



The composition and use of the Swedish car fleet

Formulation of a forecasting system

Åsa Forsman
Inger Engström

Publisher:  SE-581 95 Linköping Sweden	Publication: VTI rapport 518A	
	Published: 2005	Project code: 50740
	Project: The composition and use of the Swedish car fleet – Formulation of a forecasting system	
Author: Åsa Forsman Inger Engström		Sponsor: The emission research programme
Title: The composition and use of the Swedish car fleet – Formulation of a forecasting system		
Abstract (background, aims, methods, results) max 200 words: <p>Model-based forecasts of future emissions are important tools for finding strategies to reach the Swedish transport policy objective concerning effect on climate and air pollution. In order for the calculated emissions to reflect the actual situation, input data on kilometrage, vehicle types and vehicle ages must be representative of the region of interest. High-quality forecasts of these variables are therefore of great importance when estimating future emissions.</p> <p>This report proposes a forecasting model system for the compositions and use of the Swedish car fleet, based on experiences from existing models in the literature.</p> <p>The system operates on two levels, an aggregated (national or regional) level and a disaggregated (household or personal) level. Forecasts of the total number of cars and total kilometrage are provided from the aggregated models whereas the disaggregated models provide the distribution of cars of different types and ages and the distribution of kilometrage over the same classification of cars. These results are then combined to give the final outcomes: the number and kilometrage of cars of different types and ages.</p>		
ISSN: 0347-6030	Language: English	No. of pages: 42 + App.

Utgivare:  581 95 Linköping	Publikation: VTI rapport 518A	
Författare: Åsa Forsman Inger Engström	Utgivningsår: 2005	Projektnummer: 50470
Titel: Den svenska bilparkens sammansättning och användning – Formulering av prognosmodell	Projektnamn: Den svenska bilparkens sammansättning och användning – formulering av prognosmodell	
Uppdragsgivare: Emissionsforskningsprogrammet (EMFO)		
Referat (bakgrund, syfte, metod, resultat) max 200 ord: Modellbaserade prognoser för framtida emissioner är viktiga verktyg i arbetet med de transportpolitiska miljömålen om klimatpåverkan och luftföroreningar. För att emissionsberäkningar ska spegla faktiska förhållanden krävs att indata i form av körsträckor, fordonstyper och fordonsåldrar är representativa för den region man är intresserad av. Kvalificerade prognoser för dessa variabler är därför av stor vikt när man vill uppskatta den framtida emissionsutvecklingen. I den här rapporten formuleras ett system av prognosmodeller för den svenska bilparkens sammansättning och användning. Systemet baseras på erfarenheter från existerande modeller som beskrivs i litteraturen. Modellerna är indelade i två nivåer, dels en aggregerad nivå (nationellt eller regionalt), dels en disaggregerad nivå (hushåll eller individ). Prognoser för totalt antal bilar och totalt trafikarbete modelleras på den aggregerade nivån medan den disaggregerade nivån ger fördelningen av bilar av olika typ och årsmodell och fördelningen av trafikarbete för samma indelning.		
ISSN: 0347-6030	Språk: Engelska	Antal sidor: 42 + bilaga

Preface

Model-based forecasts of future emissions are important tools for finding strategies to reach the Swedish transport policy objective concerning effect on climate and air pollution. In order for the calculated emissions to reflect the actual situation, input data on kilometrage, vehicle types and vehicle ages must be representative of the region of interest. High-quality forecasts of these variables are therefore of great importance when estimating future emissions.

As a part of the work of improving the forecasts of future emissions, VTI has carried out a study with the aim of formulating a forecasting model system of the car fleet's composition and use. The study was founded by the emission research programme (EMFO), and the following persons at VTI have contributed to the work: Inger Engström, Åsa Forsman, Ulf Hammarström, Pontus Matstoms, and Tomas Svensson.

Linköping august 2005

Åsa Forsman, project leader

Contents

Summary	5
Sammanfattning	7
1 Introduction and aim	9
2 Orientation of existing models and model systems	11
2.1 Car ownership	11
2.2 Car fleet composition	12
2.2.1 Acquisition of new cars	12
2.2.2 Scrappage	13
2.3 Car use	13
2.4 Licence holding	14
2.5 Some specific model systems	14
2.5.1 The ALTRANS model	14
2.5.2 The COWI Cross-Country Car Choice Model	15
2.5.3 CARMOD	16
2.5.4 FACTS (Forecasting Air pollution through Car Traffic Simulation)	17
2.5.5 DVTM (Dynamic Vehicle Transaction Model)	18
2.6 Factors included in existing models	18
3 Gaps and shortcomings in existing model systems	20
3.1 Limitations of the model systems	20
3.2 Absence of motivational factors	21
4 Purchase behaviour and the decision process	22
5 A proposition of a new model system	25
5.1 Overview	25
5.1.1 Basic concepts	26
5.1.2 Co-ordination with Sampers	27
5.1.3 Summary of the sub-models	28
5.2 Description of the sub-models	29
5.2.1 Car ownership (1)	29
5.2.2 Total kilometrage (2)	30
5.2.3 Licence holdings (3)	31
5.2.4 Household car transactions (4)	31
5.2.5 Kilometrage per car (5)	34
5.2.6 Combining aggregated and disaggregated results (6)	35
5.3 Forecasts and scenarios	36
6 The steps towards a working model	37
7 Discussion	38
8 References	40

Appendix: Multinomial and nedsted logit models

The composition and use of the Swedish car fleet – formulation of a forecasting system

by Åsa Forsman and Inger Engström
VTI
SE-581 95 Linköping Sweden

Summary

Model-based forecasts of future emissions are of great importance for analysing and comparing different strategies to reach the Swedish transport policy objective concerning effect on climate and air pollution. The composition and use of the car fleet are important inputs to emission models, and the quality of the emission forecasts therefore depends on the quality of the car fleet forecasts.

Today, no Swedish national model system exists that provides all the important forecasts. The aim of this study has been to formulate such a system based on experiences with models described in the literature. The model requirements are summarised in the following list:

1. The model should provide the following outputs:
 - a. Forecasts of the total number of cars in the Swedish car fleet
 - b. Forecasts of the distribution of cars of different types and ages
 - c. Forecasts of kilometrage linked to specific car types and ages
2. The model should describe household choices with respect to car type, age, and use
3. The model should provide forecasts for different scenarios regarding
 - a. Car costs
 - b. Fuel costs
 - c. Household characteristics such as income, age, size, and other factors that influence household decisions
 - d. Demographic variables such as the proportion of the population that are of a specific age.

We did not find any single model system in the literature that fully complied with the above requirements. The proposed model system is therefore a combination of model approaches and ideas from a variety of different models. The system operates on two levels, the aggregated (national or regional) level and the disaggregated (household or personal) level (*Figure 1*). Forecasts of the total number of cars and total kilometrage are provided from the aggregated models whereas the disaggregated models provide the distribution of cars of different types and ages and the distribution of kilometrage over the same classification of cars. These results are then combined to give the final outcomes: the number and kilometrage of cars of different types and ages.

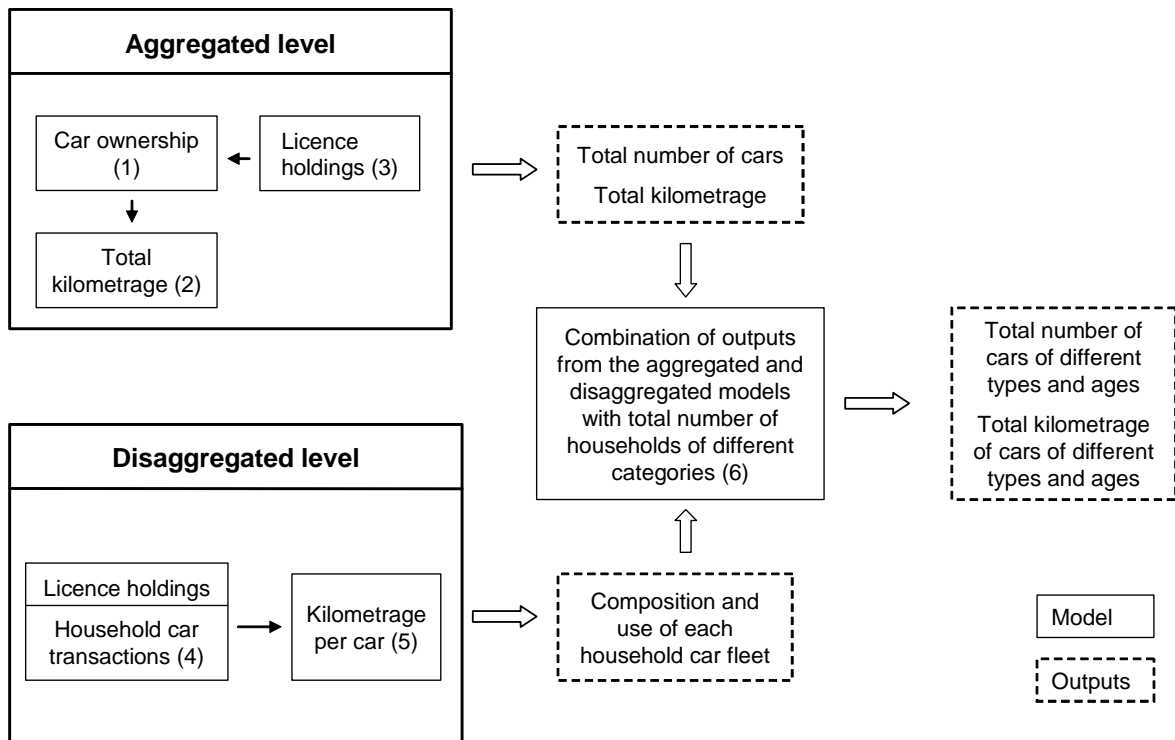


Figure 1 Overview of the proposed model system. The boxes with unbroken lines indicate models, and boxes with dashed lines indicate model outputs.

The national model system for analysis of passenger transport, Sampers, describes how people choose to travel. Important outputs are travel frequency, destination choice, travelling time, and choice of transport mode. The system also includes a car ownership model developed at VTI. Sampers does not consider different car types and consequently does not consider car use linked to specific car types. It is, for specific applications, important that the new system provides forecasts that are consistent with the forecasts from Sampers. It will therefore be possible to run the new model both with and without Sampers co-ordination.

The existing models that describe household choices rely, in general, on factors such as socio-economic and socio-demographic variables but exclude other important factors that have been shown to influence the choice process, such as attitudes and preferences. Such factors are difficult to measure but nevertheless important to include for a number of reasons. First, exclusion of essential explanatory variables may lead to model misspecification and, consequently, errors in the resulting parameter estimates. Second, forecasting the response in model outputs due to changes in motivational factors could be important in scenario calculations. Third, understanding the motivational factors is the key to changing undesirable behaviour. Moreover, even if motivational factors are not explicitly entered in the model, they implicitly influence the forecasts. For example, if motivational factors are excluded from a model, the forecasts are calculated with the implicit assumption that the motivational factors will remain the same as for the years that were used to estimate the model parameters. By including the variables explicitly, it becomes clearer how they influence the forecasts. Considering the potential for large improvements of the models, further research is needed on how motivational factors can be measured and included in the model system.

