (Fig. 15), the results seem to reflect the growth of the average number of household cars in The Netherlands. However, the stabilization in the growth of cars from 2010 (Fig. 15) cannot be explained based on the limited outcomes of the qualitative analysis.

A comparison of the findings in Table 5 with previous empirical research shows that for household composition we would have expected a positive relationship with car ownership levels (i.e. the presence of children would increase the level) (Oakil et al., 2016a). Based on data provided by Statistics Netherlands, the trend in household composition is negative, resulting in a less average number of children per household (Statistics Netherlands, 2014a; Statistics Netherlands, 2015c). No empirical results have been found regarding the parameter trends of household composition in the time horizon as been used for this study. Assuming the influence has not changed, as found in this study, we would have expected a *decrease* of household car ownership levels between 1987-2014.

Empirical research on urbanization shows that the urbanization level has slowly increased in The Netherlands in the period 1987-2014 (PBL, 2013). The trend of urbanization should therefore have a positive sign ('trend in variable'). Recent empirical research on the effect of urbanization on car levels shows there is a positive relationship, which is also supported by the findings of this study ('influence of parameter') (Oakil et al., 2016a). No empirical studies in The Netherlands have been found regarding the change of influence of urbanization on household car ownership levels ('trend in parameter'). Assuming that the influence has increased we would have expected a *higher increase* of household car ownership levels based on urbanization between 1987-2014.

With regard to education, educational levels have increased in The Netherlands between 1987-2014 (Statistics Netherlands, 2013b). Empirical studies have found evidence that the average number of cars in a household increases with higher educational levels (Flamm, 2009; Eakins, 2013). However, no evidence has been found on the positive parameter trend of education as suggested by the findings in this research. If we assume this trend be true, than we would expect that educational levels have increased the average number of cars of households between 1987-2014 in The Netherlands.

Based on the outcomes we should be critical regarding the estimated outcomes of the ORL model. Firstly, gender, age, education and working status are so-called proxies, measured on an *individual* level (i.e. based on the information on the households' reference person), whereas car ownership levels are measured on *household* level. Moreover, not all variables are most effectively measured by linear analysis. For example, the influence of the variable age on household car ownership is better described by a quadratic function, rather than a linear function. Finally, the influence of some parameters (e.g. negative influence of having children on car ownership) and trends in variables (e.g. relatively stable urbanization development) would have expected to be different in relation to aggregate data and findings from other studies.

5

Model Validation

In this chapter the ORL model outcomes will be validated. Firstly, the ORL model will be internally validated. For this purpose, a holdout validation sample will be used. The coefficients estimated by the model will be used to predict the response of the respondents in the holdout sample. After the ORL model outcomes have been internally validated, the accuracy of the model will be compared to two MNL models (using constant and changing car ownership coefficients) to prove that ORL model performs at the same level of accuracy as MNL models. The key reason to use the ORL model for this study is that the interpretation of model outcomes is more effortless compared to MNL model outcomes.

5.1. VALIDATION OF ORDERED LOGISTIC REGRESSION MODEL

It is of importance to validate the ORL modeling results in Chapter 4 to support conclusions that are based on such outcomes. Model validation can be performed on three levels: diagnostic, internal and external validation (Steyerberg & Harrel, 2002).

Diagnostic validation comprises the performance of the model outcomes on a sample size and gives indication about the validity of the estimated models (Trochim, 2006). Chapter 4 has shown that the ORL model is statistically a valid model. The estimated model showed that the inclusion of explanatory variables for car ownership would significantly improve the models compared to intercept-only models. Furthermore, a considerable part of the variance in model outcomes can be explained by the explanatory variables.

Internal validation is the type of validation that will be used in this chapter. Internal validation includes assessing whether the current state of the ORL model validly predict responses in the sample population (Steyerberg & Harrell, 2002). For this purpose a holdout sample will be used (Wang, 2005). The holdout validation sample has been constructed by randomly subtracting 20% of the total number of unique cases for every OVG and OViN dataset that has been used for this research (Appendix I). The coefficients estimated by the ORL models will be used to predict the response of the respondents in the holdout sample.

External validation comprises the evaluation whether the coefficients of the constructed ORL models can validly predict the level of car ownership of household that did not participated in the OViN and OVG surveys (Steyerberg & Harrell, 2002; Bourennane et al., 2014). In order to test the validity of such predictions responses of new respondents need to be gathered. This type of validation will not be executed due to the time constraints of this research.

As an extension to the three types of validation, this study will compare the prediction accuracy of the ORL model compared to two types of MNL models. MNL models are expected to predict outcomes better than ORL models, due to their greater flexibility (Potoglou & Susilo, 2008). This chapter will, however, show that the overall model accuracy of the ORL model, despite it much less greater flexibility, is at the same level compared to the model accuracy of the MNL models. Firstly, an MNL model will be estimated that assumes constant coefficients over time. Such a model can be obtained by only calculating the averaged value of the specific coefficients over time. The second MNL model will allow for changes in coefficients over time, which can be obtained by introducing interaction effects, whereby k-1 dummy variables are created (where k = number of years). Both models will be tested on the same validation sample as the ORL model (20% of all unique cases).

5.2. INTERNAL VALIDATION ORDERED LOGISTIC REGRESSION

The internal validity of the estimated ORL model has been tested via the following mathematical procedure. Firstly, the log (odds) for all 40,852 validation cases have been calculated by using formula 5.1 (Armstrong & Sloan, 1989):

$$logit\ (\pi_j) = \ln\left(\frac{\pi_j}{\pi_k}\right) = \alpha_j + \beta_{j1} * x_1 + \cdots + \beta_{jm} * x_m \tag{5.1}$$

$$with \quad (j = 1, \ldots, k-1)$$

In formula 5.1 *k* represents the observed response categories, being zero, one, two or three or more private cars. The α_j represents the threshold value of the associated response category *k*, whereby β_{jm} represents the coefficient value of the variable *m* for the *k* response category. The value of x_m represents here the value of the variable *m* for a specific household. For example, x_2 = household size = 3 means that the specific household consists of 3 persons. The ordinal probability of being in each of the response categories is described by π_j. The logit value of π_j is defined by $\ln\left(\frac{\pi_j}{\pi_k}\right)$, which expresses the natural logarithm of the ratio between the probability an event will occur (π_j) vis-à-vis the probability an event will not occur (π_k) (Mathew, 2015). If we specify formula 5.1 for the case of household car ownership, formula 5.2 can be derived:

$$logit\ (\pi_j) = \ln\left(\frac{\pi_j}{\pi_k}\right) = \tag{5.2}$$

$$\alpha_j + \beta_{j1} * household\ income + \beta_{j2} * household\ size + \beta_{j3} * household\ composition + \beta_{j4} * gender + \beta_{j5} * age + \beta_{j6} * \ education + \beta_{j7} * \ working\ status + \beta_{j8} * \ urbanzation\ level$$

$$with\ (j = 1, \ldots, k-1)$$

Finally, the predicted probability for every single category of household private car ownership (i.e. zero, one, two, three or more private cars) has been calculated for all the validation cases by formula 5.3:

$$\pi_j = \frac{1}{1 + e^{-(a_j + \beta_{j1} * x_1 + \cdots + \beta_{jm} * x_m)}} \quad (5.3)$$

$$with \; (j = 1, \ldots, k - 1)$$

After calculating the predicted probabilities for the car ownership levels per single household, the predicted level of car ownership has been obtained by selecting the car ownership level that was associated with the highest predicted probability. To illustrate, if the probability of having two private cars would be the highest for a specific household, the household would be associated with having two private cars. Finally, the frequencies of the predicted and observed levels of household car ownership have been determined. The frequencies have been used to estimate the model accuracy and execute a Chi-squared test. The Chi-squared test is used is used to determine whether there is a significant difference between the expected frequencies and the observed frequencies in the response categories. The results on model accuracy and Chi-squared test will be discussed in Section 5.3. Noteworthy to state is that all coefficients have been included to calculate the log (odds) values, including not-significant coefficients (see also Table 4). The inclusion of all coefficients led to substantial better model predictions compared to a model that only includes coefficients with a significance level smaller than 0.05.

5.3. INTERNAL ORL VALIDATION RESULTS

The value of the chi-squared test of 15761 given by Table 6 has a significance level of 0.000 (rounded number). Therefore the null-hypothesis of independence between the observed and predicted values can be rejected. The assumption of independent distributed data is therefore not valid, and the predicted and observed data distributions are dependent. According to the cross-tabulation provided in Table 7, 65% of all the cases has been predicted correctly by the ORL model, which means that 65% of the sample data can be explained by the explanatory variables that are included. Based on the results of the chi-squared test and cross-tabulation the predictions made by the ORL model for the number of private cars per household are therefore regarded as internally valid.

	Value	df	Asymp. Sig. (2-sided)
Chi-squared test	**15761.2753**	3	**0.000**
Number of valid cases	40582		

Table 6. Validation of ORL Model – Chi-squared test results

		Observed cases				Precision
		0	1	2	3+	
Predicted cases	0	2951	1427	15	3	67%
	1	3905	20971	5998	374	67%
	2	56	1831	2474	548	50%
	3+	0	2	12	15	52%
Sensitivity		43%	87%	29%	58%	**65%**

Table 7. Validation of ORL Model – Model precision, sensitivity and accuracy

5.4. INTERNAL VALIDATION MULTINOMIAL LOGISTIC REGRESSION

This section validates the MNL model wherein changing coefficients of car ownership determinants are included and a MNL model that uses constant (i.e. not changeable) coefficients of determinants for household car ownership. The mathematical procedure how predicted values are calculated and determined using a MNL model is slightly different compared to using an ORL model.

Firstly, the same logit functions (5.1 and 5.2) are used to calculate the logit values of the multinomial probabilities (i.e. log odds) for every response category per single household. To calculate the probabilities of belonging to a response category (zero, one, two, three or more private cars) the odds ratio needs to be calculated (π_j). To calculate the odds ratio, a different equation (5.4) is used (Stock & Watson, 2007):

$$\pi_j = \frac{e^{(a_j + \beta_{j1} * x_1 + \cdots + \beta_{jm} * x_m)}}{1 + e^{(a_1 + \beta_{11} * x_1 + \cdots + \beta_{1m} * x_m)} + e^{(a_k + \beta_{k1} * x_1 + \cdots + \beta_{km} * x_m)}} \qquad (5.4)$$

$with\ (j = 1, \ldots, k - 1)$

After calculating the predicted probabilities for the car ownership levels per single household, the predicted level of car ownership has been obtained by selecting the car ownership level that was associated with the highest predicted probability. The frequencies of the predicated and observed levels of household car ownership have been determined to estimate the overall model accuracy.

5.5. INTERNAL MNL VALIDATION RESULTS

The results show that the overall model accuracy of a MNL model that accounts for changing car ownership coefficients over time is relatively better at predicting car ownership levels compared to a model that uses constant coefficients over time. Table 8 and 9 show the predicted outcomes of the MNL models.

The MNL model that uses constant coefficient values for car ownership determinants has a total accuracy of 62%, meaning that 62% of the 40,852 cases in the holdout sample has been predicted correctly (Table 8). Nevertheless, the MNL model that does allow for changing values of car ownership variable coefficients predicts the outcomes over 4% more accurately resulting in an overall model accuracy of 66%. Moreover, when we compare the prediction of both models for all separate years between 1987 and 2014, one notices that the MNL model, allowing for alternative effects of car ownership determinants, approximates the actual data more accurately compared to a model that only uses main effects (Fig. 18). Especially, household car ownership levels in 1999, 2003, 2010 and 2014 are significantly better explained, resulting in average increase of 10% in model accuracy.

		Observed cases				Precision
		0	1	2	3+	
Predicted cases	0	1762	683	4	0	72%
	1	4949	19503	4569	251	67%
	2	195	4022	3867	621	44%
	3+	6	23	59	68	44%
Sensitivity		25%	80%	45%	7%	**62%**

Table 8. Data fit MNL model based on main effects – 1987-2014 – Model precision, sensitivity and accuracy

		Observed cases				Precision
		0	1	2	3+	
Predicted cases	0	3125	1503	14	2	67%
	1	3737	21185	6212	437	67%
	2	45	1522	2225	444	53%
	3+	5	21	48	57	44%
Sensitivity		45%	87%	26%	6%	**66%**

Table 9. Data fit MNL model including interaction effects – 1987-2014 - Model precision, sensitivity and accuracy

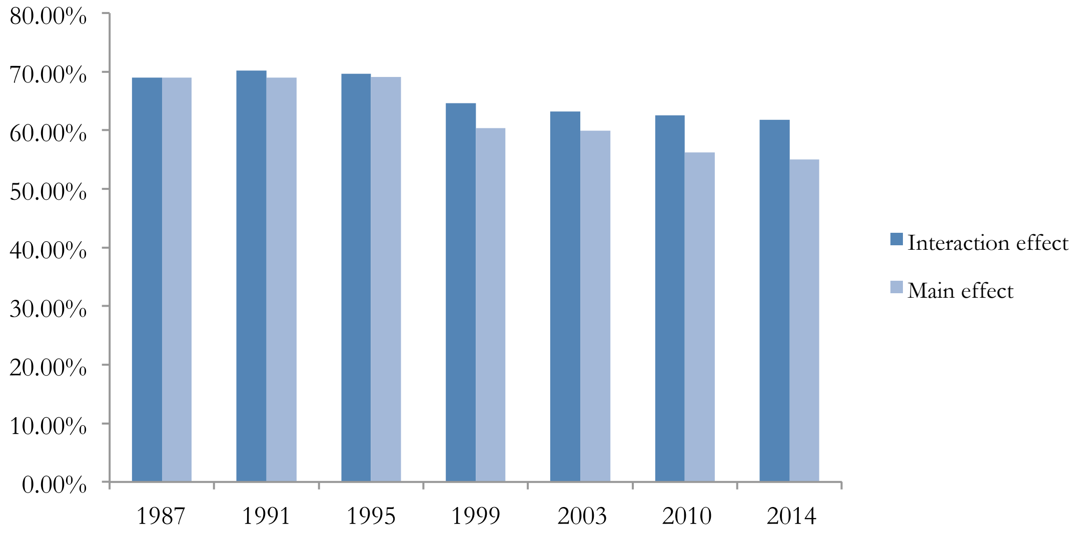


Fig. 18. Comparison MNL model prediction outcomes – interaction versus main effects model

5.6. CONCLUSION ON LOGISTIC REGRESSION VALIDATION

The overall accuracy of the ORL model that includes interaction effects does not differ much from the accuracy of both MNL models. In fact, the overall model accuracy of the ORL is even higher compared to the model accuracy of the MNL model that incorporated constant car ownership coefficients (compare 65% to 62%). Furthermore, the MNL model, which uses changing coefficient values over time, is not considerably more accurate (only 1%) compared to the ORL model. Despite the much less greater flexibility of the ORL model, car ownership household data between 1987 and 2014 has been explained on the same level of accuracy compared to both types of MNL models. In contrast to MNL models, the ORL model is much more suitable to interpret the model outcomes, since it relies on a single latent variable. In conclusion, the ORL model is considered to be the most appropriate model for this study to explore the time-dependent influence of household car ownership determinants.

6

Predictions & Policy implications

This chapter explores the prediction capabilities of the multinomial logistic regression (MNL) model. MNL models allow for alternative effects of determinants on car ownership levels, which make their application more suitable to *predict* household car ownership levels compared to ORL models. In this chapter, the MNL model that allows for changing influence of car ownership determinants on the level of household car ownership will be compared to a MNL model that uses a constant (i.e. not changeable) influence of car ownership determinants on the level of household car ownership. The aim of this analysis is to reveal whether both models differentiate in terms of prediction power for household car ownership levels. The outcomes of this analysis will be compared to the DYNAMO model, which is the main car ownership model used by Dutch ministries and political parties. Especially, in case the influence of factors is considered to be constant in the DYNAMO model, we should question ourselves what *policy implications* this might bring for the future.

6.1. PREDICTING CAR OWNERSHIP LEVELS IN 2014

The previous sections have estimated and compared the accuracy of two types of MNL models based on a validation holdout sample. Strictly speaking, such a validation procedure is not merely the same as a prediction procedure, since no new coefficients are estimated to predict outcomes for future year(s). This section will compare both MNL models (interaction effects versus main effects) based on their prediction capabilities. To this end, household data that can be used are the restructured surveys of OVG and OViN of 1987, 1991, 1995, 1999, 2003, 2010 and 2014.

To obtain relevant and most accurate results, the years between 1987 and 2010 have been used to predict household car ownership levels in the year 2014. The predictions of car ownership levels of households in 2014 are based on 174,393 households that have been covered by the OVG and OViN datasets of 1987, 1991, 1995, 1999, 2003 and 2010. The reason to choose 2014 as reference year is twofold. Firstly, it is of importance to select as many years as possible (i.e. timestamps) to execute a meaningful trend extrapolation. Secondly, with regard to the year 2014 actual data on household car ownership levels are available. The presence of actual data is of importance to compare the prediction accuracy of both MNL model outcomes.

To predict household car ownership levels in 2014 the coefficient values in both models have been extrapolated. With regard to the MNL model using constant coefficients, the averaged coefficient values between 1987 and 2010 have been extrapolated to 2014 resulting in the same average coefficient values. Concerning the MNL model that allows for interaction effects (i.e. changing effects of coefficients) a more complex procedure was necessary.

Via linear trend extrapolation of coefficient values in previous year (1987 – 2010) a new coefficient value for 2014 was estimated. The use of linear trend extrapolation is regarded to be adequate, as only six timestamps were available to extrapolate the coefficient value. In fact, polynomial trend extrapolation resulted in almost the same predicted coefficient values for 2014. Figure 19 shows an example of the procedure of linear trend extrapolation with regard to the coefficient values of household size for various household car ownership levels (see also Appendix V). The dots at the end of every trend line correspond to the new estimated coefficient for the year 2014. To determine the predicted household car ownership levels for every unique household in 2014 the same mathematical procedure and selection procedure has been used as explained in Section 5.4.

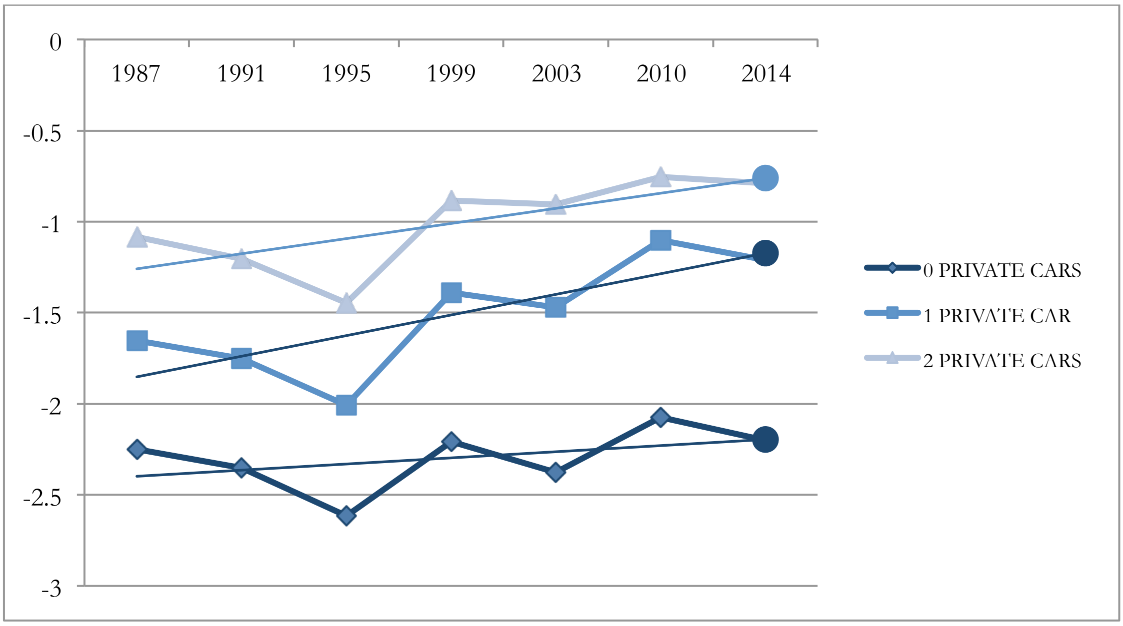


Fig. 19. Example of linear trend extrapolation – coefficients household size (1987-2014)

6.2. RESULTS PREDICTION CAR OWNERSHIP LEVELS IN 2014

Analysis of the predictions made for 2014 by both models shows that the MNL model that incorporates alternative coefficient values - via linear trend extrapolation - predicts household car ownership levels in 2014 more accurately. Based on the findings in Table 10, 62.0% of all cases are predicted correctly (N = 29,238). With regard to the predictions made by the MNL model with constant coefficient values, 54.8% of all cases in 2014 are predicted correctly (Table 11). As a result, the first model performs more than 13% better as compared to the latter in terms of prediction accuracy.

