Adding new modes to existing transport demand gravity models

Methodology for adding disaggregate discrete choice parameters to an aggregate gravity model

Master Thesis

TU Delft
dat mobility

R.J. van Kuijk

Adding new modes to existing transport demand gravity models

Methodology for adding disaggregate discrete choice parameters to an aggregate gravity model

By

Roy Josephus van Kuijk

In partial fulfilment of the requirements for the degree of

Master of Science

In Civil Engineering at
Delft University of Technology

To be defended on 3rd of July, 2018

This thesis can be retrieved electronically from: www.repository.tudelft.nl

Student number:	4002059	
Project duration:	15th of August, 2017 – 3rd of July, 2018	
Thesis committee:	Prof. dr. ir. B. van Arem	Chairman, TU Delft (CEG)
	Dr. ir. R. van Nes	Supervisor, TU Delft (CEG)
	Dr. ir. S. van Cranenburgh	Supervisor, TU Delft (TPM)
	Ir. L.J.N. Brederode	Supervisor, DAT.Mobility

Version 2 July, 2018

Sources cover pictures (used under License CC 4.0)

A car2go one-way car sharing vehicle in Seattle, Washington
Made by Aaron Rogosin Photography
http://aaronrogosin.com/wp-content/uploads/2012/12/car2go-seattle-market-1008.jpg
Transportation seamless background
Created by Macrovector - Freepik.com
https://www.freepik.com/free-vector/transportation-seamless-background_1528970.htm

Preface

This thesis – written in partial fulfilment of the requirements for the degree of Master of Science – reflects the research I have been conducting for the last nine months. I conducted this research for the transport consultancy firm DAT.Mobility. My research involved the development of a new methodology for combining different transport demand models. My motivation for developing this methodology lies in its potential value for exploring future mobility behaviour especially when this relates to new modes or mobility services.

I hope that the work reflected in this thesis provides transport planning practitioners inspiration to improve current transport demand models. The proposed methodology provides information about how these models can be extended with external utility functions from different sources. Adding these external utility functions can make transport demand models more future proof as they are able to provide indicatory forecasts about the uptake of new modes or mobility services.

Metaphorically speaking, the conduct of research and the writing of this thesis has been a long, but also beautiful trip. I am very thankful for all inspiration, guidance and knowledge support I got. First of all, I would like to thank Luuk Brederode, who has been my daily supervisor. Working with Luuk brought lots of fun and I really appreciated his interest in my work. During my internship at DAT.Mobility he provided me with lots of useful information and constructive comments on my work. Also his colleagues were really supportive and fun to work with; especially Rogier Koopal and Christiaan Palsrok helped me a lot during my internship.

My other supervisors at the TU Delft have also been really great to work with. I value their extensive knowledge about the related subjects and their open and constructive way of disusing my research progress. Rob van Nes can talk a lot about transport demand models and his passion for modelling has been a great inspiration to me. Sander van Cranenburgh helped me a lot in relation to discrete choice modelling and was very supportive by asking critical questions in order to send me in the right direction. The chairman of my thesis committee, professor Bart van Arem, kept a very good overview on my research progress and helped me with his more general view on my work. This prevented me for getting completely struck in the details of my research.

Last, but certainly not least, I would like to thank all my family and friends who have been such a great support. I was able to keep going by telling them very passionate about my research subject or complaining by them when I was not satisfied with the results so far. All those coffee meetings, lunch dates, phone calls and text conversations made me get the best out of it.

With this all being said, I would like you – as a reader – to be inspired by the research I have been conducting. Hopefully, my enthusiasm about the subject is able to inspire your future work.

Enjoy the read.

Roy van Kuijk
Delft, 25 June 2018

Executive summary

Mobility innovations and the trend of servitisation drive the development and demand for new modes and transport services such as shared cars and bicycles or demand responsive transport. Given the emergence of new modes, an important question is how these emerging modes will lead to behavioural changes and impact mobility. This is important as it potentially affects the way transport infrastructure is planned and transport policy is established. Information about future transport is thus necessary in order to prepare for impact on relevant areas, for example reflected by the 5E framework: effective mobility, efficient cities, equity, environment and economy. (van Oort, 2017)

Problem statement

It is possible to estimate new discrete choice transport demand models based on stated preference research, but this requires a significant amount of time and financial resources. Therefore, it is a welcome thought to investigate the possibility for combining discrete choice utility functions with existing transport demand gravity models such that they are able to provide indicational forecasts about the future modal splits in the presence of these new modes. For lower spatial levels in the Netherlands, many transport demand models still have an aggregate approach. This is especially interesting for transport planners for urban areas, such as Rotterdam, The Hague and Eindhoven as for those cities transport demand gravity models are intensively used.

Related to the combination of these transport demand models, three sub-problems are identified which need to be account during the methodology development:

- The existing transport demand gravity model is specified within a different modelling framework compared to the external discrete choice models.
- The communal modes of both the existing transport demand gravity model and the external discrete choice models are specified along different underlying behavioural considerations, such that the different models relate differently to the same communal modes.
- The combination of transport demand information from different sources might affect the practical validity for the modelling purpose.

These sub-problems are covered by the following problem statement:

How can transport planners add the behavioural preferences for new modes based on discrete choice models to existing transport demand gravity models in order to provide indications about the future modal split for these modes?

Research goal, research questions and research approach

The research objective is to design a methodology for adding external discrete choice utility functions to an existing transport demand gravity model. In addition, this research aims to provide indications about the practical validity of this methodology for forecasting future modal splits travel behaviour when new modes are present.

The research questions for this thesis are formulated as follows:

Main research question

How can external discrete choice utility functions be added to existing transport demand gravity models and what are the first indications for the validity of this methodology for forecasting future modal splits in the presence of new modes?

Sub research questions

- *How can the behavioural specifications of existing transport demand gravity models be expressed within the discrete choice modelling framework?*
- *How can be dealt with the different underlying behavioural considerations for the communal modes in both existing transport demand models and external models within the discrete choice modelling framework in order to add new modes from external models?*
- *Which first indications can be given about the practical validity of the developed methodology for forecasting for future modal splits in the presence of new modes?*

The research questions are answered by first exploring the backgrounds in transport demand modelling. Thereafter, the methodology is developed based on information about current model transfer methods and transport demand modelling backgrounds. Finally, the practical validity of the of the developed methodology for forecasting for future modal splits in the presence of new modes is explored my means of a case study for "Urban Mobility as a Service" in the Eindhoven-Veldhoven area.

The methodology does not aim to take any transfer bias into account. This means that the underlying behavioural considerations in the contexts of both models are considered to be the same. As this is likely not the case in practice, this poses a risk for non-valid outcomes for the case study.

Backgrounds in transport demand modelling

Two important categories of transport demand models are aggregate and disaggregate models. The main difference between these models is that aggregate models consider the general attractiveness of transport alternatives, whilst disaggregate discrete choice models consider the utility of these alternatives from an individual perspective.

A specific type of a disaggregate transport demand model is the discrete choice model by means of Multi Nomial Logit specifications within the Random Utility Maximisation modelling framework. These models have many capabilities to represent the behavioural preferences for specific attributes and individual characteristics.

Basically, the theoretical framework (Ortúzar & Willumsen, 2011) can be described by means of net utility U_{iq} which are specified for each individual and each transport alternative. The net utility consists of a systemic part which can be observed by the modeller and a unobserved utility, the error term ε_{iq} Equation 5 provides the definition of the net utility.

Equation 1: definition of the net utility

$$U_{iq} = V_{iq} + \varepsilon_{iq}$$

The systemic utility V_{jq} consists of the sum of the weighted attribute values of all trip attributes *k* for alternative *i* for decision-maker *q*. The parameters β_{ki} represent the sensitivity for the changes of the respective attribute values. The definition of the systemic utility is expressed in Equation 6.

Equation 2: definition of the systemic utility

$$V_{iq} = \sum_{k} \beta_{ik} \, x_{ikq}$$

A traveller will select the transport alternative which provide him the highest net utility. The selection of transport alternative j by decision-maker q will only occur when the condition expressed by Equation 3 is satisfied.

Equation 3: expression of individual choice based on the net utility of each alternative

$$U_{jq} \geq U_{iq}, \ \forall A_i \in A(q)$$

An important characteristic of MNL models is that the error term is independently and identically distributed for all available transport alternatives. Due to this characteristic, utility functions can only be added to different models when the variance of the error term distributions are equal and the underlying behavioural considerations are the same. When this is the case, similar transport alternatives will provide similar systemic utilities for both models.

The underlying behavioural considerations are likely to be different when the estimation and application contexts of models are different. This relates to for example the availability of transport alternatives (e.g. modes) and the information the decision-maker considers.

Methodology

In order to keep the executive summary comprehensive, the methodology is only briefly elaborated. Figure 1 provides an overview of the four methodological steps. Detailed information is provided in chapter 3.

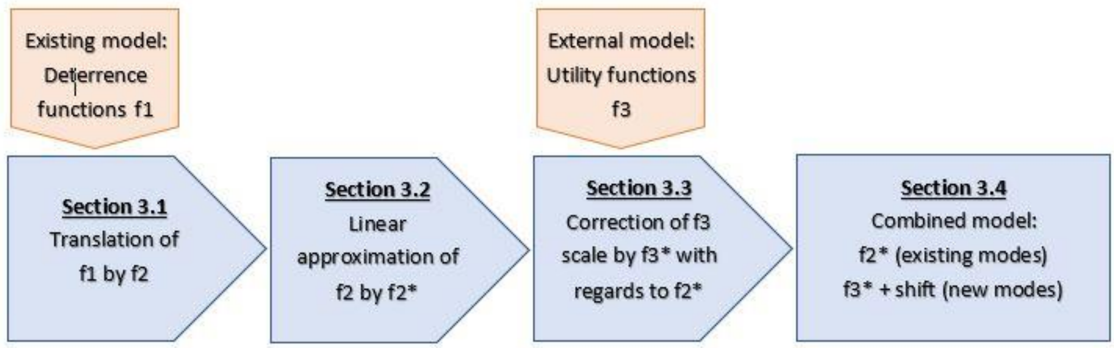


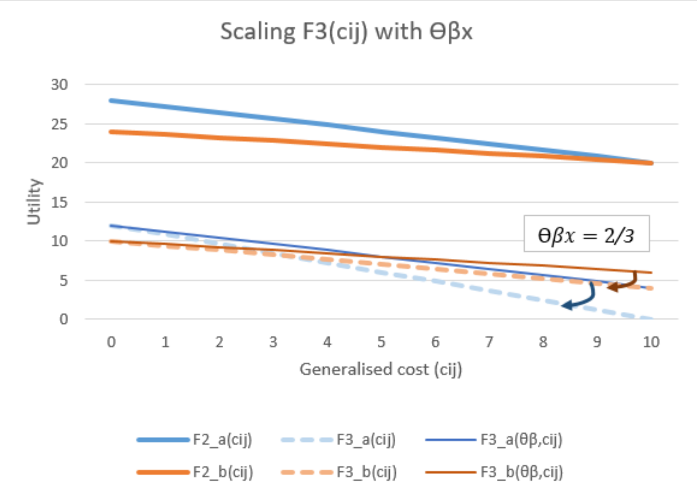
Figure 1: overview of the methodological steps

The first step consists of the translation of the deterrence functions of transport demand gravity models by means of the methodology of Bliemer (2010a); (2010b).

Thereafter, the translated utility functions from the existing transport demand model are linearized with respect to the present values of the generalised costs (which reflect the resistance for using a specific transport alternative). This provides utility function specification with equal systemic utility sensitivities for marginal changes of trip attributes for each OD-pair within the area of scope. This is similar to the discrete choice model, such that the behavioural characteristic of equal sensitivities is incorporated in the utility functions of both the existing and external transport demand model.

The third step is the most difficult one, where the underlying behavioural considerations are assumed to be the same for both models, such that difference between systemic utilities of the communal are considered to be caused by differences in the model scale. Therefore, the third step corrects the scale of the utility functions of external models by correcting the scale of the structural part and the scale of the trip-specific part of utility functions separately. By doing so, the structural systemic utility differences between two communal alternatives are equalled, which is reflected by equal distances in Figure 2 between the multiple alternatives (reflected by the blue and red line) for the multiple models. Simultaneously, Figure 3 shows the systemic utility sensitivities for marginal changes of trip attributes for both alternatives which are equalled. This is reflected by the similar slopes between the blue and red lines.

The final step determines the combined model by means of adding constant systemic utilities to the utility functions of the corrected external model. By doing so, the utility functions for the communal modes of both models are similar. This is reflected by the fit of the utility functions for the communal modes in Figure 4. It is now assumed that the model scales are similar, such that the two different models can be combined.



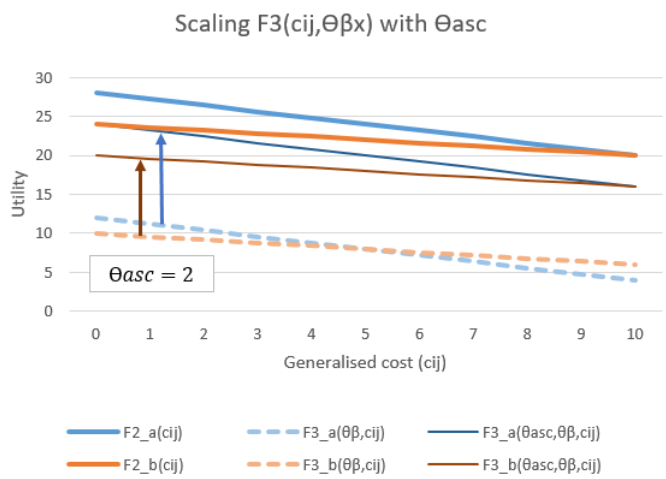
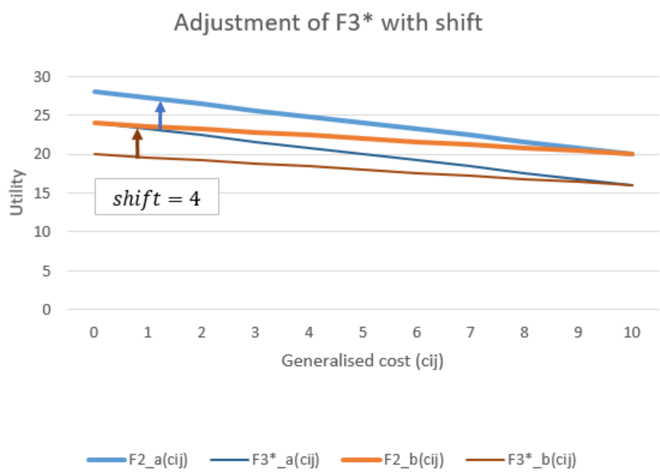


Figure 2: conceptual overview on scaling utility functions with ϴβx

Figure 3: conceptual overview on scaling utility functions with ϴASC

Figure 4: conceptual overview on shifting utility functions f3 towards f2**



Case study and the practical validity of the methodology

Based on the application of the methodology on the case study for "Urban Mobility as a Service" for the Eindhoven-Veldhoven area the practical validity of the methodology for forecasting future modal splits in the presence of new models is considered to be very limited. For the case study a utility function for demand responsive transport was added from a Chicago-related model. A utility function for shared cars was added from a Lisbon-related model. For the different model context it was found that the underlying behavioural consideration are not the same. This relates to different travel purposes, different availability of transport alternatives (modes) and differences in the aggregated modal splits of the related cities.

This was also reflected in the parameter ratio which represents the value of time (VOT). The VOT for the Chicago model was significant larger than the VOT for Eindhoven-Veldhoven. For Lisbon the VOT was significantly lower than for Eindhoven-Veldhoven. These differences in the underlying behavioural preferences were likely to cause a significant translation bias when the external model scales were corrected. The assumption of equal underlying behavioural differences between the different models was not plausible. Therefore, the modal split outcomes are not reliable for the forecasts of future modal splits in the presence of new modes.

Conclusion

The behavioural specifications of the existing transport demand gravity model can be translated to disaggregate discrete choice utility functions. The methodology of Bliemer (2010a); (2010b) is able to translation aggregate gravity models towards disaggregate discrete choice models. The non-linear translated utility functions are not considered to be directly useable for the addition of utility functions of external models. The Ordinary Least Squares method was found to be a straightforward method to approximate the translated disaggregate utility functions by minimising the squared utility differences with the linear approximation.

For the addition of external utility functions to the translated existing transport demand model the differences in underlying behavioural considerations are neglected, such that the differences between the trade-offs between modes for both models are considered to be caused by different (fictional) model scales. By correcting the scale of the external model in order to equal the fictional model scales the underlying behavioural considerations can be equalised. Consecutively, the utility functions of corrected external models can be added to the translated existing transport demand model, by adding equal additional systemic utility to the structural element of all external utility functions, such that for the communal modes the systemic utilities within the external model approximate the systemic utilities within the existing model for the same modes. For both methodological steps the Ordinary Least Squares method has been used.

Based on the application of the methodology on the case study for "Urban Mobility as a Service" for the Eindhoven-Veldhoven area the practical validity of the methodology for forecasting future modal splits in the presence of new models is considered to be very limited. The choice of the external models was found to be really important for valid results of the combined model. Significant differences in the underlying behavioural assumptions can bias the modal split outcomes, such that they do not provide realistic transport demand.

Table of Contents

List of figures

List of tables

1 Introduction

The world changes rapidly and so does the world of mobility. Compared to 15 years ago, many new transport concepts have arisen. When we go out, we can just use our mobile phones to request an Uber taxi, instead of going by bicycle to our favourite pub. For our daily commute, we don't need our brackish second-hand bicycle at the station any longer, because we easily rent a public transport bicycle ("OV-fiets") for our last mile to work. Also, when it comes to the possession of – what can best be called – transport resources, things have been changed significantly. Instead of owning a car or a bicycle, we can flex lease them or even rent them for a very short period. Many of those emerging concepts are connected to the modern trend of servitisation: the provision of added value by offering user-centred combinations of goods, services, support, self-service and knowledge. (Vandermerwe & Rada, 1988)

Emerging transport services

During the past few years, many new car- and bicycle-based services have risen. The most visible ones are probably the short-term rental cars and bicycles you can temporarily use from many places within the city. As these cars and bicycles are shared sequentially over time by multiple users, it is more appropriate to refer to shared cars and shared bicycles.

In the Netherlands, car sharing providers such as Greenweels, ConnectCar and Car2Go are expanding rapidly with over 30.000 shared cars in spring 2017 (CROW, 2017). Also, the use of shared bicycles in the Netherlands is widely spread. The shared bicycles ("OV-fietsen") which can be picked up and dropped off at most trains station were already used by 500.000 subscribers in 2017 (Nederlandse Spoorwegen, 2018). Regarding free floating bicycles – bicycles which can be parked and used everywhere within the service area – exact quantitative information is lacking. However, the city council of Amsterdam is already facing lots of difficulties with the extra public space these bicycles require. (Kruyswijk, 2017)

Another type of transport service, Demand Responsive Transport (DRT), emerged in the last two years in the Netherlands. Transdev, the holding company of multiple public transport providers introduced Breng Flex (Arnhem-Nijmegen), Bravo Flex (Helmond) and AML Flex (Amsterdam-Meerlanden) (TU Delft, 2017) for direct transport with small busses within a dense network of stops. Users of DRT can easily request a ride - which can be between any set of two stops in the network - via their smartphone and pay a fixed fee by smart card or credit card.