Den svenska bilparkens sammansättning och användning – formulering av prognosmodell

av Åsa Forsman och Inger Engström
VTI
581 95 Linköping

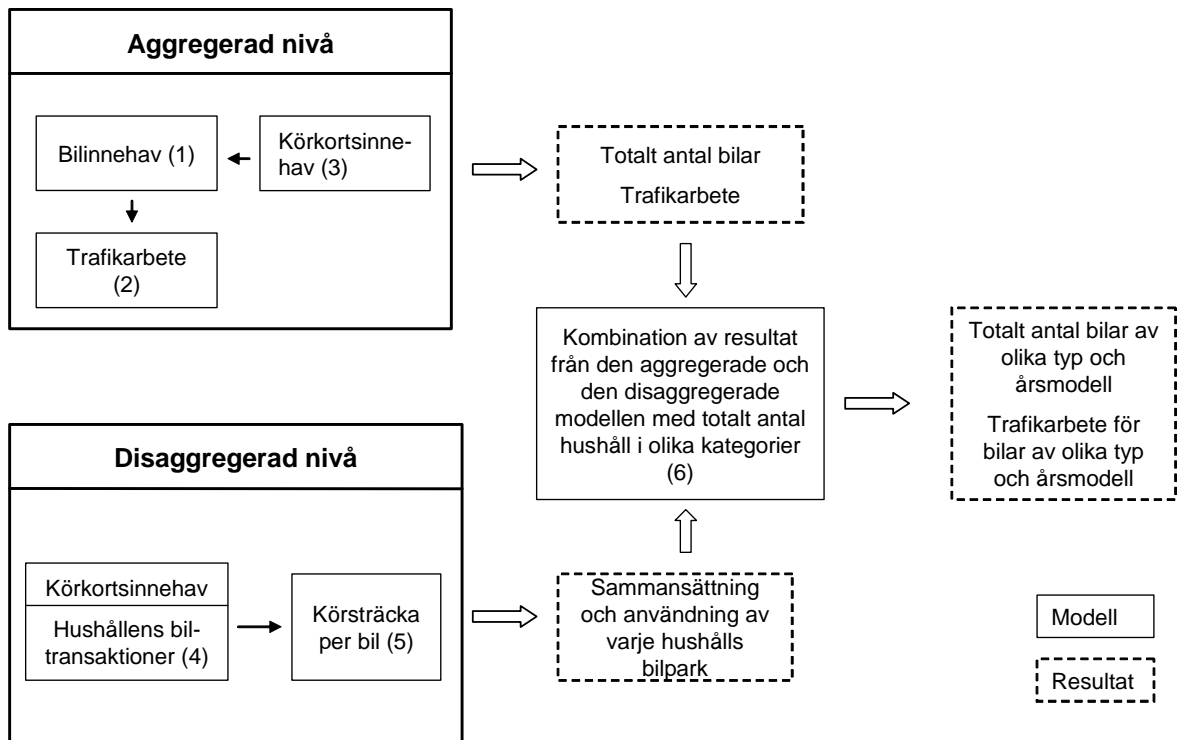
Sammanfattning

Modellbaserade prognoser av framtida emissioner är viktiga verktyg när det gäller att analysera och jämföra olika strategier för att nå de transportpolitiska miljömålen om klimatpåverkan och luftföroreningar. För att emissionsberäkningar ska spegla faktiska förhållanden krävs att indata i form av körsträckor, fordonstyper och fordonsåldrar är representativa för den region man är intresserad av. Kvalificerade prognoser för dessa variabler är därför av stor vikt när man vill uppskatta den framtida emissionsutvecklingen.

Idag finns inget nationellt modellsystem i Sverige som ger alla de viktiga prognoserna. Syftet med den här studien är att formulera ett sådant system baserat på erfarenheter av modeller som beskrivs i litteraturen. De krav som ställs på modellen sammanfattas i följande lista:

1. Modellen ska ge följande resultat:
 - a. Prognoser för det totala antalet bilar i den svenska bilparken
 - b. Prognoser för fördelningen av bilar av olika typ och årsmodell
 - c. Prognoser för det totala trafikarbetet som utförs av bilar av viss typ och årsmodell
2. Modellen ska beskriva hushållens val med hänsyn till biltyp, årsmodell och användning
3. Modellen ska ge prognoser för olika scenarios när det gäller
 - d. Kostnader för bilinnehav
 - e. Bränslekostnader
 - f. Egenskaper hos hushållet såsom inkomst, ålder, storlek och andra faktorer som påverkar dess beslut
 - g. Demografiska variabler såsom andelen av populationen som har en viss ålder.

Vi hittade inget enskilt modellsystem i litteraturen som helt uppfyllde de angivna kraven. Det föreslagna modellsystemet är därför en kombination av ansatser och idéer från ett flertal olika modeller. Systemet verkar på två nivåer, den aggregerade (nationellt eller regionalt) nivå och den disaggregerade (hushåll eller individ) nivå (*Figur 1*). Prognoser för totalt antal bilar och totalt trafikarbete fås från den aggregerade nivå medan den disaggregerade nivå ger fördelningen av bilar av olika typ och årsmodell och fördelningen av trafikarbete för samma indelning.



Figur 1 Översikt av det föreslagna modellsystemet. Boxarna med heldragna linjer beskriver modeller och boxar med streckade linjer beskriver modellresultat.

Det nationella modellsystemet för analys av persontransporter, Sampers, beskriver hur människor väljer att resa. Viktiga resultat är resfrekvens, destinationsval, restid och val av transportsätt. Systemet innehåller också en bilinnehavsmodell som är utvecklad vid VTI. Sampers tar inte hänsyn till olika sorters biltyper och därmed inte heller till användning kopplat till en speciell typ av bil. För vissa tillämpningar är det viktigt att det nya systemet ger prognoser som överensstämmer med prognoser från Sampers, det ska därför vara möjligt att köra den nya modellen både med och utan koppling till detta system.

De existerande modellerna som beskriver hushållens val, baseras i regel på socioekonomiska och sociodemografiska variabler men exkluderar andra viktiga faktorer som har visat sig ha stort inflytande över valprocessen, såsom attityder och preferenser. Här finns det utrymme för förbättringar men mer forskning behövs om hur man kan mäta dessa variabler och ta hänsyn till dem i modellerna.

1 Introduction and aim

The objectives of the Swedish transport sector are described in the transport policy bill (Kommunikationsdepartementet, 1998). The overall objective is to ensure socially and economically efficient and long-term sustainable transport resources for the public and industry throughout Sweden. There are also five subsidiary objectives; one of these is *a good environment*. The environmental objective is described in relation to long-term visions of people's health but also includes specific short-term goals in the form of maximum emissions of carbon dioxide, sulphur dioxide, nitrous oxides and volatile organic compounds (VOC).

Several measures must be taken to reduce emissions if the environmental goals are to be reached, not least the emissions from road traffic, which largely contribute to the total discharge from the transport sector. However, the effect of the measures, individually or in combination, is not evident, and model-based forecasts of future emissions are therefore of great importance for analysing and comparing different scenarios.

The EMV model (Hammarström & Karlsson, 1998) has, in recent years, been used by the Swedish Road Administration (SRA) to calculate emissions from road traffic. Recently, an EU project called ARTEMIS has developed a new model that will provide consistent emission estimates at the national, international, and regional levels. This model is currently being implemented in Sweden. Both these models calculate the emissions given a specific vehicle fleet and its use. The quality of the emission forecasts therefore largely depends on the quality of the vehicle fleet forecasts.

A number of models and model systems that provide forecasts of the car fleet's size, composition, and/or use are described in the international literature. See for example De Jong et al. (2002) for a review of existing models. However, there is no Swedish national model system that provides all the important forecasts. The national model system for analysis of passenger transport, Sampers, provides car ownership and use but does not consider different car types and consequently does not consider car use linked to specific car types (Beser & Algers, 2001). This is a severe restriction of the model since the emissions from a car largely depend on car type.

The aim of this study is to formulate a forecasting model system for the composition and use of the Swedish car fleet. The model requirements are summarised in the following list:

1. The model should provide the following outputs:
 - a. Forecasts of the total number of cars in the Swedish car fleet
 - b. Forecasts of the distribution of cars of different types and ages
 - c. Forecasts of kilometrage linked to specific car types and ages
2. The model should describe household choices with respect to car type, age, and use
3. The model should provide forecasts for different scenarios regarding
 - a. Car costs
 - b. Fuel costs
 - c. Household characteristics such as income, age, size, and other factors that influence household decisions
 - d. Demographic variables such as the proportion of the population that are of a specific age.

The second requirement is included to increase the understanding of household behaviour. The choices of car ownership and use are made by the individual household or person, even though the choices are influenced by the conditions set by the government in the form of taxes, import regulations, etc. Understanding the mechanisms behind these choices is therefore of great importance when modelling changes in the car fleet. A good understanding of the choice mechanisms also has a value in its own right, for example, when drawing up campaigns to persuade people to choose more fuel-efficient cars. This knowledge is also highly relevant for other transport objectives such as *a safe traffic*.

The proposed system will be restricted to private cars, both privately owned and company-owned. The focus of the model is on cars that are available for private driving, where the user has an influence on the choice of car. Cars that are strictly used for business (for example taxis) are not handled.

2 Orientation of existing models and model systems

The international literature on car fleet models covers a wide range, from specific models which describe one aspect of a restricted car fleet to complex model systems that describe a national car fleet's use and composition with respect to car types and ages. This section gives an overview of applied model types and a description of five specific model systems. The limitations and shortcomings of these systems are identified in section 3.

2.1 Car ownership

The models of car ownership can be divided into three basic categories. The first category includes so-called aggregated or macro-economic models which operate on a national or regional level. This category typically comprises time-series models of the total number of cars using macro-economic data such as GDP (gross domestic product) and population data as explanatory variables. The most popular functional form is an S-shaped curve; this form agrees with empirical data as well as economic theory. The S-shaped curve is supported by diffusion theories and product life cycle, which state that the take-up rate for new products is initially slow then increases as the product becomes more established, and finally the rate of increase diminishes as the market approaches saturation (Duesenberry, 1949; cited in Whelan et al., 2000). Aggregated models have been used for quite a long time; two recent applications can be found in Dargay & Gately (1999) and Whelan et al. (2000).

The second category includes cohort models, which also operate on an aggregated level. These models take into account the age structure of the population and follow the behaviour of different cohorts. The car ownership model developed at VTI is an example of such a model (Matstoms, 2002). The model is characterised by the inclination of entering or leaving car ownership status, also called entry and exit propensities. The propensities depend on factors such as income, age and petrol price.

The third and final category includes the disaggregated or micro-economic models which operate on the level where the actual choices are made, that is, on a household or individual level. Disaggregated models often explain the decisions with reference to socio-economic and socio-demographic data such as household income, size of the household, and education level of the owner. The results from individual households can be generalised to the national or regional level if the total number of households with specific characteristics are known.

The disaggregated models are further classified into static and dynamic models. Static models describe the household car fleet at a specific time point using, for example, discrete-choice models (these models are further described in 2.2.1). Dynamic models describe car transactions, either by estimating the duration between transactions (duration models) or by estimating whether a transaction has occurred during a specific time interval (discrete-time models). Dynamic models dominate in the recent literature and are generally recommended for forecasting models since a car ownership decision is a dynamic process (De Jong et al., 2002; Ramjerdi et al., 2000). In a static model, households are assumed to constantly be in equilibrium with respect to their car fleet. This

assumption is violated due to transaction costs. For example, a rise in income may not lead to an immediate car transaction.

The advantages and disadvantages of the different model categories with respect to forecasting vehicle ownership are discussed in Booz Allen & Hamilton (NZ) (2000). The main disadvantage of the aggregated models is that they omit many variables which affect car ownership at the household level. Nevertheless, these models have been found to perform reasonably well in short- to medium-term forecasting. Disaggregated models, on the other hand, can include a number of explanatory variables at the household level. However, moving from the household to the regional or national level may lead to large errors at the aggregated level. Experience has also shown that disaggregated models tend to underestimate the growth in car ownership. The cohort model seems to overcome many of the problems in the other approaches. The only observed drawback is that these models are highly data-intensive.

2.2 Car fleet composition

Vehicle type and age are important determinants for vehicle emissions. The future composition of the car fleet depends on the composition of new and scrapped vehicles.

2.2.1 Acquisition of new cars

The composition of new cars is almost exclusively modelled on a disaggregated level, using models that describe household choices. The family of logit models has played an important role for modelling the choices of car type. The most frequently used models are the multinomial and nested logits. A schematic picture of the choice structure in these models is shown in *Figure 2*. The model equations and the link between logit models and the theory of random utility maximisation are described in Appendix I; see, for example, Cramer (1991) or Train (2002) for further details. In principle, a random utility function is attached to each possible state in a set of choices, and the outcome with the highest utility is chosen. The utility of a specific vehicle may depend on both the characteristics of the vehicle and the characteristics of the household or person.

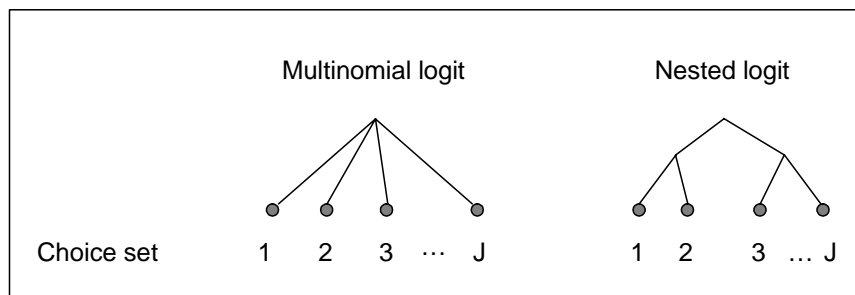


Figure 2 Schematic picture of the multinomial and nested logit models with a choice set of J possible choices. The multinomial logit treats all choices the same way whereas the nested logit first groups choices into a number of nests.