Nevertheless, these findings need to be taken into account carefully. The first shortcoming of this analysis is that the trend extrapolation in both models has been executed based on six time stamps over a period of more than two decades. Secondly, a multitude of predictions need to be made, based on a greater variety of historic data, to show whether the predicted outcomes are consistent with current findings. Finally, a method that is able to statistically compare the prediction outcomes of both models for a longer period of time is necessary to find/develop in order to ensure whether the model predictions differentiate significantly. These shortcomings will be further discussed in the reflective section of this thesis (Chapter 8).

		Observed cases				**Precision**
		0	1	2	3+	
Predicted cases	0	1627	894	27	6	64%
	1	2033	11936	3766	385	66%
	2	119	2956	4484	891	53%
	3+	2	13	40	58	51%
Sensitivity		43%	76%	54%	4%	**62.0%**

Table 10. Prediction results 2014 - MNL model including interaction effects - Model precision, sensitivity and accuracy

		Observed cases				**Precision**
		0	1	2	3+	
Predicted cases	0	693	220	10	3	75%
	1	2670	8678	1687	164	66%
	2	414	6883	6569	1092	44%
	3+	4	18	51	81	53%
Sensitivity		18%	55%	79%	6%	**54.8%**

Table 11. Prediction results 2014 - MNL model based on main effects – Model precision, sensitivity and accuracy

6.3. CAR OWNERSHIP MODELS IN THE NETHERLANDS

The comparison presented in Section 6.2 might not the be-all and end-all, the results nevertheless clearly confirm that a prediction model that incorporates changing influences of car ownership determinants, at least for the years studied in the OVG and OViN datasets, improves the prediction of household car ownership levels compared to a model that assumes constant coefficient parameters over time. Consequently, there is sufficient cause to question what implications this might bring for the application of car ownership models that are currently being used to predict future household car ownership levels in The Netherlands.

Current applications of car ownership modeling are among others static disaggregate models, which try to explain car ownership levels in The Netherlands (e.g. Potoglou & Kanaroglou, 2008a). These applications serve, however, mainly for research purposes, instead of having a practical relevance. The most important car ownership model currently used in The Netherlands is the DYNAMO model (MuConsult, 2016a). The model is widely used by Dutch ministries and political parties, in which the effects of general developments and government policy on size, composition and use of the car fleet are modeled. An important application of the model is to calculate the possession of cars among households in the short term (the next years) and long term (the next decades) (MuConsult 2016b). The rest of this section will therefore focus on the application of the DYNAMO model.

The modular structure of DYNAMO shows that for every household type the number of private cars is disaggregated into zero, one, two, and three or more private cars (Meurs & Haaijer, 2006). The explanatory variables that are used in the model are household characteristics (e.g. household income, employment status, age and household size (MuConsult, 2016a). DYNAMO uses these variables to predict car ownership levels among different types of households. An important part of the data used by DYNAMO for predicting car ownership levels originates from the OVG and OViN surveys (Meurs & Haaijer, 2006). For example, one of the earlier versions of DYNAMO (1.3) used the pooled data of OVG 1990-1998 to estimate the coefficients of household characteristics on household car ownership levels to forecast car ownership levels from 2003 till 2040 (Meurs & Haaijer, 2006).

However, no detailed insights have been found with regard to the latest version of DYNAMO (3.0). Therefore, an interview has been conducted on August 24th 2016 with dr. Rinus Haaijer, one of the developers of the DYNAMO model. Based on the outcomes of this interview the following conclusions need to be made:

- The latest version of DYNAMO uses the OVG and OViN datasets from 1990 till 2010 to predict household car ownership levels from 2016 till 2050 (zero, one, two, three or more private cars);
- The coefficients being estimated for household car ownership are the household characteristics found in the OVG and OViN, such as household income, household size and household composition;
- All datasets from 1990-2010 (OVG and OViN) have been pooled into a meta-dataset, wherein averaged coefficients (i.e. main effects) have been estimated;
- The averaged coefficients values between 1990-2010 are used to predict new levels of car ownership levels from 2016 onwards;
- Furthermore, the trends in parameters are based on different WLO scenarios till 2040 (e.g. economic trends), which together are used with the estimated coefficients to predict future car ownership levels among households.

6.4. DYNAMO MODEL FOR POLICY MAKING

According to dr. R. Haaijer the policy implications will become visisble, whenever higher accuracy of predicting household car ownership levels in the future can be obtained. According to dr. R. Haaijer the development of DYNAMO is still ongoing, but is widely used among policy makers and urban planners nowadays. The prediction of an accurate number of private cars in future years is – among others – used for predicting the total car fleet size in The Netherlands, future auto mobility and travel behavior, infrastructure building (e.g. roads and parking lots) and even election programs in The Netherlands.

The analysis from this chapter shows that incorporating *interaction effects* of coefficients from the same datasets and variables used by DYNAMO could increase the prediction accuracy of future household car ownership levels. Especially, the incorporation of interaction effects of household income and household size have the potential to increase the current prediction accuracy of the DYNAMO model (3.0), due to their relative important increase and decrease of their influence on car ownership levels over time.

Considering the relevant policy applications of DYNAMO it is necessary to further extent our notion of the possible policy *implications* of using constant instead of interaction effects to forecast household car ownership levels. To start, Figure 20 provides an overview of the current application of the DYNAMO model for policy related purposes in The Netherlands. The DYNAMO model is predominantly used for two purposes. Firstly, the DYNAMO model is used for **mid-term** and **long-term** policy-making. For this purpose both prognoses of car ownership levels and car fleet size, and indirectly car mobility prognoses via the "Landelijk Model Systeem" (LMS) are combined into mid-term (2030) and long-term (2050) scenarios on car ownership and car mobility in The Netherlands (PBL, 2015). These mid- and long-term scenarios are together bundled in the "Cahier Mobility" of the "Welvaart en Leefomgeving 2015" (WLO). These prognoses are used to provide **policy makers** (i.e. Dutch Ministries including Ministry of Infrastructure and Environment) guidance in mid-term and long-term policy-making (PBL, 2015).

Secondly, the DYNAMO model is used for **short-term** policy making by **political parties**. For this purpose, election programs of political parties are evaluated to estimate the mobility effects and financial effects (as a result of expenditures, investments and tax increase/decrease) (Zwaneveld et al., 2012). The DYNAMO model is both directly used (car ownership) and indirectly used via LMS (car mobility) to evaluate election programs of political parties. The coming sections will discuss important policy implications of the use of constant effects to forecast car ownership levels by the DYNAMO model.

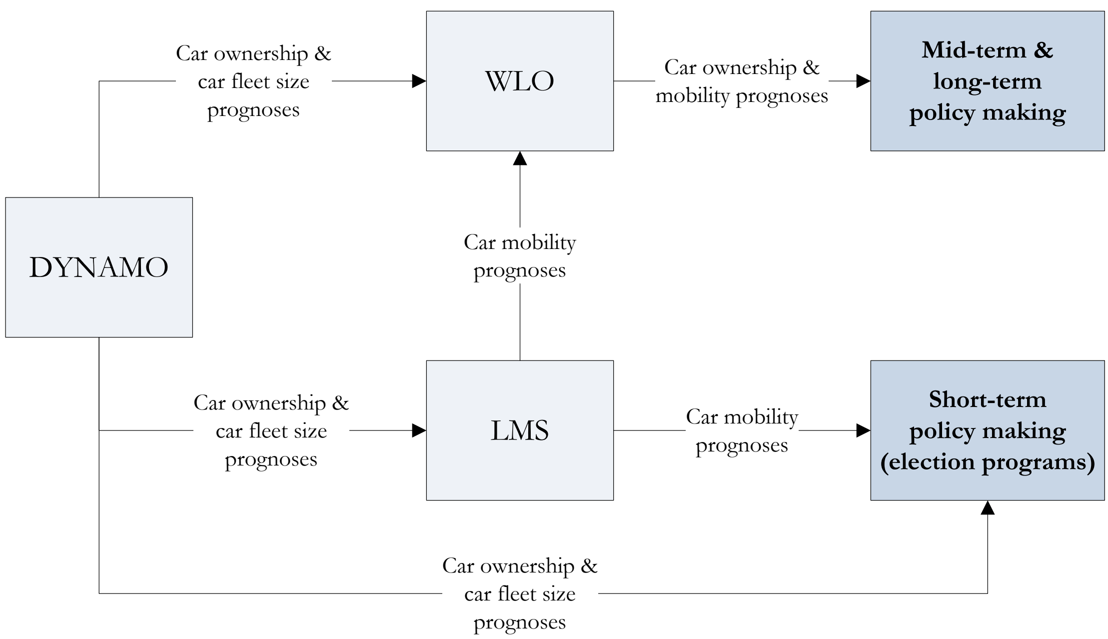


Fig. 20. Overview of application of DYNAMO in relation to WLO, LMS and policy making

6.5. MID-TERM & LONG-TERM POLICY IMPLICATIONS

The DYNAMO model uses household car ownership levels to forecast the total car fleet size in The Netherlands in future years. Not only the size of the car fleet is forecasted, but also the composition of the fleet, including the age of cars, weight and fuel type (Meurs & Haaijer, 2006). This section will exclusively focus on the model estimates regarding household car ownership levels and total car fleet size, since no investigation has been made regarding the car fleet composition in this study.

Next to forecasts regarding the total car fleet, the DYNAMO model is essential to mid-term (**2030**) and long-term (**2050**) prognoses regarding **car mobility**, which are among others used by the Ministry of Infrastructure and Environment (PBL, 2015). Car ownership levels are the driving force behind future car mobility (PBL, 2015). Therefore, future levels of car ownership are the main driving force behind **congestion** and related **construction of roads** (PBL, 2015). Moreover, future car ownership levels are the impellent for the future demand of **parking facilities** in urban areas (Snellen, van Eck, & de Jong, 2016). As a result prognoses of future car ownership levels are essential for long-term urban planning and policy making, captured in the "Meerjarenprogramma Infrastructuur, Ruimte en Transport" (MIRT) (PBL, 2015).

Future car ownership levels as forecasted by the latest version of DYNAMO (3.0) for the mid-term (2030) and long-term (2050) is presented in Figure 21.

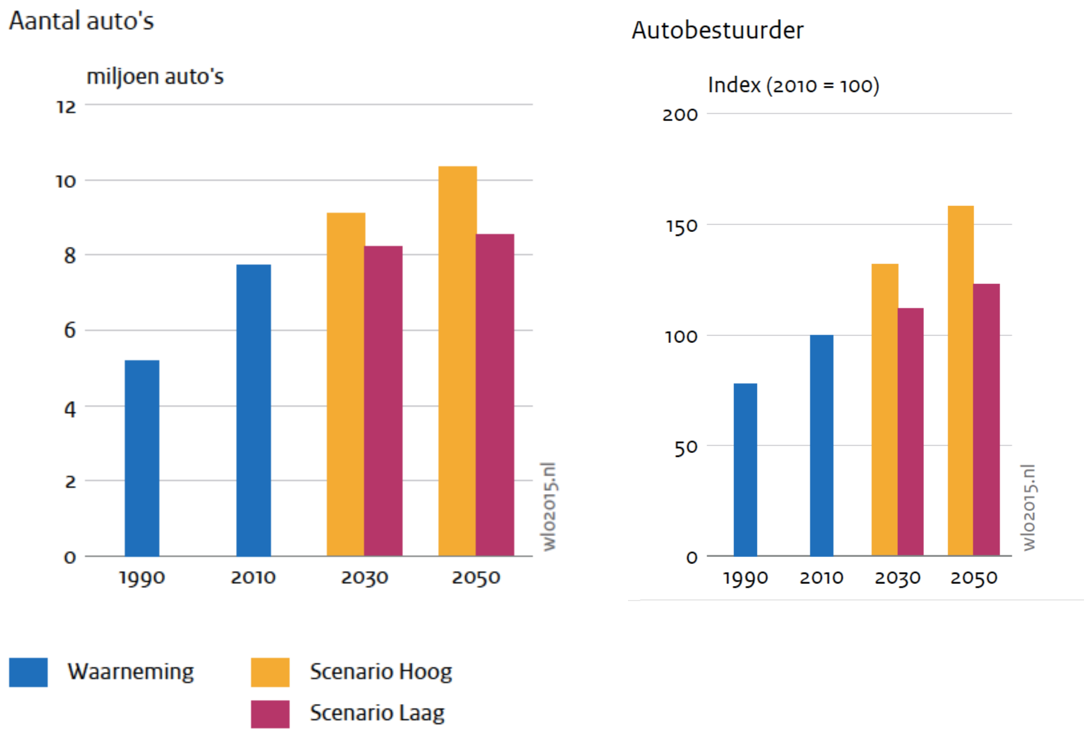


Fig. 21. Prognoses of car fleet size (left) and car mobility in driven kilomters (right) – 2030 & 2050 (PBL, 2015, p. 39)

The prognoses based on the DYNAMO model show two scenarios for the future: high and low. The difference between the two scenarios is based on differences in social and economic welfare (e.g. rise in income) in the next decades (PBL, 2015). With regard to 2030, 9 million cars are expected in the high-scenario, whereas 8.1 million cars are expected in the low-case scenario. Concerning the year 2050, 10.4 million cars (high) versus 8.5 million cars (low) are expected. Further consideration of the total car fleet estimated by DYNAMO – which is based on household car ownership levels - shows that household income influences the lion's share of the growth in the total car fleet, next to some social-demographic factors (Snellen, van Eck, & de Jong, 2016). Considering the use of main effects (instead of interaction effects) by the DYNAMO model, it is expected that the prognoses on the total car fleet are *overestimations* on the basis of (the importance of) the positive influence of household income. Assuming that the trend of household income will continue to develop in the coming decade(s). The results of this study have shown that the relative influence of household income has decreased with 10% from 1987 till 2014. To illustrate this, Figure 22 presents how the trend extrapolation has been done in DYNAMO (red line) next to the model that adjusts for interaction effects over time (green line). Notice that the red line is based on the measured points between 1990-2010, which have been used by DYNAMO (3.0). For the green line the data points between 1987 and 2014 are used. Although Figure 22 is a simplification (it presents only a single latent variable), it might be clear that based on the calculation of DYNAMO (3.0) the influence of household income is larger as compared to the interaction effects model. The exact result of the overestimation of the car fleet based on household car ownership is not clear, as it is unknown what the real parameter influence is of household income in the DYNAMO model.

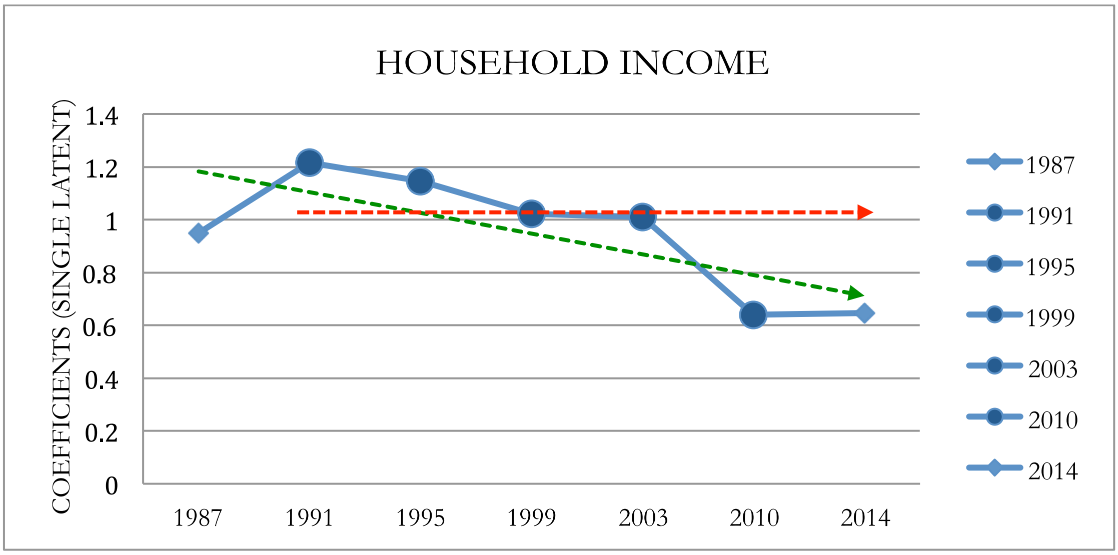


Fig. 22. Simplified trend extrapolation of household income – main effects model versus interaction effects model

Secondly, on the basis of the influence of household size, which has quite significantly changed between 1987 and 2014 (rise of 6%), we presume that the DYNAMO model *overestimates* the number of future cars based on the increasing positive influence of household size if we assume that the trend of household size will continue to decline in the coming decade(s). The same logic as described in Figure 22 can be applied, with the exception that the (parameter) influence of household size has increased instead of decreased (for household income). Though its influence is considered to be not as important as household income, it will most probably impact the car fleet size forecasts of DYNAMO (Snellen, van Eck, & de Jong, 2016).

The important question rises: what happens when car ownership levels are not correctly forecasted? To answer this question we will focus on the case of **household income**. A direct effect of the overestimation of the total car fleet in the future will impact policy-making concerning the development of **parking facilities** – both centralized and decentralized (i.e. in neighborhoods). The building of parking facilities is responsive to forecasted car ownership levels (Snellen, van Eck, & de Jong, 2016). An overestimation of private cars in The Netherlands will be reflected in parking planning and related urban planning (e.g. establishment of new neighborhoods) and could have financial implications in terms of planned investments in parking facilities.

An overestimation of the total car fleet in the mid-term (2030) and long-term (2050) will also overestimate the future car mobility activity by LMS, which is for an important part based on car ownership levels estimated by DYNAMO (PBL, 2015). Figure 21 (right) shows the average driven kilometers per car in The Netherlands for a high- and low-case scenario. In both scenarios the average number of driven kilometers will increase (PBL, 2015). An overestimation of the average car mobility will impact on **congestion** forecasts, which are defined by car mobility forecasts and road capacity (PBL, 2015). Secondly, possible **congestion charging** and planned investments to build new **roads** could in this way partially lose their effectiveness (PBL, 2016). Finally, the effectiveness of related policies such as **car purchasing taxes**, **noise regulations**, **scrapping premium schemes** are all directly, and indirectly dependent on accurate car fleet forecasts and thus affected by inaccurate forecasts of the car fleet in the future (PBL, 2008; Geilenkirchen & van Meerkerk, 2014; PBL, 2015; PBL, 2016).