What perhaps can boost the development of these transport services is the emergence of Mobility as a Service (MaaS). Albeit there is a lot of discussion about a clear definition for MaaS, within this thesis MaaS is defined as the enabler for using multiple mobility services by means of a smartphone application which takes care of trip planning, ticketing and payment. In literature, there are multiple definitions for MaaS available. These definitions are for example given by Hietanen (2014), Heikkilä (2014) and Kamargianni, Li, Matyas, and Schäfer (2016).

According the definition of MaaS in this thesis, it can be seen that multiple MaaS initiatives are currently arising in the Netherlands. Tranzer (2018) is already available for single rides for a multitude of transport. Whim (2018) aims to provide packages of mobility services and plans its launch in the Netherlands soon.

Transport modelling and transport demand models

For many years transport models are used to provide insights about current and future travelling. These transport models link the demand for travel with the supply of transport resources (modes) and transport services. The modelling practice and the "state of the art" in transport demand modelling – also referred to as travel demand modelling – has evolved significantly over time.

The development of transport demand modelling started with trip- or tour-based models, where nowadays also activity-based models are available. The first type of models determines the transport demand by means of a predetermined number of trips/tours for each zone or traveller, whilst the latter type of models determines the transport demand as the result of the need for participation in activities spread over time and space. (Bhat, Guo, Srinivasan, & Sivakumar, 2003; Ortúzar & Willumsen, 2011; Sivakumar, 2007)

Within the group of conventional trip- or tour-based transport demand models a variety of models is available. They can be described by some characteristic features: the level of disaggregation (aggregate vs. disaggregate) and the level of dynamism (static vs. dynamic). The first feature refers to the consideration of individual behaviour for modelling the transport demand. The latter refers to the feedback of the transport demand on the level of service related to the supply of transport alternatives and vice versa. (Ortúzar & Willumsen, 2011; Sivakumar, 2007)

The first transport demand models were introduced in the sixties. These models were aggregate and inspired by physical phenomena, such as the gravity model or entropy model. These models determine production and attraction of trips proportional to the size of a zone (e.g. in terms of inhabitants, available jobs, etc.). For the determination of the transport demand, the trips are distributed inversely proportional to the impedance or resistance values (e.g. monetary costs, travel time or generalised costs) for all transport alternatives such that the trip productions and attractions are matched. (Ortúzar & Willumsen, 2011; Sivakumar, 2007)

In the seventies, the reliability and the validity of the outcomes of these aggregate models were heavily discussed. The search for models which were able to respond more appropriately and directly to changing transport policies was met by the introduction of disaggregate transport demand models and became popular in the planning practice in the eighties. These models, mainly based on modelling frameworks such as random utility maximisation and constrained optimisation, use data about individual people and their preferences for transport alternatives to derive the travel demand. (Ortúzar & Willumsen, 2011; Sivakumar, 2007)

Activity based models were developed more recently, but its use in the planning practice differs strongly amongst countries. Many of these transport demand models are used in the United States, mainly driven by new legal arrangements in the nineties related to environmental and financial concerns. Within that setting, activity based models were considered to be the most appropriate tool for the forecast of complex environmental and transport efficiency related impacts. (Ettema, 1996)

Considering the trip- and tour-based transport demand models, there is a general shift towards the use of disaggregate models. (Ortúzar & Willumsen, 2011; Sivakumar, 2007) Still the current practice of transport modelling varies significantly amongst different countries and different spatial levels. Considering the Netherlands, mainly disaggregate models are used for transport demand modelling. This is the case for transport models which represent larger areas, such as the nation-wide "Landelijk Model System" (LMS), the region-wide "Nationaal Regionaal Model" (NRM) or the "Verkeersmodel Metropoolregio Amsterdam" (VMA). (Pieters, Elshout, & Herder, 2014; Rijkswaterstaat, 2017) For lower spatial levels in the Netherlands, many transport demand models still have an aggregate approach. Examples of transport demand gravity models in the Netherlands are for example the RVMK ("Regionale Verkeersmilieukaart") model for the greater Rotterdam area, the "Verkeersmodel

Haaglanden" for the greater The Hague area and the SRE ("Samenwerkingsverband Regio Eindhoven") model for the greater Eindhoven area. (Goudappel Coffeng, 2012, 2013, 2014)

Forecasting future modal splits in transport demand gravity models in the presence of new modes
Given the emergence of new modes, an important question is how these emerging modes will lead to behavioural changes and impact mobility. This is important as it potentially affects the way transport infrastructure is planned and transport policy is established. Information about future transport is thus necessary in order to prepare for impact on relevant areas, for example reflected by the 5E framework: effective mobility, efficient cities, equity, environment and economy. (van Oort, 2017)

At this moment, existing transport demand gravity modes are not able to forecast travel behaviour with regards to new modes and transport services, as they are estimated based on current mobility behaviour in the absence of these transport alternatives. De facto, these models provide no information about the (aggregated) preferences for new modes and transport services.

It is possible to estimate new discrete choice transport demand models based on stated preference research, but this requires a significant amount of time and financial resources. Examples of such estimated discrete choice models are provided in the studies of Eiró and Martínez (2014) and Frei, Hyland, and Mahmassani (2017). For that reason it is interesting to investigate the possibility for combining discrete choice utility functions with existing transport demand gravity models such that they are able to provide indicational forecasts about the future modal splits in the presence of these new modes. For this purpose it is considered to be important that the altered model is able to represent the choice probabilities ratios between the new modes and those modes in the existing models which are assumed to face significant demand changes in the presence of the new modes.

1.1 Problem identification and problem statement

Combining different models requires that the related modelling frameworks are the same. This means that if discrete choice models need to be combined with existing transport demand gravity models, one of these models need to be translated such that the travel behaviour for both models is specified within comparable modelling paradigms. Therefore, the problem of non-consistent modelling frameworks needs to be addressed first before both models can be combined. The translated model needs to provide similar modal splits as the existing model. In needs to be investigated if the translation procedure poses a translation bias. In that case the translated model is an approximation of the existing model. In order to maintain the validity of the existing transport demand model, the translation bias is not allowed to cause significant deviations between the existing model and its approximation.

Existing literature provides methodologies for the combination of multiple discrete choice models by means of updating or integrating the model parameters of two discrete choice models. This results in a new enriched model based on the information of both models. From this literature, the following methods can be distinguished: transfer scaling, Bayesian approach, combined transfer estimation approach and joint context estimation. These methodologies are mainly used for transferring models to different spatial contexts or temporal contexts. (Badoe & Miller, 1995a, 1995b; Ben-Akiva & Bolduc, 1987; Hansen, 1981; Karasmaa, 2007; Ortúzar & Willumsen, 2011)

The transfer scaling method (Badoe & Miller, 1995b) uses a data sample from context 2 to estimate a model for this context which is based on the model parameters (excluding the alternative specific constants) of context 1, and includes a scale factor for these parameters and new alternative specific constants. The Bayesian updating and combined transfer estimation methods (Badoe & Miller, 1995b; Ben-Akiva & Bolduc, 1987) estimate a new set of model parameters for context 2 based on the parameter sets of both contexts and its covariance matrices. By means of the joint context estimation method (Bradley & Daly, 1997) completely new models are estimated for both contexts based on both underlying datasets. This is done by constraining the communal parameters of both context to be equal and scaling the utilities of context 2 such that the variances of the error terms of both models are equal.

These methodologies for updating or integrating discrete choice models pose several limitations which make them unsuitable for the indications of the impact of future modes. First, there are no data set and covariance matrix available about the individual behavioural preferences for the existing transport demand gravity model. For that reason, the Bayesian updating, combined transfer estimation and joint context estimation methods cannot be used. Second, the new model needs to incorporate the currently non-existent transport alternatives. The transfer scaling method is therefore unsuitable as it is only able to transfer data from exiting transport alternatives. For these reasons, there is a need for a methodology to add utility functions from external discrete choice models to the existing transport demand model.

It is likely – as similar for the updating or integrating methodologies – this poses limitations with regards to the validity of the modelling outcomes of the combined model. Karasmaa (2007) describes two major issues for combining different discrete choice models. The first issue relates to the presence of a transfer bias. Differences in the spatial-temporal mobility contexts and the availability of modes and transport services will inherently pose a transfer bias. Travellers will make different trade-offs between transport alternatives in different situations, especially when the choice set of available transport alternatives is different. Different trade-offs between transport alternatives are also likely when the estimation procedures of the two models are different. This is the case when the estimation of gravity models is based on observed behaviour versus the estimation of discrete choice models based on stated preference (intentional) behaviour. A second issue relates to the presence of a sample bias. Stated preference research is mostly limited to a certain sample size and composition. This poses uncertainties with regards to the representation of the population by the sample.

In summary, three sub-problems are related to the need for the development of a methodology to add utility functions from external discrete choice models to the existing transport demand model.

- The existing transport demand gravity model is specified within a different modelling framework compared to the external discrete choice models.
- The communal modes of both the existing transport demand gravity model and the external discrete choice models are specified along different underlying behavioural considerations, such that the different models relate differently to the same communal modes.
- The combination of transport demand information from different sources might affect the practical validity for the modelling purpose.

Considering these sub-problems within the frame of providing indications of future modal splits in the presence of new modes and relating these to the current knowledge and practice of transport planners, the following problem statement is derived:

How can transport planners add the behavioural preferences for new modes based on discrete choice models to existing transport demand gravity models in order to provide indications about the future modal split for these modes?

1.2 Research scope

The research is scoped around aggregated gravity-based transport demand models. The distribution of the transport demand over all transport alternatives is considered to be determined by means of a single (aggregate) trip attribute for these models. For this thesis, this trip attribute is considered to be the generalised costs. The external models are all discrete choice models based on the random utility maximisation (RUM) framework and are specified as Multi Nominal Logit (MNL) models.

The specifications of the existing transport demand model are considered to be valid for its application. Context. For that reason, the change of the specifications of the existing model is considered to be non-favourable.

The proposed methodology for solving the problem of interest does not explicitly take any transfer bias into account. This means that the underlying behavioural considerations in both the context of the existing transport demand model and the external model are assumed to be the same. This assumption is not tenable in practice as both the spatial-temporal contexts and estimation contexts are different. Therefore, the impact of the transfer bias on the practical validity of the methodology will be explored when the methodology is applied on a case study.

The case study is based around the introduction of "Urban Mobility as a Service" in the Eindhoven-Veldhoven area. Detailed information about this case study is provided in section 4.1. For the case study, the existing transport demand for the Eindhoven-Veldhoven area is redistributed over the future available modes, whilst leaving the redistribution over different origins and destinations out of scope.

1.3 Research objective and questions

The research objective is to design a methodology for adding external discrete choice utility functions to an existing transport demand gravity model. In addition, this research aims to provide indications about the practical validity of this methodology for forecasting future modal splits travel behaviour when new modes are present.

The research questions for this thesis are formulated as follows:

Main research question

How can external discrete choice utility functions be added to existing transport demand gravity models and what are the first indications for the validity of this methodology for forecasting future modal splits in the presence of new modes?

Sub research questions

- *How can the behavioural specifications of existing transport demand gravity models be expressed within the discrete choice modelling framework?*
- *How can be dealt with the different underlying behavioural considerations for the communal modes in both existing transport demand models and external models within the discrete choice modelling framework in order to add new modes from external models?*
- *Which first indications can be given about the practical validity of the developed methodology for forecasting for future modal splits in the presence of new modes?*

1.4 Scientific and societal relevance

The scientific relevance of this research is that the methodology provides a way for adding external model parameters to an existing model, which is different from the existing updating or integrating methodologies. It provides a straightforward methodology to add future modes to existing models.

From a societal perspective, this research is relevant because it enables a fast method to extend existing transport demand models with new modes or mobility services. This method is beneficial as it does not require the development of completely new transport demand models. The proposed methodology is useable for conducting explorative studies, sensitivity and scenario studies regarding future mobility systems where new means for transportation are available. Extended models can help policy makers with planning for new modes such as Demand Responsive Transport, shared car and bicycle systems or more integrated concepts such as Urban Mobility as a Service.

1.5 Research methodology and thesis outline

This section provides the outline of this thesis and refers to the methodological steps for reaching the stated research goal.

In order to develop a methodology for the addition of external discrete choice utility functions to existing transport demand gravity models and to provide first indications of the practical validity of this methodology, the first need is to explore the background of transport demand modelling. Chapter 2 provides all information relevant for the design of the methodology and the assessment of the practical validity.

Chapter 3 sets out the multiple steps of the developed methodology. The developed methodology consists of the following steps which will be set out in logical order: (1) translation of gravity-based deterrence functions towards discrete choice utility functions, (2) linearization of non-linear utility functions, (3) determination of factors to correct the scale of external models and (4) the specification of the combined model.

Chapter 4 applies the developed methodology on a case study for Urban Mobility as a Service in the Eindhoven-Veldhoven area. This case study is used for the application of the methodology and the exploration of the practical validity of the methodology.

Chapter 5 is the final chapter of this thesis and provides conclusions, a discussion and recommendations. This is done by answering the research questions and reflecting on the research process of this thesis and the outcomes of the methodology application.

2 Backgrounds in transport demand modelling

Chapter 2 provides information about all contextual elements, which are relevant for the development of the methodology for the addition of external discrete choice utility functions to existing transport demand gravity models. These contextual elements also help to assess the practical validity of the methodology.

In section 2.1 the different transport demand model frameworks are discussed. Section 2.2 elaborates on the estimation context of transport demand models. Section 2.3 elaborates on the application context of transport demand models.

2.1 Transport demand modelling frameworks

This sections provides an elaboration on several trip- and tour based transport demand models. Section 2.1.1 elaborates on the general 4-step transport demand modelling approach. Section 2.1.2 provides the background of aggregate transport demand modelling frameworks. Section 2.1.3. provides the background of disaggregate transport demand modelling frameworks.

2.1.1 General 4-step transport demand modelling approach

Traditionally, four sequential steps can be distinguished within conventional transport demand modelling approach: trip generation, trip distribution, mode choice and trip assignment. (Meyer & Institute of Transportation, 2016; Ortúzar & Willumsen, 2011) Figure 5 shows the conceptualisation of this 4-stage transport demand modelling framework.

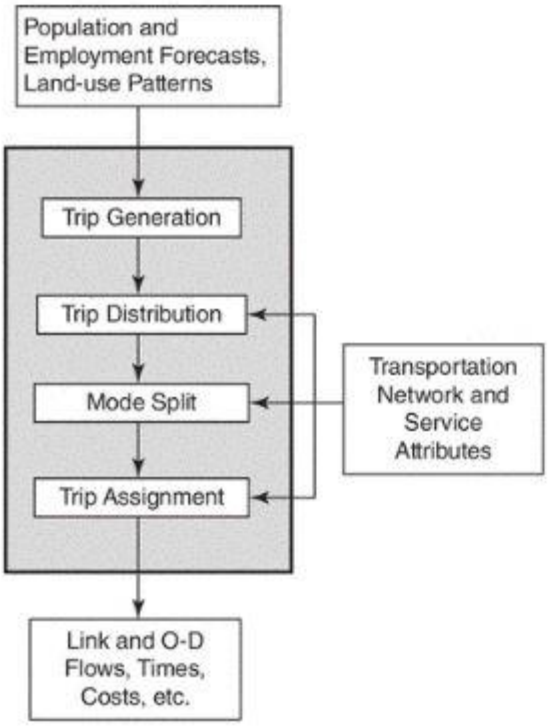


Figure 5: 4-step transport demand modelling approach (Meyer & Institute of Transportation, 2016)

In the first step the number of generated and attracted trips is estimated based on data about the population (e.g. number of inhabitants, age distributions) and the level of economic activity (e.g. number of jobs, available shopping and educational facilities). The second step distributes the trips over space, whilst the third step distributes the trips over the different available modes. For some of the models - including the SRE model (Goudappel Coffeng, 2012) - the second and third step are executed simultaneously. The final step consists of the assignment of the trip matrices to the available traffic network.

Many models calculate the impact of the assignment of trips on the level of service for the available modes. Differences in the level of service, for example caused by exceeding capacity constraints or congestion, lead to different impedance or resistance values (e.g. in terms of travel time or travel costs). When these models consider a feed-back loop going back to the trip distribution and mode choice steps, this can result in different distributions over space and modes.

2.1.2 Aggregate transport demand modelling

The earliest transport demand models were simple mathematical models. They are aggregate models as they do not consider individual travel behaviour, but the travel behaviour of groups of people within zones. This level of aggregation enables the forecast of the transport demand on a higher and more abstract level. The transport demand is quantified as a function of the size of zones and the impedance or resistance for travelling between these zones. Examples of these types of aggregate transport demand models are growth models, gravity models and entropy models. Considering the research scope, this section elaborates solely on gravity models. (Ortúzar & Willumsen, 2011; Sivakumar, 2007)

Gravity model

As a starting point, gravity models use data about the aggregate behaviour of travellers such as the total number of performed trips and the distribution of trips over modes and trip lengths. Gravity-based model are named after the underlying assumptions for the distribution of trips. Parallel to Newton's gravitational law, nearby locations will attract more transport demand than more remote locations. This can be stated more precisely, by using the more generic term impedance or resistance as travellers will also perceive difficulties reaching locations for other reasons than distance (e.g. travel costs or travel times). (Ortúzar & Willumsen, 2011)

The trips between origins i and destinations j are described by Equation 4. O_i and D_j represent the generation of trip-ends for the outbound trips from origin i and the incoming trips for destination j and are multiplied with deterrence function $f(c_{ij})$. It is likely that specific locations generate or attract not the same number of trips as defined by O_i and D_j. For that reason, balancing factors A_i and B_j come into place to regulate these trip-ends. (Ortúzar & Willumsen, 2011)

Equation 4: expression of the number of trips between origin i and destination j in a gravity-based model

$$T_{ij} = A_i \; O_i \; B_j \; D_j \; f(c_{ij})$$

The deterrence function $f(c_{ij})$ is an important part of the gravity-based model and determines how the aggregate transport demand is distributed over all transport alternatives. These deterrence functions can be specified for population segments when information is available about the aggregate travel behaviour of these segments. Deterrence functions can also be specified for different modes. Also this requires information about the aggregate travel behaviour in relation to each specified mode.

Deterrence functions can be specified in multiple ways. The most common specifications are by means of exponential, log-normal or top-log-normal functions. Depending the purpose of the model and the characteristics of the transport demand a specific deterrence function type can be chosen such that the actual travel behaviour is approximated appropriately. (Ortúzar & Willumsen, 2011)

2.1.3 Disaggregate transport demand modelling

Advances in transport demand modelling resulted in a shift away from these aggregate models and led to the rise of disaggregate models. These models use disaggregate data about individual travels by individuals between the zones in the study area. The fundamental difference between aggregate and disaggregate models is that the disaggregate models view the individual (or household or firm) as the decision-making unit. One possibility for disaggregate transport demand modelling is by means of discrete choice modelling.