Both the multinomial and nested logits describe the choice between a finite number of discrete alternatives. In a multinomial logit, the odds of choosing one state over another are independent of all other alternatives in the choice set. This property is known as independence from irrelevant alternatives (IIA). If this

property does not hold but the choice set can be grouped into subsets (also called nests) where the property holds within each subset, then the nested logit can be used.

The multinomial logit has, for example, been used by Birkeland and Jordal-Jørgensen (2001) to model the choice of car type. The choice set consisted of 50 alternative types.

De Jong (1996) used a nested logit model to describe the choice of car type. The set of alternatives was divided into two nests, one with diesel and one with non-diesel cars. This division gave a significant improvement in the model fit as compared to a multinomial logit. Another application of the nested logit is presented in Mohammadian & Miller (2003). They used six nests, one for each class of car, and each nest comprised 4 classes of vintages.

2.2.2 Scrappage

Both aggregated and disaggregated approaches have been used to model scrapping of cars. In aggregated models, the changes in cohorts of cars of specific ages are followed; sometimes the models also take into account car type. The changes in the cohorts are either described as the number of scrapped cars or as survival rates for cars of different ages. In the ALTRANS model (further described in 2.5.1), scrap page is modelled with a generalised linear model approach (Kveiborg, 1999). The number of scrapped cars of a specific fuel type, weight, and vintage is explained by a linear combination of variables such as price indices for fuel, repair costs, and income.

An example of a disaggregated model of household behaviour is found in De Jong et al. (2001). They used a multinomial logit model to discriminate between five choices: keep, scrap, or sell for private owners, and keep or scrap for car dealers. The model was fitted to data from a stated preference (SP) survey of households that owned an old car (7 years or older) or had recently scrapped a car. These households received a questionnaire about intentions to scrap, intentions to replace and the relative importance of factors influencing these decisions.

2.3 Car use

The modelling of car use has traditionally not received as much attention as modelling of car ownership (Wall, 1991). One explanation is, according to Wall, that the kilometrage per vehicle for a long time was relatively stable and it was more important to model car ownership. However, at least for Swedish conditions, Wall argues that the stability of kilometrage per vehicle was due to a balancing of factors that contributed to upward and downward trends. In the future, these factors may not balance each other and it is therefore important to pay attention to car use modelling as well.

Wall (1991) used a log linear regression model to describe the average kilometrage per vehicle on an aggregated level. Average kilometrage was explained by fuel price, total private consumption, proportion of women car users, and proportion of pensioner car users.

Regression models have also been used to explain car use on a disaggregated level. De Jong (1996) and Steg et al. (2001) used linear regression models to explain car use for a specific car and for a specific household, respectively. Golob et al. (1997) instead used a structural equation approach to model kilometrage given car type.

2.4 Licence holding

Although licence holding is not needed as input to emission models, it is highly relevant for modelling car ownership and use and is therefore included in this overview.

Cohort models are often used to explain the development of licence holdings on an aggregated level. Applications can be found in HCG & TØI (1990), Byrsjö (1998), and Christensen & Gudmundsson (2003). Variables such as age, sex, and income are used to explain entry and exit propensities.

On the disaggregated level, licence holdings have been modelled with discrete-choice models. One example is the model described in HCG & TØI (1990) where the licence distribution within a two-adult household is described by a multinomial logit model. There are four choice alternatives: either the head of the household, the partner, both, or neither have a licence. Additional adults in the household are included in a separate model.

2.5 Some specific model systems

Five different model systems are described in this section. The presentation of the systems includes only the parts that are relevant for this study, that is, the parts that pertain to car ownership, car use and the composition of the car fleet. The other parts of the models are described in the cited references.

2.5.1 The ALTRANS model

ALTRANS is a transport model system developed by the National Environmental Research Institute in Denmark (Christensen & Gudmundsson, 2003). This system focuses mainly on public transportation, but car ownership and use are also included.

The system is divided into three main parts.

Part I: A GIS-based geographical model for travel times and distances. The model is based on detailed information of the public transport system such as timetables and routes. Distances by car can readily be calculated since the road structure is included in the model.

Part II: A micro-economic model of individual travel behaviour. Mode choice, passenger kilometres travelled by each individual and each mode, and the number of cars possessed by each family are important outputs from this model. In addition, this part includes a sub-model of licence holdings.

Part III: A car fleet and emission model. This model calculates the total emissions from the road and railway. Since the emissions differ for different types of cars and for cars of different ages, the composition of the car fleet is forecasted using a macro-economic approach. The cars are divided into 5 different types (3 sizes of petrol-driven cars and 2 sizes of diesel-driven cars) and 20 age classes.

The number of cars owned by each family is estimated by a nested logit model with two nests. The nests divide the choices of having no car and having one or two cars (these are the only options). The variables used to explain the choices are: the costs of car ownership, socio-economic variables for the household and its individual members, accessibility of public transport, and number of licence

holders. The sub-model of licence holdings operates on both the individual and the national levels. An individual's probability of holding a licence is estimated by a binary logit model with the explanatory variables sex, age, income, occupation, and degree of urbanisation. These probabilities are used to assign licence holdings to individuals but also as inputs to a cohort model. The cohort model is used to forecast the number of licence holders on a national level, taking into consideration changes in the age structure and income.

The model for the composition of the car fleet is based on the relation

$$Stock(t) = Stock(t-1) + N(t) - Sc(t), \quad (1)$$

where $Stock(t)$ is the number of cars at time t , $N(t)$ is the addition of new cars between time $t-1$ and time t , and $Sc(t)$ is the number of scrapped cars during the same period. The size of the total stock of cars at different time points is provided from the travel behaviour model. The number of scrapped cars is calculated separately for cars of different types and ages; a linear regression model is used for each group of cars. The addition of new cars can now be calculated from (1). The new cars are distributed among the 5 types, and the distribution is assumed to be the same as for the base year of the forecasts.

The total kilometrage of the car fleet is calculated in the travel behaviour model and distributed among different types and ages according to the following assumptions. First, all cars of the same age are driven the same distance, regardless of type. Second, the kilometrage differs between cars of different ages, but the relative kilometrage is constant in time. The distribution of driving distance for cars of different ages is calculated from surveys made at the Danish Road Directorate.

2.5.2 The COWI Cross-Country Car Choice Model

The Danish consultancy group COWI A/S has developed a car choice model that can be modified to match the conditions of several different countries. The model was applied to Swedish conditions in a project that aimed at analysing the effect of introducing various CO₂-differentiated registration taxes. Both the model and the study are described in Sand Jespersen et al. (2002). The model considers only the composition of the car fleet; use is not included.

The main structure of the model is shown in *Figure 3*, and the three modules are described below.

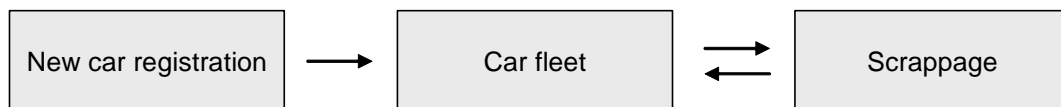


Figure 3 Basic model structure of the COWI Cross-Country Car Choice Model (from Sand Jespersen et al., 2002).

The module for new car registrations includes a number of discrete choices made by the car buyers. The first choice is between a privately owned and a company-owned car, and this choice is decided by a binary logit model. The model input is the cost of having the car as a private car minus the cost of having the car as a company car. Given the choice of a privately owned car, a nested logit model describes the car buyer's choice of car model. The nesting variable is car type (estate, hatchback, saloon, etc.). If a company-owned car is chosen, the buyer's

decision is modelled with a multinomial logit. The buyer can be either the company or an employee depending on who makes the decision about car type. The models for both private and company cars take into account socio-economic conditions of the buyer and car characteristics, although the exact variables differ between the two models.

The car choice models are disaggregated since they model the buyer's choices. The scrapping module, on the other hand, consists of an aggregated model that estimates the survival rates of different car types and ages. The rates are based on both empirical survival curves for the Swedish car fleet and the effect that changed prices for new cars have on scrapping. These price elasticities are estimated on data from England and provided for three types of petrol-driven cars (based on engine sizes) and diesel-driven cars.

The car fleet module uses outputs from the other two modules in order to keep track of the total car fleet. The composition of the car fleet in the base year is taken from the Swedish vehicle register. Each year the fleet is updated by the number of new cars and the number of scrapped cars. Used imported cars are not included in the model. In total, the car fleet is divided into 312 groups of cars, based on the following variables:

- Age (0 to 25+)
- Engine size (3 groups)
- Fuel type (diesel and petrol)
- Make (Swedish or not)

2.5.3 CARMOD

CARMOD is a model system of the dynamics of the Australian car fleet, developed by the Bureau of Transport and Communications Economics. The model has been used to assess the effects of different measures on greenhouse gas emissions from the car fleet (BTCE, 1996).

CARMOD is an aggregated model system that only considers the size of the car fleet as a function of time and total population. The number of vehicles per person at time t ($MVPER(t)$) is described by the following relation:

$$MVPER(t) = \frac{k}{1 + ae^{-bt}} \quad (2)$$

where k is the saturation level, and a and b are parameters that determine the shape of the curve. The parameters k , a , and b are estimated to fit data from the period 1945 to 1994.

The system also includes a scrapping model with a structure similar to that of the model for vehicles per person. The proportion of cars of a specific vintage, i , surviving to an age of T years, $S_i(T)$ is modelled as

$$S_i(T) = \frac{1}{a_i + (1 - a_i)e^{b_i T}} \quad (3)$$

where the size of the model parameters, a_i and b_i , depends of the vintage of the car. The results from the scrapping model are combined with the total number of vehicles per person and a forecast of the total population to provide a forecast of

the total number of cars of different ages. The model does not include different car types.

The average vehicle kilometres per car is assumed to be constant; however, the total kilometrage is distributed among vehicles of different ages according to the results from a survey of motor vehicle use.

2.5.4 FACTS (Forecasting Air pollution through Car Traffic Simulation)

FACTS is a Dutch model that can be used to forecast car ownership, car use and emissions under alternative economic and demographic scenarios (De Jong et al., 2002). It is classified as a heuristic simulation method and based on two principal assumptions: a constant money budget and maintenance of mobility level. The assumption of a constant money budget implies that the percentage of income spent on travel is fixed for homogenous groups of households. In this case, households are grouped by age, income and composition. Maintenance of mobility level can, for example, have the effect that a household chooses to buy a cheaper car if the fuel price increases.

The module in FACTS that describes car ownership and use is called SMAK and comprises the following steps.

- Company cars are treated separately from private cars. These cars are allocated to households based on a probability distribution that is specific for each household group.
- The available budget for private car ownership and private car use is determined.
- The number of kilometres per household is determined by drawing from a specific probability distribution per household class. The household will be considered to own a car
 - if it has a mobility need above a predetermined minimum value, and
 - if it has a sufficiently large car budget.
- If the above prerequisites are fulfilled, car ownership is determined based on the type of household, mobility needs, and whether the household has been allocated a company car. For example, a single household with a company car cannot own a private car.
- The next step is to choose the car type. A total of 18 car types are distinguished based on the following variables:
 - Fuel type; 3 classes: petrol, diesel, LPG (Liquefied Petroleum Gas)
 - Weight; 3 classes: ≤ 950 kg, 951–1149 kg, ≥ 1150 kg
 - Age; 2 classes: ≤ 5 years old, > 5 years old

The car choosing procedure involves several steps that will not be described here. In principle, the household chooses the most expensive car type it can afford (corrected for the number of kilometres driven).

The above steps are described in more detail in De Jong et al. (2002). This report also includes a discussion of the advantages and disadvantages of FACTS. One disadvantage is that the basic assumptions of a constant money budget for car ownership and use and maintenance of mobility level are at odds with economic theory.

2.5.5 DVTM (Dynamic Vehicle Transaction Model)

A model system of car-holding duration, type choice, and use is described in De Jong (1996). All sub-models in the system operate on a disaggregated level.

Car holding is described by a duration model which estimates the time between purchase of a car and its replacement. This model can be used to forecast whether a vehicle is replaced or not during a specific time period. Other transactions, such as acquisition and disposal, are not included, and only one car per household is considered. Given a replacement, the type choice is modelled by a nested logit model with separate nests for diesel and non-diesel cars. A total of 1,000 make/model/age-of-car combinations were included in the model, but each household could choose from a subset of 20 cars. The choices were influenced by the following factors:

- income and cost variables,
- comparison of previous and new car attributes (to account for such things as brand loyalty),
- attributes of alternatives on the car market.