Research on car purchasing taxes (**BPM**) to reduce CO_2-emmissions by conventional cars shows that by the estimations of DYNAMO car ownership levels will be reduced in the long-term (Geilenkirchen & van Meerkerk, 2014). However, an overestimation of the car fleet in the mid-term (2030) and long-term (2050) could make this policy less effective both in financial gains from the tax as well as the targeted reduction of CO_2-emissons (Geilenkirchen & van Meerkerk, 2014). With regard to the effectiveness of road taxes (**MRB**), it is anticipated that an overestimation of the car fleet will have smaller implications, due to the smaller effect of the MRB on the total car fleet (MuConsult, 2003; Geilenkirchen & van Meerkerk, 2014). Another policy based on the externalities of car use, **noise disturbance**, is made on a long-term strategic level by the outcomes of the LMS model, which in turn depends on the outcomes of the DYNAMO model (Mourik, 2008). An overestimation of the total car fleet in the future could increase the likelihood of less effective policies related to noise-reduction at the strategic level. Finally, the volume of **car scrapping premiums** is dependent on the financial means and allocation choices made by the Dutch government (van Dam, de Groot, & Verwest, 2006). As a result of an overestimation of the number of cars in the future, initiatives such as scrapping premiums might be discouraged by the government to be used due to a lack of anticipated financial means to support such incentives (van Dam, de Groot, & Verwest, 2006). On the other hand, premiums can be lowered by anticipation, which in turn increase the likelihood consumers are discouraged to provide their old car for demolition and receive a premium to trade for a younger (and thus cleaner) car.

Concluding, an overestimation of the size of the total car fleet in mid- and long-term forecasts could lead to overestimations in the necessity of parking facilities, and will impact on congestion forecasts, which are defined by car mobility forecasts and road capacity. Moreover, overestimations of the car fleet could lead to less effective policies (e.g. taxes, noise regulations and car scrapping premiums) made by Dutch Ministries that are dependent on accurate car ownership forecasts. The important assumption of this analysis is that trends of the parameter influence as found for the last three decades will continue to develop. Further research should make clear what the effects are in absolute terms for car ownership and car mobility regarding the foreseen overestimation of the influence of household income, and to a lesser extent household size (overestimation) by DYNAMO. If such effects are significant in absolute terms on car fleet forecasts, Dutch Ministries could be provided with i) better insight whether current policies related to car ownership levels are (still) **effective** and ii) financial insight in the **cost savings** that might be achieved or **extra investments** that are necessary for mid-term and long-term policy revisions, including adaptations in MIRT (PBL, 2015).

6.6. SHORT-TERM POLICY IMPLICATIONS

Next to mid-term and long-term policy making by Dutch ministries, the DYNAMO model is used to evaluate short-term policy making (**2020**) by political parties. **Election programs** of political parties are evaluated to estimate the mobility effects and financial effects as a result of expenditures, investments and tax increase/decrease (Zwaneveld et al., 2012). On the following points the election programs are evaluated by DYNAMO that are (indirectly) related to car ownership levels (Zwaneveld et al., 2012):

I. Discount of taxation of automobile travel from home to the workplace;
II. Taxation of private cars on the basis of average number of driven kilometers;
III. Changes in road taxes (MRB);
IV. Changes in car purchasing taxes (BPM);
V. Taxation of congestion;
VI. Investments or cuts in building new roads.

Points II, III, IV have been directly measured by DYNAMO, whereas points I, V and VI have been calculated by the combination of DYNAMO and LMS (Zwaneveld et al., 2012). To estimate the qualitative short-term impact of using main effects instead of interaction effects by DYNAMO, the current versions of political party programs will be used comprising the period between 2013 and 2017 (Zwaneveld et al., 2012). The reason is that for the majority of political parties no new party programs have been published from 2017 onwards - with the exception of D66 and PVV. Nevertheless, it is assumed that most perspectives on the proposed measures will remain the same for the period till 2020 (Zwaneveld et al., 2012).

Table 12 gives an overview of the political measures related to car ownership proposed by nine political parties in The Netherlands. VVD, PvdA, PVV, CDA, SP, D66, ChristenUnie and SGP all propose a **discount on the taxation** in relation to **work** travel by car (see 'I'). The number indicates the maximum number of eurocents per driven kilometer per car related to work purposes that will not be taxed (Zwaneveld et al., 2012).

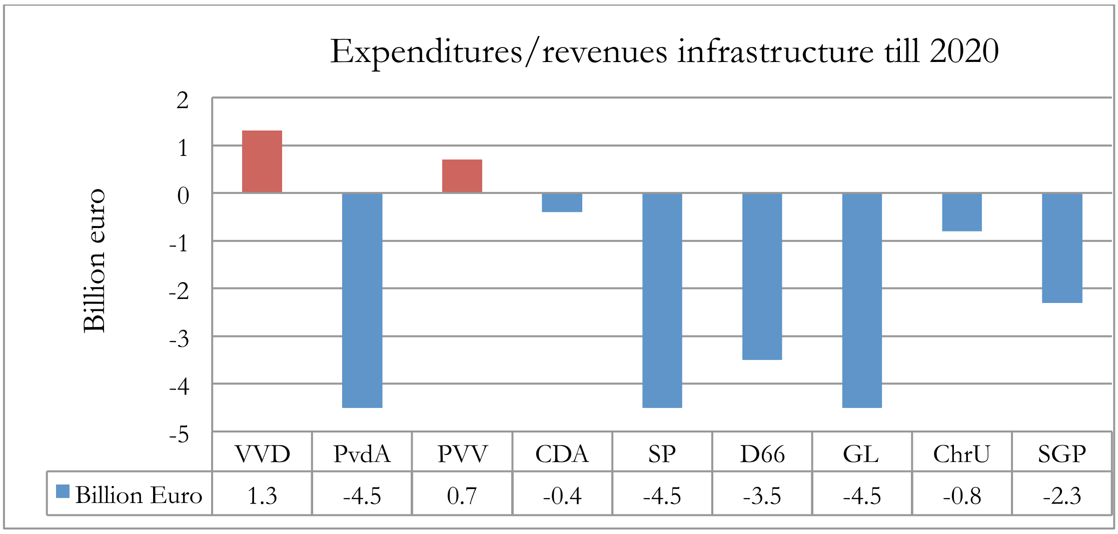
Five of the nine political parties - PvdA, SP, D66, Groenlicht en ChristenUnie – propose a tax on the number of **driven kilometers** by car (in Dutch: 'kilometerheffing') (Zwaneveld et al., 2012). The number relates to the number of eurocents that need to be paid per driven kilometer by car (see 'II'). However, the same parties propose the abolition of road taxes – **MRB** (see 'III'). Furthermore, the PvdA, SP and Groenlinks propose an increase of the car purchasing tax (**BPM**), whereas the PVV, D66 and ChristenUnie propose a decrease of the BPM (Zwaneveld et al., 2012). The VDD and SGP propose no adjustments in BPM tax (see 'IV'). Additionally, the political parties PvdA, D66 and Groenlinks propose taxation on congestion (Zwaneveld et al., 2012). The **congestion charge** is carried on top of the kilometer tax (if applicable) and applies to all vehicles, including cars. Finally, the VVD and PVV propose more investments in **roads**, whereas all other political parties propose budget cuts in investing in new roads (see 'VI').

	VVD	PvdA	PVV	CDA	SP	D66	GL	ChrU	SGP
I. Discount work travel (cent/km)	**19**	**19**	**19**	**13**	**19**	**0**	**0**	**10**	**19**
II. Kilometer tax (cent/km)	**0**	**4**	**0**	**0**	**6**	**5**	**10**	**8**	**0**
III. MRB tax	**Yes**	**No**	**Yes**	**Yes**	**No**	**No**	**No**	**No**	**Yes**
IV. BPM tax	**0**	**+**	**–**	**0**	**+**	**–**	**+**	**–**	**0**
V. Tax on congestion	**No**	**Yes**	**No**	**No**	**No**	**Yes**	**Yes**	**No**	**Yes**
VI. Road investments	**+**	**–**	**+**	**–**	**–**	**–**	**–**	**–**	**–**

Table 12. Measures proposed by political parties on taxation of car ownership and car mobility 2013-2017 based on Zwaneveld et al (2012) - own depiction

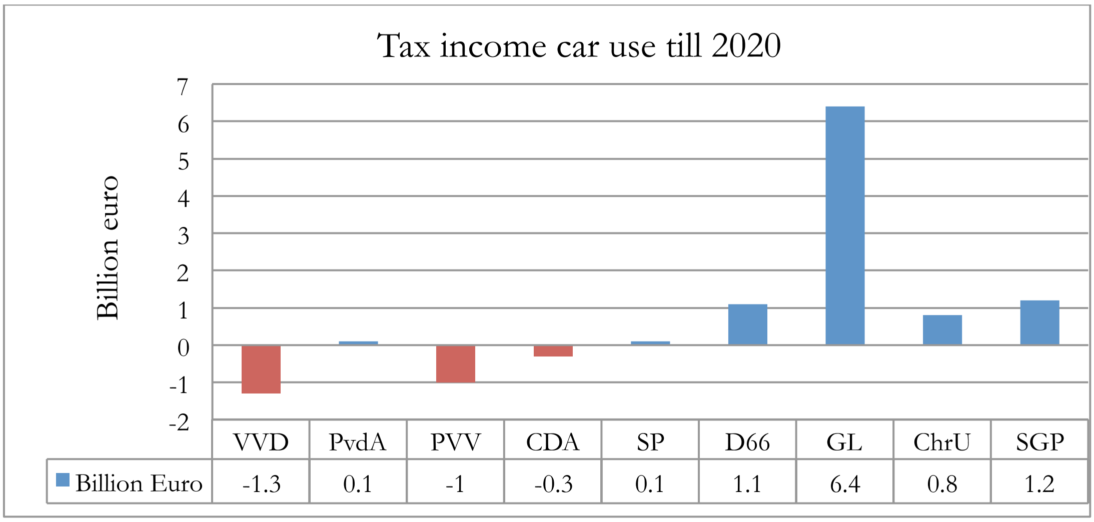
The DYNAMO model together with the LMS model has calculated the impacts of the mobility programs by the political parties of 2013-2017. The effects are measured in terms of financial impact and car mobility. The effects are compared to the "base-case" scenario that is projected for 2020 based on the implementation of the 'Begrotingsakkoord 2013' (Zwaneveld et al., 2012). Figure 23 shows the change in expenditures or generated revenues induced by the political programs relative to the base-case. For example, the program of the VVD would "spend" 1.3 billion by tax reliefs (Fig. 24), whereas Groenlinks will generate 4.5 billion euros (Fig. 23) in revenues due to a tax income of 6.4 billion euro (Fig. 24). Moreover, the effects of the political programs are evaluated in terms of the change in car use and associated congestions (Fig. 25). Calculations from DYNAMO combined with LMS show the following outcomes (Zwaneveld et al., 2012):

- Discount of taxation of automobile travel from home to the workplace (I) have a significant impact on car use and congestion. This is the main explanation for the increase in car use and congestion in the VVD, PVV and CDA programs;
- The most important explanation for the reduced car use and reduced congestion in the programs of PvdA, SP, D66, Groenlicht and SGP is the introduction of a kilometer tax (II);
- Effects of MRB (III), BPM (IV), congestion tax (V) and road investments (VI) are limited compared to the other measures in terms of car use and congestion.



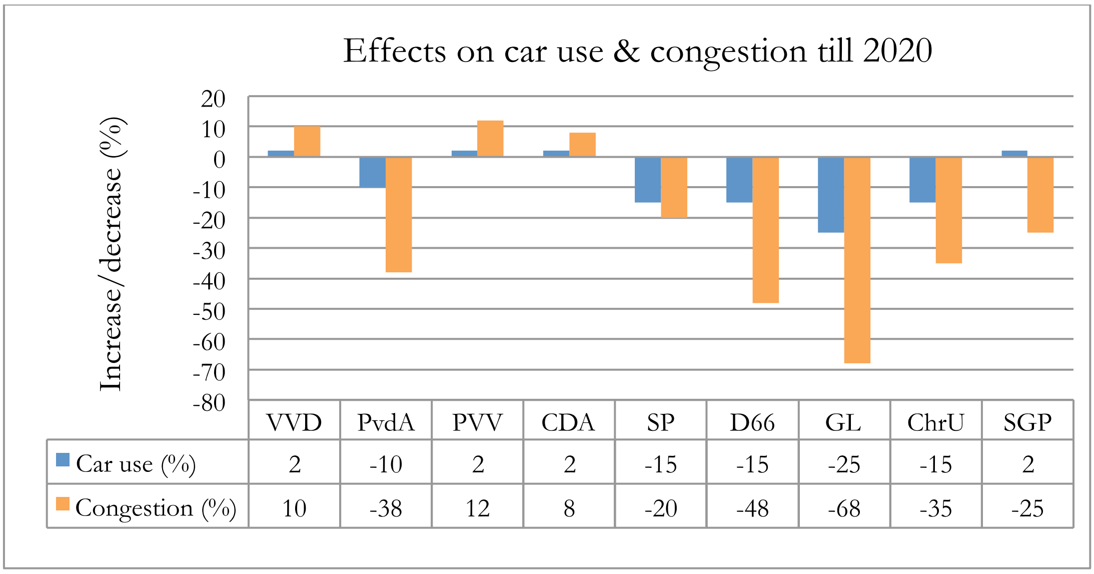
	VVD	PvdA	PVV	CDA	SP	D66	GL	ChrU	SGP
Billion Euro	1.3	-4.5	0.7	-0.4	-4.5	-3.5	-4.5	-0.8	-2.3

Fig. 23. Extra investments/divestments proposed by political party programs till 2020 based on Zwaneveld et al (2012) - own depiction



	VVD	PvdA	PVV	CDA	SP	D66	GL	ChrU	SGP
Billion Euro	-1.3	0.1	-1	-0.3	0.1	1.1	6.4	0.8	1.2

Fig. 24. Extra tax incomes generated by political party programs till 2020 based on Zwaneveld et al (2012) - own depiction



	VVD	PvdA	PVV	CDA	SP	D66	GL	ChrU	SGP
Car use (%)	2	-10	2	2	-15	-15	-25	-15	2
Congestion (%)	10	-38	12	8	-20	-48	-68	-35	-25

Fig. 25. Effects on car use and congestion generated by political party programs till 2020 based on Zwaneveld et al (2012) - own depiction

To provide a case to evaluate the effects of DYNAMO using main effects instead of interaction effects for the outcomes of political programs, we will focus on the influence of **household income**. As described in Section 6.5 the total car fleet estimated by DYNAMO shows that household income influences the lion's share of the growth of the total car fleet (Snellen, van Eck, & de Jong, 2016). Given that household income has a positive relationship with the number of cars owned by a household, and is supposed to have an overestimated influence in DYNAMO (Section 6.5), total car fleet projections could be overestimated in the current model. If we assume that the (parameter) influence of household income will continue to decline till 2020, the question is how the impact of the political measures will alternate in terms of financial effects and car mobility effects.

Firstly, a correction for the overestimated car ownership levels till 2020 by DYNAMO will most probably reduce the calculated effects of car use and congestion development by the LMS model (PBL, 2015). As a consequence, this could partially affect the plans of the VVD and PVV to invest in new roads (VI). Secondly, as a result of the reduction in car use, the tax revenues of both Groenlinks and D66 are supposed to decline, since both parties rely on a kilometer tax, but do not use a tax discount on car travel (Table 12). In addition, these parties introduce a congestion tax, which gains are likely to decrease due to decreased car mobility (de Borger & Mayeres, 2007). On the other hand, it is expected that for the VVD, PVV and CDA the tax expenses will reduce (concerning the tax discount). Furthermore, for the PvdA, SP, ChristenUnie and SGP it is not clear whether tax gains will increase, since these parties use combinations of both a tax discount on car travel, introduce a tax on driven kilometers or even use a congestion tax (Table 12).

With regard to the *costs of car ownership* - i.e. the costs of the acquisition/ownership of cars including MRB (III) and BPM (IV) - the overestimation of the influence of household income by DYNAMO might also indirectly affect estimated car ownership and mobility effects till 2020. For this purpose, we will focus on income elasticity and price elasticity regarding the costs of car ownership (i.e. purchasing costs, MRB, BPM, etc.). Note that this is different compared to price elasticity related to the *costs of car use* - e.g. fuel taxes (van Essen & Schroten, 2008). Based on the observed decreasing influence of household income on car ownership levels, we would expect an increase in income inelasticity of cars during the past decades, meaning that the responsiveness of the number of cars demanded has reduced in response to changes in household income (OECD, 2002).

Based on Slutsky's equation of the price elasticity of demand we would further expect that the price inelasticity of demand has increased as well (Tobie & Houthakker, 1950):

$$EP = K(X) * ET + (1 - X) * ES \tag{6.1}$$

Where:
EP = elasticity of the demand of private cars
K (X) = proportion of household's income spent on private cars
ET = income elasticity of demand for private cars
X = good (in this case the private car)
ES = substitution elasticity of demand for private cars

Equation 6.1 shows that the change of price elasticity of the demand for cars depends in part upon the magnitude of the income elasticity of the demand for cars. Substitution means the substitution of a car with a different fuel type or other related transport options, such as public transport (de Borger & Mayeres, 2007). Empirical research shows that the substitution of private cars is often relatively inelastic, as it becomes necessary for consumers to replace cars in the course of time. Moreover, once a car has been bought, consumers become accustomed to car use, so that car ownership becomes a necessity rather than a luxury (Dargay, 2001). Secondly, the proportion of a household's income spent on vehicles, including private cars has been relatively constant over the past two decades (Statistics Netherlands, 2016g). One of the important reasons is that car ownership prices have increased, whereas the relative purchasing power of households has also increased in the past decades (Statistics Netherlands, 2015f). Based on these findings and Slutsky's equation on elasticity we would assume that the price inelasticity of the demand of private cars has increased in the past decades.

An overestimation of the influence of household income on the number of private cars by DYNAMO, could therefore underestimate the increase of price inelasticity of demand based on the period of 1990-2010, for which the main effects of household income are estimated. As a result, the effectiveness of fixed taxes - such as the MRB and BPM – could be affected (de Borger, Mulalic, & Rouwendal, 2016). This is illustrated in Figure 26. A high price-inelasticity of demand (2) is less susceptible for fixed taxes and therefore less effective compared to a lower price-inelasticity of demand (1). As a result, an overestimation of the influence of household income in the DYNAMO model could indirectly lead to an overestimation of car ownership levels and mobility effects on the basis of the influence of fixed taxes. Since the DYNAMO model calculates and uses price-elasticity to estimate the effects of fixed taxes, including MRB and BPM, car mobility prognoses for political parties programs could change (de Jong, Kouwenhoven, Geurs, 2010). This could apply to PvdA, D66 and Groenlinks using the BPM tax, which is considered to be more influential as the MRB tax on car ownership (see Section 6.5). Based on household income this finding could also be relevant for the discussion as to whether or not to shift the burden from fixed taxes on car ownership (e.g. BPM) to variable costs of car use (e.g. kilometer tax) (de Borger & Mayeres, 2007; Fosgerau, & Jensen, 2013; de Borger, Mulalic, & Rouwendal, 2016). However, little empirical research has been done to the combined effect of a decrease in car purchasing prices and a simultaneous rise in variable car costs (van Essen & Schroten, 2008).