Discrete choice models and Random Utility Maximisation (RUM)

Discrete choice models represent choice behaviour made by individual travellers. The framework for discrete choice models can be presented by a set of general assumptions. These assumptions are related to the following items (Ben-Akiva & Bierlaire, 1999):

1. decision-maker – defining the decision-making entity and its characteristics;
2. alternatives – determining the options available to the decision-maker;
3. attributes – measuring the benefits and costs of an alternative to the decisionmaker;
4. decision rule – describing the process used by the decision-maker to choose an alternative.

The decision-maker is considered to be an individual (i.e. a traveller). The way this individual makes a decision regarding travelling is determined by the decision rule. These decision rules can be the same for a group people, when a discrete choice model is not specified to represent the heterogeneity of preferences. This could be done by means of specifying alternative distributions of behavioural parameters or by introducing parameters in relation to person-specific attributes.

Decision-makers are able to choose between a set of alternatives. This is considered to be the choice set. This choice set is a finite set of alternatives which the modeller assumes to be relevant (i.e. available and is considered in practice) for the decision-maker.

Each alternative in the choice set is characterised by a set of attributes. In the case of transport demand models these often relate to determinants for travel behaviour such as travel time, travel costs, etc.

The decision-rule relates to the evaluation process of the different alternatives for each individual. In the context of transport demand models, it determines how the decision-maker decides which travel option it will choose. In discrete choice models, decisions are often made based on Random Utility Maximisation (RUM). The theoretical framework for RUM is described by Domencich, McFadden, and Associates (1975) and Williams (1977).

Basically, the theoretical framework can be described as follows. The set of transport alternatives A is available. The alternative $A_j \epsilon A$ is characterised by a net utility U_{iq} which is specified for decision-maker *q*. (Ortúzar & Willumsen, 2011) Equation 5 provides the definition of the net utility which consists of the sum of the systemic part V_{iq}, which can be observed by the modeller, and the error term ε_{iq}

Equation 5: definition of the net utility

$$U_{iq} = V_{iq} + \varepsilon_{iq}$$

The systemic utility V_{jq} consists of the sum of the weighted attribute values of all available attributes *k* for alternative *i* for decision-maker *q*. The parameters θ_{ki} represent the sensitivity for the changes of the respective attribute values. (Ortúzar & Willumsen, 2011) The formulation of the systemic utility is expressed in Equation 6.

Equation 6: definition of the systemic utility

$$V_{iq} = \sum_{k} \beta_{ik} \, x_{ikq}$$

RUM models assume, as does the economic consumer theory, that the decision-maker has a perfect discrimination capability. In addition, both the decision-maker and the analyst are assumed to have incomplete information. Therefore uncertainty must be taken into account. This is done by means of the random utility part ε_{jq}, which is better known as the error term. (Ben-Akiva & Bierlaire, 1999)

Four sources of uncertainty can be taken into account. For this thesis, elaboration on these uncertainty factors is not necessary; it is only important to understand what is in these error terms as distinguished by Manski (1977):

1. unobserved alternative attributes;
2. unobserved individual characteristics (also called “unobserved taste variations”);
3. measurement errors;
4. and proxy, or instrumental, variables.

The decision-maker chooses the alternative which represents the highest net utility (see Equation 7). Because the systemic utilities are already determined, the actual decision is determined by the values of the error terms (i.e. random utility parts) of the alternatives. (Ortúzar & Willumsen, 2011) This is reflected in Equation 8.

Equation 7: expression of individual choice based on the net utility of each alternative

$$U_{jq} \geq U_{iq}, \; \forall A_i \in A(q)$$

Equation 8: expression of individual choice between two alternatives based on the systemic utilities and error terms

$$V_{jq} - V_{iq} \geq (\varepsilon_{iq} - \varepsilon_{jq})$$

One of the most frequently used RUM models is the so-called Multi Nomial Logit (MNL) model. It is specified by means of the assumption that the error terms of the utility functions are independent and identically Gumbel (Extreme Value type I) distributed. By specifying the error term distributions in this way, the probability that decision-maker *q* chooses alternative *i* from choice set A can be calculated by Equation 9. (Ben-Akiva & Bierlaire, 1999)

Equation 9: expression of choice probabilities in Multinomial Logit

$$P(i|C_q) = \frac{e^{V_{iq}}}{\sum_{j \in C_q} e^{V_{jq}}}$$

This provides the MNL model the property of Independence from Irrelevant Alternatives (IIA). This implies that the ratio of choice probabilities of any two alternatives is unaffected by the systematic utilities of any other alternatives. The assumption on the IIA property can be a problem in situations where the choice probabilities are likely to be correlated as they share a significant part of similar attributes. This is often illustrated by the red bus/blue bus paradox in a mode choice context. (Ortúzar & Willumsen, 2011)

From Equation 8 it follows that the absolute values of the systemic utility are not relevant for the determination of choice probabilities. What matters are the systemic utility differences between the different alternatives. This implicates that the addition of systemic utility constants with equal values for all transport alternatives will not change the choice probabilities.

MNL models need to be specified by means of a scale factor. The scale factor regulates the utility ratios between the different transport alternatives. Generally, the modeller does not define a scale factor such that the model scale takes the value of 1. In that case, the utility functions of the model are assumed to have an error term with a variance of $\pi^2/6$.

However, it is possible to estimate a MNL model with similar behavioural specifications whilst the scale factor is different. Scaling means that the complete utility function is multiplied with a constant. Scaling a MNL maintains the choice probabilities for all scale factors larger than 0. Scaling a MNL model with a constant λ changes the variance of the error terms with a factor λ^2.

The IID assumption of the error terms in the utility functions of MNL models pose problems for the combination of utility functions from different MNL models. When the underlying behavioural considerations are the same, utilities need to be equal for similar transport alternatives within both models. Otherwise the error term variances of both models are different as the utilities for the same transport alternative for both models are not normalised to the same value.

In situations where the error components of transport alternatives are likely to be correlated there is no compliance with the prerequisite IID constraint. For those situations, a nested logit (NL) model can be estimated. A NL model specifies different scale factors for different groups of alternatives. These groups are called nests. This means that the variances of the error terms for the alternatives in different nests are different. (Ortúzar & Willumsen, 2011) Figure 6 shows an example of a conceptual structure of a NL model.

Sometimes the nested logit structure is confused with a hierarchy of choices; in the example of Figure 6 this would mean that someone first chooses between car-based or public transport and thereafter a specific mode is chosen. This is however not the essence of the NL model. (Train, 2009)

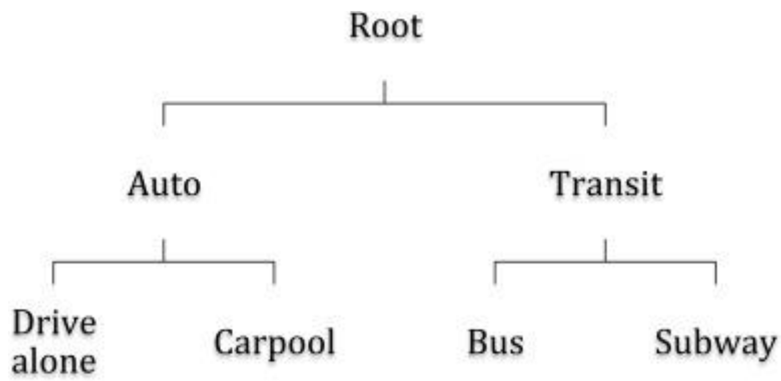


Figure 6: conceptual example of a nested logit model (Bunch & Rocke, 2016)

Because the net utilities have different error term distributions, attribute changes for transport alternatives in nest *k* will not affect the ratios of choice probabilities in nest *i*.

Sometimes the nested logit structure is confused with a hierarchy of choices; in the example of Figure 6 this would mean that someone first chooses between car-based or public transport and thereafter a specific mode is chosen. This is however not the essence of the NL model.

Choice probabilities for alternatives in the NL model are constructed by two elements: the probability for choosing nest *k* and the probability for choosing alternative *i* within nest *k*. The probability for choosing a nest is calculated by the so-called log sum. Equation 10 provides the specification of this log sum. (Ortúzar & Willumsen, 2011)

Equation 10: definition of the log sum over all alternatives within a nest

$$LS_q = ln\left(\sum_{i\in k} exp\left(\frac{V_{iq}}{\mu_k}\right)\right)$$

The probability for choosing nest *k* can be calculated by treating the log sums of the different nest as they were utilities of different alternatives. The only thing what needs to be accounted for is the difference in scale parameter between the nest. This means that all different log sums are multiplied with their own scale parameter before the choice probabilities are calculated. Within nest *k*, choice probabilities are calculated as is the case in MNL, only that the systemic utilities are divided by the nest scale parameter μ_k. (Ortúzar & Willumsen, 2011)

Equation 11 shows how the conditional probability of choosing one alternative given a chosen nest can be calculated. The left part provides the choice probabilities within a nest, where the right part provides the choice probabilities between nests.

Equation 11: expression of conditional choice probabilities in Nested Logit

$$P(i|k) = \frac{exp\left(\frac{V_i}{\mu_k}\right)}{\sum_{i\in k} exp\left(\frac{V_i}{\mu_k}\right)} \quad \frac{exp\,\mu_k\,ln\left(\sum_{i\in k} exp\left(\frac{V_i}{\mu_k}\right)\right)}{\sum_{k\in K} exp\,\mu_k\,ln\left(\sum_{i\in k} exp\left(\frac{V_i}{\mu_k}\right)\right)}$$

A different type of model is the so-called Mixed logit (ML) model. These models are characterised as highly flexible that can approximate any random utility model. One of its main benefits is that it can account for random taste variation within a population. (Train, 2009)

In general, a ML models functions as follows: one or multiple parameters of the underlying MNL are changed according specified distribution functions. This means that the choice probabilities for a ML model can be calculated by taking the integral of the weighted logit probability according the distribution function of the varied parameters. When these varied parameters are defined by *β*, the standard logit probability can be used together with the distribution function of *β* for choosing alternative *i* by decision-maker *q* can be given as in Equation 12. (Train, 2009)

Equation 12: expression of choice probabilities in Mixed Logit

$$P_{iq}(\beta) = \int L_{iq}(\beta) f(\beta) \delta(\beta)$$

These models can provide a better representation of individual behaviour within a group. The main disadvantage is that it takes more computation time as the choice probabilities for each possible combination of β's needs to be calculated.

Another model is the Latent Class model which is first described by Lazarsfeld and Henry (1968). Latent class choice models are designed to capture unobserved heterogeneity. The underlying assumption is that the heterogeneity is generated by discrete constructs. These constructs are not directly observable and therefore represented by latent classes. (Ben-Akiva & Bierlaire, 1999) An example of these latent discrete constructs can be lifestyle types or typologies for mobility behaviour.

Latent class models consist of two sub models: a class membership model and a class-specific choice model. Equation 13 represents the general latent class model specification. The left part determines the choice for a specific alternative, where the right part determines the probability decision-maker *q* belong to latent class *s*.

In this specification, *S* refers to the number of latent classes. The model is estimated by means of β_s which is a parameter vector of the class-specific choice model and Θ which is the parameter vector for the class membership model. X_n represents a vector of the attributes of alternative or decision-maker *n* and C_s is the available choice set for class *s*. (Ben-Akiva & Bierlaire, 1999)

Equation 13: expression of choice probabilities in Latent Class

$$P(i|C_q) = \sum_{s=1}^{S} P(i|X_n; \beta_s; C_s) P(s|X_n; \theta)$$

2.2 Estimation contexts of transport demand models

There are essential differences between the different types of data used for the estimation transport demand models. This section elaborates on the most relevant data characteristics for estimating transport demand models. Section 2.2.1 considers aggregate and disaggregate data. Section 2.2.2 considers the use of stated preference or revealed preference data.

2.2.1 Aggregate or disaggregate data

Data about aggregate travel behaviour and data related to disaggregate trade-offs between alternatives generally not provide the same behavioural information. As behavioural specifications are different, the estimated will also be different.

Aggregate travel behaviour data only reveals the general preference for transport alternatives within a given mobility context – including the supply of mobility services – together with presumed valuations of the alternative's attributes. This implicates that aggregated transport demand models are suitable to provide more strategic forecasts as they do not consider heterogeneity of behavioural preferences within the considered population.

The behaviour forecasted by its respective models reflect the overall behaviour over a large group of people. This can become problematic when there are local or temporal differences in the sensitivities for marginal changes of trip attributes or when people for sub-populations develop different travel preferences. For those situations, aggregate data is not able to reflect these specific preferences.

Data related to disaggregate trade-offs between alternatives provides more information than aggregated data as it describes travel behaviour from an individual's perspective. Also socio-demographic attributes can be used for the estimation of a disaggregate model.

2.2.2 Revealed Preference (RP) and Stated Preference (SP) data

When a model is based on revealed-preference data, this means that it is based on observed choices and behaviour. This is preferable when you want to model mobility behaviour as close to reality. The downside of the use of revealed-preference data is that it is less likely that the behavioural specifications will hold when the mobility context changes (e.g. introducing congestion charges or implementing new payment systems for public transport). A specific example of a significant change within the mobility context is the introduction of new modes or new transport services, as revealed-preference data does not show what people will choose when both present and future modes are available. (Ortúzar & Willumsen, 2011)

Stated-preference data is opposed to revealed-preference data, as it is based on hypothetical behaviour. This behaviour is mainly derived from questionnaires and/or hypothetical choice sets. This means that the mobility context can be changed such that a wider range for valid model application can be established. The downside of stated-preference data is that it only shows the intention for mobility behaviour, contrary to empirically determined behaviour. This means that respondents of mobility questionnaire provide different behaviour than they would in reality, e.g. by providing socially-acceptable answers or by being more sceptical about new modes than they will be when these are actually available. (Ortúzar & Willumsen, 2011) In addition, stated-preference data does not reflect specific elements of the mobility context in reality. This means that even if the data reflects the underlying behavioural preferences well, its validity will always be an issue. (Ben-Akiva et al., 1994)

Regarding stated-preference questionnaires it is well known that the context and format of the hypothetical setting affect the response. (Ben-Akiva et al., 1994) Despite an abundance of studies, there is no consensus about the underlying causes of this hypothetical bias. Possible explanations for the hypothetical bias can be the simplification of the choice situation, such that respondents imagine the non-specified choice set characteristics. Another reason can be that respondents do not understand the choice context completely. (Loomis, 2011; Murphy, Allen, Stevens, & Weatherhead, 2005)

2.3 Application contexts of transport demand models

This section provides information about the application context. The context in which people behave is very important for the way they evaluate transport alternatives. In section 2.3.1 an elaboration is given on the spatial-temporal mobility context. Section 2.3.2 elaborates on the availability of modes and the level of service.

2.3.1 Spatial-temporal mobility context

Based on the model dimensions of the SRE model (Goudappel Coffeng, 2012), which is subject to the case study later on, spatial-temporal differences can be easily described. This list aims to provide a general overview of the spatial-temporal mobility context.

- **Spatial dimension**
  Location, urban density and size
- **Temporal dimension**
  Year, season, month, day and period of day data is gathered for model estimation
- **Mode availability**
  Mode types including their range of level of services in terms of e.g. travel costs and travel times.
- **Network characteristics**
  Network speeds, (public transport) frequencies and network granularity.
- **Socio-economic demographics**
  Levels of income, age distributions, employment and car possession rates.

For each context, the (model) dimensions are different. This implicates that each context has a different set of possible trade-offs. This means that we cannot plainly transfer models from one context to the other, as the specified mobility behaviour intends to be only valid in the original context.

The context characteristics determine the factors for making a trade-off. It is likely that someone in the hilly surroundings of Paris will make a different trade-off when it comes to biking than someone in the flat city of Amsterdam. Apart from the characteristics in itself, also the heterogeneity and the available range of its corresponding attributes play an important role. An estimated model for urban travelling is likely not able to forecast nation-wide travel behaviour as the possible travel time and distance ranges will be completely different.

2.3.2 Level of service and the availability of modes

Modes are characterised by the quality they offer to travellers for transport. The level of service (LOS) refers to the quality of traffic networks to accommodate transport. The LOS influences the mode-specific trip attribute values, for example in terms of waiting time, travel time and costs.

Modes and mobility services do not provide a constant level of service. One reason for this is that all transport modes are dependent on the underlying network. For example, bicyclists have to wait for traffic lights and congestion can occur on motorways. For mobility services it is even more complex. Conventional public transport is scheduled-based, which means that the total travel time by public transportation is dependent on the time of departure. Also new modes can face a dynamic LOS. The requests for demand responsive transport can fluctuate over time, such that waiting times for a vehicle to arrive can be different for each request. Also differences in request allocation strategies can pose varying ranges of waiting and travel times. For shared modes, such as shared cars or bicycle, the availability of these modes is an important determinant. When these bicycle or cars are not available in an area this poses longer accessing or waiting times to use these modes. For commercial mobility services, the company's profitability drives the LOS. For example, the taxi service Uber uses surge pricing when the demand for taxi services is high. (Hall, Kendrick, & Nosko, 2015)

Transport demand models for which the LOS does not remain constant during a model run are referred to as dynamic transport demand models. Although this is a better representation of reality, it is not frequently used for transport demand modelling on larger areas. As the transport demand affects the LOS and vice versa, many computations have to be conducted to provide a dynamic level of service. Therefore many transport demand models are static, as is the case for the SRE model (Goudappel Coffeng, 2012) which is subject to the case study. This does not mean that the variation of the level of service is completely left out of scope for this model. For car traffic, after a first model run is conducted, the assigned loads on the car network determine updated LOS. After the LOS is updated, renewed impedance or resistance values (e.g. in terms of generalised costs) are determined for car trips, such that the transport demand can be distributed for a new iteration. (Goudappel Coffeng, 2012)

2.4 Conclusion

The transport demand modelling frameworks, estimation contexts and application contexts determine the behavioural specifications of transport demand models. Regarding the transport demand modelling frameworks, aggregate gravity models cannot be combined with disaggregate discrete choice models as the modelling frameworks are different. Aggregate gravity models consider the general attractiveness of transport alternatives, whilst disaggregate discrete choice models consider the utility of these alternatives from an individual perspective.

The underlying behavioural considerations are not likely to be the same when the estimation and application contexts of models are different. This relates to for example the information about the transport alternatives the traveller considers and the availability of transport alternatives. When a transport planner considers information from transport demand models within a different context, he needs to be cautious about the validity of the behavioural specifications in this context.

3 Methodology

Chapter 3 sets out the four steps of the methodology. Figure 1 reflects the logical order of these steps. The methodology is developed in order to add external discrete choice utility functions to an existing transport demand gravity model.

Section 3.1 elaborates on the translation of gravity-based deterrence functions towards discrete choice utility functions. Section 3.2 provides information about the linearization of non-linear translated utility functions. In section 3.3, an elaboration is provided about the determination of factors to correct the scale of external models. Section 3.4 provides information about the specification of the combined model. In section 3.5 a brief conclusion of this chapter is provided.