Car use is described by the following regression model:

$$\ln(k_i) = \alpha \ln(y_i - c_i) + \beta v_i + \gamma \mathbf{Z}_i + e_i \quad (4)$$

where i indicates the decision-maker, k_i is annual kilometrage in the car, y_i is annual net household income, c_i is fixed car costs, v_i is variable car costs, \mathbf{Z}_i is a vector of attributes of the decision-maker, e_i is the error term, and α , β , γ are model parameters.

2.6 Factors included in existing models

An overview of variables used to explain car ownership, car type, and car use on the aggregated and disaggregated level is presented in *Table 1* and *Table 2*, respectively. The list is far from exhaustive, but it covers the principal variable types. Most models use variables listed in the tables or closely related variables. The trends in aggregated time series data are typically explained by demographic and macro-economic factors. When household behaviour is modelled, the most commonly used variables are socio-demographic and socio-economic household data and car characteristics and costs.

In addition to the variables listed in *Table 1* and *Table 2*, motivational factors have been considered in a couple of studies. Steg et al. (2001) included a variable that described environmental awareness when modelling car use. Wu et al. (1999) emphasized the importance of including psychological and sociological factors when modelling car ownership in developing countries. They included use-value attitudes and sign-value attitudes in a model of car ownership in China.

Table 1 Overview of variables used in aggregated models of car ownership, car type, and use.

Variable type	Examples of variables
Demographic	Population size Proportion of inhabitants in different age groups Proportion of cars owned by pensioners Proportion of cars owned by women Number of licence holders
Macro-economic	Gross domestic product (GDP) Consumer price index Fuel price index Total private consumption index Income index Purchase price index

Table 2 Overview of variables used in disaggregated models of car ownership, car type, and use.

Variables related to	Variable type	Examples of variables
Household	Socio-demographic	Size of household Number of children Age Gender Living area, urban or rural Housing form
	Socio-economic	Income Employment status Education level
	Other	Number of licence holders Service station availability (relevant for alternative fuel cars)
Household vehicle fleet		Number of owned cars Average market price of household fleet Average age of household fleet
Car	Fixed costs	Purchase price Annual vehicle tax Insurance Operating costs
	Variable costs	Fuel price Operating costs
	Car characteristics	Engine size Fuel-type Age Other characteristics such as van, sports car, or SUV Brand Model

3 Gaps and shortcomings in existing model systems

The orientation of existing models in the previous section showed the wide range of models that have been applied in the area. A lot of effort has been spent on elaborating the causes of household choices and overall trends. However, it is not as simple as copying an existing model system. This conclusion is based on two observations. First, none of the studied systems are complete with respect to our requirements of a new system described in the introduction. Second, the disaggregated choice models generally exclude important factors that affect the choice process, such as attitudes and preferences.

3.1 Limitations of the model systems

The five model systems described in section 2.5 all have shortcomings in relation to the requirements of the new model system. The main limitations of each system are described below.

The ALTRANS model

The disaggregated car ownership model in ALTRANS is static whereas the general recommendation in the recent literature is to use a dynamic model (see section 2.1). Moreover, the system does not include an explicit car type choice model; the distribution of new cars is assumed to be the same as for the base year of the forecasts. Thus, changes in household characteristics do not lead to changes in composition of the car fleet. Another limitation is that all cars of the same age are assumed to be driven the same distance, regardless of type.

The COWI Cross-Country Car Choice Model

The COWI model also includes a static ownership model. The use of the car fleet is not included in the system.

CARMOD

This system differs from the first two since it only operates on an aggregated level, thus, household behaviour is not modelled. The total number of cars depends only on total population and time, which limits the use of the model as a scenario explorer. Car types are not included.

FACTS (Forecasting Air pollution through Car Traffic Simulation)

The limitations of FACTS are thoroughly discussed in De Jong et al. (2002). A couple of the main objections are mentioned here. As in the ALTRANS and the COWI models, FACTS uses a static car ownership model. Moreover, the two basic model assumptions, constant money budget for car ownership and use and maintenance of mobility level, are at odds with economic theory. Finally, the choice of car type is influenced only by car cost (not by other attributes of the household).

DVTM (Dynamic Vehicle Transaction Model)

DVTM, as described by De Jong (1996), only included replacements of cars; no other transactions, such as acquisitions and disposals, were included, which, of course, severely limited the use of the model. Since 1996, those other transactions

have been introduced, according to De Jong et al. (2002). To our knowledge, the extended model has not been published and we can therefore not evaluate the merits of these improvements.

The five systems briefly described in this report have been chosen to show the diversity of operational systems. A number of other systems also exist, but we have not been able to find any other system that fully matches our requirements.

3.2 Absence of motivational factors

Most of the disaggregated models take into account socio-economic and socio-demographic factors that are easily observed (Choo & Mokhtarian, 2004). However, the models cannot explain why two neighbours with the same socio-economic and socio-demographic status own completely different types of cars, or why one neighbour takes the car and the other takes the bus to work. The underlying motivations that lead to these differences in behaviour are difficult to measure but nonetheless important to understand for three main reasons. First, the lack of important explanatory variables may lead to model misspecification and, consequently, errors in the resulting parameter estimates. Second, forecasting the response in model outputs due to changes in motivational factors could be important in scenario calculations. Third, understanding the motivational factors is the key to changing undesirable behaviour. The problem of model misspecification can for some cases be solved by using a flexible model specification (Train, 2002). A flexible model structure may, however, lead to other problems such as overparametrisation.

A few papers acknowledge the importance of including motivational factors in the models, as mentioned in section 2.6. Steg et al. (2001) report from a literature study which revealed that factors such as emotions evoked by car use, social norms, personal norms, and awareness of the problems caused by car use were related to car use and travel mode choice. Kitamura et al. (1997) found that attitudinal variables were more important than socio-economic and neighbourhood descriptors for explaining trip frequency of different modes. However, their results were based on a study with very low response rate (17.6%), so the merits of these results are highly questionable.

4 Purchase behaviour and the decision process

The experiences of including motivational factors in car ownership models are very limited. We have therefore studied the literature of car purchasing behaviour and the decision process.

For understanding the behaviour involved in purchasing an item, the theory of planned behaviour can be used, according to Caprara et al. (1998). This is a model that predicts a person's behaviour via one's intention to engage in that behaviour. The intention is predicted by three parameters: attitude towards the behaviour; subjective norm, which means the person's perception of the social pressure to engage in the behaviour; and perceived behavioural control, which means the person's perception of how difficult or easy it would be to perform the behaviour (Figure 4).

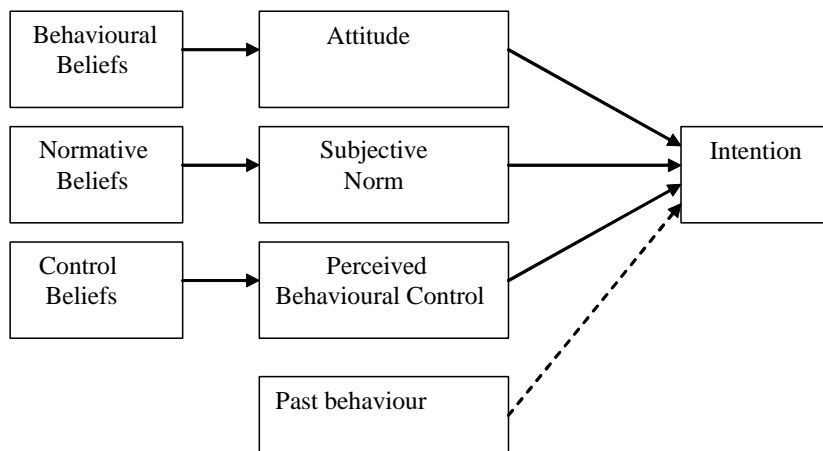


Figure 4 Theory of planned behaviour.

These three predictors are functions of other variables. Attitude is a product of a person's behavioural beliefs, that is, the perceived consequences of the behaviour. Subjective norm is a product of normative beliefs, which means perception of other important persons' preferences about whether one should engage in the behaviour and one's motivation to comply with those preferences. Perceived behavioural control is determined by control beliefs, that is, the subjective probability that a certain event should occur and whether this would facilitate or hinder the intended behaviour. When looking at consumer behaviour there has been an addition to the model, which is past behaviour. This parameter will measure self-reported habits and could, for example, explain brand loyalty.

When purchase intention of a car (an Opel) was tested with this model, it was shown that the parameter attitude was most important to discriminate between users and non-users and the main difference among the two groups regarded behavioural beliefs, that is, those issues that are considered as possible outcomes of the behaviour. There was also a difference in normative beliefs. More specifically, this result showed that users believed more than non-users that the purchase of an Opel would have outcomes such as strong, durable car without problems, reliable car, providing safety, high-level of comfort and good quality-cost relationship. The users also believed that their relatives (children, partner etc.) would like it and they were more inclined to comply with their approval. They also believed that the purchase would be facilitated by a good price.

Taken together it was attitude towards purchasing the Opel and past behaviour that were the best predictors of the purchasing intention. Perceived control and subjective norm were less relevant even if these also were significant (Caprara et al., 1998).

Choo & Mokhtarian (2004) suggested other factors that affect purchase behaviour. They found that travel attitudes, personality, lifestyle and mobility significantly affected an individual's vehicle type choice.

Attitudes can also affect mode choice. Car users often have positive attitudes towards car use and do not see the negative effects. The feeling of freedom is stronger than the concern of greenhouse gas emissions. Why should I sacrifice for the collective interest when others do not? Why shouldn't I take the car when everyone else do? Discussions like this make car users continue to use their cars (Golob & Hensher, 1998; Tertoolen et al., 1998). This positive attitude towards driving leads to a frequent use of the car and there is no reflection of whether it is good to use the car or not (Boe et al., 1999).

Literature about factors affecting purchase behaviour is limited; there is more about the decision process to purchase a car. Marell et al. (1996) have studied the decision process that leads to replacement. They found that there are a lot of models with information about factors affecting the choice of which type of car to buy but there is almost no information about the decision-making process leading up to the purchase. Every car owner has an aspiration level defining a minimal acceptable quality standard for the car. The car owner also has a current level, which means that they judge the current quality of the owned car. If there is a discrepancy and the discrepancy is strong enough between those two levels, it will lead to replacement. Several factors affect the aspiration and current level such as economy, taste marketing and other socio-demographic factors, but the relation among those aspects has to be investigated further (Marell et al., 1996; Punj & Brookes, 2002).

When the car owner has decided to replace the car he/she starts seeking information about model, size, engine etc., but often this is not the starting point. The decision to replace a car consists of several "sub-decisions". Some of these sub-decisions have already been made in advance, before the decision of replacement is made. A lot of consumers make product-related decisions in advance and store them in memory for later use; such stored decisions are called pre-decisional constraints. These pre-decisional constraints could be either marketer-related or household-related. Marketer-related refers, for example, to brand choice (Saab or Volvo) or which dealer to go to. The consumer could, for example, have decided that the next time he or she buys a car it will be a Volvo. Household-related constraints are not brand-dependent; it is more about the price of the car or the size, etc. These pre-decisional constraints influence the rest of the decision process and the information search, but the nature of their influence depends on why the current vehicle is being replaced and on the circumstances at the time. It seems like consumers with marketer-related pre-decisional constraints simplify their purchase decisions more than those consumers with household-related constraints. The first group may decide which car to buy at the same time as they decide to replace the car. Those with household-related pre-decisional constraints seem to make a more complex and engaging search with more considerations. This shows a bit of the internal dynamics of the consumer decision process and different processing strategies (which sometimes seems to simplify the decision, and sometimes does not).

When it comes to searching for information, the individual's beliefs, earlier experience and cognitive abilities have a role. According to TRB (1996), the purchaser's motivation (which is influenced by his or her beliefs and experiences) shapes how much effort goes into collecting information and where the information is sought. The information is plentiful and available from many different sources, and sometimes it is difficult to get the right information. If the product is expensive and important there is greater interest in getting information, and cars seem to be important. Consumers' cognitive abilities are varied, and therefore they search for, transform and use information in very different ways and so different purchasers will make different decisions. Their memory, knowledge and beliefs are different. People use new information in their existing beliefs. If they know nothing about the topic it could be hard to understand or it could be misunderstood. Buying a car can be complex since there are numerous alternatives, and therefore many people eliminate a lot of alternatives and thereby minimize complex and difficult calculations and trade-offs.