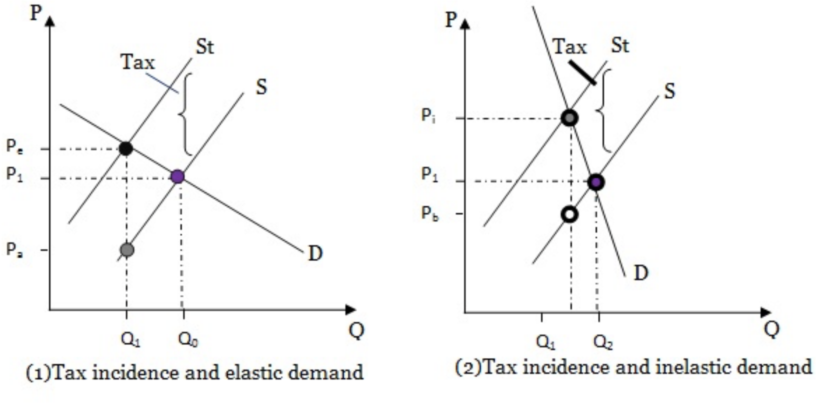


Fig. 26. Effectiveness of fixed tax incidences and elasticity of demand (Source: Justdan, 2016)

The different outcomes that might be the result due to an overestimation of the influence household income by DYNAMO should be nuanced on the basis of the following points (the first six points also apply to Section 6.5):

- The analysis is only based on household disposable income, leaving all other variables constant;
- The analysis is qualitative, and gives only a *direction* of possible policy implications;
- The (parameter) influence of household income is assumed to be overrepresented by DYNAMO on the basis of 1990-2010 data (see Fig. 22);
- The overestimated (parameter) influence of household income is assumed to have an effect on car ownership and car mobility prognoses by DYNAMO. However, effects of other variables (e.g. household size) could partially diminish the effect of household income on total car fleet prognoses and indirectly car mobility activity;
- The relation between car ownership levels (DYNAMO) and car mobility (LMS) is in reality more complex, which does not always imply 1-to-1 relations as sometimes seems to be suggested in this analysis (e.g. Zwaneveld et al., 2012);
- The trend in the (parameter) influence of household income is assumed to continue to decline in the short-, mid-, and long term. This is uncertain, especially for the long-term (2050);
- Based on the Slutsky's equation the relation between income and price elasticity is dependent on the proportion of the household's income spent on private cars and substitution elasticity, which could change over time (e.g. car sharing as substitution);
- The political party programs are the most accurate programs evaluated by DYNAMO, but can be outdated. Moreover, the proposed measures in the political party programs are assumed to remain the same for the period till 2020 (Zwaneveld et al., 2012).

Despite these limitations, the main purpose of this analysis is to provide better insight in the possible policy implications for political parties in the short-term (2020). In conclusion, an overestimation of car ownership levels might impact the current foreseen effectiveness of tax-based policies - both financially and car mobility-related. On the basis of tax discount on car travel and kilometer tax, it is supposed that tax revenues of both Groenlinks and D66 are expected to decline, whereas it is expected that for the VVD, PVV and CDA tax expenses will reduce. Secondly, on the basis of price inelasticity of cars, fixed taxes (e.g. BPM and MRB) might lose some effectiveness in reducing car ownership levels and car mobility. Further research should make clear what the effects are in absolute terms for car ownership and car mobility regarding the foreseen overestimation of the influence of household income and other important variables such as household size by DYNAMO. If such effects are significant in absolute terms on car fleet forecasts, Dutch political parties could be provided with i) better insight whether current tax-based policies on car mobility effects are (still) effective and ii) financial insight in the cost savings that might be achieved or extra investments that are necessary for effective short-term policy making.

7

Conclusion

The current stabilization of car use in industrialized countries, also referred to as 'peak car', seems to come hand in hand with a decline in car ownership growth. In The Netherlands, the growth in household car ownership has slowly decreased from 3.1% to 0.6% in 2001-2015. There is much unknown on the causes of this trend. Factors considered include a decline of car ownership by young adults, increased urbanization or economic trends. Also the role of traditional factors, such as income levels, is suggested to have weakened over the years.

7.1. GAP IN EXISTING RESEARCH

The decline in car ownership levels has led to the awareness among researchers that factors influencing car ownership levels might have changed over time. The changing influence of factors on car ownership could partly explain the current observed phenomena such as car ownership saturation and peak car. To the best of the authors' knowledge, there is, however, an absence of studies that investigate the changing influence of factors on car ownership over time.

To contribute to existing approaches that study the influence of factors on car ownership this study will use Netherlands as case study. The reason to choose The Netherlands is twofold. Firstly, The Netherlands is considered to have and apply state of the art knowledge about car ownership (modeling) for research and policy purposes. Secondly, The Netherlands provides one of the most resourceful databases regarding car ownership in the world. The aim of this research is to provide whether and to what extent the influence of such factors, on a household level, have changed over time. The research question in this study is formulated as follows:

To what extent has the influence of determinants on households' choice on the number of private cars to own changed in The Netherlands in recent decades?

Furthermore, understanding how factors have contributed to the households' choice of the number of private cars to own is of importance to urban planners and decision makers, since car ownership levels are associated with urban sprawl and automobile travel. Consequently, one should question what implications the results of this study might bring for the application of car ownership models that are currently being used to predict future car ownership levels in The Netherlands. Especially, in case the influence of factors is considered to be constant over time periods comprising multiple years, it becomes of importance what policy implications this might bring for the future. This study aims, therefore, as an extension to the first research objective, to provide more insight in the policy implications of predicting household car ownership levels based on constant versus changeable influence of determinants over time. For this purpose, special attention will be given to the most dominant car ownership model currently used in The Netherlands: DYNAMO. The model is widely used by Dutch ministries and political parties, in which the effects of general developments and government policy on size, composition and use of the car fleet are modeled.

7.2. STUDY DESIGN

To reveal the changing influence of factors on household car ownership, the statistical method of ordered logistic regression (ORL) is used. The method has previously been successfully applied to determine the effects of factors on household car ownership levels. In this study, the dependent variable household car ownership is specified into four alternative levels: zero, one, two, and three or more cars. The ORL model has primarily an exploratory/explanatory purpose to reveal the changing influence of determinants on household car ownership levels. A MNL model has been used later in this study to endogenously determine the number of cars owned by households in The Netherlands, based on constant versus changeable influences of determinants over time.

To perform the logistic regression analysis, data on 203,630 households from the national Traffic Survey in The Netherlands on household mobility is used. The Traffic Survey has been carried out under various names (OVG/MON/OViN) from 1985 till 2014. The datasets record all trips and trip stages for one day among participants and includes demographic as well as economic characteristics of the respondents. The annual mobility surveys use significant sample sizes – up to 50,000 unique households spread across The Netherlands. About 80% of the mobility in The Netherlands is covered by the surveys. The annual surveys that are used in this study are the surveys of OVG-1987, OVG-1991, OVG-1995, OVG-1999, OVG-2003, OViN-2010 and OViN-2014. The selection of years concerned a trade-off between accuracy and time efficiency/computing power. The substantial datasets facilitate to investigate relatively small subgroups within the population, such as households owning three or more cars.

7.3. INFLUENCE OF CAR OWNERSHIP DETERMINANTS

Based on previous studies on car ownership determinants and the data that has been made available for this research, the influence of **household income, size, composition** (i.e. presence of children)**, gender, age, education, working status** and **level of urbanization** on household car ownership has been investigated between 1987 and 2014. Descriptive analysis of the datasets shows that most households obtain one or more cars (87% in 2014). Households that have no car have decreased between 1987 and 2014 (68%), whereas households that own two or more cars have increased (65%). Household income levels have steadily increased, while the average household size has decreased over time (2 person-households increased with 35%, whereas 4-person households decreased with 67%). The presence of children in households has gradually declined between 1987 and 2014 (41% in 1987 compared to 29% in 2014). With regard to the age of the households' reference person, the total share of younger age-classes (i.e. 18-39 years) has decreased with 50%, while older age-classes (50+ years) have increased with 64% between 1987 and 2014. Further, the group that followed higher education (secondary and tertiary) has grown between 1987 and 2014. Especially, the group that attended higher professional education (HBO/University) has significantly increased with 136% between 1987 and 2014. The employment status of the households' reference person shows a shift of fulltime employment (>30 hours per week) towards part-time employment (<30 hours per week). In 1987 around 53% of the reference persons was employed fulltime, compared to 43% in 2014. Finally, the share of households living in different urbanized regions has been relatively constant between 1995 and 2014 (highly to very-high urbanized regions accounted for 40% in this period).

Ordered logistic regression results show that the influence of household income, size, composition, gender, age, education, working status and urbanization level on car ownership, in general, **have changed between 1987 and 2014**. Especially, the relative influence of **household income** and **household size** are substantial, contributing to more than 60% of the total influence on household car ownership in all years studied. Whereas the influence of household income on car ownership decreased over time (from 38% in 1987 to 28% in 2014), the influence of household size has increased (from 29% in 1987 to 35% in 2014). Both household income and household size show a positive relationship with the number of cars owned by a household. The same positive relationship is observed for education, working status and urbanization with household car ownership. The relative influence of these variables on household car ownership has gradually grown with an average total of 2% between 1987 and 2014. With regard to household composition a relatively strong negative relationship with car ownership is observed, which has remained relatively stable between 1987 and 2014 (around 10%). Also age and gender show predominantly a negative relationship with the number of cars owned by a household. Their influence on car ownership has decreased over 1% between 1987 and 2014, and was relatively small in 2014.

Based on a qualitative analysis of the combined results of the descriptive analysis and ordered logistic regression analysis we expect that household composition, education and urbanization levels have increased the average number of household cars for the period studied. The effect of household income, household size, gender, age and working status on the increase/decrease of the average number of cars is not clear - due to the contradicting development of trends in variables vis-à-vis influence of parameters. Furthermore, based on the limited outcomes of the qualitative analysis, the stabilization in the growth of cars (especially in the last couple of years) cannot necessarily be explained.

With regard to the internal validity of the model outcomes, a holdout sample has been used by subtracting 20% of the total number of unique households in the datasets (N = 40,852). Analysis shows that 65% of all the cases has been predicted correctly by the model suggesting that 65% of the sample data can be explained by the explanatory variables that are included. Moreover, comparison shows that an ORL model is highly appropriate compared to MNL models to explain household car ownership levels. In fact, the overall model accuracy of the ORL, despite it much less greater flexibility, is even higher compared to the model accuracy of the MNL model that incorporates constant car ownership coefficients (compare 65% to 62%).

7.4. POLICY IMPLICATIONS

The notion of the relative contribution of car ownership determinants on the households' choice of the number of private cars to own is of importance to urban planners and decision makers, since car ownership is a key element in the study and simulation of urban systems. An extension of this research is to provide more insight in the policy implications of predicting household car ownership levels based on models that use a constant versus changeable influence of determinants over time. By means of multinomial logistic regression, which is considered to be the most appropriate method to predict car ownership levels, a first comparison has been made between both types of prediction models.

For this comparison, the years of 1987, 1991, 1995, 1999, 2003 and 2010 have been used to predict household car ownership levels in the year 2014. The predictions of car ownership levels of households in 2014 are based on 174,393 households that have been covered by the OVG and OViN datasets between 1987 and 2010. The reason to choose 2014 as year of reference is to select as many years as possible (i.e. timestamps) to execute a meaningful trend extrapolation, and compare the predicted outcomes based on actual data that has been made available for 2014.

Results show that a MNL model that allows for changing influences of the studied car ownership determinants, incorporating alternative coefficient values, predicts household car ownership levels in 2014 more accurately. Overall, 62.0% of all cases are predicted correctly (N = 29,238). Concerning the MNL model that assumes no changing influences of the car ownership determinants, using constant coefficient values, predicts 54.8% of all cases in 2014 correctly. As a result, the first model performs on average more than **13%** better as compared to the latter in terms of prediction accuracy.

A comparison with the most important car ownership model in The Netherlands, DYNAMO, shows that the results of this study could possibly further extend the prediction accuracy of the model - supported by an interview with dr. R. Haaijer (developer of DYNAMO). According to dr. R. Haaijer, the latest version of DYNAMO (3.0) estimates main (i.e. constant) effects of household coefficients (income, size, etc.) using OVG and OViN data from 1990-2010 to predict household car ownership levels till 2050. To this end, the model does not allow for changing influences of car ownership determinants, since no interaction effects are estimated. The accurate prediction of household car ownership levels in the future is, however, of importance for policy makers in the mid- and long-term, and political parties in the short-term in The Netherlands. DYNAMO's predictions of the number of private cars owned by households are, among others, used for predicting the total car fleet size in the Netherlands, future auto mobility and travel behavior, infrastructure building (e.g. roads and parking lots) and even election programs in The Netherlands.

To qualitatively investigate the short-term (2020) to long-term (2050) policy implications of using main effects by DYNAMO, this study has exclusively focused on the influence of household income on car ownership levels. Considering the use of main effects by the DYNAMO model, it is expected that the prognoses on the total car fleet are overestimations on the basis of household income assuming that the trend of the (parameter) influence of household income will continue to decline in the coming decade(s). Firstly, overestimation of the size of the total car fleet in **mid-term** (2030) and **long-term** (2050) forecasts for **Dutch Ministries** (e.g. Ministry of Infrastructure and Environment) could lead to overestimations in car mobility, and will impact on congestion forecasts. Investments in parking facilities, possible congestion charging and planned investments to build new roads could in this way partially lose their effectiveness. Moreover, overestimations of the car fleet could lead to less effective policies (e.g. taxes, noise regulations and car scrapping premiums) made by Dutch Ministries that are dependent on accurate (as possible) car ownership forecasts.

Secondly, an overestimation of car ownership levels on the basis of household income could impact the effectiveness of tax-based policies (financially and car mobility-related) for **political parties** in the **short-term** (2020). Concerning tax discount on car travel and the kilometer tax, it is supposed that tax revenues from current political programs of Groenlinks and D66 are expected to decline, whereas the tax expenses suggested in the programs of VVD, PVV and CDA are expected to decline till 2020. In addition, fixed taxes on car ownership, including BPM and MRB tariffs, might lose their effectiveness in reducing car ownership levels and car mobility due to a higher price inelasticity of cars than currently is anticipated by DYNAMO.

7.5. FUTURE RESEARCH RECOMMENDATIONS

The logistic regression analysis has proven to be a suitable method to *explore* the changing influence of important car ownership determinants on household car ownership levels over time. The analysis could, however, be further improved, by including other factors that have been excluded from the national Traffic Survey. For example, attitudinal/psychological factors are considered to become increasingly important factors affecting household car ownership levels. Secondly, determinants could be investigated more thoroughly by specifying the effects of unique categories on car ownership levels. For this study, effects on household car ownership have only been estimated on factor-level. Complementary, some factors might be extended in their categorization to reveal new outcomes. For example, household composition, now being categorized into "children" and "no children" could be extended to family-types, such "single", "couple", "couple with children", "extended family", etc. In the future we might find from these extensions that "the influence of new couples with children has significantly decreased on the probability of owning one or more cars". Thirdly, the effects of some factors can be investigated more thoroughly by using other types of analysis. For example, the influence of the variable age on household car ownership is better described by a quadratic function, rather than a linear function as used in this study.

Secondly, this study has shown that incorporating *interaction effects* of coefficients - from the same datasets and variables used by DYNAMO - could increase the prediction accuracy of future household car ownership levels. With regard to the relevant short-, and long-term policy applications of DYNAMO more research is necessary to further extend our notion of the capabilities of DYNAMO using interaction effects. Especially, the interaction effects of household income and household size have the potential to increase the current prediction accuracy of the DYNAMO model (3.0). Accordingly, if the effects are significant on car fleet forecasts and car mobility, Dutch Ministries and political parties could be provided with i) better insight whether current policies related to car ownership levels and car mobility are (still) effective and ii) gain financial insight in the cost savings that might be achieved or extra investments that are necessary for short-term and long-term policy revisions. Concluding, this study recommends implementing a module in DYNAMO that allows for interaction effects over time to evaluate the predicted outcomes of the model in terms of car ownership levels and related car mobility activity.

Thirdly, the findings of this study are related to the case of **The Netherlands**. However, the current debate on the changing influence of factors on car ownership levels applies to multiple industrialized countries. The continuous decline (in the growth) of car ownership levels in Australia, Belgium, Canada, France, Germany, Japan, Sweden, UK and USA could demonstrate that the influence of car ownership determinants has changed over time (Millard-Ball & Schipper, 2011). It is therefore recommended for these countries to perform an analysis based on longitudinal disaggregate-household data to confirm whether the changing influence of determinants follows a similar pattern as found in The Netherlands. With regard to the policy outcomes of this study, it is recommended to investigate the (sensitivity of) effects of the changing influence of car ownership determinants (if applicable) on car fleet estimations and car mobility forecasts. This is especially important for countries like the UK and USA that formulate short-term and long-term mobility policies based on car ownership models similar to DYNAMO (Zwaneveld et al., 2012).

8

Reflection

This chapter provides a reflection on the research goals that have been adapted during the execution of this study. Furthermore, the limitations of this study, which have impacted the outcomes and depth of this research, will be discussed.

8.1. REFLECTION ON RESEARCH GOALS

The research goals that have been formulated at the beginning of this research have been adapted several times in accordance to the research progress. At the beginning of the study, there was a specific interest in the explanation of the current decline in car ownership levels and car use currently observed in The Netherlands and other industrialized countries. Nevertheless, previous studies that investigated the relation between important factors and the decline in car ownership levels and car use made explicit that explanations are difficult to compose due to the complexity of the matter. However, as time proceeded, it became clear to the author that the current debate on factors influencing car ownership levels and car use has reached consensus in the sense that the only explanation of the current observed trends must be found in the changing influences of factors. This notion has partly defined the purpose of this study, since no studies, among others in The Netherlands, have tried on large-scale (in time, geography and number of factors) to find the (changing) patterns of influence on car ownership levels.

Secondly, the moment the outcomes were analyzed, it was not straightforward clear how the results of this study could be used to create added value for policy makers. The first suggestion was to forecast household car ownership levels in future years using the results of this study. However, based on the data and explanatory variables available for this research such forecasts would most probably be relatively inaccurate. Furthermore, prediction outcomes could not be verified by the existence of real data. Analysis of the current debate showed the importance of endogenous car ownership models to forecast new levels of car ownership, using factors that currently are being questioned with regard to their effect and relevancy. This finding led to the question what implications the results of this study might bring for the endogenous prediction of household car ownership levels on a basic level. To understand what the implications might be, the idea was created to use as much as possible information extracted from the datasets to compare two prediction models that endogenously predict household car ownership levels based on constant and alternative effects of coefficients respectively.

8.2. LIMITATIONS OF STUDY

Throughout this thesis most limitations regarding data, number of variables and types of analyses have already been discussed. This brief section tries to elaborate further on some of the limitations, which are considered to be important.

Firstly, the **research methods** that have been used for this study have proven to be sufficient for explanatory and exploratory purposes. However, further research can be in two domains. Firstly, the ORL and MNL models are both static models. Nevertheless, household decisions regarding car ownership levels tend to be more dynamic over time (de Jong & Kitamura, 2009). Dynamic car ownership models are still in their infancy and face a number of teething troubles. Currently, the most problematic issue is that dynamic models assume, based on utility theory, that for every point in time the household chooses the best set of private cars with the "highest utility". However, in practice, households do not transact cars so often – among other things due to transaction costs. As a result, dynamic models often suffer from including the right household decision processes, whereby the process of household car ownership is accurately formulated (de Jong et al., 2009). However, when these problems are overcome, dynamic models are considered to have greater potential to explain car ownership levels compared to static ORL and MNL models. Furthermore, it is assumed that dynamic are considered to be more sophisticated forecasting models for longer periods of time, especially in case radical changes are expected in the future (e.g. car sharing and rise of electric vehicles).