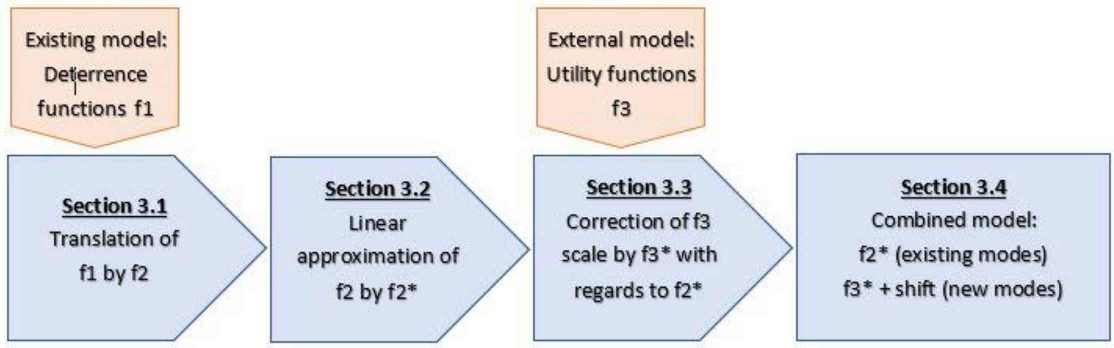


Figure 7: overview of the methodological steps

3.1 Translation of deterrence functions towards discrete choice utility functions

The first step of the methodology consists of the translation of the deterrence functions from the existing gravity model towards discrete choice utility functions.

In the problem analysis (section 2.1) was stated that aggregate gravity models cannot be combined with disaggregated discrete choice models as both units of modelling are different. Aggregate gravity models consider the general attractiveness of transport alternatives, whilst disaggregate discrete choice models consider the utility of these alternatives from an individual perspective.

Therefore one of the models needs to be translated such that both modelling frameworks are the same. Let f_1 be a deterrence function of the existing gravity model and let f_2 be the translated utility function from f_1. The translation can be in both ways, but this methodology only considers the translation of the aggregate gravity model towards a disaggregate discrete choice model. This is the easiest translation procedure as the same explanatory variable of the aggregate form is used in the disaggregate form. The other way around, the explanatory variables of the disaggregate form need to be aggregated towards a single explanatory variable before the translation procedure can be executed.

A method for the translation of f_1 towards f_2 is already available and provides identical model splits. The description of this methodology, including evidence of its validity is given by Bliemer (2010a); (Bliemer, 2010b). Table 1 sets out the utility functions following the translation of multiple deterrence function specifications with the generalised costs (c_{ijm}) being its only explanatory variable.

Table 1: specification of MNL utility functions based as a result of the translation of gravity-based deterrence functions

Deterrence function typology	Utility function Vijm(α_m,β_m,c_{ijm})
Exponential: $f_m(c_{ijm}) = \boldsymbol{\alpha_m} \exp(\beta_m\, c_{ijm})$	$V_{ijm} = \ln(\alpha_m) + \beta_m\, c_{ijm}$
Log-normal: $f_m(c_{ijm}) = \boldsymbol{\alpha_m} \exp(\beta_m ln^2(c_{ijm} + 1))$	$V_{ijm} = \ln(\alpha_m) + \beta_m \ln^2(c_{ijm} + 1)$
Top-lognormal: $f_m(c_{ijm}) = \boldsymbol{\alpha_m} \exp(\beta_m ln^2(\frac{\boldsymbol{c_{ijm}}}{\boldsymbol{\gamma_m}}))$	$V_{ijm} = \ln(\alpha_m) + \beta_m \ln^2(\frac{c_{ijm}}{\gamma_m})$

To visualise the translation of deterrence functions let *q*(c_{ijm}), *r*(c_{ijm}) and *s*(c_{ijm}) represent respectively exponential, log-normal and top-lognormal deterrence functions. Figure 8 and Figure 9 provide the deterrence functions q(c_{ijm}), r(c_{ijm}), s(c_{ijm}) and its translated utility functions *q'*(c_{ijm}), *r'*(c_{ijm}), *s'*(c_{ijm}) respectively. The used parameters for the deterrence functions are α_q = 3, β_q = -2, α_r = 5, β_r = -1,5, α_s = 5 β_s = -3 γ_s = 2. Note that in Figure 9 r'(c_{ijm}) and s'(c_{ijm}) are non-linear, which is caused by the logarithmical scale for c_{ijm}. This means that the sensitivities for marginal changes of c_{ijm} are not equal within the area of scope for the translated log-normal and top-lognormal deterrence functions.

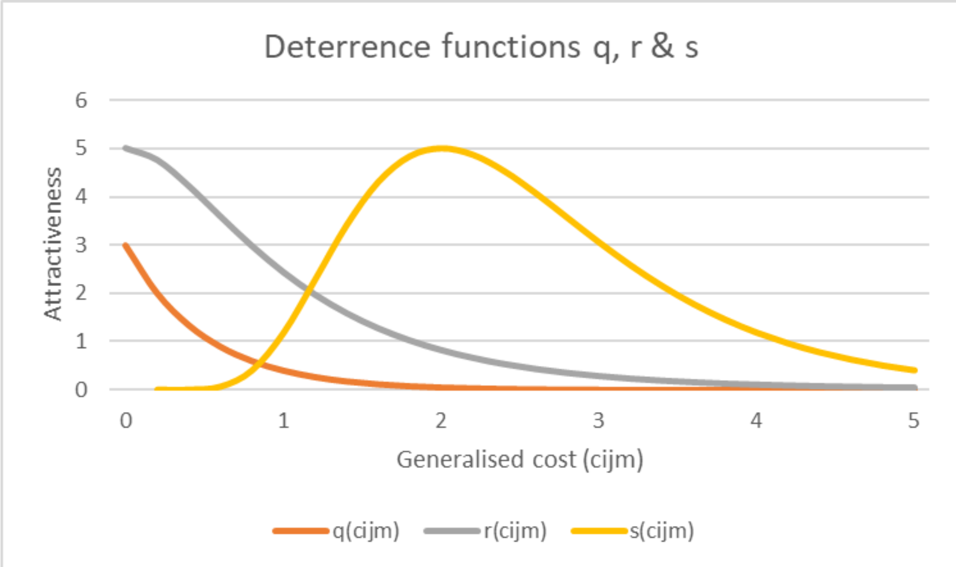


Figure 8: example of deterrence functions q, r & s

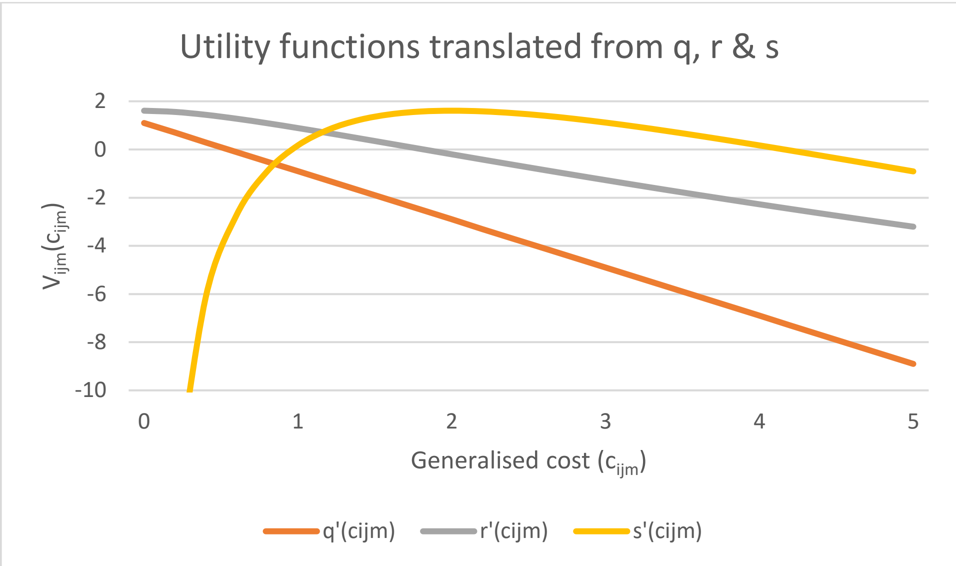


Figure 9: example of utility functions q', r' & s' by translating q, r & s

3.2 Linearization of non-linear translated utility functions

The second step of the methodology consists of the linearization of non-linear utility functions which resulted from the previous translation step.

According the problem analysis (section 2.1), there is a need for a discrete choice model translated from the existing gravity model, such that the systemic utility sensitivities for marginal changes of trip attributes are equal for any origin-destination pair (OD-pair). This uniform systemic utility sensitivity within the complete area of scope is needed as the external models which will be added to the existing model are also specified for equal utility sensitivities for marginal changes of trip attributes for any OD-pair. This makes it more plausible that the underlying behavioural considerations are the same, such that the presence of a transfer bias can be limited.

For this methodological step, let f_1 refer to a log-normal deterrence function of the existing gravity model and f_2 to a non-linear utility function as a result of the translation of f_1. Let f_{2*} refer to the linearized utility functions f_2 with regards to the present values of c_{ijm} within the area of scope.

Consider the mode-specific utility function $f_{2*,m}(c_{ijm}) = ASC_m + \beta_m c_{ijm}$. When both ASC_m and β_m are fixed parameters, the utility function is linear towards c_{ijm} and is represented by a straight line in a graph with c_{ijm} and the (systemic) utility V_{ijm} being its dimensions. The linear utility function $f_{2*,m}(c_{ijm})$ is not able to fit exactly the non-linear utility function $f_{2,m}(c_{ijm})$. The structural error between both functions – defined as the linearization bias - will cause a modal split bias when f_{2*} is used for transport demand modelling instead of f_2. This implicates a trade-off between the transfer bias versus a linearization bias.

This methodological step aims to find the best fit of $f_{2*,m}(c_{ijm})$ towards $f_{2,m}(c_{ijm})$ by minimising the squared utility differences of these functions for all present values of c_{ijm} within the area of scope. Linearization is done by means of the Ordinary Least Squares (OLS) method (McCallum, Hughes-Hallett, & Gleason, 1994). This method uses a set of explanatory values for which the two functions need to be fitted. The original function $f_{2,m}(c_{ijm})$ is determined by c_{ijm} values. This means that the set of explanatory values can be considered as a scatterplot with all present c_{ijm} values within the area of interest. This set is mode-specific as $f_{2,m}(c_{ijm})$ also has a mode-specific specification.

The methodology of OLS aims to minimize the sum of the squared errors between the two functions. Minimising is done by altering the ASC_m and β_m of $f_{2*,m}(c_{ijm})$. Equation 14 represents how the squared error is specified. The sum of the squared errors is based on the set of c_{ijm} for all origins i and destinations j in the study area and are mode-specifically specified. The set of c_{ijm} values is not weighted for the determined use of the related transport alternatives. This is methodological choice as it implicates equal importance of all transport alternatives. Alternatively, the set of c_{ijm} values is weighted for the determined use of the alternatives, when the modeller focusses on minimising the bias for the most important (i.e. most used) transport alternatives.

Equation 14: definition of the squared error of the utility differences between the utility function of the current model (f2) and its linearized approximation (f2)*

$$SSE_m = \sum_i \sum_j (f_{2*}(c_{ijm}) - f_2(c_{ijm}))^2$$

According the first sub-problem, a significant linearization bias needs to be avoided. This requires the assessment of the linearization bias for the conduct of this methodological step. The following step-by-step approach is developed for the assessment of the impact of the linearization bias on the modal split.

- **Consider the study area, the characteristics of transport alternatives and the modelling purpose**
  This provides background information about for which transport alternatives the accuracy of the modal share outcome are considered to be most important.
- **Inspect the functions of $f_{2,m}(c_{ijm})$ and $f_{2*,m}(c_{ijm})$ for each mode visually and determine for which values of c_{ijm} a significant linearization bias is present.**
- **Provide hypotheses about where to expect errors in the modal split outcome for $f_{2*,m}(c_{ijm})$ in relation to $f_{2,m}(c_{ijm})$.**
  This methodology provide hypotheses about the impact. These are discussed later on.
- **Define a tolerance range for which the modal share outcomes are allowed to deviate and define the percentage of transport alternatives for which the tolerance range is accepted to be exceeded.**
  The tolerance range is typically the absolute %-difference between the modal share for $f_{2,m}(c_{ijm})$ and $f_{2*,m}(c_{ijm})$.
- **Calculate the errors in the modal share outcomes**
- **Assess whether these errors are significant**
  This is based on the application context and the accepted tolerance levels (if defined). Attention should especially be given to those transport alternatives for which modal split errors were expected.

Therefore, only hypotheses are stated for the linearization of the translated non-linear utility functions. The hypotheses in this section are formulated as quantitative hypotheses, rather than qualitative hypotheses. This means that the hypotheses will not be assessed by either accepting or rejecting them. It will predominantly provide a starting point for the discussion about the practical validity of the proposed methodology.

By assumption that $f_{2*,m}(c_{ijm})$ approximates $f_{2,m}(c_{ijm})$ the best for values of c_{ijm} in the middle of the range of present c_{ijm} values, it is likely that the linearization bias will impact the modal split outcomes for outer range values of c_{ijm}. *In* addition, discrete choice models are characterised by large sensitivities of the modal split for marginal changes *of* c_{ijm} when the respective mode *m* has a significant market share.

Therefore, the linearization bias is expected to have a significant impact on the modal split only for specific OD-pairs. This is reflected by two qualitative hypotheses:

1. There is a relevant deviation of the modal split for OD-pairs characterised by modes with generalised cost value in the lower range.
2. There is a relevant deviation of the modal split for OD-pairs characterised split by a significant market share for one of the modes available.

3.3 Determination of factors to correct the scale of external models

The third step of the methodology consists of the correction of the scale of external models in order to enable the addition of external utility functions – representing the new modes – to the translated linear utility functions which represent the existing transport demand model.

For indicational forecasts about the future modal split in the presence of new modes it is most important that the combined model is:

- Able to represent the choice probabilities ratios between the new modes and those modes in the existing models which are assumed to face significant demand changes in the presence of the new modes
- Based on external models which comprise similar utility differences between different OD-pairs for the same mode. In this case, both the existing and external model provide similar sensitivities with respect to the other modes in the existing model.

The problem is that utility functions from different discrete choice models cannot be add together directly. Within the MNL framework, as elaborated in section 2.1.3, this is only allowed when the scale factors of the different models are the same and the underlying behavioural considerations for each utility function are similar.

For the different models the scale factors are not specified and therefore implicitly have the value of 1. Given the contexts of the different models, it is likely that the underlying behavioural considerations are not the same.

For this methodology it is necessary to neglect the differences in underlying behavioural considerations. By doing so, the differences between the model specifications for the communal modes – the modes which are present in both models – can be explained by fictional differences in the model scales. These differences cause different choice probabilities for both models when only the communal modes are considered.

By means of this approach, the external model scale can be corrected. Because the scale factors of the different models are the same, this approach de facto alters the sensitivities for marginal changes of the trip attributes for each mode and the structural systemic utility differences between the different modes for the external model.

The correction of the external model scale invokes therefore a transfer bias. Enabling the addition of the external utility functions is namely at cost of altering the behavioural specifications of the external model. The behavioural specifications of the external model will namely be structurally different, in order for the model scales of both models to be the same.

Given the most important requirements of the combined model, making a trade-off between these features of the external model is undesirable. A trade-off between these features can be prevented when the underlying behavioural considerations are equal or at least highly congruous between both models. Therefore, it is important to assess the contexts of both models. The following model specifications need to be inspected and compared in order to provide indications about the differences between the underlying behavioural considerations:

- The number and types of included trip attributes
  The inclusion of similar trip attributes in both models is an indication for similar underlying behavioural considerations. Both models assume that choices are based on similar characteristics of an alternative.

- Absolute parameter values
  For similar model scales and similar attribute units, the parameters for similar trip attributes can be compared. When these parameter values are (almost) similar, this is an indication for similar underlying behavioural considerations.
- Parameter ratios
  For different model scales and similar attribute units, parameter ratios can be compared. When these ratios are (almost) similar, this is an indication for similar underlying behavioural considerations.
- Hypothetical representation of the error term
  Depending on the estimation procedure of both models or differences in the (sampled) populations for the model estimation, the error terms can represent different behavioural considerations. For example by unobserved alternative attributes or unobserved taste variations within the population. This will not impact the variance of the error term, but will impact the model parameters.

The aim of the methodology is to equal the model scales of both models, by assumption that the model scales are different. The correction of the external model scale requires communal modes, which enable the comparison of behavioural considerations between both models. For both models, the set M with the available modes is defined. For both models, an equal set M_c is defined consisting of all communal modes. For the utility functions of both models, the communal modes are denoted by *c*.

Let $f^*_{2,c}$ represent the utility functions for modes *c* of the existing model according the previous methodological steps and let $f_{3,c}$ represent the utility functions for modes *c* of the external model. The utility functions f_3 for which the model scale is corrected are denoted by utility functions f_{3^*}.

For equal model scales it is known that the utilities for communal modes are the same for each OD-pair. In addition, equal choice probabilities for both models are reflected by equal utility differences between multiple communal modes for both models. Therefore, for the approximation of choice probabilities of the existing model, two or more communal modes are needed.

Therefore, this methodological step aims to minimise the differences *of the systemic utility differences between all combinations of communal modes* between both models. The difference of underlying behavioural preferences in practice is reflected in the structural systemic utility differences between communal modes (structural element) and sensitivities towards marginal changes of trip attributes (trip-specific element). The structural systemic utility differences are captured by the ratios of the *ASC* from multiple communal modes. The sensitives towards marginal changes of trip attributes are reflected by the summed product of the trip attributes *x* and the related parameters β_x.

The differences *of the systemic utility differences between all combinations of communal modes* between both models can best be minimised by means of an extra degree of freedom, rather that solely one correction factor for the complete model scale. Instead of correcting the complete model scale, the structural elements and the trip-specific elements of the external model can be corrected separately.

When only a single communal mode is available, there is no possibility to correct for structural systemic utility differences between multiple modes. For this situation the methodology aims to minimize the utility differences between communal modes by solely correcting the differences in the systemic utility sensitivities towards marginal changes of trip attributes.

First, correction factors are introduced which are the same for all utility function of the external model. The correction factors θ_{ASC} and $\theta_{\beta X}$ are multiplied respectively with the structural elements and trip-specific elements of all utility functions f_{3*}. The respective expression of these utility functions is given in Equation 15.

Second, the minimisation procedure of the differences *of the (systemic) utility differences between all combinations of communal modes* between both models for all combinations of origins *i* and destinations *j* within the area of scope is conducted by means of the Ordinary Least Squares (OLS) method (McCallum et al., 1994). By altering θ_{ASC} and $\theta_{\beta X}$ the sum of the squared differences are minimised as visualised by Equation 16. In this equation *a* and *b* represent the communal modes within M_c.