It is therefore important to consider how the information is framed and presented because that affects how it is perceived. Whether the available information is deemed good enough and worth paying attention to depends on internal factors, that is, if it is perceived that the information will help in meeting desired goals and if it is consistent with the individual's prior beliefs. It also depends on external factors, such as the design, density and clarity of the information, for example.

Another factor that affects the decision process could have to do with personality and is called cognitive closure. Houghton & Grewal (2000) observed that different people need and want to have different kinds of information. Some people want to come to cognitive closure quickly, which means that they want a firm answer to a question and do not want ambiguity. This type of person and consumer will make a quick decision on few pieces of information, and he or she will avoid information that is incongruent with previously formed beliefs because that type of information could reopen the closed decision process. This means that people with a high need for cognitive closure who also have a brand preference are likely to maintain a positive image of the brand as they are unlikely to search for an alternative. They would also avoid and ignore negative information about the brand, as this information would frustrate closure. This would lead to a focus on information with stereotypical evidence instead of diagnostic evidence. So people with a high need for cognitive closure do not search for information as much as people with a low need for cognitive closure (Houghton & Grewal, 2000).

The conclusion of this section is that there are a number of factors other than socio-economic and socio-demographic ones that affect consumers' decision process in buying a new car. Some of them have been presented here, but there are others. So even if the neighbours look the same from the outside, they have different beliefs, experiences, attitudes etc., and they handle the decision process in different ways due to these factors. But how these factors relate to each other and what weight they have in the decision must be further analysed.

5 A proposition of a new model system

In section 3, we concluded that none of the described model systems fully comply with our requirements as described in the introduction of this report. A new model system is therefore proposed. The formulation of the system is based on experiences and ideas from existing models described in section 2.

The national model system for analysis of passenger transport, Sampers, describes how people choose to travel. Important outputs are travel frequency, destination choice, travelling time, and choice of transport mode. The system also includes a car ownership model developed at VTI. It is, for specific applications, important that the new system provides forecasts that are consistent with the forecasts from Sampers. It will therefore be possible to run the new model both with and without Sampers co-ordination. An improved car ownership model could also replace the one included in the current Sampers version.

5.1 Overview

The proposed model system operates on two levels: the aggregated (national or regional) level and the disaggregated (household or personal) level (*Figure 5*). Forecasts of the total number of cars and total kilometrage are provided from the aggregated models whereas the disaggregated models provide the distribution of cars of different types and ages and the distribution of kilometrage over the same classification of cars. These results are then combined to give the final outcomes, the number and kilometrage of cars of different types and ages.

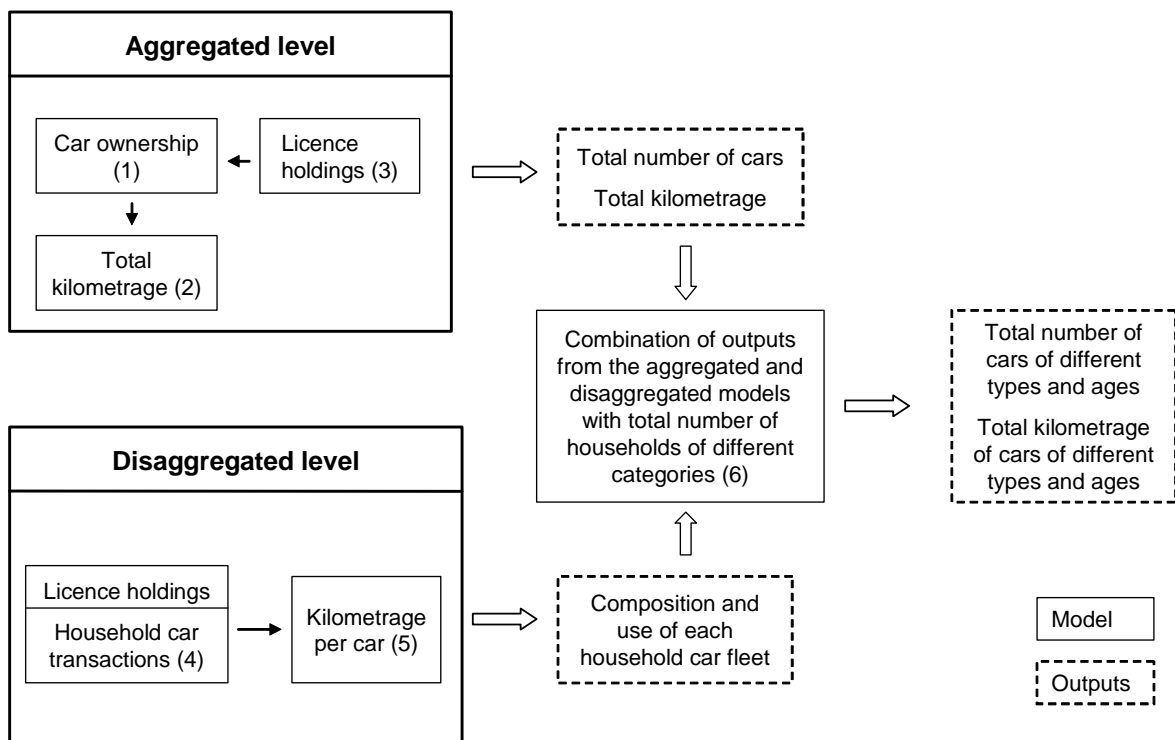


Figure 5 Overview of the proposed model system. The boxes with unbroken lines indicate models, and boxes with dashed lines indicate model outputs.

The main idea behind the two-level system is to overcome the disadvantages that are linked to each type of model category. The disaggregated models have been

shown to underestimate the growth in car ownership (Booz Allen & Hamilton (NZ), 2000). This observation is primarily based on static models that do not capture the long-term trend of increasing car ownership. The problem should be less pronounced for dynamic transactions models, but, since they are often adjusted to short time-series, there is a great risk that they also will underestimate ownership growth. Furthermore, small errors at the household level can add up to large errors at the national or regional level. An aggregated car ownership model is therefore used to control the level of the total number of cars. The strength of the disaggregated models is that they take into account variables that affect decisions at the household level and thereby contribute to the understanding of household behaviour. Car transactions and car type choices are therefore modelled at the disaggregated level. The same reasoning applies to car use. Total kilometrage of all cars is estimated at the aggregated level. The kilometrage of a specific car is assumed to depend on both characteristics of the car and the car owner and is therefore modelled at the disaggregated level.

The two-level system also enables efficient use of data since different data sources can be used for different levels. For example, aggregated models can be based on individual-based registers such as the Swedish vehicle register while disaggregated models can be based on household surveys.

A similar model approach has been used in Australia. The overall level of car ownership was described by an aggregated model, and the cars were then distributed within urban areas using disaggregated models (see Grouenhout & Bell, 1985; RCA & Travers Morgan, 1981; Travers Morgan, 1982; all cited in Booz Allen & Hamilton (NZ), 2000).

5.1.1 Basic concepts

Car types

The basic constraint is that the classification of car types should be consistent with the classification used in the ARTEMIS model. The list of car types in ARTEMIS is not finalised so a definite classification cannot yet be made. What we do know is that petrol and diesel cars, each divided into three classes of cubic capacity, will be included. Alternative fuel type cars will also be included, but the exact categories are not determined.

In addition to the above categories, we will also distinguish between cars of different sizes/functions since this may be important for household choices. A large car with a relatively small engine or a small car with a large engine may be similar with respect to emissions but appeal to different types of households. To start with, we will adopt the classification that Euro NCAP uses when they present their results of crashworthiness: super minis, small family cars, large family cars, executive cars, small MPVs, large MPVs, small off-rovers, large off-rovers, and roadsters. Depending on the composition of the Swedish car fleet, we may need to merge some classes in order to achieve a sufficiently large number of cars in each class.

In summary, the model will distinguish car types defined by the following variables, although not all combinations of these variables will exist:

- Fuel type (at least 3 classes)
- Cubic Capacity (3 classes)
- Size/functions (9 classes).

Households

The majority of the disaggregated models in the literature are household based, that is, the decisions are assumed to be made by the household as a unit. The household concept is rarely defined, but the following definition, frequently used by Statistics Sweden (SCB), is probably generally applicable; this definition says that persons who live in the same dwelling and who share the housekeeping belong to the same household. This definition is, however, impractical to use since such households cannot be identified in the Swedish national registration of the population. Only the more restricted *family* can be identified; a family consists of a married couple with or without children, and unmarried couple with children under the age of 18, or a single person with or without children. Unmarried couples without children are defined as singles. Since forecasts of the population also are restricted to families, the disaggregated models in the proposed model system will be based on families instead of households. Hereafter, the words will be used synonymously.

Model time steps

The model will operate on an annual time step. State variables such as the number of cars a household owns will be evaluated at the end of each year whereas kilometrage will represent the total distance driven throughout the year.

Motivational factors

The term motivational factors is here used in a wide sense to cover factors such as attitudes, norms, and preferences. The conclusion from section 4 was that further research is needed before we can point out the most important factors to include in the model system. For now we will only describe where the motivational factors are important to consider.

Privately owned and company-owned cars

Both privately owned and company-owned cars that are available for private driving are handled in the system. Company-owned cars that are strictly used for business purposes will not be included at this stage.

5.1.2 Co-ordination with Sampers

Forecasts of the total number of cars and total kilometrage from Sampers are used as a basis for national and regional transportation planning by the Swedish Road Administration (SRA) and the Swedish Institute for Transport and Communications Analysis (SIKA). Emission rates are often calculated as complements to these outputs. It is important for the credibility of these applications that the emission calculations are based on data that are consistent with the forecasts from Sampers. Co-ordination with the Sampers system could also improve the quality of our forecasts since that system also includes other transportation modes. It is, on the other hand, also valuable to have a system that works independently of other systems.

To meet both standpoints, it will be possible to run the system both with and without Sampers co-ordination. Three different modes are suggested: co-ordination at the aggregated level (mode 1), co-ordination at both the aggregated and disaggregated levels (mode 2), and independent mode (mode 3). The co-ordination affects three of the sub-models, number (1), (2), and (5) in *Figure 5*.

The implications of the different modes are explained in relation to the summary of these sub-models in the next section.

5.1.3 Summary of the sub-models

This section includes a brief summary of each sub-model, and a more detailed description is provided in the next section. The model number corresponds to the numbers in *Figure 5*.

Car ownership (1)

Car ownership at the aggregated level is described by a cohort model. The model is essentially the same as the one developed at VTI (Matstoms, 2002), which is also included in Sampers. Here, we suggest an improvement of the model by entering, for example, licence holdings as an explanatory variable.

The version of the model currently implemented in Sampers will be used in the co-ordination modes.

Total kilometrage (2)

Total car kilometrage from Sampers is used here if the model is run in either of the co-ordination modes.

In the independent mode, total kilometrage is estimated in two steps. First, the average number of kilometres per car is calculated by a log-linear regression model following the ideas from Wall (1991). Second, the total kilometrage is obtained by multiplying the average car kilometrage by the total number of cars provided from sub-model (1).

Licence holdings (3)

Aggregated licence holdings are, like car ownership, described with a cohort model. Licence holdings are included only to provide inputs to the car ownership model (1) and are therefore used only when the system is run in the independent mode.

Household car transactions (4)

This sub-model is the most comprehensive part of the model system. It is a dynamic transaction model describing household behaviour.

Dynamic models are generally recognised as superior to static models when it comes to describing a household's car purchasing behaviour (Kitamura & De Jong, 1992; Ramjerdi et al., 2000). The main reason is that static models are based on the assumption that households are constantly in equilibrium with respect to their car fleet. This assumption is violated due to the costs involved in buying a new car.

A transaction model can either estimate the duration between transactions (duration models) or whether or not a transaction has occurred during a specific time interval (discrete-time models). There is no clear preference for either of the models in the literature. Kitamura & De Jong (1992) proposed a discrete-time frequency model rather than a duration model. However, their main motivation for this was that the duration models did not take into account time-varying explanatory variables. This limitation has been overcome in later versions of these models, so that reason for proposing a discrete-time model no longer holds.

Nevertheless, the report by Kitamura & De Jong (1992) is valuable since it describes how the transaction models can be linked to economic theory.

We will apply a discrete-time model. This way, we only need to know whether or not the household has changed its car fleet from time T to time T+1; we do not need to know how long a household has owned a specific car.