Secondly, regarding the **data** that has been used for this research some important amendments can be made. Firstly, due to time and budget constraints only the datasets OVG-1987, OVG-1991, OVG-1995, OVG-1999, OVG-2003, OViN-2010 and OViN-2014 datasets have been restructured into one meta-dataset. This study could be further improved by adding other datasets in the given time horizon (1987-2014) or even include datasets that extend the time horizon. Secondly, as pointed out in Chapter 3 the datasets do not fully represent the Dutch population. Especially, the years 1999 and 2003 show consistent deviations from other years studied (e.g. household composition, gender, education and urbanization). The reason for these deviations is that the study comprises a long period of time, which has led to new adaptations in the sample. For example, in 1999 the research method was changed to the so-called "Neu Kontiv Design" to increase the response percentage from 40% to 70% from 1999 onwards (Kadrouch & Moritz, 1998). In 2004, the method changed again to carry out more surveys on workdays and in small provinces. Though weights have been assigned in the datasets, it was regarded to be highly inefficient to analyze a substantial database (millions of cases) that reflects the entire population of The Netherlands. As a result, the results might be affected by an overrepresentation of relatively small subpopulations in the dataset. Descriptive statistics already showed that not all developments of the determinants (e.g. urbanization levels) are equally presented in the population.

Thirdly, we should not underestimate the effect of the exclusion of certain **variables**, such as psychological/attitudinal factors, in this research. Recent studies on psychological determinants on car ownership reveal interesting insights in the psychology of (non)-car owners. The study of Belgiawan et al. (2011) revealed that primarily convenience but also prestige and social orderliness are significant determinants that might create "anti-car" trends among new generations. Besides, environmental concern is a determinant that gains increasing attention from researchers, and has already led to switching behavior of consumers towards electric cars or other modes of transport such as public transport (e.g. Lieven, Mühlmeier, Henkel, & Waller, 2011). Despite such "soft" variables are difficult to measure, it is a considerate shortcoming of this study. To illustrate, the variance of car ownership levels presented by the datasets has only been limitedly explained by the estimated models - 65% as shown in the validation sample.

Fourthly, the **predictive** modelling part of this study faces shortcomings in terms of the number of time stamps for trend extrapolation, the number of predictions and, the variety and multitude of historic data used for the predictions. The estimated model is therefore considered to be too limited and needs further extension to create meaningful predictions. However, it has never been the goal of this study to predict household car ownership levels. The contribution of this study for future research, should be found in the novelty of the proposed prediction model, that allows for alternative car ownership coefficient values over time, which, based on the results of this study, could have more implications than currently is foreseen.

Fifthly, many assumptions have been made with regard to the formulation of **policy implications** by means of a case study on the overestimated (parameter) influence of household income by DYNAMO (Section 6.6). The policy implications should be regarded as highly qualitative – they only *indicate* what might be encountered in short-term and long-term policymaking. As a result, the formulated policy implications need more in-depth revision, both quantitatively and based on more accurate information from policymakers (e.g. updates of programs from political parties).

Bibliography

Armstrong, B. G., & Sloan, M. (1989). Ordinal Regression Models for Epidemologic Data. *American Journal of Epidemiology, 129*(1), 191-204.

Ben-Akiva, M., & Lerman, S. R. (1974). Some estimation results of a simultaneous model of auto ownership and mode choice to work. *Transportation*, *3*(4), 357-376.

Beirão, G., & Cabral, J. S. (2007). Understanding attitudes towards public transport and private car: A qualitative study. *Transport policy*, *14*(6), 478-489.

Belgiawan, P. F., Schmöcker, J. D., & Fujii, S. (2011). Psychological Determinants for Car Ownership Decisions. In *Proceedings of the 16th International Conference of Hong Kong Society for Transportation Studies (HKSTS), Hong Kong.*

Belgiawan, P. F., Schmöcker, J. D., Abou-Zeid, M., Walker, J., Lee, T. C., Ettema, D. F., & Fujii, S. (2014). Car ownership motivations among undergraduate students in China, Indonesia, Japan, Lebanon, Netherlands, Taiwan, and USA. *Transportation*, *41*(6), 1227-1244.

Berri, A. (2009). A cross-country comparison of household car ownership: a cohort analysis. *IATSS research*, *33*(2), 21-38.

Bhat, C. R. (1997). Work travel mode choice and number of non-work commute stops. *Transportation Research Part B: Methodological*, *31*(1), 41-54.

Bhat, C. R., & Pulugurta, V. (1998). A comparison of two alternative behavioral choice mechanisms for household auto ownership decisions. *Transportation Research Part B: Methodological*, *32*(1), 61-75.

Böcker, L., van Amen, P., & Helbich, M. (2016). Elderly travel frequencies and transport mode choices in Greater Rotterdam, The Netherlands. *Transportation*, 1-22.

Böhm, C., & Jacopini, G. (1966). Flow diagrams, turing machines and languages with only two formation rules. *Communications of the ACM*, *9*(5), 366-371.

Bouhuijs, I., & Engelen, L. (2015). *Koopkracht en Inkomen Werkend Nederland, 2012/2013.* Retrieved from Statistics Netherlands website: https://www.cbs.nl/nl-nl/maatwerk/2015/49/-/media/e82919e486594023ae8ee2a2e6f46840.ashx.

De Borger, B., & Mayeres, I. (2007). Optimal taxation of car ownership, car use and public transport: Insights derived from a discrete choice numerical optimization model. *European Economic Review*, *51*(5), 1177-1204.

De Borger, B., Mulalic, I., & Rouwendal, J. (2016). Substitution between Cars within the Household. *Transportation Research Part A: Policy and Practice*, *85*, 135-156.

Bourennane, H., Couturier, A., Pasquier, C., Chartin, C., Hinschberger, F., Macaire, J.-J., & Salvador-Blanes, S. (2014). Comparative Performance of Classification Algorithms for the Development of Models of Spatial Distribution of Landscape Structures. *Geoderma, 219-220*, 136-144.

Centraal Planbureau voor de Leefomgeving (PBL) (2008). *Effecten van omzetting van de aanschafbelasting op personenauto's in een kilometerprijs.* Retrieved from PBL website: http://www.pbl.nl/sites/default/files/cms/publicaties/cpb_document_166.pdf.

Centraal Planbureau voor de Leefomgeving (PBL) (2013). *Ontwikkeling stedelijk gebied in beeld.* http://www.pbl.nl/nieuws/nieuwsberichten/2013/ontwikkeling-stedelijk-gebied. Accessed on July 8th, 2016.

Centraal Planbureau voor de Leefomgeving (PBL) (2015). *Cahier Mobiliteit.* Retrieved from PBL website: http://www.wlo2015.nl/wp-content/uploads/PBL_2015_WLO_Mobiliteit_1686.pdf.

Centraal Planbureau voor de Leefomgeving (PBL) (2016). *Kansrijk mobiliteitsbeleid.* Retrieved from PBL website: http://www.pbl.nl/sites/default/files/cms/publicaties/PBL&CPB_2016_Kansrijk_Mobiliteitsbeleid_2337.PDF.

Cirillo, C., & Xu, R. (2011). Dynamic discrete choice models for transportation. *Transport Reviews*, *31*(4), 473-494.

van Dam, F., de Groot, C., & Verwest, F. (2006). *Krimp en ruimte. Bevolkingsafname, ruimtelijke gevolgen en beleid.* Retrieved from PBL website: http://www.pbl.nl/sites/default/files/cms/publicaties/Krimp_en_ruimte.pdf.

Dargay, J. M. (2001). The effect of income on car ownership: evidence of asymmetry. *Transportation Research Part A: Policy and Practice*, *35*(9), 807-821.

Dargay, J. M. (2002). Determinants of car ownership in rural and urban areas: a pseudo-panel analysis. *Transportation Research Part E: Logistics and Transportation Review*, *38*(5), 351-366.

Delbosc, A., & Currie, G. (2014). Changing demographics and young adult driver license decline in Melbourne, Australia (1994–2009). *Transportation*, *41*(3), 529-542.

Eakins, J. (2013). *The Determinants of Household Car Ownership: Empirical Evidence from the Irish Household Budget Survey* (No. 144). Surrey Energy Economics Centre (SEEC), School of Economics, University of Surrey.

van Essen, H. & Schroten, A. (2008, November). *Hogere prijzen, minder reizen?* Paper presented at Colloquium Vervoersplanologisch Speurwerk 2008*,* Santpoort. Den Haag: PBL.

Fenger, J. (1999). Urban air quality. *Atmospheric environment*, *33*(29), 4877-4900.

Flamm, B. (2009). The impacts of environmental knowledge and attitudes on vehicle ownership and use. *Transportation research part D: transport and environment*, *14*(4), 272-279.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 39-50.

Fosgerau, M., & Jensen, T. C. (2013). A green reform is not always green. *Transportation Research Part C: Emerging Technologies*, *30*, 210-220.

Fujii, S., & Kitamura, R. (2003). What does a one-month free bus ticket do to habitual drivers? An experimental analysis of habit and attitude change. *Transportation*, *30*(1), 81-95.

Geilenkirchen, G., & van Meerkerk, J. (2014). *Milieueffecten van de begrotingsafspraken 2014.* Retrieved from PBL website: http://www.pbl.nl/sites/default/files/cms/publicaties/PBL_2014_Milieueffecten-van-Begrotingsafspraken-2014-(Motie%20Klaver)_1353.pdf.

Golob, T. F. (1990). The dynamics of household travel time expenditures and car ownership decisions. *Transportation Research Part A: General*, *24*(6), 443-463.

Goodwin, P., & van Dender, K. (2013). 'Peak car'—themes and issues. *Transport Reviews*, *33*(3), 243-254.

Groot, W. (2012). *Over brandstofprijzen en automobiliteit.* Retrieved from Kennisinstituut voor Mobiliteit (KiM) website: https://english.kimnet.nl/binaries/kimnet-english/documents/reports/2012/01/04/regarding-fuel-prices-and-automobility-a-brief-analysis-of-price-and-cost-elasticities/over-brandstofprijzen-en-automobiliteit.pdf.

de Groote, J., van Ommeren, J. N., & Koster, H. (2015). Car Ownership and Residential Parking Subsidies: Evidence from Amsterdam.

Hague Consulting Group (1989). *Resource Papers for Landelijk Model, vol. 2*. The Hague: Hague Consulting Group.

Handy, S., Cao, X., & Mokhtarian, P. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research Part D: Transport and Environment*, *10*(6), 427-444.

Janssens, D., Cools, M., Miermans, W., Declercq, K., & Wets, G. (2011). Onderzoek Verplaatsingsgedrag Vlaanderen 4.2 (2009-2010), Verkeerskundige interpretatie van de belangrijkste tabellen (Analyserapport). *Opdracht van de Vlaamse Overheid, Departement Mobiliteit en Openbare Werken. Hasselt: Instituut voor Mobiliteit.*

Johansson-Stenman, O. (2002). Estimating individual driving distance by car and public transport use in Sweden. *Applied Economics*, *34*(8), 959-967.

Johnstone, N., Serret Y., & Bureau C. (2009). *The Determinants of Car Ownership and Use – OECD.* http://www.oecd.org/env/consumption-innovation/43372475.pdf. Accessed on July 4th, 2016.

de Jong, G. (1990). An indirect utility model of car ownership and private car use. *European Economic Review*, *34*(5), 971-985.

de Jong, G., Fox, J., Daly, A., Pieters, M., & Smit, R. (2004). Comparison of car ownership models. *Transport Reviews*, *24*(4), 379-408.

de Jong, G. C., & Kitamura, R. (2009). A review of household dynamic vehicle ownership models: holdings models versus transactions models. *Transportation*, *36*(6), 733-743.

de Jong, G., Kouwenhoven, M., Geurs, K., Bucci, P., & Tuinenga, J. G. (2009). The impact of fixed and variable costs on household car ownership. *Journal of Choice Modelling*, *2*(2), 173-199.

de Jong, G., Kouwenhoven, M., Geurs, K. (2010). Minder BPM, meer auto's? *Tijdschrift Vervoerswetenschap, 46,* 18-24.

Justdan (2016). *Elasticity and Taxes.* https://justdan93.wordpress.com/tag/elasticity-of-supply/. Accessed on September 14th, 2016.

Kadrouch, S. & Moritz, G. (1998). *Redesign onderzoek verplaatsingsgedrag OVG: verschillen tussen het onderzoek verplaatsingsgedrag OVG en het Neu KONTIV Design NKD.* CBS, Heerlen.

Kitamura, R. (2009). A dynamic model system of household car ownership, trip generation, and modal split: model development and simulation experiment. *Transportation*, *36*(6), 711-732.

Kuhnimhof, T., Zumkeller, D., & Chlond, B. (2013). Who made peak car, and how? A breakdown of trends over four decades in four countries. *Transport Reviews*, *33*(3), 325-342.

le Vine, S., & Jones, P. (2012). On the move: making sense of car and train travel trends in Britain.

Lieven, T., Mühlmeier, S., Henkel, S., & Waller, J. F. (2011). Who will buy electric cars? An empirical study in Germany. *Transportation Research Part D: Transport and Environment*, *16*(3), 236-243.

Manting, D. (2014). Ruimtelijk beleid kan niet zonder demografische analyse.

Mathew, A.T. (2015). *Evaluating Logistic Regression Models.* http://www.r-bloggers.com/evaluating-logistic-regression-models/. Accessed on July 18th, 2016.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

Meurs, H. (1990). Trip generation models with permanent unobserved effects. *Transportation Research Part B: Methodological*, *24*(2), 145-158.

Meurs, H. (1993). A panel data switching regression model of mobility and car ownership. *Transportation Research Part A: Policy and Practice*, *27*(6), 461-476.

Meurs, H., & Haaijer, R. (2006, September). *Dynamo: een nieuw dynamisch automarkt model voor Nederland.* Paper presented at Colloquium Vervoersplanologisch Speurwerk 2006*,* Amsterdam. Amersfoort: MuConsult.

Millard-Ball, A., & Schipper, L. (2011). Are we reaching peak travel? Trends in passenger transport in eight industrialized countries. *Transport Reviews*, *31*(3), 357-378.

Miller, E. J., Hunt, J. D., Abraham, J. E., & Salvini, P. A. (2004). Microsimulating urban systems. *Computers, environment and urban systems*, *28*(1), 9-44.

van Mourik, H. (2008). *Toets op het verkeersmodel Landelijk Model System.* Retrieved from Kennisinstituut voor Mobiliteitsbeleid (KiM) website: https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2008/03/03/toets-op-verkeersmodel-landelijk-modelsysteem/toets-op-verkeersmodel-landelijk-modelsysteem.pdf.

MuConsult (2003), *Fiscale maatregelen en mobiliteit. Inventarisatie van bestaande maatregelen. Eindrapport.* MuConsult, Amersfoort.

MuConsult (2016a). *Dynamo automarktmodel.* http://www.muconsult.nl/producten/18-over-ons/producten/45-dynamo-automarktmodel. Accessed on May 12th, 2016.

MuConsult (2016b). *MuConsult levert nieuwe versie automarktmodel Dynamo.* http://www.muconsult.nl/nieuws/nieuws-archief/131-muconsult-levert-nieuwe-versie-automarktmodel-dynamo. Accessed on August 24th, 2016.

Nolan, A. (2010). A dynamic analysis of household car ownership. *Transportation research part A: policy and practice*, *44*(6), 446-455.

Oakil, A. T. M., Ettema, D., Arentze, T., & Timmermans, H. (2014). Changing household car ownership level and life cycle events: an action in anticipation or an action on occurrence. *Transportation*, *41*(4), 889-904.

Oakil, A. T. M., Manting, D., & Nijland, H. (2016a). Determinants of car ownership among young households in the Netherlands: The role of urbanisation and demographic and economic characteristics. *Journal of Transport Geography*, *51*, 229-235.

Oakil, A. T. M., Nijland, L., & Dijst, M. (2016b). Rush hour commuting in the Netherlands: Gender-specific household activities and personal attitudes towards responsibility sharing. *Travel Behaviour and Society*, *4*, 79-87.

OECD (2002). *Income elasticity of demand.* https://stats.oecd.org/glossary/detail.asp?ID=3233. Accessed on September 14th, 2016.

Oi, W. Y., & Shuldiner, P. W. (1963). An Analysis of Urban Travel Demand. *Northwestern University Press, Evanston, IL.*

Ortúzar, J. D. D., & Willumsen, L. G. (2002). Modelling transport. John Wiley & Sons. *West Sussex, England.*

Potoglou, D., & Kanaroglou, P. S. (2008a). Modelling car ownership in urban areas: a case study of Hamilton, Canada. *Journal of Transport Geography*, *16*(1), 42-54.

Potoglou, D., & Kanaroglou, P. S. (2008b). Disaggregate demand analyses for conventional and alternative fueled automobiles: a review. *International journal of sustainable transportation*, *2*(4), 234-259.

Potoglou, D., & Susilo, Y. (2008). Comparison of vehicle-ownership models. *Transportation Research Record: Journal of the Transportation Research Board*, (2076), 97-105.

Rijksinstituut voor Volksgezondheid en Milieu (RIVM) (2013). *Autobezit in grote steden relatief laag.* http://www.zorgatlas.nl/beinvloedende-factoren/fysieke-omgeving/dichtheid-personenauto-s/. Accessed on July 8th, 2016.

Rubin, O., Mulder, C. H., & Bertolini, L. (2014). The determinants of mode choice for family visits–evidence from Dutch panel data. *Journal of Transport Geography*, *38*, 137-147.

Scott, D. M., & Axhausen, K. W. (2006). Household mobility tool ownership: modeling interactions between cars and season tickets. *Transportation*, *33*(4), 311-328.

Smits, W., & de Vries, R. (2013). *Sociaaleconomische trends 2013 – Veranderende beroepsloopbanen van mannen en vrouwen.* https://www.cbs.nl/nr/rdonlyres/e80f22b6-eda2-4361-99f2-a0fd9ae9e78e/0/20131003v4art.pdf. Accessed on July 8th, 2016.

Snellen, D., van Eck, J. R., & de Jong, A. (2016). *Demografische verandering en verstedelijking.* Retrieved from PBL website: http://www.pbl.nl/sites/default/files/cms/publicaties/PBL2016_Demografische-verandering-en-verstedelijking.pdf.

Statistics Netherlands (1983). *Typologie van de nederlandse gemeenten naar urbanisatiegraad op 28 februari 1971.* http://www.volkstellingen.nl/nl/documentatie/1971/hoofdstuk_7_indelingen_van_nederland/urbanisatie.pdf. Accessed on June 29th, 2016.

Statistics Netherlands (2008). *Wat is mijn besteedbaar inkomen*? https://www.cbs.nl/nl-nl/achtergrond/2008/50/wat-is-mijn-besteedbaar-inkomen-. Accessed on September 10th, 2016.

Statistics Netherlands (2013a). *Car ownership rate down among young people.* https://www.cbs.nl/en-gb/news/2013/20/car-ownership-rate-down-among-young-people. Accessed on May 30th, 2016.

Statistics Netherlands (2013b). *Beroepsbevolking; behaalde onderwijs naar persoonskenmerken 2001-2012.* http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=71822NED&D1=0&D2=a&D3=0&D4=1,3-9&D5=0-2,4,8-10&D6=0&D7=a&HDR=T,G6,G2,G5,G3&STB=G1,G4&VW=T. Accessed on July 7th, 2016.

Statistics Netherlands (2014a). *Huishoudens; kindertal, leeftijdsklaase kind, region, 1 januari.* http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=71487NED. Accessed on July 7th, 2016.

Statistics Netherlands (2014b). *Arbeidsdeelname; paren met of zonder kinderen 1992-2013.* http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=71854NED&D1=a&D2=a&D3=a&HDR=G2,G1&STB=T&VW=T. Accessed on July 7th, 2016.

Statistics Netherlands (2015a). *Transport en mobiliteit 2015*. http://download.cbs.nl/pdf/2015-transport-en-mobiliteit-2015.pdf. Accessed on April 6th, 2016.