Equation 15: expression of utility functions f3 with a structural and trip-specific element*

$$f_{3*} = \theta_{ASC}\ ASC_m + \theta_{\beta x}\ \sum_{x=1}^{A_m} \beta_x\ x_{mij}$$

Equation 16: definition of the sum of squares of the utility differences when there are multiple communal modes

$$SSE\Delta V = \sum_{a=1}^{M_c}\sum_{b=1}^{M_c}\sum_{i=1}^{m}\sum_{j=1}^{n}((f_{3*,aij} - f_{3*,bij})(f_{2*,aij} - f_{2*,bij}))^2$$

$$with\ i \neq j\ and\ a \neq b$$

When only a single communal mode is available, the methodological procedure is slightly different. As it is not possible to correct the differences *of the (systemic) utility differences between all combinations of communal modes* between both models, solely the (systemic) utility differences between $f^*_{2,c}$ and $f_{3*,c}$ are minimised. As this only relates to the correction of the trip-specific element, correction factor θ_{ASC} is fixed to 1. The definition of the sum of squares of the (systemic) utility differences is given in Equation 17.

Equation 17: definition of the sum of squares of the utility differences when there is a single communal mode

$$SSE\Delta V = \sum_{i=1}^{m}\sum_{j=1}^{n}(f_{3*,aij} - f_{2*,bij})^2$$

$$with\ i \neq j\ and\ a \neq b$$

The next figures provide an example of the impact of this methodological step for communal modes *a* and *b*. The example is presented as a series of sequential impacts. In reality, the OLS methodology alters the values of θ_{ASC} and $\theta_{\beta x}$ iteratively in order to determine the sum of the squared utility differences.

In Figure 10 the impact of $\theta_{\beta x}$ on the slope of utility functions $f^*_{3,c}$ is shown. In addition, Figure 11 shows the impact of θ_{ASC} on the structural distance between utility functions of $f^*_{3,c}$. In this example, the correction of $f^*_{3,c}$ resulted in similar slopes of the related utility functions and similar structural distances between them in comparison to $f^*_{2,c}$. By this correction, the absolute (systemic) utility differences between the communal modes are similar such that both $f^*_{3,m}$ and $f^*_{2,m}$ forecast the same modal split

This example is a simplification from reality, as the set of the summed product of β_{mija} and X_{mija} is replaced by a single β related to the generalised cost as the only trip attribute value. External utility functions mostly have multiple explanatory variables (attributes) such that these function cannot be considered as a line but as a multi-dimensional plane. For *n* number of explanatory variables, the solution space for the values of θ_{ASC} and $\theta_{\beta x}$ to minimise the SSE ΔV can be visualised as a multi-dimensional plane with *n-1* dimensions in a space with *n* dimensions. (McCallum et al., 1994)

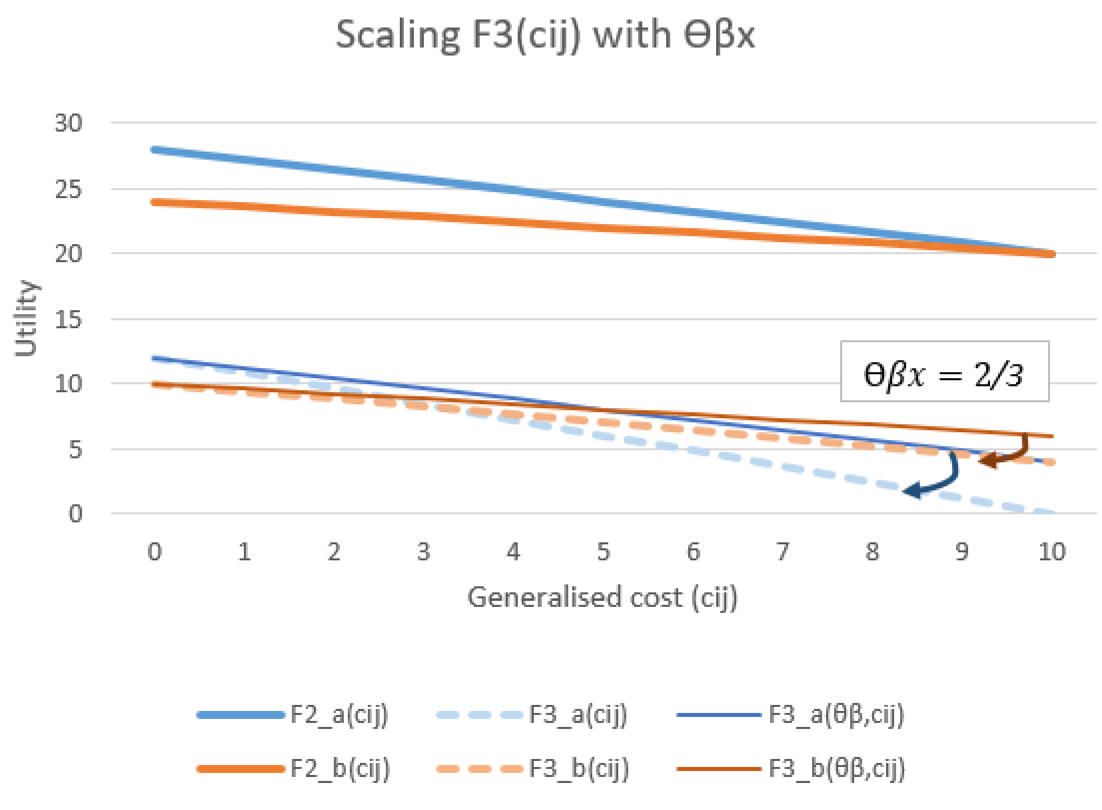


Figure 10: conceptual overview on scaling utility functions with Θβx

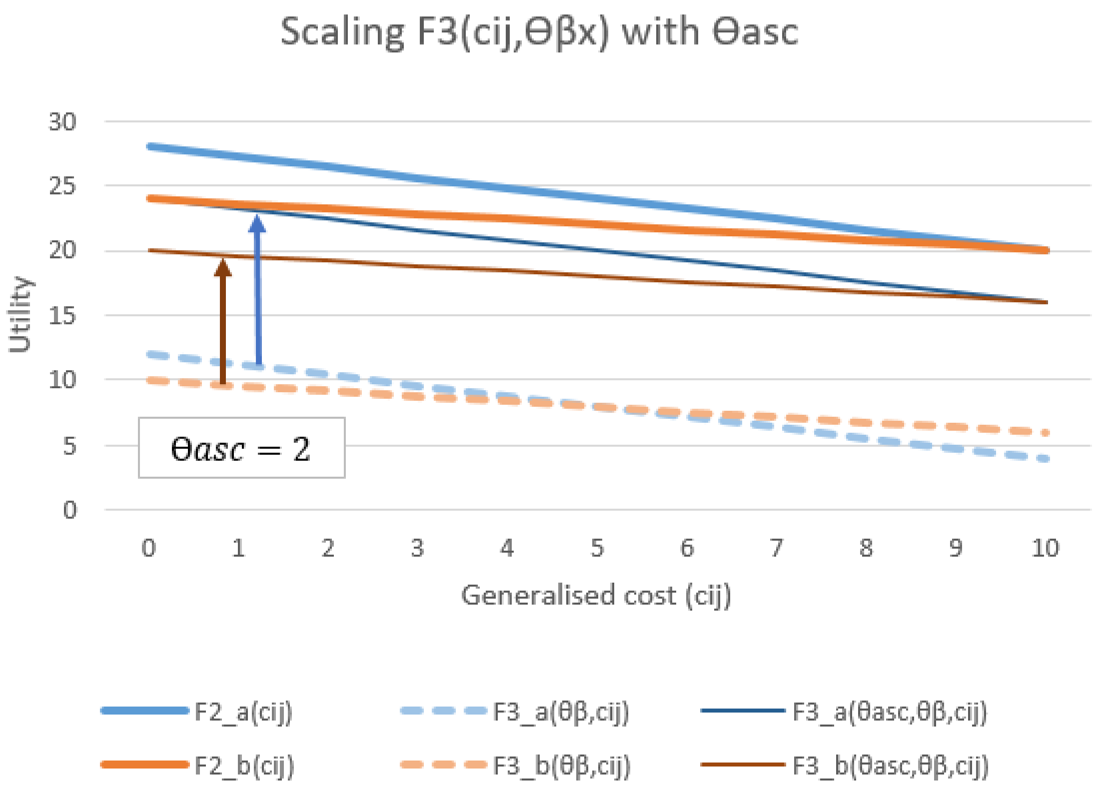


Figure 11: conceptual overview on scaling utility functions with ϴASC

The sizes of the correction factors θ_{ASC} and $\theta_{\beta X}$ need to be interpret as a measure for the impact size of the transfer bias. The structural systemic utility differences between modes and the systemic utility sensitivities towards marginal changes of trip attributes are enlarged by the multiplication of the alternative-specific constants and the summed product of the trip attributes *x* and the related parameters β_x with the respective correction factors.

The parameters β_x for all trip attributes *x* are multiplied with the same correction factor $\theta_{\beta X}$. This means that the parameter ratios remain similar. Parameter ratios relate to important behavioural considerations as they provide information about the trade-off of trip attributes with respect to other trip attributes. One of the most important parameter ratios is the value of time (VOT). The VOT provides information about the amount of money a traveller is prepared to pay in order to save a unit of time.

3.4 Specification of the combined model

The fourth step of the methodology consists of the addition of external utility functions – representing the new modes – to the translated linear utility functions by adjusting the alternative-specific constants of the external utility functions uniformly.

For similar behavioural considerations and similar model scales the systemic utilities for communal modes need to be the same. In order to equal the utilities of the communal modes, an equal amount of systemic utility is added to all utility functions of the external model. This additional systemic utility constant is defined as the shift factor. When the systemic utilities of the communal modes in the existing model are approximated by the external model, the utility functions are considered to be interchangeable. The utility functions of the non-communal modes of the external model represent similar trade-offs with the communal modes, independent of the origin of the utility functions of these modes.

For this methodological step Let $f_{2*,c}$ and $f_{2*,m}$ represent the utility functions for respectively the communal modes *c* and the non-communal modes *s* in the existing model and let $f_{3*,c}$ and $f_{3*,m}$ represent the utility functions for respectively the communal modes *c* and the non-communal modes *s* in the external model. The resulting set of utility functions for the combined model is represented by f_4.

The size of the shift value is determined by having the best approximation of the ASCs of $f^*_{3,c}$ towards the ASCs of $f^*_{2,c}$. The sensitivities of the systemic utility towards marginal changes of the trip attributes for $f^*_{3,c}$ already approximate the trip-specific sensitivities of $f^*_{2,c}$ as a result of the correction of the external model scale. Therefore, there is no need to determine the shift value by means of minimising the systemic utility differences between $f^*_{3,c}$ and $f^*_{2,c}$ for all present OD-pairs.

Thus, the shift value is determined by minimising the difference between the ASC_c of $f^*_{2,c}$ and the ASC_c of $f^*_{3,c}$ and the shift value for all communal alternatives. This minimising procedure is – similar to the previous methodological step – conducted by means of the Ordinary Least Squares (OLS) method (McCallum et al., 1994). The shift value is altered such that the smallest value of the sum of the squared errors is obtained. Equation 18 reflects the respective sum of the squared errors.

Equation 18: definition of the squared error of the ASC and the shifted ASC of f2 and f3* respectively*

$$SSE\ \Delta ASC = \sum_{c}^{M_c} ((ASC_{f_{3*,c}+shift}) - ASC_{f_{2*,c}})^2$$

Figure 12 provides an example of the methodological procedure. It elaborates on the previous example of fictive sets of $f^*_{2,c}$ and $f^*_{3,c}$ for communal alternatives *a* and *b* in section 3.3. In this example the ASC_c for $f^*_{2,c}$ is 28 for alternative *a* and 24 for alternative *b*. For $f^*_{3,c}$ the corrected ASC_c is 24 for alternative *a* and 20 for alternative *b*. According to these values, a shift of $f^*_{3,c}$ with a value of 4 would provide the best approximation of the utility functions $f^*_{2,c}$.

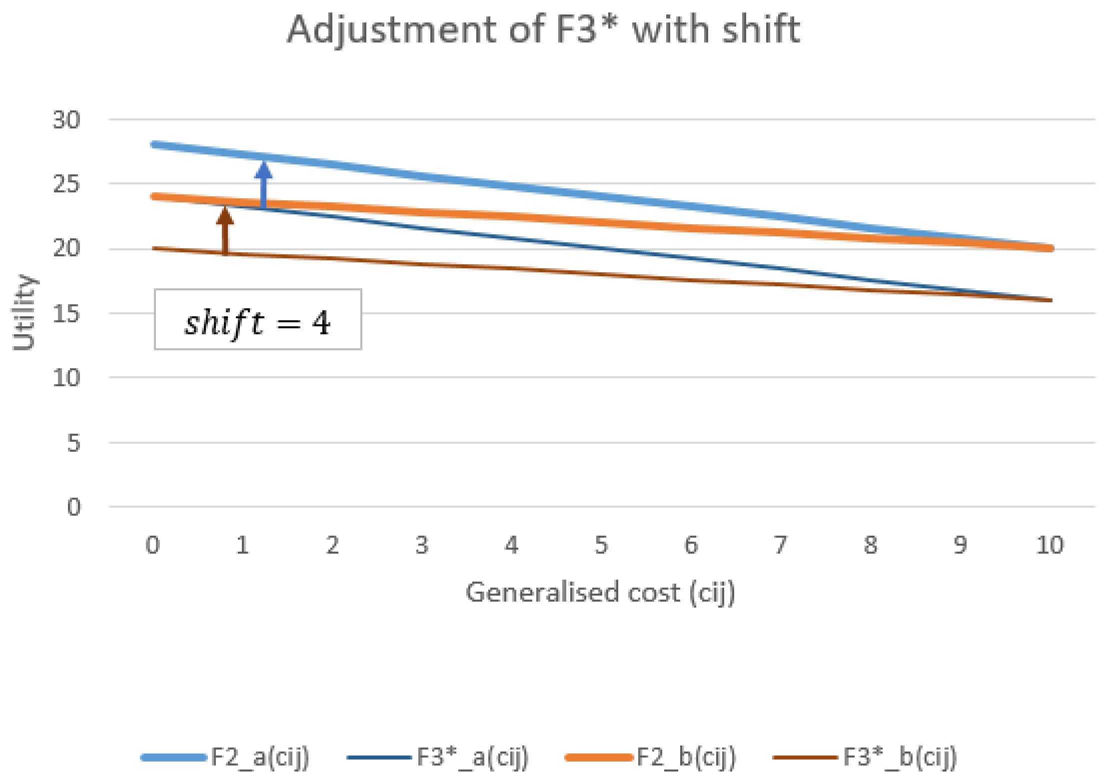


Figure 12: conceptual overview on shifting utility functions f3 towards f2**

In this example $f^*_{3,c}$ can be shifted such that they exactly correspond with the utility functions of $f^*_{2,c}$. The number of communal alternatives impacts the ability to approximate the systemic utilities closely. When there are more communal alternatives it is more difficult for each alternative to approximate the respective utility functions. This implicates that it is recommendable to assess the need to include all communal alternatives. When some of the communal alternatives are excluded from the methodology, this could improve the approximation of the systemic utilities of the remaining communal alternatives. In that case, the choice probabilities between the communal alternatives and the non-communal (existing and new) alternatives could be better represented, potentially improving the validity of the combined model f_4. However, this is likely to be at cost of less valid choice probabilities for the excluded modes.

After the determination of this shift value, the utility functions $f^*_{2,m}$ representing the new modes are shifted with this value and are added to the utility functions f^*_2, such that these utility functions are referred to as f_4 within the combined transport demand model. The representation of f_4 in terms of the utility functions of $f^*_{2,m}$, $f^*_{2,c}$ and ($f^*_{2,m}$ + *shift*) is visualised by Equation 19.

By doing so, f_4 represents the same choice probabilities of the new alternatives (new modes) *n* in relation to the communal alternatives. Similarly, f_4 represents the similar choice probabilities between the communal alternatives *c* and non-communal alternatives *m* of the existing transport demand model. When the assumption of the underlying behavioural considerations of both models being equal is valid, model f_4 would provide a realistic representation of future choice probabilities when both existing and new modes are present.

Equation 19: set of utility functions of model f4

$$f_4 = \begin{cases} f_{2*,c} \\ f_{2*,m} \\ f_{3*,m} + shift \end{cases}$$

3.5 Conclusion

The methodology consists of four methodological steps: (1) the translation of deterrence functions towards discrete choice utility functions, (2) the linearization of non-linear translated utility functions, (3) the determination of factors to correct the scale of external models and (4) the specification of the combined model.

These methodological step specify multiple utility functions in order to specify the utility functions of the combined model. An overview of these utility functions is given in Table 1. Information about these utility functions can be found in the related sections.

Table 2: overview of the utility functions specified within the methodology

Function	Function description	Section
f_1	Deterrence functions of gravity model	3.1
f_2	Translation of f1 as a utility function	3.1 / 3.2
f_{2*}	Linear approximation of f_2	3.2 / 3.3 / 3.4
f_3	Utility functions of external model	3.3
f_{3*}	Scaled version of f_3 towards the model scale of f_{2*}	3.3
f_4	Combined model with utility functions f2* for existing modes and shifted f3*for new modes	3.4

4 Application of methodology

Chapter 4 applies the methodology in a case study for Urban Mobility as a Service in the Eindhoven-Veldhoven area. This is done such that to explore the practical validity of the methodology for forecasting future modal splits in the presence of new models.

Section 4.1 provides information about the case study and the selected models used for the methodology application. The next sections provide information about the execution of the different methodological steps for this case study. Section 4.2 relates to the translation of deterrence functions towards discrete choice utility functions. Section 4.3 relates to the linearization of non-linear translated utility function. Section 4.4 relates to the determination of factors to correct the scale of external models. Section 4.5 relates to the specification of the combined model. Section 4.6 provides the modal splits of the combined model. Finally, in section 4.7 the conclusions are provided.

4.1 Case study

In section 4.1.1 the case study is introduced. Section 4.1.2 provides an elaboration about the contexts of the transport demand models which will be combined for this case study. In section 4.1.3 the contexts of these models are compared.

4.1.1 Introduction

The case study for the application of the methodology relates to the concept of Urban Mobility as a Service (Urban MaaS) in the Eindhoven-Veldhoven (EHV-VHV) area in the year 2020.

Eindhoven is the 5th largest city of the Netherlands with over 227.000 inhabitants and is situated in the south of the country. It accommodates a university of technology and is surrounded by the 2nd largest airport of the Netherlands. One of its satellite towns is Veldhoven with over 44.000 inhabitants. (CBS, 2018) It is directly attached to the city of Eindhoven with the national highway A2 defining the boundary between the two places. The prevailing urban transport demand model for the larger Eindhoven area is the so-called the SRE ("Samenwerkingsverband Regio Eindhoven") model. (Goudappel Coffeng, 2012)

As stated in the introduction, future travel behaviour will likely be impacted by the emergence of Mobility as a Service. In this thesis, MaaS is considered to be the enabler for using multiple transport services by means of a smartphone application which takes care of trip planning, ticketing and payment. This general definition needs to be specified more, such that it becomes a tangible concept to apply in the case study.

The specific MaaS concept for this case study is called Urban MaaS. It aims to provide citizens of the EHV-VHV area to travel anywhere within the area with a multitude of transport options. Travellers can use the Urban MaaS app for getting real-time information about the transport options available. Apart from using their own car or bicycle, they can use public transport (which mainly consists of city buses and high frequency operated bus lines), demand responsive transport (DRT) and shared cars. The latter two modes are currently not represented in the SRE model. It is assumed that the availability of these modes will have a significant impact on the travel behaviour in the EHV-VHV area. Therefore, external models need to be combined with the SRE model, such that utility functions for these modes can be add to this model.