The number of cars per household at the end of each year is described with a Markov chain model. Whether a transaction has occurred and the type of transaction are modelled given the results from the Markov model. The following transactions are possible: purchasing an additional car, replacement, and disposal. If a car is disposed of, it can either be sold or scrapped. If the household buys a new car, the car type and age are determined in a separate model.

This sub-model is not influenced by Sampers co-ordination, and the same version of the model is used in all modes.

Kilometrage per car (5)

A modified version of the regression model in DVTM will be used to estimate the annual kilometrage of a specific car type (De Jong, 1996). The kilometrage is explained by characteristics of both the car and the household.

Results from the regression models are combined with outputs from Sampers if the system is run in co-ordination mode 2. Sampers gives total car use per person but the kilometrage is not linked to a specific car type. The total kilometrage for a household calculated by Sampers will therefore be distributed between the household cars according to the results from the regression model.

Combining aggregated and disaggregated results (6)

The results from the aggregated and disaggregated models are here combined to provide the final outputs from the model system, that is, total number of cars of different types and ages and total kilometrage for the same division of cars. The calculations include sample enumeration of the household outcomes.

5.2 Description of the sub-models

This section includes descriptions of all six sub-models of the forecasting system. The version of the system described here is the version that can be run independently of the Sampers system.

5.2.1 Car ownership (1)

VTI has developed a regional car ownership model (Matstoms, 2002) which was completed in 1999. This model can provide forecasts on both the national and municipal levels. Although it is an independent model, today it is primarily used as an integrated part of the Sampers system (Beser & Algers, 2001). The model relies on estimates of individuals' entry and exit propensities to and from car ownership. This refers to the proportion of the population that had no car but acquired one during a given year, and, conversely, to those who gave up car ownership status during that time. Non-linear regression models for the entry and exit propensities were estimated on rich statistical material, covering all adults in Sweden and their car ownership status 1980–1995. The most important variables are the individual's age, sex and region of residence. Besides these variables, the models also depend on income, petrol price, increase of GNP, and the proportion of leased-out cars. The last variable is included in order to incorporate company-

owned cars in the model. The model equations are described in detail in Matstoms (2002). As a part of the proposed model system, we suggest that the model will be improved by including licence holdings as an additional explanatory variable. The total number of licence holders in different groups classified by age and sex will be provided from the licence-holdings model (3).

The geographical dimension is handled in the following way. The almost 300 municipalities in Sweden are divided into a certain number of homogenous groups, and for each group the model is estimated separately. That means that all models have the same basic form, but municipalities in different groups have different coefficients.

Given the above types of models, the future number of car owners, by age, sex and municipality, can easily be expressed by a simple recursive relation. A detailed description of the number of car owners the first year and a corresponding description of the population for all coming years are then assumed.

5.2.2 Total kilometrage (2)

Total kilometrage is calculated as the average number of kilometres per car times the total number of cars. The average number of kilometres per car is described by a log-linear regression model following the ideas from Wall (1991). The general form of the model is

$$\ln Y = \ln \beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \dots + \beta_p \ln X_p + \ln \varepsilon \quad (5)$$

where Y is the dependent variable, X_1, \dots, X_p are explanatory variables, $\beta_0, \beta_1, \dots, \beta_p$ are regression parameters and ε is an error term. The parameters can be estimated with ordinary least squares methods.

Edwards et al. (1999) have developed a model for calculating the total kilometrage of different types of vehicles in Sweden (motorcycles, cars and light trucks, heavy trucks, etc). Total kilometrage for each year from 1950 onwards have been calculated with this model. These data, divided by the total number of cars will be used as the dependent variable. The choice of explanatory variables very much follows what Wall (1991) suggested, but we have made some modifications to comply with the data that are available from official sources. The basic set of variables will be fuel price index, income index, number of cars per 1000 inhabitants, proportion of cars that are owned by pensioners, and proportion of cars that are owned by women. The resulting set of variables that will be included in the model will be a result from the model development.

A couple of potential problems with this model approach were highlighted in Wall (1991). First, he experienced some problems with the parameter estimation, due to multi collinearity in the explanatory variables. However, since we will not use the exact variables he did, it is not possible to say beforehand whether the same problem will occur. If it does, we can either modify the model specification or use another estimation technique. The second problem was that the residuals followed a pattern that is typical for positively auto correlated data. If this is the case, the parameter estimates may be biased. Again, we do not know if this will occur in our model, but if it does, it can be handled by modelling the errors in (5) as an ARMA process; see for example Makridakis et al. (1998).

5.2.3 Licence holdings (3)

Licence holding is, like car ownership, modelled with a cohort approach. However, licence holding is more stable over a person's lifetime than car ownership. Once a licence has been acquired it is often kept throughout the rest of the lifetime, or at least into old age. Exit propensities are therefore negligible for most ages, and entry and exit propensities will not be modelled separately. Instead, the proportion of persons in an age cohort that change licence-holding status from one year to the next, hereafter called propensity of change, will be modelled. For young age classes, the propensity of change will almost equal entry propensities.

We will follow the work by Byrsjö (1998) where he modelled entry propensities based on data from the Swedish register of licence holders. He showed that the following log-linear model gave good results for both men and women:

$$\pi = \alpha KKI^\beta e^{\gamma AGE} \varepsilon \quad (6)$$

The variables *KKI* and *AGE* denote price index for taking a licence and age; α , β , and γ are model parameters, and ε is the error term. He applied the model for ages 18–50. Cedersund & Henriksson (2005) used the same model for the age class 18–24. Since entry propensities in this age group largely differ from propensities in older groups, the model estimated separately on this group showed improved fit. We will therefore use model (6) with separate parameter estimates for the age class 18–24 and 25–50, and for men and women.

Model (6) always gives positive propensities, which implies that the cohort as a whole increases the number of licence holders from one year to the next. In age cohorts over 50, we will allow for the possibility of a decreasing number of licence holders by using the following linear regression model:

$$\pi = \alpha + \beta KKI + \gamma AGE + \varepsilon \quad (7)$$

Here too, separate models will be estimated for men and women.

5.2.4 Household car transactions (4)

The car transaction model describes changes in a household's car fleet with respect to its size and composition. This is the most complex of the models in the proposed system and comprises in itself a number of sub-models. An overview of the model is shown in *Figure 6*. All characteristics of the family and its car fleet at the base year of the calculation are assumed to be known, either from register data or from surveys.

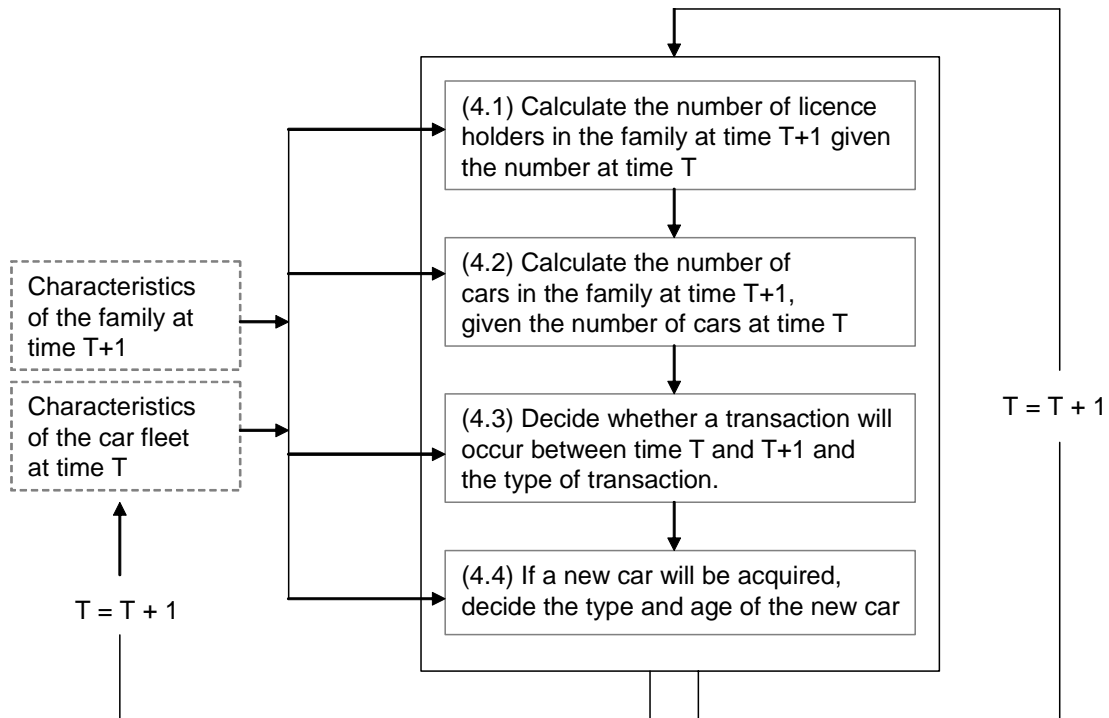


Figure 6 Overview of the car transaction model.

Licence holdings (4.1)

The number of licence holders at each time step is calculated and used as an important input to the car ownership model (4.2). The probability that a person changes his or her licence-holding status is taken from the corresponding propensity in the aggregated cohort model (3). This is the same procedure that is used in the ALTRANS model (Christensen & Gudmundsson, 2003). Whether a person changes licence-holding status or not thus depends on the same explanatory variables that are used in model (3).

Car ownership (4.2)

The status of car ownership at the end of each year is described by a Markov chain. The possible states are that the family owns 0, 1, 2, or 3 cars. Let Z_t be the random variable that describes the number of cars the family owns at time t , and let P_{ij} denote the transition probability from state i to j , that is, the probability that a family owns j number of cars at time $t+1$ and i number of cars at time t ;

$$P_{ij} = P(Z_{t+1} = j | Z_t = i) \quad (8)$$

A basic property of the Markov chain is that the probability of being in a certain state at the next time period depends only on the current state and not on previous states. This can be expressed as

$$P(Z_{t+1} = j | Z_t = i, Z_{t-1} = i_{t-1}, \dots, Z_1 = i_1, Z_0 = i_0) = P(Z_{t+1} = j | Z_t = i) \quad (9)$$

and is called the Markovian property. If this property does not hold, the definition of the states can be redefined and include the number of cars at both the current and one or several of the previous time periods.

For simplicity, we will assume that only one transaction will occur during each time period. Possible transactions are purchasing an additional car, disposing of a car, or replacing a car. This assumption restricts the transition matrix in the following way:

$$\begin{pmatrix} P_{00} & P_{01} & 0 & 0 \\ P_{10} & P_{11} & P_{12} & 0 \\ 0 & P_{21} & P_{22} & P_{23} \\ 0 & 0 & P_{32} & P_{33} \end{pmatrix}$$

For example, it is not possible to change from zero to two cars during one year ($P_{02} = 0$). If this assumption turns out to be unrealistic, the model can easily be extended to include several transactions. The point of restricting the number of transactions is to decrease the number of parameters that need to be estimated and thereby decrease the data demand.

The household and car fleet characteristics will be taken into account when estimating the transition probabilities by using a method developed by Aalen et al. (2001). This means that the probabilities will be different for different households.

Type of transaction (4.3)

The Markov model in (4.2) describes only how many cars a family owns. Depending on whether the number has changed or not, the next step is to decide if a transaction has occurred. This is done according to the scheme in *Table 3*.

Table 3 *Actions taken by the household for each possible combination of car ownership at time t and t+1.*

Number of cars owned at		
time t	time t+1	Action taken
0	0	Do nothing
0	1	Decide whether to purchase a private or a company car
1	0	Decide whether to sell or scrap the car
1	1	Decide whether to keep the old car or replace the car; given replacement, decide whether the replacement car will be a private or a company car
1	2	Decide whether to purchase a private or a company car
2	1	Decide which car to dispose of, decide whether to sell or scrap the car
2	2	Decide whether to keep both cars or replace one of the cars; given replacement, decide which car to replace and decide whether the replacement car will be a private or a company car
2	3	Decide whether to purchase a private or a company car
3	2	Decide which car to dispose of, decide whether to sell or scrap the car
3	3	Decide whether to keep all three cars or replace one of the cars; given replacement, decide which car to replace and decide whether the replacement car will be a private or a company car

Each of the decisions that are described in *Table 3* is modelled with a binary or multinomial logit depending on the number of possible choices. For example, if a household will dispose of a car during the year, the choice between selling and scrapping the car is modelled by a binary logit. If the household owns three cars at time t and two cars at time $t+1$, a multinomial logit is used to model which car will be disposed of. The reason for deciding whether the new car should be privately or company-owned is that the car type choice can be quite different depending on who will own the car.