Statistics Netherlands (2015b). *Inkomensklassen; particuliere huishoudens naar diverse kenmerken*. http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=70958NED&D1=0&D2=0-2&D3=0-7&D4=0-1,8&D5=a&HDR=T,G3,G1,G4&STB=G2&VW=T. Accessed on July 7th, 2016.

Statistics Netherlands (2015c). *Bevolking; kerncijfers*. http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=37296ned&D1=52-55&D2=a&HDR=G1&STB=T&VW=T. Accessed on July 7th, 2016.

Statistics Netherlands (2015d). *Gemiddeld inkomen; particuliere huishoudens naar diverse kenmerken*. http://statline.cbs.nl/Statweb/publication/. ?DM=SLNL&PA=70843ned&D1=a&D2=0-1,8-9,15,18&D3=0-4&D4=a&HDR=G1,G2,T&STB=G3&VW=T. Accessed on August 23 th, 2016.

Statistics Netherlands (2015e). *Documentatierapport Onderzoek verplaatsingen in Nederland (OViN)*. Retrieved from Statistics Netherlands website: https://www.cbs.nl/NR/rdonlyres/E9A10D87-28AD-4540-9A16-7167DB83B838/0/ovinmicrodata.pdf.

Statistics Netherlands (2015f). *Koopkracht gestegen na vier jaar daling*. https://www.cbs.nl/nl-nl/nieuws/2015/37/koopkracht-gestegen-na-vier-jaar-daling. Accessed on September 14th, 2016.

Statistics Netherlands (2016a). *Motorvoertuigenpark; type, leeftijdsklasse, 1 januari*. http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=82044ned&D1=a&D2=a&D3=a&HDR=T&STB=G1,G2&VW=T. Accessed on April 6th, 2016.

Statistics Netherlands (2016b). *Bevolking; kerncijfers*. http://statline.cbs.nl/statweb/publication/?vw=t&dm=slnl&pa=37296ned&d1=0-2,8-13,19-21,25-35,52-56,68&d2=0,10,20,30,40,50,60,64-65&hd=151214-1132&hdr=g1&stb=t. Accessed on May 10th, 2016.

Statistics Netherlands (2016c). *Voertuigbezit van personen en huishoudens*. https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/voertuigbezit-van-personen-en-huishoudens. Accessed on May 25th, 2016.

Statistics Netherlands (2016d). *Bevolking; geslacht, leeftijd, burgelijker staat en region, 1 januri*. http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=7461BEV&D1=0&D2=a&D3=101-120&D4=4-66&HDR=T,G3&STB=G1,G2&VW=T. Accessed on July 7th, 2016.

Statistics Netherlands (2016e). *Bevolkingspyramide*. https://www.cbs.nl/nl-nl/visualisaties/bevolkingspiramide. Accessed on July 7th, 2016.

Statistics Netherlands (2016f). *Beschikbaar inkomen van huishoudens stijgt*. https://www.cbs.nl/nl-nl/nieuws/2016/12/beschikbaar-inkomen-van-huishoudens-stijgt. Accessed on September 10th, 2016.

Statistics Netherlands (2016g). *Bestedingen; consumptie huishoudens*. http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=70076ned&D1=0-2&D2=0,2-3,16,22,24,28&D3=16,33,50,67,84,101,118,135,152,169,186,203,220,237,254,271,288,305,322,339,344-364&HDR=T&STB=G1,G2&VW=T. Accessed on September 14th, 2016.

Steyerberg, E., & Harrell, F. (2002). *Validation of predictive regression models*. University Medical Center Rotterdam. Rotterdam, Netherlands.

Stock, J. H., & Watson, M. W. (2007). *Introduction to Econometrics* (2nd ed.). Boston, Massachusetts: Pearson Addison Wesley.

SWOV (2013). *SWOV Fact Sheet – Mobility on Dutch roads*. SWOV, Leidschendam.

SWOV (2016). *Mobility of persons*. https://www.swov.nl/UK/Research/Cijfers/ Toelichting-gegevensbronnen/Personenmobiliteit-UK.html. Accessed on May 24th, 2016.

Tobie, J., & Houthakker, H. S. (1950). The effects of rationing on demand elasticities. *The Review of Economic Studies*, *18*(3), 140-153.

Traa, M., Geilenkirchen, G., & Hilbers, H. (2014). *Kortetermijnramingsmodel voor het bezit en gebruik van personenauto's in Nederland*. PBL, Den Haag.

Trochim, W. M. K. (2006). *Threats to Conclusion Validity*. http://www.socialresearch methods.net/kb/concthre.php. Accessed on July 12th, 2016.

Uncles, M. D. (1987). A beta-logistic model of mode choice: Goodness of fit and intertemporal dependence. *Transportation Research Part B: Methodological*, *21*(3), 195-205.

van der Waard, J., Jorritsma, P., & Immers, B. (2013). New drivers in mobility; what moves the Dutch in 2012?. *Transport Reviews*, *33*(3), 343-359.

Vermunt, J. K., & Magidson, J. (2005). Latent GOLD® choice 4.0 user's manual. *Statistical Innovations Inc., Belmont, MA.*

Wrigley, N. (1990). Dirichlet-logistic models of spatial choice. *Spatial Choices and Processes, edited by MM Fischer, P. Nijkamp and Y. Papageorgiou, North-Holland, Amsterdam.*

Wang, Y. (2005). A Multinomial Logistic Regression Modeling Approach for Anomaly Intrusion Detection. *Computers & Security, 24*, 662-674.

Worldwide Inflation Data (2016). *Historische inflatie Nederland – CPI inflatie.* Retrieved from http://nl.inflation.eu/inflatiecijfers/nederland/historische-inflatie/cpi-inflatie-nederland.aspx.

Wu, N., Zhao, S., & Zhang, Q. (2016). A study on the determinants of private car ownership in China: Findings from the panel data. *Transportation Research Part A: Policy and Practice*, *85*, 186-195.

Zwaneveld, P., Hilbers, H., Weijschede- van der Straaten, W., Zondag, B., van Meerkerk, J., & Verrips, A. (2012). Retrieved from PBL website: https://www.cpb.nl/sites/default/files/publicaties/download/cpb-pbl-achtergronddocument-keuzes-kaart-2013-2017-bereikbaarheidseffecten.pdf.

Appendix I – Data Structuring

This Appendix describes on a step-by-step basis the procedures that have been executed to prepare and structure the data. The Appendix will discuss how variables have been created and re-categorized. The section will also describe which data has been extracted from the surveys to execute the logistic regression analyses and validation analysis.

STEP 1

As a first step of the data structuring procedure an extra variable for the OVG 1987 and 1991 has been created to count the number of children per household, which has not been provided by the datasets. The creation of this variable is essential to obtain insight in the household composition (i.e. whether or not a household obtains children). By using various functions in Excel (e.g. IF, COUNTIF, VLOOKUP) the number of children per household has been calculated (see Fig I).

BZ	CA	CB	CC	CD	CE	CF
Check if Unique Child	ID #	multiplication factor	count cases	correct for #	vlookup	children per household
0	150010	0	84063	0	#N/A	0
0	150010	0	84063	0	#N/A	0
0	150010	0	84063	0	#N/A	0
0	150010	0	84063	0	#N/A	0
0	150010	0	84063	0	#N/A	0
0	150010	0	84063	0	#N/A	0
0	150037	0	84063	0	1	1
0	150037	0	84063	0	1	1
0	150037	0	84063	0	1	1
1	150037	150037	1	1	1	1
0	150045	0	84063	0	=VLOOKUP(CA12;'Vlookup-sheet'!$A$1:$B$7267;2;FALSE)	
0	150045	0	84063	0	VLOOKUP(lookup_value; table_array; col_index_num; [range_lookup])	
0	150045	0	84063	0	#N/A	0
0	150045	0	84063	0	#N/A	0
0	150045	0	84063	0	#N/A	0

Fig. I. Example of Excel procedure to create extra variable "children per household" – OVG 1987

STEP 2

Hereafter, for all datasets only unique household members have been selected for regression analyses and validation analysis. Some members have been asked to fill the questionnaire more than once. Nevertheless, this research is only interested in characteristics of households and individual persons so that extra cases per individual do not provide additional information for analysis. The procedure to clear extra cases per unique household members has been executed by SPSS.

STEP 3

The third step comprised the selection of a unique person per head of the family for all seven datasets (OVG 1987, 1991, 1995, 1999, 2003 & OViN 2010 & 2014). The procedure, which has been performed by Excel, started selecting unique persons (no replications) of the household and the person of reference that was assigned as "head of the family" for the "position in household" variable. However, for the OVG 1987 dataset alone, this procedure resulted that only 381 cases of the total of 25052 unique members that were assigned as head of the family were female which would led to a distorted distribution of men and women in the dataset. The same issue would apply for the OVG 1991 dataset. Therefore another more detailed procedure had been executed. Firstly, for the OVG 1987 and 1991 datasets "children", "extra family members" and "unknowns" were cleared. Next, all unique households were selected (i.e. singles and households with only one single row of data input).

B	C	D	K	L	M	N	O	P	Q
volgnr	urbgr	hhgr	maatpa	children per househol	household with childre	Selection of unique household ID's for non -singles			
150010	32	2	3	0	0	2			
150010	32	2	2	0	0	2			
150088	32	2	5	0	0	2			
150088	32	2	3	0	0	2			
150150	35	2	8	0	0	1			
150231	35	4	3	1	1	2			
150231	35	4	2	1	1	2			
150274	35	2	7	0	0	2			
150274	35	2	7	0	0	2			
150282	35	4	2	2	1	=COUNTIF(B:B;B11)			
150312	35	4	3	2	1	COUNTIF(range; criteria)			
150312	35	4	8	2	1	2			
150320	35	2	3	0	0	2			
150320	35	2	2	0	0	2			
150355	35	3	3	0	0	2			
150355	35	3	2	0	0	2			

Fig. II. Example of Excel procedure to identify "selection of unique household ID's for non-singles" – OVG 1987

Hereafter, for the selection of "head of the family" and "partner", which all consisted of two rows of data entry, randomly either males or females were selected. This procedure has led to a more realistic distribution of men and women in the dataset for analysis purposes. Considering the other datasets (from 1995 onwards) this procedure was not necessary, because these datasets also registered unique household identification numbers, which after selecting resulted in a relatively equal distribution of men and women.

STEP 4

All the variables the datasets comprise that are no use for this study have been cleared. Such variables do not need to be analyzed by the ORL model, but clearing them will increase the calculation speed of the model (every dataset consists of an order of magnitude of thousands unique cases). In the end, the datasets were reduced to following variables: year of reference, identification number of household, household income, household size, household composition, age of reference person, education of reference person, gender of reference person, working status of reference person, level of urbanization and number of cars counted per household. For the OVG 1987 and 1991 the same procedure has been executed with the exception that the variable "household composition" was substituted by "household with children".

STEP 5

As a fifth step all children (0-17 years) have been cleared from the datasets. Hence, children under 18 in The Netherlands are legally not permitted to buy or obtain a car. This group is therefore not relevant to be analyzed during the research.

STEP 6

Subsequently, there has been searched for "unknowns" for all entries in all the datasets. Every entry with an "unknown" has been cleared from the list for all variables of interest (i.e. list-wise deletion). To illustrate, in the 1987 and 1991 OVG datasets also provided the option of "unknown gender" next to the option "male" and "female". Because this research is only interested in the effects of gender on car ownership levels among households such unknowns have been removed. Also human errors, such as unique persons that filled in the dataset twice, were corrected for. Finally, within every dataset there has been searched for missing values. No missing values have been found via the explore-procedure of SPSS in the datasets (see Fig III).

	Cases		
	Valid	Missing	Total
	N	N	N
Household income	163048	0	163048
Household size	163048	0	163048
Children in household	163048	0	163048
Gender	163048	0	163048
Age	163048	0	163048
Education	163048	0	163048
Working status	163048	0	163048
Urbanization	163048	0	163048
Number of private cars	163048	0	163048

Fig. III. SPSS procedure to identify missing values

STEP 7

The next steps discuss the most important changes that have been made to the variables of interest. As a first step the household income has been measured by using a two-step procedure. Firstly, the income for every unique household has been calculated by taking the average number of income per category. For example, when a household earns between 20,000 euro and 30,000 euro it is assumed that the household earns 25,000 euro. This measure was necessary to create a new categorization of income. Secondly, for the 1987/1991/1995/1999 OVG datasets the currency "gulden" has been converted to "euro", using an exchange rate of 2.2.

STEP 8

Considering the household composition for the 1995/1991/2003 OVG and 2010/2014 OViN datasets the number of children per household have been counted by using the "HHSam" (proxy for household composition) and "HHlft1/HHlft2/HHlft3" (proxies for counted cases of children) variables. The procedure to calculate the number of children for the 1987/1991 OVG datasets is described in STEP 1.

STEP 9

With regard to working status the OVG 1987/1991/1995 have been re-categorized into three categories: "no employment", "employment less than 30 hours per week" and "employment more than 30 hours per week". It is assumed that individuals that fall into the categories student/scholar, retired or WAO (law for disability insurance) are non-employed. Considering the OVG 1999/2003 and OViN 2010/2014 the categories "employment less than 12 hours per week" and "employment 12-30 hours per week" have been merged into "employment less than 30 hours per week", because this "employment less than 12 hours per week" is not provided by the other datasets.

STEP 10

The variable urbanization for the years 1987 and 1991 has been re-categorized. Both datasets use an outdated system to measure urbanization (i.e. rural communities, urbanized rural communities and urbanized municipalities). The three categories have been re-organized into the following five categories based on Statistics Netherlands: i) very high-density areas (≥ 2,500 addresses per km^2); ii) high density areas (1,500-2,500 addresses per km^2); iii) moderately high density areas (1,000-1,500 addresses per km^2); iv) low density areas with (500-1,000 addresses per km^2); and v) very low density areas (<500 addresses per km^2). This alternation has been done on the basis of documentation on censuses of Statistics Netherlands (Statistics Netherlands, 1983).

STEP 11

The number of cars owned by a household has been re-categorized for the OVG datasets from 1987 till 2003. All categories comprising motor, scooter, bicycle or other vehicles have been merged into the category "zero cars" next to the already existing categories of 1, 2 and 3 and more private cars. Considering the other variables including the size of the household, gender, age and educational level of the reference person no further amendments needed to be made.

STEP 12

After restructuring the data and re-categorizing several variables 20% of the cases has been extracted from the original dataset. These cases are used as a "validation population" to test the predicting accuracy of the ORL model during the validation procedure. The selection of 20% of the total number of cases has been carried out via an SPSS procedure that selects cases based on random order.

STEP 13

Finally, by translating SPSS files into Excel files and vice versa some information can be lost on explanatory and dependent variables. For the reason that labels and values, which are normally assigned in the SPSS files, are lost once the file has been translated into an Excel file. As a consequence, the values and labels of the variables have been (manually) assigned to the relevant variables in the adapted SPSS datasets.

Appendix II – Descriptive Statistics (1)

This Appendix gives a detailed overview of some descriptive statistics of the eight independent and the dependent variable household car ownership in the years 1987, 1991, 1995, 1999, 2003, 2010, and 2014. The total number of cases, minima, maxima, means and standard deviations are given.

I. 1987 - DESCRIPTIVE STATISTICS

Table I. Descriptive statistics – OVG 1987

	N	Minimum	Maximum	Mean	Std. Deviation
1. Household income	6623	1	3	2.25	.775
2. Household size	6623	1	6	2.78	1.307
3. Household composition	6623	1	2	1.41	.492
3. Gender	6623	1	2	1.55	.498
4. Age	6623	1	7	3.92	1.394
5. Education	6623	1	4	2.38	.938
6. Working status	6623	1	3	1.86	.946
7. Urbanization	6623	1	5	2.40	.956
8. Number of private cars	6623	0	3	.95	.599
Total cases	**6623**				

II. 1991 - DESCRIPTIVE STATISTICS

Table II. Descriptive statistics – OVG 1991

	N	Minimum	Maximum	Mean	Std. Deviation
1. Household income	5955	1	3	2.28	.639
2. Household size	5955	1	6	2.71	1.289
3. Household composition	5955	1	2	1.36	.481
4. Gender	5955	1	2	1.53	.499
5. Age	5955	1	7	3.96	1.351
6. Education	5955	1	4	2.51	.932
7. Working status	5955	1	3	1.91	.944
8. Urbanization	5955	1	5	2.43	.941
9. Number of private cars	5955	0	3	1.01	.576
Total cases	**5955**				

III. 1995 - DESCRIPTIVE STATISTICS

Table III. Descriptive statistics – OVG 1995

	N	Minimum	Maximum	Mean	Std. Deviation
1. Household income	46014	1	3	2.55	.699
2. Household size	46014	1	6	2.61	1.294
3. Household composition	46014	1	2	1.34	.475
4. Gender	46014	1	2	1.56	.496
5. Age	46014	1	7	4.09	1.381
6. Education	46014	1	4	2.65	.949
7. Working status	46014	1	3	1.83	.918
8. Urbanization	46014	1	5	3.07	1.340
9. Number of private cars	46014	0	3	1.03	.616
Total Cases	**46014**				

IV. 1999 - DESCRIPTIVE STATISTICS

Table IV. Descriptive statistics – OVG 1999

	N	Minimum	Maximum	Mean	Std. Deviation
1. Household income	38316	1	4	3.07	.961
2. Household size	38316	1	6	2.34	1.269
3. Household composition	38316	1	2	1.29	.452
4. Gender	38316	1	2	1.35	.478
5. Age	38316	1	7	4.44	1.459
6. Education	38316	1	4	2.63	1.009
7. Working status	38316	1	3	2.14	.937
8. Urbanization	38316	1	5	2.91	1.318
9. Number of private cars	38316	0	3	1.00	.689
Total Cases	**38316**				

V. 2003 - DESCRIPTIVE STATISTICS

Table V. Descriptive statistics – OVG 2003

	N	Minimum	Maximum	Mean	Std. Deviation
1. Household income	18971	1	4	3.18	.939
2. Household size	18971	1	6	2.33	1.266
3. Household composition	18971	1	2	1.29	.453
4. Gender	18971	1	2	1.35	.478
5. Age	18971	2	7	4.65	1.391
6. Education	18971	1	4	2.68	1.013
7. Working status	18971	1	3	2.09	.936
8. Urbanization	18971	1	5	2.86	1.300
9. Number of private cars	18971	0	3	1.06	.705
Total Cases	**18971**				

VI. 2010 - DESCRIPTIVE STATISTICS

Table VI. Descriptive statistics – OVG 2010

	N	Minimum	Maximum	Mean	Std. Deviation
1. Household income	23361	1	6	4.10	1.427
2. Household size	23361	1	6	2.57	1.253
3. Household composition	23361	1	2	1.32	.467
4. Gender	23361	1	2	1.53	.499
5. Age	23361	1	7	4.51	1.378
6. Education	23361	1	4	2.92	.918
7. Working status	23361	1	3	1.99	.899
8. Urbanization	23361	1	5	3.05	1.322
9. Number of private cars	23361	0	3	1.26	.722
Total Cases	**23361**				

VII. 2014 - DESCRIPTIVE STATISTICS

Table VII. Descriptive statistics – OVG 2014

	N	Minimum	Maximum	Mean	Std. Deviation
1. Household income	23353	1	6	4.12	1.438
2. Household size	23353	1	6	2.49	1.212
3. Household composition	23353	1	2	1.29	.455
4. Gender	23353	1	2	1.53	.499
5. Age	23353	1	7	4.67	1.385
6. Education	23353	1	4	2.99	.899
7. Working status	23353	1	3	1.95	.900
8. Urbanization	23353	1	5	3.01	1.311
9. Number of private cars	23353	0	3	1.25	.735
Total Cases	**23353**				

Appendix III – Descriptive Statistics (2)

This Appendix gives an overview of some descriptive statistics of the eight independent and the dependent variable household car ownership in the years 1987, 1991, 1995, 1999, 2003, 2010, and 2014. In this appendix the measured quantities expressed in percentages for every sub-category are presented by means of bar charts. The figures give a clear picture how much every sub-category is presented in the datasets from 1987 onwards till 2014. This information has been used for the descriptive analyses in Chapter 4. Subcategories that are not presented by the datasets have been left out of the visualization.