DRT is conceptualised as public transport on request and is operated by mini-busses offering door-to-door transport. After raising a request, the traveller waits at its origin. The mini-bus will directly transport the traveller, without any transfers. However, the mini-bus can deviate from the shortest route by picking up (or dropping of) other passengers. The fare system of DRT consists of a single fare, independent of the distance travelled within its service area.

The shared car system is conceptualised as a station-based one-way system. Users of the Urban MaaS app get information about the nearest located station with shared cars available. The system is developed such that within 250 metres at least two stations are available. Travellers open the shared car with their smartphone or smart card and can directly use the car. When approaching the destination, travellers have to park the car at a different station, such that there will be some egress time involved. The use of shared cars is paid by minute. The fare system does not differentiate its fares – driven by demand fluctuations – within the considered period-of-day.

For this case study, the trip attributes for both DRT and shared car are considered to be static and relate to the level of service of the car network. Table 3 states the respective values for the attributes of DRT and shared car.

Table 3: specification of trip attributes for DRT and shared car in case study

Trip attribute	Attribute value
DRT – in-vehicle time (IVT)	1,4 x private car IVT
DRT – waiting time	10 minutes
DRT - number of transfers	0 transfers
DRT – costs	3,5 euro
Shared car – access+egress time	3 minutes
Shared Car – in-vehicle time	1 x private car IVT
Shared car - costs	0,31 euro/min

The concepts of DRT and shared cars are in line with the respective concepts in the external models which are used for this case study. For DRT, this relates to the discrete choice model of Frei et al. (2017). For shared cars, this relates to the discrete choice model stated by Martínez, Correia, Moura, and Mendes Lopes (2017), which originates from an earlier estimation procedure of Eiró and Martínez (2014). These models will be combined with the SRE model.

For this case study, only the non-peak period (i.e. rest-of-day) is considered for purpose "other", which mainly consists of social-leisure trips. This subset of the overall travel demand is chosen as travellers are considered to be potentially susceptible for the use of Urban MaaS for these specific trips. The underlying hypothesis is that travellers have more fixed preferences ("habits") for commute, education-related and shopping-related trips. The case study considers both users groups which are specified within the SRE model, namely the car-available (CA) and car-unavailable (CU) group. The CA group has always access to use a private car, whilst this is not the case for the CU group. Therefore, the terminology can be confusing, as the CU group also relates to people who have limited access to a private car.

4.1.2 Contexts of the transport demand models

This section provides information about how the three transport demand models can be characterised and are estimated. This relates to the modelling frameworks and the estimation of the respective models introduced in section 2.1.

SRE model for Eindhoven-Veldhoven

The SRE model (Goudappel Coffeng, 2012) is the prevailing urban transport demand model for the greater Eindhoven area, which is a trip-based aggregated gravity model. It provides travel forecasts for three modes: private car, public transport and bicycle.

The model is structured such that the assignment of private car trips impacts the level of service (LOS) of the car network which affects the generalised costs for the private car alternatives. This implicates that the forecasted transport demand in the first iteration affects the LOS, such that travellers could can choose different alternatives within next iterations. For the methodology application this feed-back loop is considered out of scope.

The transport demand model of the SRE model is specified by means of log-normal deterrence functions which are different for each combination of model dimensions. The generalised costs reflect the impedance or resistance value for transport and is composed by several elements. For private car and bicycle these are the travel time and travel costs. For public transport these are the access time, the waiting time, the in-vehicle time, the egress time, the number of transfers and the travel costs.

Table 4 provides the estimated parameters of the SRE model. The parameter values of the deterrence functions are large. This relates to the strong preference for car transport for the CA user group and the area of scope which considers the transport alternatives for all destinations within the Netherland.

Table 4: specification of the gravity model parameters in the SRE model for purpose "other"

Parameter	Car	Public Transport	Bicycle
Car available (CU)			
α_m	20241	1048	27912
β_{cijm}	-1,30	-0,82	-1,19
Car unavailable (CU)			
α_m	1897	983	8835
β_{cijm}	-1,60	-0,89	-1,42

The parameters of the deterrence functions are deducted from the national travel research in the Netherlands ("Onderzoek verplaatsingen in Nederland; OViN"), where thousands of participants are questioned about their travel behaviour. Therefore, the SRE model is estimated based on revealed preference data. The parameters aim to fit the modal splits and trip length distributions from the OViN research. (CBS, 2015)

That means that there is a single deterrence function available for each combination of mode, purpose and class membership of CA/CU. As a very big set of possible transport options is considered, a single deterrence function will pose problems for very long trips (more concisely: trips with high generalised cost values) and very short trips. The long trips have a low prevalence such that it is very difficult to forecast the use of specific alternatives at the long range. For the short trips it is expected that there are many situations where specific particularities of local zones have a significant impact on mobility behaviour, rather than the single value of the generalised costs.

Regarding the application context, the most important socio-demographic information is given in Table 5. The age distribution of the population is given in Table 6. The age distribution of Eindhoven shows a vast amount of young people. This can be explained by the presence of multiple higher education institutions.

The general modal split of Eindhoven as a percentage of the total number of trips – being independent of trip purpose or car availability – amount 42% for car, 5% for bus, 40% for bicycle and 13% walking.

Considering this modal split, it becomes clear that the share of bicycle trips is very high. This can be explained by the high urban densities of the city, the penetration rate of bicycles (i.e. bicycle ownership rates are high) and the quality and density of the biking infrastructure. The modal split for buses is higher than the modal split which is determined by the transport demand model. This can be explained as the transport demand model also considers Veldhoven for which the accessibility by public transport is assumed to be lower.

Table 5: socio-demographic information of the Eindhoven population

Socio-demographic information	Eindhoven
Percentage male	51%
Average household size	1,9 person/HH
Gross monthly income	2425€/person
Percentage of people having a car available	68,4%
Average cars per household	0,9 car/HH

Table 6: age distribution of the Eindhoven population

Age distribution	Eindhoven
18-24	16%
25-44	35%
45-64	29%
65+	20%

Demand Responsive Transport in Chicago

Frei et al. (2017) conducted stated preference research towards the attitudes of car transport versus conventional public transport and flexible public transport ("flex transit"). Both a mixed logit and a multinomial logit model were estimated in this research. Respondents filled in a questionnaire and each respondent was requested to make at least six trade-offs between the three transport alternatives having their current commute trip as a reference. Frei et al. (2017) conceptualises flexible public transport as public transport by mini-vans and gives travellers the opportunity to either have a pick-up and drop-off at home or at pre-defined hubs. When the traveller desires to be picked-up at home, waiting times were assumed to be longer. The latter conceptualisation of flexible public transportation is congruent with the definition of DRT in this case study.

Participants were selected from all over the greater Chicago area. The reference commute trips varied from very short distances (2< miles) to long distances (>25 miles) such that the estimated has a relatively broad application range. The alternative's attributes were varied along the specifications stated in Table 7 (Frei et al., 2017).

Table 7: overview on the variation of trip attributes in Frei et al. (2017)

	Private car	Public transport	Flex transit
In-vehicle time	1,5; 2; 2,5x Google car time	0,5; 1,0; 1,5x Google public transport time	1,8; 2,4; 3x Google car time
Travel costs	0,5; 1,0; 1,5 x reference costs	Short/medium: $1, $2 or $4 Long: $2, $4 or $6	Short/medium: $1, $ 2 or $4 Long: $2, $4 or $6
Walking time	1,3 or 5 minutes	0,5; 1,0; 1,5x Google walk time	1,3,5 minutes
Wait time	NA	0 minutes	Short/medium: 1, 3 or 5 minutes Med/long: 4, 7 or 12 minutes
Headway	NA	5, 12, 25 minutes	20, 30 or 60 minutes
Number of transfers	NA	Short: 0 or 1 transfer Medium/long: 0, 1 or 2 transfers	0 or 1 transfer

Table 8: socio-demographic information of the population sample of Frei et al. (2017)

Socio-demographic information	Sample Chicago
Percentage male	57%
Average household size	2,38 person/HH
Percentage of people having a car available	53%
Car Ownership	67%

Table 9: age distribution of the population sample of Frei et al. (2017)

Age distribution	Sample Chicago
18-24	7%
25-44	64%
45-64	26%
65+	3%

The group of participants completed 1997 choice experiments. Some socio-demographic information about this population sample is given in Table 8 and

Table 9 respectively. Most outstanding is the presence of participants in the age category 24-44, meanwhile having low participant rates for the age categories 18-24 and 65+.

The estimated models of Frei et al. (2017) determines choice probabilities on a multitude of explanatory variables. Apart from the attributes of each transport alternatives this relates to contextual information (e.g. weather circumstances), current commute characteristics (e.g. home-work distance), travel-related organisational memberships (e.g. carsharing member) and socio-demographic information (e.g. age and gender). In terms of statistical fit, the multinomial logit model has an adjusted rho-squared value of 0,320. The related mixed logit model has an adjusted rho-squared value of 0,423 indicating the latter model to be a better representation of the respondent's choices.

For simplicity's sake, the related dummy and categorical variables are left out of scope for the methodology application. This means that for these variables the base values are assumed. In future research, population synthesize methods can be used to generate more realistic population characteristics. Information about the base values is stated in appendix I. This appendix also provides the full estimation results for both the mixed logit and multinomial logit model.

The model parameters of the multinomial logit model which are used for this case study are stated in Table 10. All model parameters are significant at a level of confidence of 90%. This parameter set is slightly adjusted from the original parameter set of the estimated multinomial logit model. The parameter for costs is converted such that it can be used for Euro-based fares.

Table 10: estimated parameters from Frei et al. (2017)

Parameter	Unit of measurement of explanatory variable	Car	Public Transport	DRT
ASC		0	-0,81	-1,816
$\beta_{distance}$	Kilometre			
$\beta_{access\ time}$	Minute		-0,084	
$\beta_{in\text{-}vehicle\ time}$	Minute	-0,068	-0,063	-0,064
$\beta_{waiting\ time}$	Minute		-0,084	-0,107
$B_{transfer}$	# transfer		-0,2215	-0,2215
$\beta_{egress\ time}$	Minute		-0,084	
β_{costs}	Euro ($1 = €0,81)	-0,1847	-0,1847	-0,1847

The modelling context relates to the Greater Chicago Area. The most important socio-demographic information is given in Table 11. The age distribution of the population is given in Table 12. The age distribution of Chicago shows a vast amount of people between the age of 25 and 44. In the estimation sample the percentage of this age class group was even higher. This can mainly be explained by the related trip purpose commuting as this age group predominantly represents the working people. The same explanation seems to be valid for the overrepresentation of the male population in the sample.

Table 11: socio-demographic information of the Chicago population in comparison to the sample

Socio-demographic information	Chicago	Sample Chicago
Percentage male	49%	57%
Average household size	NA	2,38 person/HH
Percentage of people having a car available	NA	53%
Car Ownership	NA	67%

Table 12: age distribution of the Chicago population in comparison to the sample

Age distribution	Chicago	Sample Chicago
18-24	11%	7%
25-44	45%	64%
45-64	30%	26%
65+	14%	3%

The general modal split of Chicago as a percentage of the total number of trips for the purpose commuting amounts 52,7% for car, 27,9% for public transport, 6,8% for walking and 12,6% of other transport alternatives.

Shared cars in Lisbon

Eiró and Martínez (2014) used both revealed preference and stated preference data in relation to a multitude of existing and new transport alternatives. The existing alternatives consist of: private car, motorcycle, bus, metro and a combination of public transport modes. The new transport alternatives related to the shared use of modes such as: carpooling, carsharing, express minibus and shared taxi. This was research was conducted to gain a better understanding about the future trade-offs between transport alternatives when new (shared) alternatives are available. The research was performed within the Greater Lisbon area with most trip distances up to 15 kilometres. The estimated model is quite general as it is not specified for a specific trip purpose.

Revealed preference data from an online panel was combined with interview data from 1000 respondents in order to mitigate the online sample bias. The interviews provided both revealed and stated preferences. Participants were asked to provide all information about their trips of the previous day. In addition, participants needed to select alternatives from future alternative sets consisting of a private modes, public modes and future modes. These fictive choices were based on all trips the participant had made the previous day. Unfortunately, Eiró and Martínez (2014) did not explicate the ranges for which the alternative's attributes were varied. General information about the population sample is given in Table 13.

Table 13: socio-demographic information of the population sample of Eiró and Martínez (2014)

Socio-demographic information	Sample Lisbon
Percentage male	40%
Average household size	2,66 person/HH
Gross monthly income	2031 euro
Average age	45,7 year

Based on the research of Eiró and Martínez (2014) a nested model was constructed by Martínez et al. (2017). This nested model consists of a nest with private modes (car, taxi and motorcycle) and public modes (bus, walk, subway, multi-modal public transport and car sharing). The quality of fit of this model is indicated by an adjusted rho-squared value of 0,37 with all model parameters significant at a level of confidence of 90%.

Only the nest with public modes is considered for the methodology application. It is not expected that doing this poses additional problems to the validity of the methodology and the outcomes of the combined model; The proposed methodology adds utility functions to an existing model within the MNL modelling framework. This implies that the methodology itself does not establish whether an MNL model or a different kind of modelling framework provides the most valid outcomes on future mobility behaviour.

The shared car alternative is conceptualised similarly to the shared car concept in this case study, namely being a station-based one-way system. The bus alternative is considered to be the only communal alternative as the urban public transport system of Eindhoven-Veldhoven is based on buses.

The two beforementioned alternatives establish the model which will be added to the existing transport demand model. Table 14 provides the estimated parameters of these alternatives and – as a reference - the estimated parameters of the excluded alternatives.

Table 14: estimated parameters from Martinez et al. (2017)

Parameter	Unit of measurement of explanatory variable	Bus CA (= Public Transport in EHV-VHV for Car Available)	Bus CU (= Public Transport in EHV-VHV for Car Unavailable)	Walk	Heavy Public Transport (underground + rail)	Bus + Heavy Public Transport	Carsharing for Car Available	Carsharing for Car Unavailable
ASC		-0,218	0,344	-1,140	-0,270	-0,706	-1,498	-1,19
$\beta_{distance}$	Kilometre							
$\beta_{access\ time}$	Minute	-0,051	-0,051		-0,053	-0,051	-0,082	-0,082
$\beta_{in\text{-}vehicle\ time}$	Minute	-0,024	-0,024	-0,005	-0,015	-0,015	-0,026	-0,026
$\beta_{waiting\ time}$	Minute	-0,028	-0,028		-0,045	-0,028		
$B_{transfer}$	# transfer	-0,199	-0,199		-0,150	-0,150		
$\beta_{egress\ time}$	Minute	-0,051	-0,051		-0,053	-0,051	-0,082	-0,082
β_{costs}	Euro	-0,498	-0,498		-0,498	-0,498	-0,229	-0,229

Note that the utility functions of the alternatives are specified for both CA and CU population groups. This is done as public transport card holders gain an extra utility of 0,562 for bus and an extra utility of 0,308 for car sharing. This impact is taken into account as the model itself is very general as it is not specified for a specific trip purpose. In order to bring in some heterogeneity in preferences, the CU population group is considered to have a public transport card.

The modelling context consists of The Greater Lisbon Area which has nearly three million inhabitants of which 550.000 live within the city of Lisbon. Everyday more than 400.000 people travel towards Lisbon for work or study. (Instituto Nacional de Estatística, 2012)

The most important socio-demographic information is given in Table 15. The age distribution of the population is given in

Table 16. Most important characteristic of the population of Lisbon is the average household size to be relatively high with 2,62 persons per household. When the socio-demographics of the population are compared with the characteristics of the sample group it can be seen that the sample under represents males and people with a higher income.

Table 15: socio-demographic information of the Lisbon population in comparison to the sample

Socio-demographic information	Lisbon[1]	Sample Lisbon
Percentage male	48%	40%
Average household size	2,62 person/HH	2,66 person/HH
Gross monthly income	2289 euro	2031 euro
Average age	46,4 year	45,7 year

Table 16: age distribution of the Lisbon population

Age distribution	Lisbon
18-24	9%
25-44	37%
45-64	32%
65+	22%

The general modal split of Lisbon as a percentage of the total number of trips for all trip purposes together amounts 43,7% for car, 21,4% for bus, 13,8% for walking and 19,4% of walking and 1,7% of other transport alternatives.

4.1.3 Comparison of the contexts of the case study models

This analysis aims to determine to which extent the model estimation contexts and the model application contexts of the different models are similar. This information is important as it needs to be compared to the methodology assumption that the underlying behavioural preferences for each model are the same.

Estimation contexts

The three considered models are all estimated with a different modelling framework. The SRE model is an aggregated gravity-based model and uses information about trip generation rates, modal splits and trip length distributions from Dutch travel research (CBS, 2015). The model is calibrated such that the trip generation per zone and empirical traffic flow count data is approximated as close as possible. The Chicago model is a multinomial logit model and it estimation is based on a stated-choice experiment. The Lisbon model is a nested logit model which is estimated on both (revealed preference) travel diaries as stated-choice experiments.

The use of a nested logit model is not considered to be a problem as the IID assumption holds for all alternatives within a nest. The proposed methodology can still be used, by only considering the communal alternatives who share the same nest as the shared cars alternative. This means that the private car alternative cannot be considered as a communal alternative.

The purposes for which the three models are estimated are not the same. The SRE model considers the purpose “other”, while the Chicago model is estimated for the purpose commute. The Lisbon model is estimated independent for the type of purpose. It is much likely that these different purpose result in different behavioural considerations. This is an indication that the underlying behavioural considerations for the three models are not the same.

Also, the number of considered alternatives (modes) is different for the three models. The SRE model only considers private car, public transport and bicycles. This model is developed for policy making and these three modes can be considered to be the most relevant for transport policy and infrastructure investments. The Chicago model only considers choices between private car, public transport and DRT. The aim of the related study was to determine the trade-offs between these three modes. The Lisbon model considers the most alternatives with private car, road-bound and rail-bound public transport, walking, shared car and other alternatives (taxi, multi-modal public transport and motorcycle). As the participants for the two external modes consider different alternative sets, this means that de facto the behavioural considerations of these participants are not the same.

Unfortunately, there is too little information known to make a sound comparison between the population of the SRE model and the sampled populations from the external models. What is known is that the percentage of young (age: 18-24) and old (age: 65+) people op the sampled population of the Chicago is model is low. This makes sense, as this model is estimated for the purpose “commuting”.

Table 17 considers considering the current modal splits of the sampled population. From this table, it can be seen that the use of private car is relatively low for Chicago and Lisbon. A possible explanation for this is that these cities are bigger and the population numbers are higher, such that higher urban densities in these cities make it more difficult to use a private car. According these sampled modal splits it is also striking that public transport is more important for the external models than for the SRE model. This can be explained that cycling is considering to be very normal in the Netherlands such that many non-private car trips are conducted by bicycle. In addition, it is remarkable that the number of trips on foot is much higher for the Lisbon sample in comparison to the other models.