Type and age of the new car (4.4)

If the family decides to purchase a car, it can choose between a number of types defined by the following variables: fuel type (at least 3 classes), cubic capacity (3 classes), and size/function (9 classes), although not all combinations of these classes will be available. In addition, the age of the car needs to be decided. The choice of what car to purchase will be modelled by a nested logit model. This is a commonly used approach to model car type choice; see for example Page et al. (2000) or Mohammadian & Miller (2003). A four-level nested logit will be used, the layers representing fuel type, cubic capacity, size/function, and age; although, not necessarily in that order. To restrict the number of choices, we will assume that if a car is a replacement for another car, the replacement car will be newer than the previous car. Again, this assumption can be relaxed if it turns out to contradict the household's behaviour.

Separate car choice models will be estimated for privately owned and company-owned cars.

When the aggregated and disaggregated results have been combined (as described in section 5.2.6), the total car fleet will be checked so that the demand for used cars does not exceed the supply. If this is the case, the household's car fleet needs to be adjusted.

Explanatory variables

As shown in *Figure 6*, the car transactions are explained by the characteristics of the car fleet at time t and the characteristics of the family at time $t+1$. For each model (4.2)–(4.4), we will start with the same set of explanatory variables and gradually remove the ones that do not contribute to explaining the respective behaviours. The starting set of variables will essentially be the ones listed in *Table 2* plus a number of motivational factors.

5.2.5 Kilometrage per car (5)

It is important to calculate kilometrage for a specific car rather than for a specific person since car use for different types of cars is a key output from the model system. These two measures might be quite different since one person can drive several cars and, conversely, one car can be driven by several persons. However, how much a car is used of course depends on both the type of car and the family that has access to the car. We will therefore model car use with respect to both attributes of the car and of the family.

The model formulation is in principle the same as was used in the Dutch model system DVTM (De Jong, 1996); see equation (4) in section 2.5.5. This model included only car costs, variable and fixed, as car attributes. The distance a car is

driven may also depend on other attributes; for example, for long trips not only fuel efficiency but also the size of the luggage space and how comfortable the car is may be important. We also know that newer cars are driven longer distances than older cars. Two variables have been added to the model which are not in the original formulation by De Jong (1996): car type (T) and age of the car (a). Car type is entered as a categorical variable with the same nine types that are used throughout the whole model system. The different engine size classes are indirectly included as variable car costs. It is not evident how the age of the car should be entered in the model; this must be further investigated. We therefore use the general formulation $f(a)$, which means that any transformation of a could be used. For family i and car j the kilometrage is modelled as

$$\ln(k_{ij}) = \alpha \ln(d_j - c_i) + \beta v_i + \gamma \mathbf{Z}_j + T_i + f(a_i) + e_{ij} \quad (10)$$

where k_{ij} is annual kilometrage of the car, d_j is annual net family income, c_i is fixed car costs, v_i is variable car costs, \mathbf{Z}_j is a vector of attributes of the family, T_i is the car type, $f(a_i)$ is a function of the age of the car, e_{ij} is the error term, and α , β , γ are model parameters.

The following list of attributes of the family is based on factors that have been shown to be important for car use in the two studies by De Jong (1996) and Steg et al. (2001).

- Type of family (single with children, single without children, couple with children, couple without children)
- Age
- Gender
- Education level
- Employment
- Environmental awareness.

We will also take into account whether the car is privately or company-owned and motivational factors.

5.2.6 Combining aggregated and disaggregated results (6)

The disaggregated transaction model (4) gives results for each type of household that enters the model in the form of the probability of owning a car of a specific type and age. These probabilities are then weighted by the distribution of the household types to represent the total population. The procedure is called sample enumeration and is described in Train (2002). The total number of cars of a specific type and age is finally calculated by multiplying the adherent probability by the total number of cars from model (1).

The kilometrage of each car and household type is handled in a similar way and the results are generalised to represent the distribution of kilometrage for different car types and ages. These results are combined with the total kilometrage from model (2) to calculate the final outcome, which is total kilometrage for each car class.

5.3 Forecasts and scenarios

The model system can be used to calculate forecasts for a wide variety of scenarios. Scenarios that are made by altering one or several of the explanatory variables are easily implemented. Examples of such scenarios are: an increase in household incomes, a larger proportion of elderly persons, and change of attitudes. Other scenarios can be studied if they can be redefined as changes in the explanatory variables. One example is the introduction of CO₂-differentiated vehicle taxes. A change in the registration tax can be represented as an increased purchase price, and a change in the annual vehicle tax can be represented as a change in the annual fixed costs.

Forecasts of the population, such as the number and composition of families, are not conducted within the system. When such forecasts are needed for scenario calculations, they can be obtained from sources such as Statistics Sweden or Inregia AB.

6 The steps towards a working model

The implementation of the proposed model system is not straightforward. We here briefly describe the research and development that is required to obtain a working model.

The literature overview in section 4 showed that motivational factors played an important role in the decisions related to car ownership and use. However, further research is needed to clarify which factors should be included in the model. We believe that this research should be based on the theory of planned behaviour. This is a well-established theory in the behavioural sciences and it was shown by Caprara et al. (1998) that this theory is also applicable to car purchasing behaviour. The theory has also been used to explain why young persons do or do not choose to acquire a driver's licence (Engström & Nolén, 2005). There are also well-established methods to measure the different factors included in the model (attitudes, social norms, etc). As a first step, we will conduct a small-scale questionnaire study to investigate how these factors correlate to car purchasing and car use behaviour and how they interact with each other. From this study we can draw conclusions about what factors to include in the model system.

The next step is to make a large-scale survey where motivational, socio-economic, and socio-demographic variables are investigated, as well as car purchasing and car use behaviour. These data will be the basis for estimating the parameters in the disaggregated models. Moreover, these data could also be used to estimate the prevalence of specific attitudes, norms, etc. in the population, which is important to know for base case scenarios. Preferably, data such as family income, car type, and kilometrage should be taken from existing registers and not asked in the questionnaire in order to minimise measurement errors. However, this means that several registers must match, and this is only allowed after an examination of the study by the ethical board at Statistics Sweden.

To estimate the transaction model we need data from at least two consecutive years from the same families. The best way to do this is through a panel survey with two waves where the same sample receives questionnaires two years in a row. An alternative would be to send out the questionnaire once and ask about car ownership at both the current and previous turn of the year. However, such a procedure would lead to errors in the data due to the difficulty of remembering the situation one year ago. Two years is enough to estimate the parameters in the model, but additional years would improve the quality of the model.

The aggregated models can be estimated directly on data from official registers, and no new data collection is needed. The registers that will be used are: the vehicle register, the national registration of the population, the register of licence holdings, and total income statistics compiled by Statistics Sweden.

When all sub-models are developed and implemented, the system should be tested as a whole. This is done in two ways. First, the model should be validated to see how well the system performs in a forecasting situation. This could be done on previous data, although we will not know all inputs at that time. Second, a sensitivity analysis should be conducted to investigate how sensitive the model is to changes in the input variables (Saltelli et al. 2000). Even though we have the model specification of all sub-models, it is very difficult to predict how the system will behave as a whole. It may turn out that the model outputs are extremely sensitive to changes in a specific input or that the response to a specific range of the input variables behaves in an erratic way.

7 Discussion

A new forecasting model system for the composition and use of the Swedish car fleet is proposed in this report. The system is formulated to comply with the requirements described in the introduction. That is, it provides forecasts of the total number of cars in the Swedish car fleet, the number of cars of different types and ages and kilometrage of cars with the same classification. The system also describes household choices with respect to car type, age and use. Finally, it can provide forecasts for scenarios based on car costs, fuel costs, household characteristics, demographic variables, and other factors.

In the literature overview, we did not find any single system that could be directly adopted. We therefore formulated a new system, based on a combination of model formulations and ideas from a variety of different models.

A deficiency in the existing models is that they generally rely only on objective factors such as socio-economic and socio-demographic variables, although a number of studies have shown that motivational factors such as attitudes and norms also have a strong impact on a household's behaviour. One reason for not including such variables may be that they are difficult to observe, and it is not obvious how the factors should be included in the model. On the other hand, even if motivational factors are not explicitly entered in the model, they implicitly influence the forecasts and model fits. For example, if motivational factors are excluded from a model, the forecasts are calculated with the implicit assumption that the motivational factors will remain the same as for the years that were used to estimate the model parameters. By including the variables explicitly, it becomes apparent how they influence the forecasts. It also becomes clearer how much of the behaviour can be affected by other factors, for example economic incentives. Considering the large potential improvements of the models, there is a strong need for further research on how motivational factors can be measured and included in the model system.

The proposed system is entirely focused on car ownership and car use; no other transport modes are considered. We have chosen to concentrate the efforts on the car fleet and primarily on the household's behaviour related to car ownership and use. To develop a new transportation model in parallel to Sampers would also be of questionable value. Instead, we describe how the new system can be coordinated with Sampers and in that way take into account the interaction between different transport modes.

We have not discussed different geographical scales in this report, but more or less assumed that the national scale is the level of interest. There is, however, no principal difference between the regional and the national scales. The model could just as easily be adjusted to a region as to the whole country, as long as the data are available at the corresponding level. A regional version of the aggregated car ownership model already exists. The only problem is that a car can be driven in more than one region, so the kilometrage for a region may be under- or overestimated if the net contribution differs between the regions. The extent of the problem depends on the size of the regions; the larger the region, the smaller the problem. Of course, this problem also exists on the national level; a car registered in Sweden can be driven in another country, and foreign cars can be driven in Sweden. The net contribution to the total kilometrage is, however, probably negligible.

It is difficult to find a classification of cars with respect to size/function in which both are well defined and important in relation to household choices. We adopted the classification scale used by Euro NCAP, which has the advantage that all cars that have been tested are classified according to this scale. However, cars that are not tested, mostly older cars, are not classified, and this must be done before the model can be developed.

8 References

- Aalen, O.O., Borgan, O. & Fekjaer, H: **Covariate Adjustment of Event Histories Estimated from Markov Chains: The Additive Approach.** Statistical Research Report No. 2. Department of Mathematics, University of Oslo. Oslo. 2001.
- Agresti, A: **Categorical Data Analysis.** John Wiley & Sons. New York. 1990.
- Beser, M. & Algers, S: **SAMPERS – the New Swedsh National Travel Demand Forecasting Tool.** National Transportation Models – Recent Developments and Prospects. Springer. 2001.
- Birkeland, M. E. & Jordal-Jørgensen, J: **Energy Efficiency of Passenger Cars.** Paper for European Transport Conference 2001, PTRC. Cambridge. 2001.
- Boe, O., Fujii, S. & Gärling, T: **Empirical Tests of a Model of Automobile Choice Incorporating Attitude, Habit, and Script.** Paper presented at the Urban Transport systems conference, Lund University, Sweden, June 7–8. 1999.
- Booz Allen & Hamilton (NZ): **Model for Forecasting Vehicle Ownership in New Zealand.** Transfund New Zealand Research Report 161. Wellington. 2000.
- BTCE: **Costs of Reducing Greenhouse Gas Emissions from Australian Cars: An Application of the BTCE CARMOD Model.** Working paper 24. Bureau of Transport and Communications Economics. Canberra. 1996.
- Byrsjö, J: **Benägenheten att ta körkort. Faktorer som påverkar Sveriges befolkning att ta körkort.** VTI notat 80-1998. Statens väg- och transportforskningsinstitut. Linköping. 1998.
- Caprara, G.V., Barbaranelli, C. & Guido, G: **Empirical Investigations of Determinants of Purchase Intention According to the Theory of Planned Behavior.** Ricerche di Psicologia. Vol 22. No 3. pp 147–168. 1998.
- Cedersund, H-Å & Henriksson, P: **Modell för att skatta ungdomars körkortstagande.** VTI rapport 511 (manuscript). VTI. Linköping. 2005.
- Choo, S. & Mokhtarian, P. L: **What Type of Vehicle do People Drive? The Role of Attitude and Lifestyle in Influencing Vehicle Type Choice.** Transportation Research Part A. Vol 38. pp 201–222. 2004.
- Christensen, L. & Gudmundsson, H: **Modelanalyser af mobilitet og miljø. Slutrapport fra ALTRANS og AMOR II.** DMU report 447. Danmarks Miljøundersøgelser, Miljøministeriet. 2003.
- Cramer, J. S: **The LOGIT Model : an Introduction for Economists.** Edward Arnold. London. 1991.
- Dargay, J. M. & Gately, D: **Income's Effect on Car and Vehicle Ownership, Worldwide: 1960-2015.** Transportation Research Part A. Vol 33. pp 101–138. 1999.
- De Jong, G. C: **A Disaggregate Model System of Vehicle Holding Duration, Type Choice and Use.** Transportation Research Part B. Vol 30. No 4. pp 263–276. 1996.
- De Jong, G. C., Fox J, Pieters, M., Vonk, L. & Daly, A: **Audit of Car Ownership Models.** Report 01192. RAND. Santa Monica. 2002.
- De Jong, G. C., Vellay, C. & Fox, J: **Vehicle Scrappage: Literature and a New Stated Preference Survey.** Paper for European Transport Conference 2001, PTRC. Cambridge. 2001.