Table VIII. Descriptive statistics – household income distributions

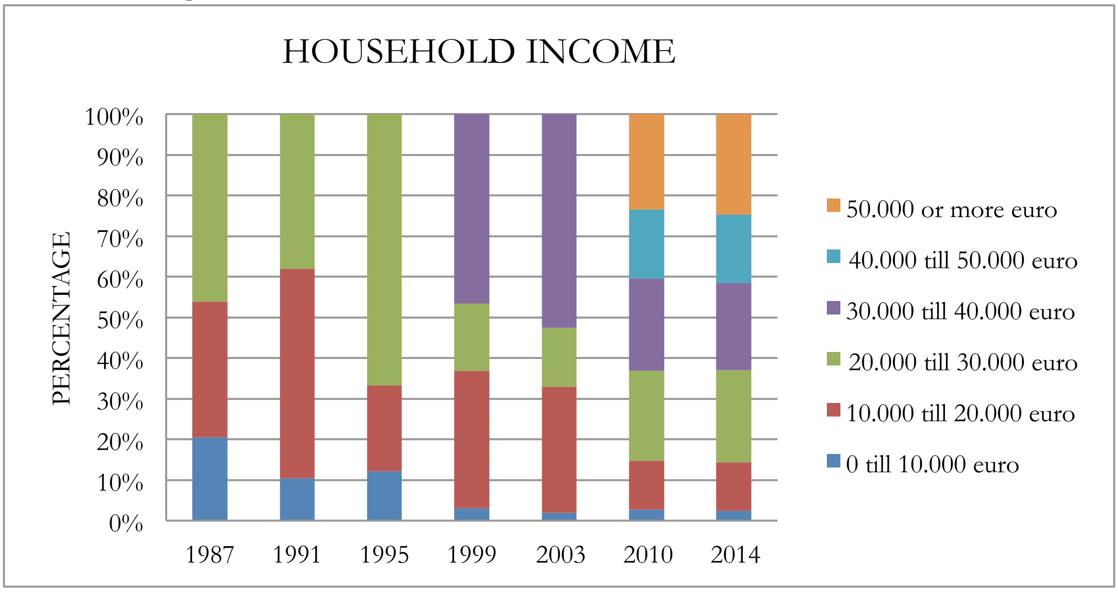


Table IX. Descriptive statistics – household size distributions

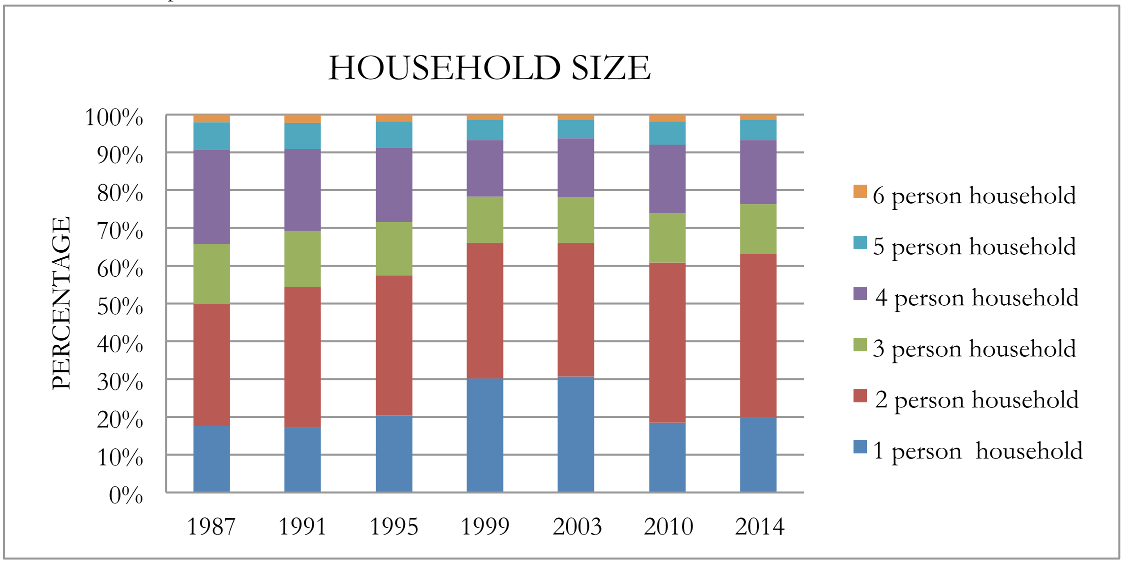


Table X. Descriptive statistics – household composition distributions

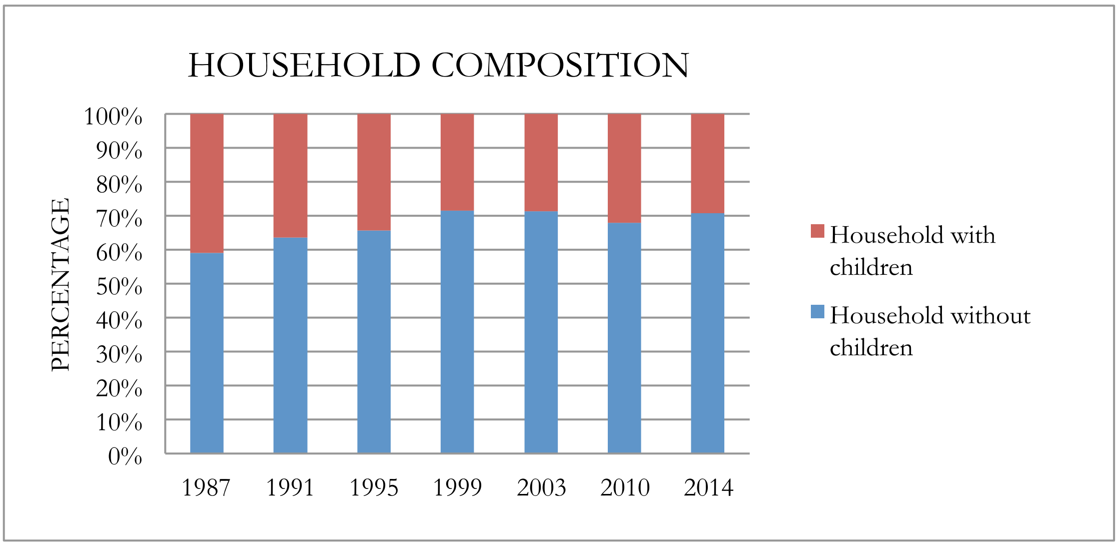


Table XI. Descriptive statistics – gender distributions

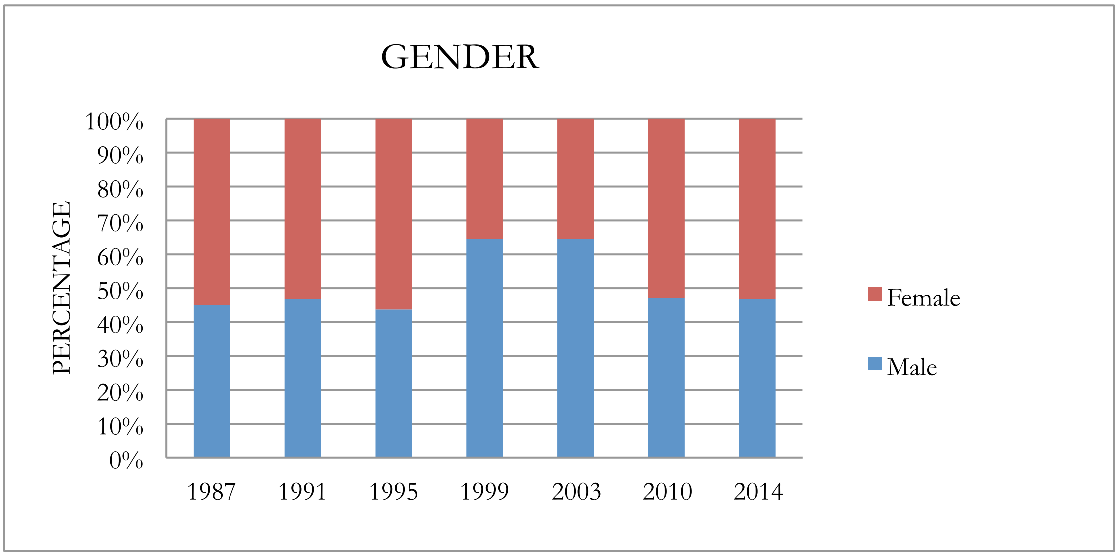


Table XII. Descriptive statistics – working status distributions

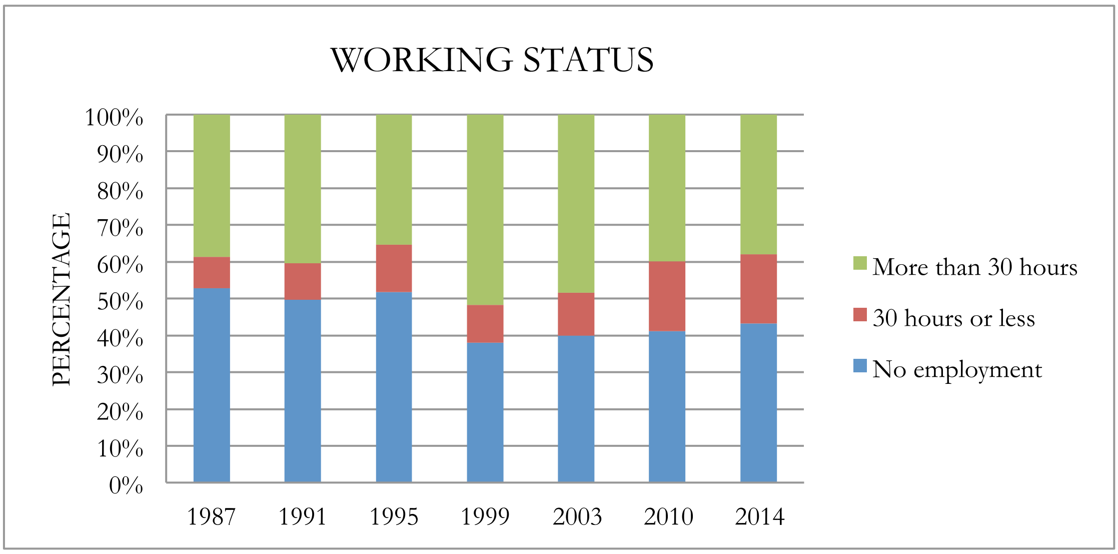


Table XIII. Descriptive statistics – urbanization distributions

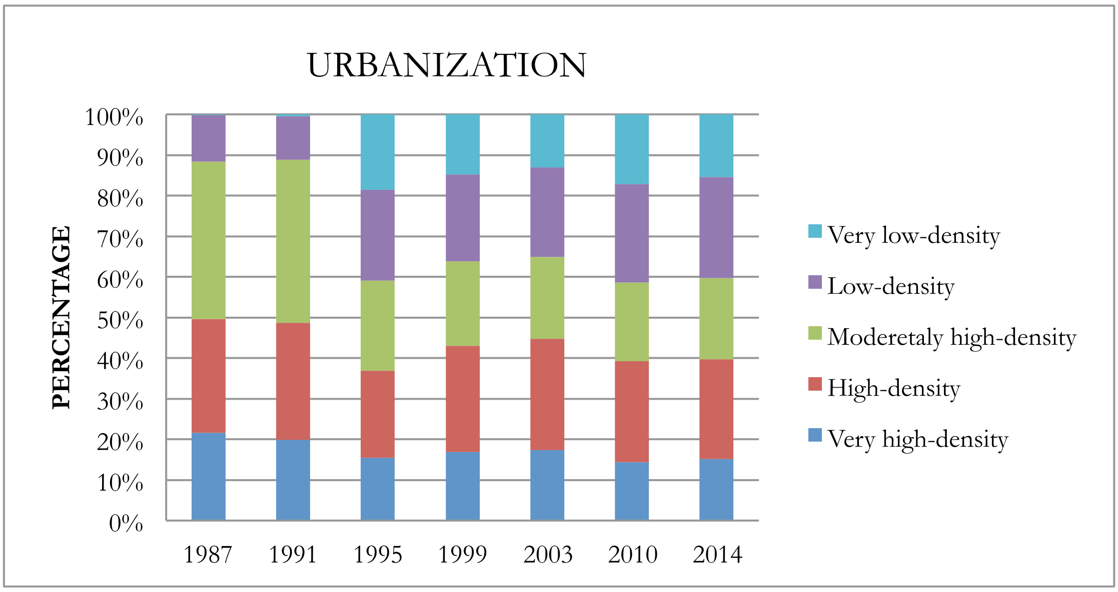


Table XIV. Descriptive statistics – age distributions

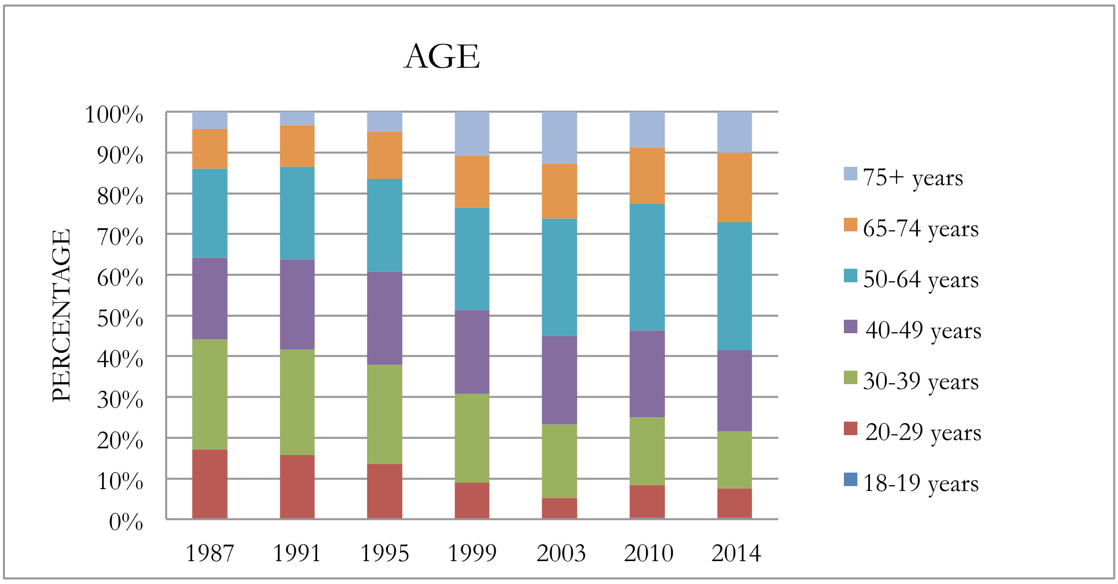
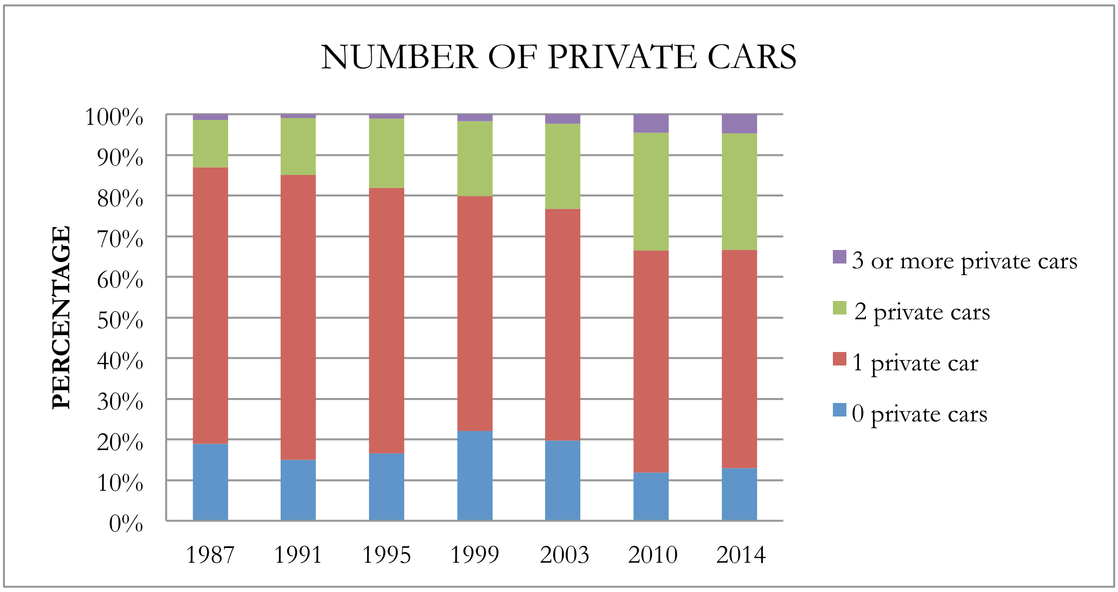


Table XIV. Descriptive statistics – private car distributions



Appendix IV – Model Fitting

This Appendix gives information on the model fitting, goodness of fit and (pseudo) R-square of the ORL model that has been used for this research. The results will be shown per model fitting test for all explanatory variables that have been included in the ORL model.

I. ORDER LOGIT REGRESSION – MODEL FITTING INFORMATION

The model fitting information of the ORL models gives answer to the question whether the model does improve the ability to predict the model outcomes. The -2 Log Likelihood compares what the actual outcome is compared to the probability that the model predicts the outcome. The significance level of the Chi-Square test of 0.000 shows that there is a statistically improvement of the model by including all the explanatory variables in the model.

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	192869.251			
Final	117895.025	74974.226	56	**0.000**

Table XV. ORL model fitting information – Explanatory variables on household car ownership

II. ORDER LOGIT REGRESSION – GOODNESS-OF-FIT

The Goodness-of-Fit table gives answer to the question whether the observed data is consistent with the model that is fitted. The null-hypothesis is that the model is a good fit. However, the significance level of the Pearson chi-square of 0.000 means that the data does not fit very well with the model, which gives a less positive picture. However, one needs to be very careful interpreting the results of the Goodness-of-Fit table. In large sample sizes we may find statistical significance, whereas findings are small and uninteresting (i.e. findings are not substantively significant) (Fornell & Larcker, 1981). Because sample sizes of the datasets differentiate between 5,955 and 46,014 cases, it is assumed that the Model Fitting Information gives enough information that the model is improving the outcome, regardless of the Goodness-of-Fit information. Moreover, the deviance shows a significance level of 1.000 meaning that null-hypothesis will be maintained resulting that the model is a good fit.

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	156535.324	96979	**0.000**
Deviance	83044.339	96979	**1.000**

Table XVI. ORL Goodness of Fit – Explanatory variables on household car ownership

III. ORDER LOGIT REGRESSION – PSEUDO R-SQUARE

The Pseudo R-Square (Cox and Snell, Nagelkerke and McFadden) is significant. For example, the Nagelkerke Pseudo R-Square of 0.423 is considered to be a relatively high. The estimated Cox and Snell and McFadden R-Squares of 0.369 respectively 0.224 are also considerate. Concluding, all three indicators show that a considerate proportion of the variance in the model outcome is explained by the explanatory variables on household car ownership.