Table 17: comparison of the modal splits from the population samples of the three models

Modal split of the sampled population	Source	Car	Bus	Bicycle	Metro	Walk	Other
Eindhoven-Veldhoven		53,7%	1,2%	45,7%	-	-	-
DRT - Chicago	Frei et al. (2017)	24,6% (33,4%)	49,5% (66,6%)	-	-	7,9%	18,0%
Shared car - Lisbon	Martínez et al. (2017)	38,9%	23%	-	7,8%	26,3%	4%

Estimation and modelling results

Also, the estimated parameters of the different models can provide useful information about the underlying behavioural preferences of these modes. The direct comparison of model parameters is not possible as it is not known if the error term distributions are the same. Therefore, only the ratio of model parameters can be assessed.

One of the most important parameter ratio’s is the Value of Time (VOT). The values for VOT are predetermined in the SRE model by CBS (2015). The VOT of the external models can be determined by calculating the ratio of $\beta_{\text{in-vehicle time}} / \beta_{\text{costs}}$, where the parameter for in-vehicle time is multiplied with 60 to correspond to a VOT with specified dimensions in Euro / Hour. For the Nested Logit model of Lisbon these parameters need to be multiplied with the corresponding scale factor.

Table 18: overview on the value of time (VOT) for the different transport demand models

VOT (euro/hour)	Private Car EHV-VHV	Private Car Chicago	Private Car Lisbon	Public Transport EHV-VHV	Public Transport Chicago	Public Transport Lisbon	Bicycle EHV-VHV	DRT Chicago	Shared Car Lisbon
Scale β_{costs}		1	1		1	1,951		1	1,951
β_{costs}		-0,18468	-0,256		-0,18468	-0,498		-0,18468	-0,229
Scale $\beta_{\text{in-vehicle time}}$		1	1		1	1,951		1	1,951
$\beta_{\text{in-vehicle time}}$		-0,068	-0,026		-0,064	-0,024		-0,063	-0,026
VOT	6,282	22,09227	6,09375	5,616	20,79272	2,891566	6,282	20,46784	6,812227

Table 18 provides the VOT for the different models. What stands out from this table is that the VOT values in the Chicago model are much higher than for the other modes. This can be explained as its respective research was focussed on the purpose “commuter”. Also, the VOT for public transport in Lisbon is also considered to be different in comparison to the SRE model. This could be explained by the different compositions of the (sampled) populations. These different VOT values strongly suggest that the underlying behavioural considerations are not the same. This will also have implications for the methodology application. These implications are further considered in section 4.2.2.

Table 19 considers the modelling outcomes of the three models. From this table it can be considered that, for the sampled population, DRT is a significant alternative in comparison to private car and public transport within Chicago. Both private car and public transport lose market share, whilst this impact is relatively bigger for public transport. For the sampled population in Lisbon, it can be seen that the uptake of shared cars is limited. All other modes in Lisbon lose some of its market share in favour of shared cars.

Table 19: comparison of the modal splits as a modelling outcome of the three models

Modal split of the sampled population	Car	Bus	Bicycle	Metro	Walk	Other	DRT	Shared Car
Eindhoven-Veldhoven	53,7%	1,2%	45,7%	-	-	-	-	-
DRT - Chicago	28% (33,7%)	55% (66,3%)	-	-	-	-	17%	-
Shared car - Lisbon	38,3%	22,5%	-	7,5%	25,5%	3,8%	-	2,4%

Application contexts

Also, the actual populations of the three respective cities are considered. Table 20 and Table 21 provide socio-demographic information about the respective populations. What stands out from Table 20 is that the average household size in Lisbon is much bigger in comparison to Eindhoven. No data was found on the average household size in Chicago, but information from its population sample provide us a first indication that households in Chicago are also larger. Also, the levels of income are different for Lisbon in respect to Eindhoven. Table 21 point out that relatively many young people (age 18-24) live in Eindhoven, whilst the percentage of elder people (age 65%) is relatively low for Chicago.

Table 20: comparison of the socio-demographic information of the three city's populations

Information	Eindhoven	Chicago	Lisbon
Percentage male	51	49	48%
Average household size	1,9	NA	2,62
Gross monthly income	2425 euro	NA	2289 euro

Table 21: comparison of the age distributions of the three city's populations

Age distribution	Eindhoven	Chicago	Lisbon
18-24	16%	11%	9%
25-44	35%	45%	37%
45-64	29%	30%	32%
65+	20%	14%	22%

Table 22 provides information about the actual modal splits for the complete populations of the three cities. From this table it becomes clear that the modal split for private car in Chicago is larger than for the other cities. This could be explained by the considered purpose of commuting. Large differences can be found when Eindhoven is compared to the other cities for biking and public transport. Cycling is an important mode of transportation in the Netherlands and this is reflected in the modal split for Eindhoven.

Table 22: comparison of the modal splits of the complete population of the three cities

Modal split of the complete population	Source	Car	Bus	Bicycle	Metro	Walk	Other
Eindhoven (without Veldhoven)	Gemeente Eindhoven (2015)	42,0%	5,0%	40,0%	-	13,0%	-
DRT – Chicago (purpose: commute)	U.S. Census Bureau (2015)	52,7%	27,9%			6,8%	12,6%
Shared car – Lisbon	Instituto Nacional de Estatística (2012)	43,7%	21,4%	-	13,8%	19,4%	1,7%

4.2 Translation of a gravity model into non-linear utility functions

The deterrence functions are specified for mode *m* (car, public transport, bicycle) and user group *p* (car available or car unavailable) estimate transport demand based on the generalised costs c_{ijm} for all combinations of origins *i* and destinations *j* in the SRE study area.

The parameters of the log-normal deterrence functions for purpose “other” for all available modes in the SRE model and both population groups are stated in Table 23. In Figure 13, Figure 14 and Figure 15 figure the deterrence functions are visualised in a graph.

Table 23: set of estimated parameters for the purpose "other" in the SRE model

Mode *m* and group *p*	α_{mp}	β_{mp}
Car, CA	20241	-1,30
Public Transport, CA	1048	-0,82
Bicycle, CA	27912	-1,19
Car, CU	1897	-1,60
Public Transport, CU	983	-0,89
Bicycle, CU	8835	-1,42

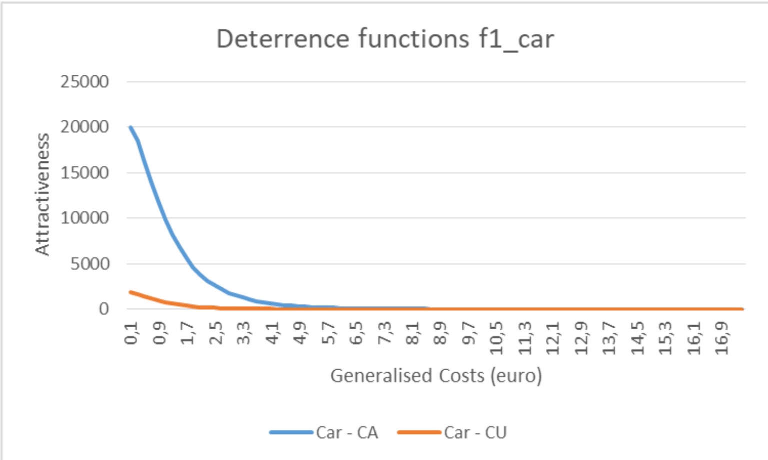


Figure 13: visualisation of the deterrence functions for car in the SRE-model

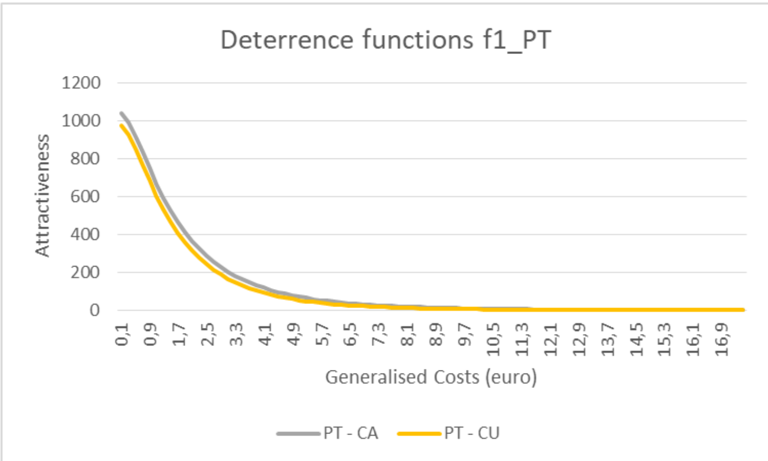


Figure 14: visualisation of the deterrence functions for public transport in the SRE-model

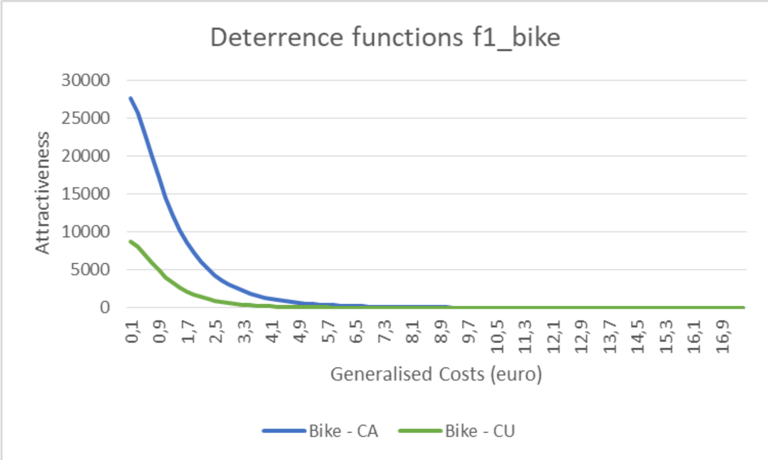


Figure 15: visualisation of the deterrence functions for bicycle in the SRE-model

From these figures it becomes clear that the attractiveness of the private car and bicycle alternative are significantly different for both population groups. This makes sense, as private cars are less available to the CU population group. Apparently, many people from this user group will substitute car transportation by biking, unless public transportation is considered to be more attractive.

The translation of these deterrence functions provide the utility functions f_2. as provided in Equation 20. As the utility functions are translated from log-normal deterrence functions, the utility functions are not specified to be linear for changes of the generalised costs c_{ijm}.

Equation 20: set of utility functions f2 in the SRE model for purpose other

$$f_2 = \begin{cases} f_{2,car,ca} = 9,92 - 1,3\ ln(c_{ijm}+1)^2 \\ f_{2,pt,ca} = 6,95 - 0,82\ ln(c_{ijm}+1)^2 \\ f_{2,bicycle,ca} = 10,24 - 1,19\ ln(c_{ijm}+1)^2 \\ f_{2,car,cu} = 7,55 - 1,6\ ln(c_{ijm}+1)^2 \\ f_{2,pt,cu} = 6,89 - 0,89\ ln(c_{ijm}+1)^2 \\ f_{2,bicycle,cu} = 9,09 - 1,42\ ln(c_{ijm}+1)^2 \end{cases}$$

4.3 Linearization of the non-linear utility functions

For the linearization the set of c_{ijm} for all origins *i* and destinations *j* in the study area are considered, as the linearized utility functions only have to forecast the modal split within the study area. Figure 16, Figure 17 and Figure 18 reflect the distribution of the generalised costs for private car, public transport and bicycle respectively.

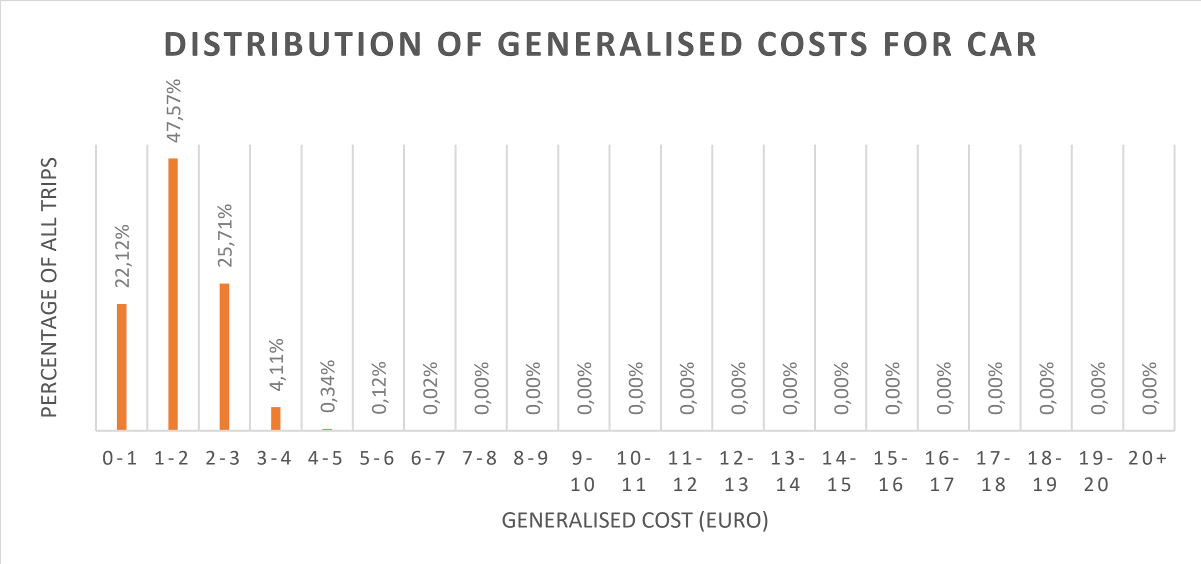


Figure 16: distribution of the generalised costs within the study area for car

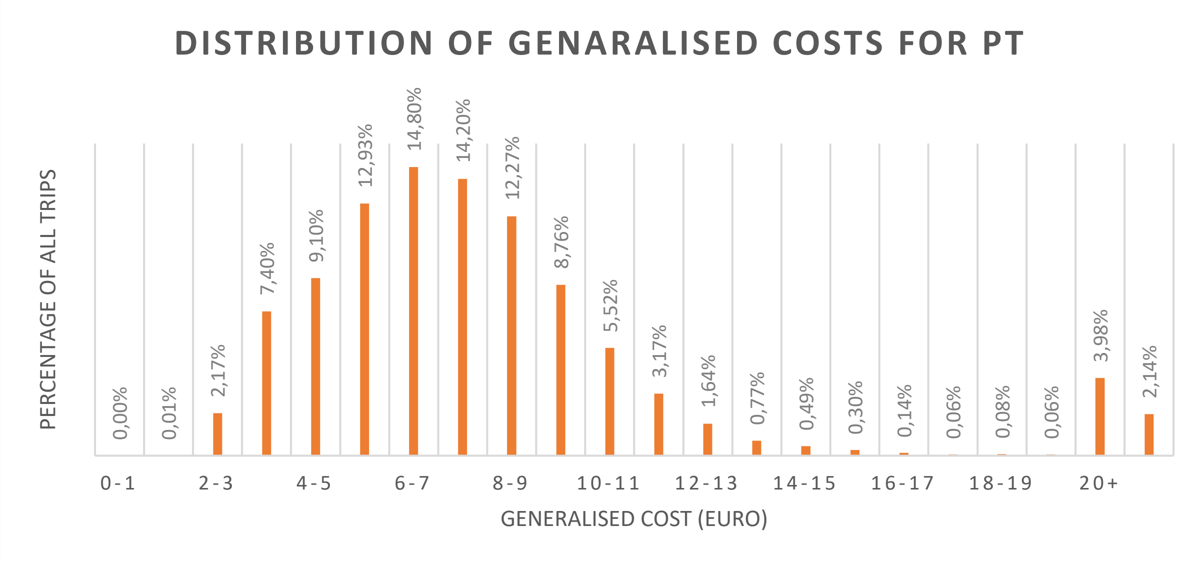


Figure 17: distribution of the generalised costs within the study area for public transport

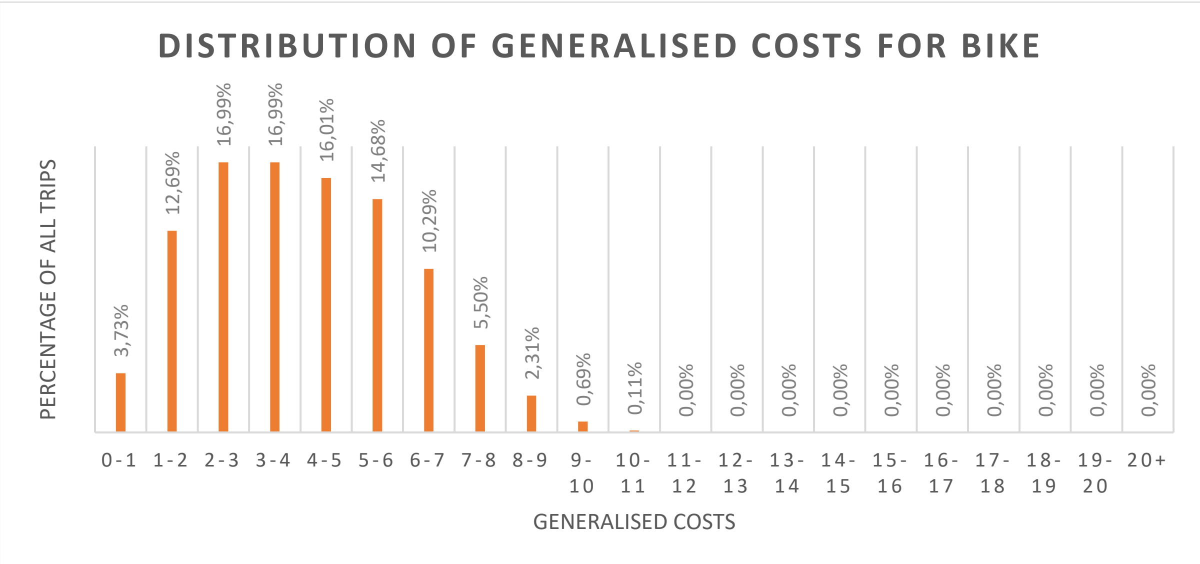


Figure 18: distribution of the generalised costs within the study area for bicycle

According these distributions, no private car alternatives are present in the study area with generalized costs values over 6,80 euro. For bicycle the maximum value for the generalised costs amounts 11,60 euro. For public transport however, high values (>20 euro) are determined from the SRE model. When specific zones are too far from the nearest public transport stop, the related alternatives are considered to have an infinite generalised cost value. This relates to more than 2% of the available public transport alternatives within the study area. Therefore, the impedance values for public transport for these zones with a low accessibility by public transport are not considered for this linearization step. This is legitimate, as for these OD-pairs public transport is not perceived to be a serious alternative by the general population. For this case study, the maximum generalised cost value is considered to be 17,60 euro.

Minimising the squared errors between $f_{2,mp}(c_{ijm})$ and $f_{2*,mp}(c_{ijm})$ resulted in the linearized utility functions $f_{2*,mp}(c_{ijm})$ as reflected in Equation 21. Appendix III provides the detailed results of the Ordinary Least Squares method and the approximation exercise.