- Edwards, H., Nilsson, G., Thulin, H. & Vorwerk, P: **Trafikarbetet uttryckt i fordonskilometer på väg i Sverige 1950-1997**. VTI rapport 439. Statens väg- och transportforskningsinstitut. Linköping. 1999.
- Engström, I. & Nolén, S: **Ungdomars inställning till att ta körkort 2002-2004**. VTI rapport 510. VTI. 2005.
- Golob, T. F., Bunch, D. S. & Brownstone, D: **A Vehicle Use Forecasting Model Based on Revealed and Stated Vehicle Type Choice and Utilisation Data**. Journal of Transport Economics and Policy. Vol 31. pp 69–92. 1997.
- Golob, T. F. & Hensher, D. A: **Greenhouse Gas Emissions and Australian Commuters' Attitudes and Behaviour Concerning Abatement Policies and Personal Involvement**. Transportation Research D. Vol 3. pp 1–18. 1998.
- Hammarström, U. & Karlsson, B. O: **EMV – a PC program for calculating exhaust emissions from road traffic: Program description and user guide**. VTI meddelande 849A. Swedish National Road and Transport Research Institute. Linköping. 1998.
- HCG & TØI: **A Model System to Predict Fuel Use and Emissions from Private Travel in Norway from 1985 to 2025**. Hague Consulting Group. The Netherlands. 1990.
- Houghton, D., C. & Grewal, R: **Please, Let's Get an Answer – Any Answer: Need for Consumer Cognitive Closure**. Psychology & Marketing. Vol 17. No 11. pp 911–934. 2000.
- Kitamura, R. & de Jong, G. C: **Dynamic Vehicle Transaction Models: Report of the Feasibility Study**. Report 130-2. Hague Consulting Group. 1992.
- Kitamura, R., Mokhtarian, P. L. & Laidet, L: **A Micro-Analysis of Land Use and Travel in Five Neighborhoods in the San Francisco Bay Area**. Transportation. Vol 24. pp 125–158. 1997.
- Kommunikationsdepartementet: **Transportpolitik för en hållbar utveckling**. Regeringens proposition 1997/98:56. 1998.
- Kveiborg, O: **Forecasting Developments of the Car Fleet in the Altrans Model**. Paper presented at the conference Trafikdage på AUC. August 30–31st, 1999.
- Makridakis, S., Wheelwright, S. C. & Hyndman, R. J: **Forecasting: Methods and Applications**. John Wiley & Sons, Inc. 1998.
- Marell, A., Davidsson, P. & Gärling, T: **The Role of Aspiration level for Decisions to Replace Automobiles**. 7th World Conference of Transport Research. Elsevier Science Ltd. Sydney, Australia. 1996.
- Matstoms, P: **Modeller och prognoser för regionalt bilinnehav i Sverige**. VTI rapport 476. Linköping. 2002.
- Mohammadian, A. & Miller, E. J: **Empirical Investigation of Household Vehicle Type Choice Decisions**. Transportation Research Record. Vol 1854. pp 99–106. 2003.
- Page, M., Whelan, G. & Daly, A: **Modelling the Factors which Influence New Car Purchasing**. Paper presented at the European Transport Conference 2000. PTRC, Cambridge. 2000.
- Punj, G. & Brookes, R: **The influence of pre-decisional constraints on information search and consideration set formation in new automobile purchases**. International Journal of Research in Marketing. Vol 19. pp 383–400. 2002.

- Ramjerdi, F., Rand, L. & Saetermo, I-A: **Models for Car Ownership, Transactions and Vehicle Type.** Bullentin nr 187:2. Teknik och samhälle. Trafikplanering. Lunds universitet. Lund. 2000.
- Saltelli, A: **Sensitivity Analysis.** Wiley. Chichester. 2000.
- Sand Jespersen, M., Jordal-Jørgensen, J., Kristensen, N. & Matstoms, P: **Impacts from CO2 differentiated vehicle taxes on CO2 emissions from passenger cars.** 5187. Naturvårdsverket. Stockholm. 2002.
- Steg, L., Geurs, K. & Ras, M: **The Effects of Motivational Factors on Car Use: a Multidisciplinary Modelling Approach.** Transportation Research Part A. Vol 35. pp 789–806. 2001.
- Tertoolen, G., van Kreveld, D. & Verstraten, B: **Psychological Resistance Against Attempts to Reduce Private Car Use.** Transportation Research A, Vol 32. pp 171–181. 1998.
- Train, K. E: **Discrete Choice Methods with Simulation.** Cambridge University Press. 2002.
- TRB: **Shopping for Safety: Providing Consumer Automotive Safety Information.** 248. Transportation Research Board. 1996.
- Wall, R: **Bilansvändningens bestämningsfaktorer.** VTI meddelande 648. Statens väg- och transportforskningsinstitut. Linköping. 1991.
- Whelan, G. A., Wardman, M. & Daly, A. J: **Is There a Limit to Car Ownership Growth? An Exploration of Household Saturation Levels Using Two Novel Approaches.** Paper presented at the European Transport Conference 2000. PTRC, Cambridge. 2000.
- Wu, G., Yamamoto, T. & Kitamura, R: **Vehicle Ownership Model that Incorporates the Causal Structure Underlying Attitudes toward Vehicle Ownership.** Transportation Research Record. Vol 1676. pp 61–67. 1999.

Multinomial and nested logit models

The multinomial and nested logit models are briefly described here. For further details see, for example, Cramer (1991) or Train (2002).

The multinomial logit models the choice of any finite number of alternatives. For clarity of presentation we will start by introducing the special case in which the choice set is restricted to two alternatives ($Y = 0$ or $Y = 1$). This is referred to as the bivariate logit. In this model, the probability of household i choosing $Y = 1$ (hereafter denoted $\pi_1(\mathbf{x}_i)$ where \mathbf{x}_i is a vector of explanatory variables) is described by a sigmoid curve with the functional form

$$\pi_1(\mathbf{x}_i) = \frac{\exp(\boldsymbol{\beta}'\mathbf{x}_i)}{1 + \exp(\boldsymbol{\beta}'\mathbf{x}_i)} \quad (\text{I.1})$$

where $\boldsymbol{\beta}$ is a vector of model parameters. When the parameters have been estimated from observed data, the probability, $\pi_1(\mathbf{x}_i)$, can be calculated for any households described by \mathbf{x}_i . This outcome is of interest in studies where the aim is to estimate the distribution of choices for a population of households. In other studies, it may be of interest to classify individual households. The probability of choosing state 1 is then compared to a predefined limit. If the probability exceeds this limit then the household is said to choose state 1, else, the household is said to choose state 0. This limit is often set to 0.5 but other values are sometimes chosen.

One reason for the popularity of the logit models in economic applications is the close connection to the theory of random utility maximisation. For each possible state, household i attaches a random utility denoted U_{i0} for the choice $Y_i = 0$ and U_{i1} for $Y_i = 1$. The utilities can be formulated as

$$U_{i0} = V_{i0} + \varepsilon_{i0} \quad (\text{I.2a})$$

$$U_{i1} = V_{i1} + \varepsilon_{i1} \quad (\text{I.2b})$$

where V_i is a systematic component that is a linear function of the explanatory variables \mathbf{x}_i and ε_i is a random component. Utility maximisation implies that the state with the highest utility is chosen, that is, state 1 is chosen, if the probability of $U_{i1} > U_{i0}$, $P(U_{i1} > U_{i0})$ exceeds 0.5. If the errors ε_{i0} and ε_{i1} are independent and follow a type I extreme value distribution, the probability $P(U_1 > U_0)$ is described by the bivariate logit model

$$\pi(V_i) = P(Y_i = 1) = P(U_{i1} > U_{i0}) = \frac{\exp(V_i)}{1 + \exp(V_i)} \quad (\text{I.3})$$

We now return to the multinomial logit. For household i , the probability of choosing alternative j can be expressed as

$$\pi_j(\mathbf{x}_{ij}) = \frac{\exp(\boldsymbol{\beta}'\mathbf{x}_{ij})}{\sum_{h \in C_i} \exp(\boldsymbol{\beta}'\mathbf{x}_{ih})} \quad (\text{I.4})$$

where C_i is the set of response alternatives. Different households can thus have different possible outcomes to choose from. Moreover, the explanatory variables can include both characteristics of the choices and characteristics of the chooser. This specification of the model was first proposed by McFadden in 1974 (see Agresti, 1990; chapter 9.2.3).

In a multinomial logit, the odds of choosing a over b are independent of the other alternatives in the choice set. This property is known as independence from irrelevant alternatives (IIA). If this property does not hold, another model specification should be considered.

If the IIA property does not hold for the entire choice set, but the choices can be grouped into subsets (or nests) where the property holds within each subset, then the nested logit model can be applied. The probability for household i to choose alternative j in nest k is calculated as

$$\pi_{jk}(\mathbf{x}_{ijk}) = \frac{\exp\left(\frac{\boldsymbol{\beta}'\mathbf{x}_{ijk}}{\lambda_k}\right) \left(\sum_{m \in B_k} \exp\left(\frac{\boldsymbol{\beta}'\mathbf{x}_{imk}}{\lambda_k}\right)\right)^{\lambda_k - 1}}{\sum_{n=1}^N \left(\sum_{m \in B_n} \exp\left(\frac{\boldsymbol{\beta}'\mathbf{x}_{imn}}{\lambda_n}\right)\right)^{\lambda_n}} \quad (\text{I.5})$$

where B_k is the choice set of nest k , and N is the total number of nests. The parameter λ_k is a measure of dependence between alternatives in nest k .

VTI är ett oberoende och internationellt framstående forskningsinstitut som arbetar med forskning och utveckling inom transportsektorn. Vi arbetar med samtliga trafikslag och kärnkompetensen finns inom områdena säkerhet, ekonomi, miljö, trafik- och transportanalys, beteende och samspel mellan människa-fordon-transportsystem samt inom vägkonstruktion, drift och underhåll. VTI är världsledande inom ett flertal områden, till exempel simulatorteknik. VTI har tjänster som sträcker sig från förstudier, oberoende kvalificerade utredningar och expertutlåtanden till projektledning samt forskning och utveckling. Vår tekniska utrustning består bland annat av körsimulatorer för väg- och järnvägstrafik, väglaboratorium, däckprovningsanläggning, krockbanor och mycket mer. Vi kan även erbjuda ett brett utbud av kurser och seminarier inom transportområdet.

VTI is an independent, internationally outstanding research institute which is engaged on research and development in the transport sector. Our work covers all modes, and our core competence is in the fields of safety, economy, environment, traffic and transport analysis, behaviour and the man-vehicle-transport system interaction, and in road design, operation and maintenance. VTI is a world leader in several areas, for instance in simulator technology. VTI provides services ranging from preliminary studies, highlevel independent investigations and expert statements to project management, research and development. Our technical equipment includes driving simulators for road and rail traffic, a road laboratory, a tyre testing facility, crash tracks and a lot more. We can also offer a broad selection of courses and seminars in the field of transport.

vti

HUVUDKONTOR/HEAD OFFICE

LINKÖPING

POST/MAIL SE-581 95 LINKÖPING

TEL +46 (0)13 20 40 00

www.vti.se

BORLÄNGE

POST/MAIL BOX 760

SE-781 27 BORLÄNGE

TEL +46 (0)243 446 860

STOCKHOLM

POST/MAIL BOX 6056

SE-171 06 SOLNA

TEL +46 (0)8 555 77 020

GÖTEBORG

POST/MAIL BOX 8077

SE-402 78 GÖTEBORG

TEL +46 (0)31 750 26 00