Pseudo R-Square

Cox and Snell	**0.369**
Nagelkerke	**0.423**
McFadden	**0.224**

Table XVII. Pseudo R-square of ORL Model - Explanatory variables on household car ownership

Appendix V – MNL Model Estimates

This Appendix shows the parameter estimates of the constructed MNL model to predict household car ownership levels in 2014, based on data from 1987, 1991, 1995, 1999, 2003 and 2010. In the most left column the car ownership level is presented (zero, one or two private cars). The second column provides the coefficient estimates. The fifth column presents their significance level. The coefficients of 1991-2010 can be calculated by summing the main effect (1987) with the specific interaction effect of the year of reference. The coefficient value of household income in 2010, for the category of zero cars, is calculated as follows: *INCOME + INCOME_2010 = -2.279 + 1.110 = -1.169.*

Table XXII. Ordered logistic regression models estimates

						95% Confidence Interval for Exp(B)	
Number of private cars		B	Std. Error	Wald	Sig.	Lower Bound	Upper Bound
0	Intercept	**11.999**	.257	2187.221	**.000**		
	INCOME	**-2.279**	.268	72.609	**.000**	.061	.173
	INCOME_1991	**-.367**	.415	.780	**.377**	.307	1.564
	INCOME_1995	**.349**	.297	1.387	**.239**	.793	2.537
	INCOME_1999	**.383**	.277	1.923	**.166**	.853	2.523
	INCOME_2003	**.463**	.284	2.658	**.103**	.911	2.773
	INCOME_2010	**1.110**	.270	16.928	**.000**	1.788	5.146
	SIZE	**-2.250**	.119	355.526	**.000**	.083	.133
	SIZE_1991	**-.102**	.188	.292	**.589**	.625	1.306
	SIZE_1995	**-.366**	.131	7.762	**.005**	.536	.897
	SIZE_1999	**.041**	.129	.100	**.751**	.809	1.341
	SIZE_2003	**-.128**	.137	.877	**.349**	.673	1.150
	SIZE_2010	**.176**	.130	1.835	**.176**	.924	1.540
	COMPOSITION	**3.806**	.322	139.767	**.000**	23.935	84.559
	COMPOSITION_1991	**1.715**	.584	8.623	**.003**	1.769	17.460
	COMPOSITION_1995	**1.497**	.357	17.584	**.000**	2.220	8.996
	COMPOSITION_1999	**.154**	.346	.199	**.656**	.592	2.301
	COMPOSITION_2003	**.261**	.366	.507	**.477**	.633	2.660
	COMPOSITION_2010	**-.697**	.346	4.062	**.044**	.253	.981
	GENDER	**.286**	.248	1.329	**.249**	.819	2.164
	GENDER_1991	**-.285**	.383	.557	**.456**	.355	1.591
	GENDER_1995	**-.155**	.267	.336	**.562**	.508	1.445
	GENDER_1999	**.124**	.265	.220	**.639**	.674	1.903
	GENDER_2003	**.129**	.275	.219	**.640**	.663	1.950
	GENDER_2010	**-.139**	.257	.293	**.588**	.526	1.440
	AGE	**-.106**	.111	.916	**.338**	.724	1.118
	AGE_1991	**-.123**	.169	.534	**.465**	.635	1.231

	AGE_1995	**-.153**	.120	1.646	**.200**	.679	1.084
	AGE_1999	**.146**	.116	1.587	**.208**	.922	1.451
	AGE_2003	**.157**	.120	1.714	**.191**	.925	1.479
	AGE_2010	**-.030**	.114	.070	**.791**	.777	1.212
	EDUCATION	**.281**	.132	4.532	**.033**	1.023	1.716
	EDUCATION_1991	**-.223**	.206	1.178	**.278**	.534	1.197
	EDUCATION_1995	**-.216**	.142	2.306	**.129**	.610	1.065
	EDUCATION_1999	**-.534**	.139	14.773	**.000**	.447	.770
	EDUCATION_2003	**-.559**	.144	15.128	**.000**	.432	.758
	EDUCATION_2010	**-.527**	.139	14.394	**.000**	.449	.775
	WORKINGSTATUS	**-.357**	.137	6.848	**.009**	.535	.914
	WORKINGSTATUS_1991	**.126**	.212	.352	**.553**	.748	1.720
	WORKINGSTATUS_1995	**.043**	.148	.084	**.772**	.780	1.396
	WORKINGSTATUS_1999	**-.197**	.147	1.803	**.179**	.616	1.095
	WORKINGSTATUS_2003	**-.304**	.154	3.912	**.048**	.546	.997
	WORKINGSTATUS_2010	**-.485**	.144	11.291	**.001**	.464	.817
	URBANIZATION	**-.694**	.117	35.412	**.000**	.397	.628
	URBANIZATION_1991	**.006**	.192	.001	**.974**	.691	1.465
	URBANIZATION_1995	**.046**	.122	.142	**.706**	.824	1.331
	URBANIZATION_1999	**.059**	.121	.240	**.624**	.837	1.344
	URBANIZATION_2003	**-.041**	.123	.108	**.742**	.754	1.223
	URBANIZATION_2010	**-.116**	.120	.940	**.332**	.703	1.126
1	Intercept	**9.574**	.248	1491.320	**.000**		
	INCOME	**-1.474**	.264	31.214	**.000**	.136	.384
	INCOME_1991	**.011**	.408	.001	**.978**	.455	2.250
	INCOME_1995	**.538**	.293	3.377	**.066**	.965	3.041
	INCOME_1999	**.309**	.272	1.290	**.256**	.799	2.324
	INCOME_2003	**.337**	.279	1.453	**.228**	.810	2.421
	INCOME_2010	**.805**	.265	9.194	**.002**	1.329	3.761
	SIZE	**-1.655**	.108	236.263	**.000**	.155	.236
	SIZE_1991	**-.095**	.171	.307	**.579**	.650	1.272
	SIZE_1995	**-.354**	.119	8.885	**.003**	.556	.886
	SIZE_1999	**.266**	.115	5.372	**.020**	1.042	1.635
	SIZE_2003	**.183**	.119	2.349	**.125**	.950	1.518
	SIZE_2010	**.553**	.113	23.803	**.000**	1.392	2.171
	COMPOSITION	**3.434**	.299	131.454	**.000**	17.229	55.730
	COMPOSITION_1991	**1.482**	.556	7.113	**.008**	1.481	13.075
	COMPOSITION_1995	**1.413**	.333	18.012	**.000**	2.140	7.893
	COMPOSITION_1999	**-.282**	.320	.776	**.378**	.403	1.412
	COMPOSITION_2003	**-.143**	.333	.184	**.668**	.451	1.665
	COMPOSITION_2010	**-1.396**	.314	19.788	**.000**	.134	.458
	GENDER	**.244**	.242	1.023	**.312**	.795	2.050
	GENDER_1991	**-.387**	.373	1.076	**.300**	.327	1.411

	GENDER_1995	**-.444**	.260	2.909	**.088**	.385	1.068
	GENDER_1999	**-.386**	.258	2.234	**.135**	.410	1.127
	GENDER_2003	**-.475**	.267	3.178	**.075**	.369	1.048
	GENDER_2010	**-.521**	.248	4.408	**.036**	.365	.966
	AGE	**-.248**	.109	5.140	**.023**	.630	.967
	AGE_1991	**.011**	.167	.004	**.947**	.729	1.402
	AGE_1995	**-.054**	.118	.208	**.648**	.752	1.194
	AGE_1999	**.213**	.114	3.475	**.062**	.989	1.547
	AGE_2003	**.228**	.118	3.738	**.053**	.997	1.583
	AGE_2010	**.142**	.112	1.618	**.203**	.926	1.435
	EDUCATION	**.463**	.128	13.083	**.000**	1.236	2.042
	EDUCATION_1991	**-.362**	.200	3.288	**.070**	.471	1.030
	EDUCATION_1995	**-.330**	.138	5.737	**.017**	.548	.942
	EDUCATION_1999	**-.471**	.134	12.297	**.000**	.480	.812
	EDUCATION_2003	**-.426**	.138	9.490	**.002**	.498	.856
	EDUCATION_2010	**-.456**	.134	11.661	**.001**	.488	.823
	WORKINGSTATUS	**-.076**	.132	.332	**.564**	.716	1.200
	WORKINGSTATUS_1991	**.000**	.205	.000	**.998**	.669	1.495
	WORKINGSTATUS_1995	**-.095**	.143	.436	**.509**	.687	1.204
	WORKINGSTATUS_1999	**-.213**	.142	2.273	**.132**	.612	1.066
	WORKINGSTATUS_2003	**-.329**	.147	5.008	**.025**	.539	.960
	WORKINGSTATUS_2010	**-.501**	.137	13.286	**.000**	.463	.793
	URBANIZATION	**-.297**	.112	6.966	**.008**	.596	.926
	URBANIZATION_1991	**-.024**	.185	.016	**.899**	.679	1.405
	URBANIZATION_1995	**.038**	.118	.106	**.745**	.825	1.309
	URBANIZATION_1999	**.027**	.116	.054	**.816**	.818	1.290
	URBANIZATION_2003	**-.037**	.118	.097	**.756**	.764	1.215
	URBANIZATION_2010	**-.108**	.115	.891	**.345**	.716	1.124
2	Intercept	**3.600**	.250	206.556	**.000**		
	INCOME	**-.220**	.269	.670	**.413**	.473	1.360
	INCOME_1991	**.023**	.414	.003	**.956**	.454	2.304
	INCOME_1995	**.506**	.298	2.878	**.090**	.924	2.977
	INCOME_1999	**.121**	.278	.191	**.662**	.655	1.946
	INCOME_2003	**.165**	.285	.337	**.561**	.675	2.062
	INCOME_2010	**.061**	.271	.051	**.822**	.625	1.806
	SIZE	**-1.084**	.108	101.437	**.000**	.274	.418
	SIZE_1991	**-.120**	.171	.492	**.483**	.634	1.240
	SIZE_1995	**-.363**	.119	9.328	**.002**	.551	.878
	SIZE_1999	**.199**	.115	2.991	**.084**	.974	1.528
	SIZE_2003	**.179**	.119	2.254	**.133**	.947	1.509
	SIZE_2010	**.331**	.113	8.607	**.003**	1.116	1.738
	COMPOSITION	**2.350**	.299	61.708	**.000**	5.832	18.836
	COMPOSITION_1991	**1.343**	.554	5.877	**.015**	1.293	11.344

COMPOSITION_1995	**1.253**	.333	14.210	**.000**	1.825	6.721
COMPOSITION_1999	**-.068**	.319	.046	**.830**	.499	1.746
COMPOSITION_2003	**.010**	.332	.001	**.976**	.527	1.936
COMPOSITION_2010	**-.472**	.313	2.283	**.131**	.338	1.151
GENDER	**.415**	.245	2.864	**.091**	.937	2.446
GENDER_1991	**-.316**	.377	.703	**.402**	.348	1.527
GENDER_1995	**-.421**	.264	2.553	**.110**	.391	1.100
GENDER_1999	**-.485**	.262	3.441	**.064**	.369	1.028
GENDER_2003	**-.678**	.270	6.300	**.012**	.299	.862
GENDER_2010	**-.425**	.251	2.862	**.091**	.400	1.070
AGE	**-.274**	.111	6.048	**.014**	.612	.946
AGE_1991	**-.037**	.169	.049	**.826**	.692	1.342
AGE_1995	**-.103**	.120	.731	**.392**	.713	1.142
AGE_1999	**.107**	.116	.858	**.354**	.887	1.397
AGE_2003	**.085**	.120	.507	**.476**	.861	1.378
AGE_2010	**.092**	.113	.655	**.418**	.878	1.369
EDUCATION	**.429**	.129	10.971	**.001**	1.191	1.979
EDUCATION_1991	**-.146**	.202	.527	**.468**	.582	1.282
EDUCATION_1995	**-.155**	.139	1.235	**.266**	.652	1.126
EDUCATION_1999	**-.338**	.136	6.194	**.013**	.546	.931
EDUCATION_2003	**-.234**	.140	2.791	**.095**	.602	1.041
EDUCATION_2010	**-.219**	.135	2.638	**.104**	.616	1.046
WORKINGSTATUS	**-.025**	.134	.036	**.849**	.750	1.267
WORKINGSTATUS_1991	**.007**	.207	.001	**.975**	.670	1.512
WORKINGSTATUS_1995	**.030**	.145	.042	**.838**	.775	1.369
WORKINGSTATUS_1999	**-.116**	.144	.649	**.421**	.672	1.180
WORKINGSTATUS_2003	**-.211**	.149	1.996	**.158**	.604	1.085
WORKINGSTATUS_2010	**-.208**	.139	2.235	**.135**	.618	1.067
URBANIZATION	**-.054**	.114	.222	**.637**	.758	1.185
URBANIZATION_1991	**-.034**	.188	.032	**.858**	.670	1.397
URBANIZATION_1995	**-.039**	.120	.109	**.742**	.760	1.215
URBANIZATION_1999	**-.002**	.118	.000	**.988**	.792	1.257
URBANIZATION_2003	**-.040**	.120	.109	**.741**	.760	1.216
URBANIZATION_2010	**-.064**	.117	.303	**.582**	.746	1.179

Appendix VI – Correction for Inflation

Based on the disposable household income that is used in this study, the assumption has been made that household disposable income has been adjusted for inflation. However, we cannot exclude this has not been the case. Therefore, additional analysis has been performed to investigate the effects of an inflation correction. In this analysis we assume that the disposable household income has not been corrected (Fig. IV). To adjust for inflation the official (yearly) figures of the consumer price index (CPI) have been used (Fig. III). For the inflation adjustment the year 1987 has been chosen as year of reference.

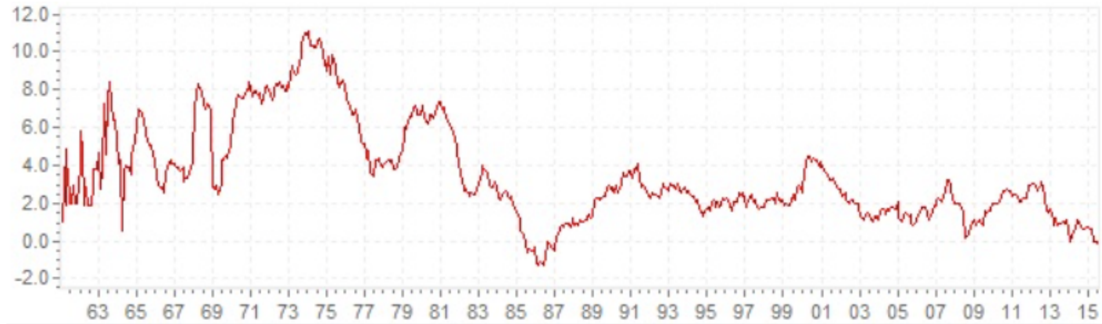


Fig. III. Historical inflation figures (CPI) in The Netherlands (1963-2015) (Worldwide Inflation Data, 2016)

Comparison of the relative effects estimated by the ORL model shows that adjustment of inflation influences to a certain extent the outcomes of composition, gender, age, education, working status and urbanization (Fig. IV and Fig. V). All relative importances are still in the same bandwidth between 1987 and 2014. Furthermore, the relative importance of household size has increased with 6% without inflation correction, and increased with 4.5% with inflation correction. The most important variable, household income, has decreased with 10% between 1987 and 2014 without inflation correction, and decreased over 6% with inflation correction. Concluding, household size and household income show the *same trends of increase and decrease but less strongly* (roughly 25-30% less compared to no inflation correction between 1987-2014).

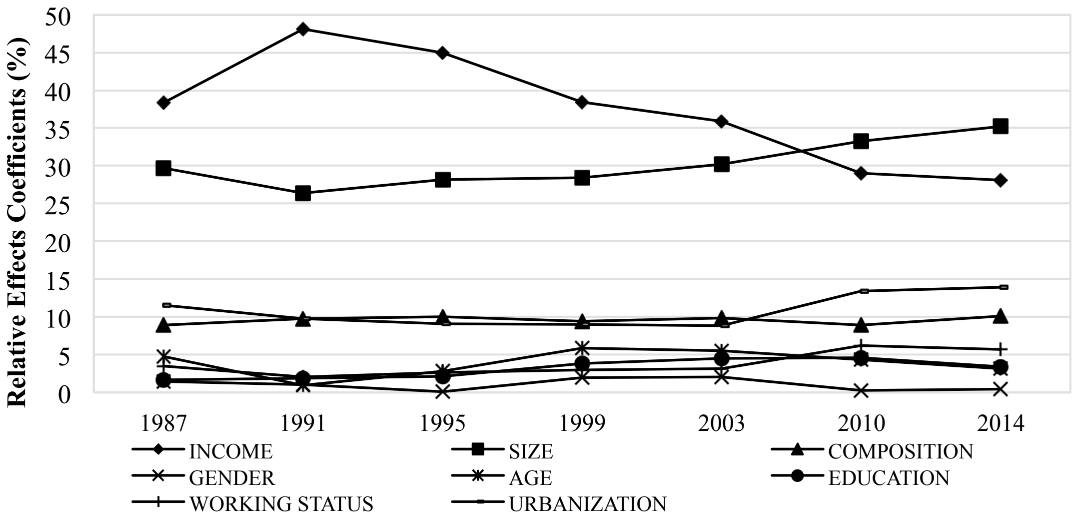


Fig. IV. Relative effects without inflation adjustment

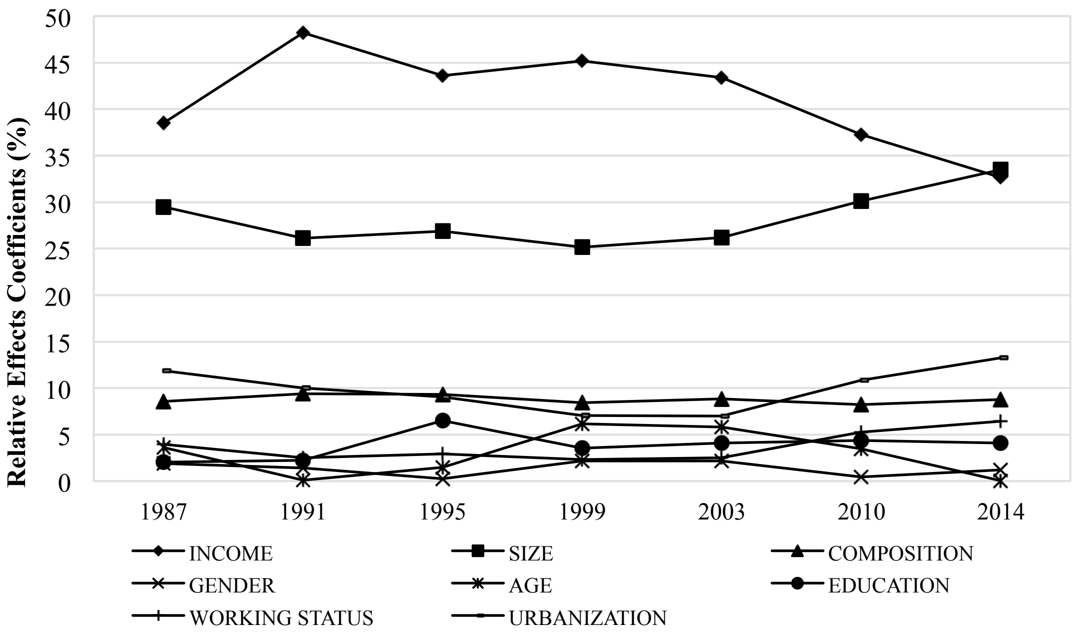


Fig. V. Relative effects with yearly inflation adjustment (CPI)