Equation 21: set of linearized utility functions f2 in the SRE model for purpose other*

$$f_{2*} = \begin{cases} f_{2*,car,ca} = 10,19 - 0,92\ ln(c_{ijm})^2 \\ f_{2*,pt,ca} = 6,95 - 0,42\ ln(c_{ijm})^2 \\ f_{2*,bicycle,ca} = 10,24 - 0,74\ ln(c_{ijm})^2 \\ f_{2*,car,cu} = 7,55 - 1,13\ ln(c_{ijm})^2 \\ f_{2*,pt,cu} = 6,89 - 0,45\ ln(c_{ijm})^2 \\ f_{2*,bicycle,cu} = 9,09 - 0,89\ ln(c_{ijm})^2 \end{cases}$$

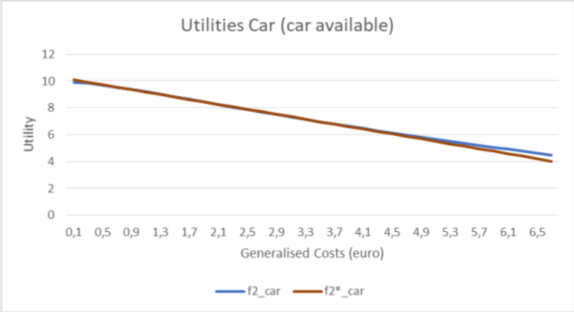
The functions $f_{2,mp}(c_{ijm})$ and $f_{2*,mp}(c_{ijm})$ are compared in Figure 19-RFigure 24, per alternative (mode) and user group. For each of the utility functions $f_{2*,mp}(c_{ijm})$ a bias is visible in relation to $f_{2,mp}(c_{ijm})$. From these figures it seems plausible that the generalised cost range determines the extent of this bias. For public transport the largest bias can be seen. The biases are smaller for bicycle and private car.

In general, the linearization bias is expected to have a significant impact on the modal share only for specific transport alternatives. This was reflected in two qualitative hypotheses:

- There is a relevant deviation of the modal split for OD-pairs characterised by modes with generalised cost value in the lower range.
- There is a relevant deviation of the modal split for OD-pairs characterised split by a significant market share for one of the modes available.

In order to reflect on these hypotheses a general overview is given on the impact of the linearization bias on the modal split. Table 24 and Table 26 provide information about the number of OD-pairs for which a relevant deviation of the model split between the two utility functions is determined. The modal split for each origin-destination combination shows a relevant bias when the absolute deviation of the modal split of at least one alternative is larger than 5 percent point.

The beforementioned tables with the total number of OD-pairs with a relevant deviation in the modal split cannot be interpreted directly. These totals and their prevalence within the study area are highly dependent on the size of the study area itself. Because there is interest in the relation with short distances, Table 25 and Table 27 display the OD-pairs with a deviation and a short distance as a percentage of all available OD-pairs with a deviation.



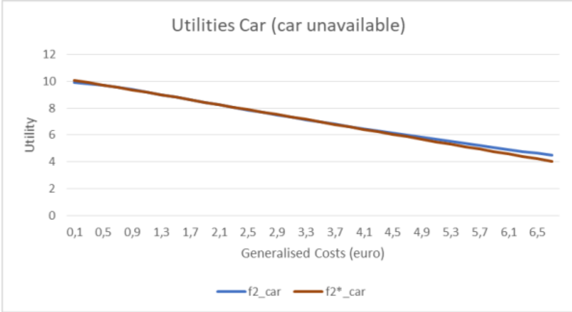
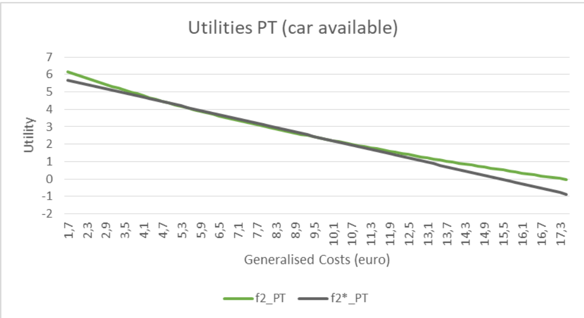


Figure 19 - left: visualisation of the translated and linearized utility functions for car (CA)

Figure 20 - right: visualisation of the translated and linearized utility functions for car (CU)



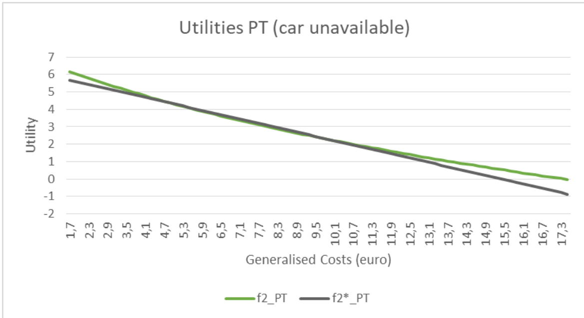
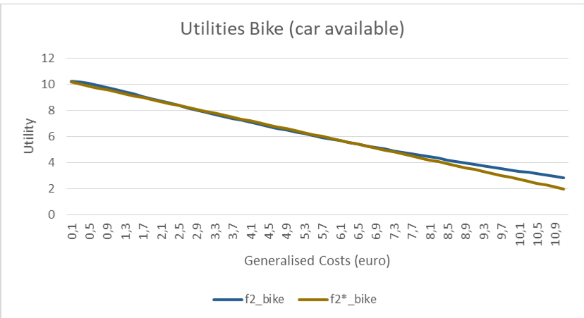


Figure 21 - left: visualisation of the translated and linearized utility functions for public transport (CA)

Figure 22 - right: visualisation of the translated and linearized utility functions for public transport (CU)



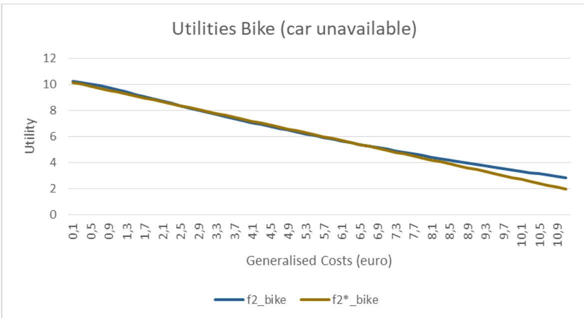


Figure 23 - left: visualisation of the translated and linearized utility functions for bicycle (CA)

RFigure 24 - right: visualisation of the translated and linearized utility functions for bicycle (CU)

From these tables it becomes clear that most of the OD pairs which show a relevant deviation of the modal split are situated close to each other. This is a strong indication that a linearization bias occurs for distances shorter than 2 kilometres. From that perspective, at least the first hypothesis seems to be valid.

Table 24: OD-pairs with a significant deviation of the modal split for CA

Significant deviation of the modal split for Car Available			
Origin-Destination	Cells with deviation	Total number of cells	difference
EHV-EHV	20922	445556	4.6%
EHV-VHV	463	162992	0,2%
VHV-VHV	9400	59292	15,9%
VHV-EHV	451	162992	0,3%

Table 25: OD-pairs with a short distance and a significant deviation of the modal split for CU

Percentage of short distances for significant deviations for Car Available			
Origin-Destination	Cells with deviation	Total number of cells with a short distance	%
EHV-EHV	19519	20922	93,3%
EHV-VHV	346	463	74,8%
VHV-VHV	9039	9400	96,1%
VHV-EHV	329	451	73,0%

Table 26: OD-pairs with a significant deviation of the modal split for CU

Significant deviation of the modal split for Car Unavailable			
Origin-Destination	Cells with deviation	Total number of cells	%
EHV-EHV	20578	445556	4,6%
EHV-VHV	363	162992	0,2%
VHV-VHV	9640	59292	16,3%
VHV-EHV	373	162992	0,2%

Table 27: OD-pairs with a short distance and a significant deviation of the modal split for CU

Percentage of short distances for significant deviations for Car Unavailable			
Origin-Destination	Cells with deviation	Total number of cells with a short distance	%
EHV-EHV	15960	20578	77,6%
EHV-VHV	198	363	54,6%
VHV-VHV	7508	9640	77,9%
VHV-EHV	187	373	50,1%

A closer watch is given to four OD-pairs situated along a short distance from each other. The four combinations relate to the references zones Eindhoven Burghplan en Veldhoven Cobbeek and one of their surrounding zones. Information about the reference zones is provided in Appendix V.

Table 28 provides information about the relevant deviations in the modal splits for these OD-pairs indicated with a value of 1 for a relevant deviation and a value of 0 if there is no relevant deviation.

From this table is becomes clear that also these modal splits show relevant deviations. Except for the OD-pairs of Veldhoven Cobbeek with Veldhoven South-West and vice versa for the car unavailable user group. A possible explanation for this could be that private car is a less important alternative for the car unavailable user group in comparison to the car available user group. This means that a bias in the linearized utility function for the car unavailable user group will have a smaller impact on the deviation of the modal split.

Considering the abovementioned, it is likely that the linearization of non-linear utility functions causes significant deviations in the modal split for OD-pairs along short distances. This means that users of the linearization method should keep in mind that modelling outcomes can be affected for these OD-pairs.

In relation to the practical validity this does not seem to be an issue for the SRE model. Similar to the SRE gravity model, the approximation f2* considers large ranges of generalised costs for each mode due to the specification of deterrence functions for nation-wide transport. This implicates that for extreme values – both high and low – the deterrence functions are likely to misrepresent the transport demand. From this perspective, the misrepresentation of the modal split – for reasons of generic deterrence function – was already in the SRE model, such that an additional linearization bias does not make the practical validity for OD-pairs along short distances any worse.

In order to evaluate the second hypothesis, four OD-pairs with significant market shares for public transport are considered: trip combinations between the train station and the airport and the train station and the industrial area of ASML in the south of Veldhoven.

Table 28 provides information about the relevant deviations in the modal splits for these OD-pairs indicated with a value of 1 for a relevant deviation and a value of 0 if there is no relevant deviation.

Table 28: overview of the significant deviations in the modal split for the references zones

Origin	Destination	Typology	Car Available	Car Unavailable
Station	Airport	Good PT	1	0
Airport	Station	Good PT	1	0
Station	ASML	Good PT	0	0
ASML	Station	Good PT	0	0
Eindhoven Burghplan	Eindhoven South	Short distance (1,63km)	1	1
Eindhoven South	Eindhoven Burghplan	Short distance (1,63km)	1	1
Veldhoven Cobbeek	Veldhoven South-West	Short distance (1,96km)	0	1
VHV South-West	Veldhoven Cobbeek	Short distance (1,96km)	0	1

From Table 28 it becomes clear that there are only significant deviations in the modal split for the car available user group for the zonal combination “station – airport” and vice versa. This is not as expected. Therefore, a closer look is given on the modal splits of both $f_{2,mp}(c_{ijm})$ and $f_{2*,mp}(c_{ijm})$.

According Table 29 and Table 30 it becomes clear that the bias of the linearized utility function of public transport does not affect the market share of public transport significantly. For the specific case of “station – airport” and vice versa, the demand for car transportation has risen significantly at cost of the demand for bicycle transportation. It is likely that this is not the result of a bias of the linearized utility function for public transport, but the result of a bias of the linearized utility function for bicycle.

Resuming the beforementioned, there are currently no indications of a linearization bias which leads to relevant modal split deviations for OD-pairs with a significant market share for public transport. However, relevant modal split deviations for competitive public transport connections are still likely to occur. In future research, more effort should be given to assess the presence and extent of the impact of the linearization bias.

Table 29: overview of the modal splits from f2 and f2 and its deviations for car unavailable*

Car Available		Current modal split			Approximated modal split			Deviation of approximation		
Origin	Destination	Car	PT	Bicycle	Car	PT	Bicycle	Car	PT	Bicycle
Station	Airport	15,3%	13,1%	71,7%	21,5%	12,6%	65,8%	6,3%	-0,5%	-5,8%
Station	ASML	71,8%	4,9%	23,4%	72,5%	5,3%	22,2%	0,7%	0,5%	-1,2%
Eindhoven South-East	Eindhoven South	58,8%	0,0%	41,2%	52,9%	0,0%	47,1%	-5,8%	0,0%	5,8%
Veldhoven West	Veldhoven South-West	59,6%	0,3%	40,1%	54,7%	0,3%	45,0%	-4,9%	0,0%	4,9%
Airport	Station	15,5%	11,9%	72,6%	21,9%	11,3%	66,8%	6,4%	-0,6%	-5,8%
ASML	Station	72,5%	3,9%	23,6%	73,4%	4,1%	22,5%	0,9%	0,2%	-1,1%
Eindhoven South	Eindhoven South-West	58,8%	0,0%	41,2%	52,9%	0,0%	47,1%	-5,8%	0,0%	5,8%
VHV South-West	Veldhoven West	59,4%	0,5%	40,0%	54,5%	0,6%	44,8%	-4,9%	0,1%	4,8%

Table 30: overview of the modal splits for f2 and f2 and its deviations for CU*

Car Unavailable		Current modal split			Approximated modal split			Deviation of approximation		
Origin	Destination	Car	PT	Bicycle	Car	PT	Bicycle	Car	PT	Bicycle
Station	Airport	2,0%	48,6%	49,4%	3,2%	49,4%	47,4%	1,2%	0,8%	-2,1%
Station	ASML	34,1%	34,7%	31,3%	33,7%	37,5%	28,8%	-0,4%	2,9%	-2,5%
Eindhoven South-East	Eindhoven South	32,4%	0,0%	67,6%	26,5%	0,0%	73,5%	-5,9%	0,0%	5,9%
Veldhoven West	Veldhoven South-West	33,0%	1,2%	65,8%	28,0%	1,1%	70,9%	-5,0%	-0,1%	5,1%
Airport	Station	2,1%	45,7%	52,2%	3,5%	46,0%	50,5%	1,3%	0,4%	-1,7%
ASML	Station	36,8%	29,4%	33,8%	37,3%	30,9%	31,8%	0,5%	1,5%	-2,0%
Eindhoven South	Eindhoven South-West	32,4%	0,0%	67,6%	26,5%	0,0%	73,5%	-5,9%	0,0%	5,9%
VHV South-West	Veldhoven West	32,6%	2,6%	64,9%	27,5%	2,8%	69,7%	-5,0%	0,2%	4,8%

4.4 Determination of factors to correct the scale of external models

The next step corrects the scale factors of the external models such that these approximate the model scale of the linearized translated utility functions of the SRE model.

Regarding communal modes, the Chicago model represents the modes private car and public transport, whilst The Lisbon solely represents the mode public transport. This means that the conduct of this methodological step is different for the two models.

By means of the Ordinary Least Squares (OLS) method the differences *of the (systemic) utility differences between all combinations of communal modes* are minimised for both the CA and CU user group model specifications of the Chicago model. The same method is used to minimise the (systemic) utility differences between the utility functions of public transport for $f_{2*,m}$ and $f_{3*,m}$. For the latter minimisation procedure, the correction factor θ_{ASC} is fixed with a value of 1.

Table 31-Table 34 provide the correction factors for both external models and for both the CA and CU user groups. Appendix IV provides the parameters of the corrected models f_{3*} in comparison with the parameters of the original external models f_3 and the parameters of the linearized translated SRE model f_{2*}.

Table 31: correction factors of f3 for the Chicago model and the CA group*

Correction factor	Value
θ_{ASC}	4,99
$\theta_{\beta x}$	4,23

Table 32: correction factors of f3 for the Chicago model and the CU group*

Correction factor	Value
θ_{ASC}	2,03
$\theta_{\beta x}$	4,25

Table 33: correction factors of f3 for the Lisbon model and the CA group*

Correction factor	Value
θ_{ASC}	1
$\theta_{\beta x}$	4,56

Table 34: correction factors of f3 for the Lisbon model and the CU group*

Correction factor	Value
θ_{ASC}	1
$\theta_{\beta x}$	4,95

The sizes of the correction factors θ_{ASC} and $\theta_{\beta X}$ need to be interpret as a measure for the impact size of the transfer bias. This means that for all models the systemic utility sensitivities for marginal changes of trip attributes are almost 5 times larger than the current sensitivities. This provides a strong indication that the underlying behavioural considerations are not the same.

4.5 Specification of the combined model

Table 35 provides the shift values for the utility functions from the different external models. The parameters of the resulting combined models for both the car available and car unavailable user group are provided in Table 36 and Table 37.

Table 35: shift factor of f3 for the Chicago model and the CA group*

Shift factor	Value
Chicago – CA	10,3
Chicago – CU	7,3
Lisbon – CA	6,6
Lisbon - CU	5,9

Table 36: specification of the new model parameters for car available (CA)

Parameter for Car Available	Unit of measurement of explanatory variable	Private Car	Public Transport	Bicycle	DRT	Shared Car
ASC		10,19	6,39	10,23	1,26	5,11
$\beta_{generalised\ costs}$	Euro	-0,92	-0,42	-0,74	0	0
$\beta_{access\ time}$	Minute				0	-0,37
$\beta_{in\text{-}vehicle\ time}$	Minute				-0,27	-0,12
$\beta_{waiting\ time}$	Minute				-0,45	0
$B_{transfer}$	# transfer				-0,94	0
$\beta_{egress\ time}$	Minute				0	-0,37
β_{costs}	Euro				-0,78	-1,04

Table 37: specification of the new model parameters for car unavailable (CU)

Parameter for Car Unavailable	Unit of measurement of explanatory variable	Private Car	Public Transport	Bicycle	DRT	Shared Car
ASC		7,88	6,28	9,08	4,21	4,75
$\beta_{generalised\ costs}$	Euro	-1,13	-0,45	-0,89	0	0
$\beta_{access\ time}$	Minute				0	-0,41
$\beta_{in\text{-}vehicle\ time}$	Minute				-0,27	-0,13
$\beta_{waiting\ time}$	Minute				-0,45	0
$B_{transfer}$	# transfer				-0,94	0
$\beta_{egress\ time}$	Minute				0	-0,41
β_{costs}	Euro				-0,78	-1,13

4.6 Modal splits of the combined model

This section provides the modal split outcome from the resulting combined model and reflects on these modelling outcomes in relation to the contexts of the case study models. The final results are presented starting with the aggregate modal split result followed by the disaggregate modal split results.

Table 38: modal split outcomes of the current SRE model and the new model

	Current SRE (CA)	Current SRE (CU)	SRE with future modes (CA)	SRE with future modes (CU)
Car	61,0%	38,0%	60,7%	37,0%
Public Transport	0,005%	2,7%	0,5%	2,6%
Bicycle	39,5%	59,2%	38,4%	58,3%
DRT	NA	NA	0,0%	0,0%
Shared Car	NA	NA	0,4%	2,1%

Aggregate modal split

Table 49 provides the aggregated modal split in comparison to the modal split of the current SRE model. This table shows that private cars and bicycles are slightly less used according to the combined model in favour of the new modes. For public transport an increase can be seen for the user group car available. This increase is not visible for the car unavailable user group. An increase of the modal split for public transport is not the intended result of the methodology. As it is assumed that the utility function for public transport remains the same, the utility for public transport alternatives should not increase. It is plausible that the linearization bias causes this slight increase.

The aggregated modal splits show that shared car is able to gain some market share. On the aggregated level, DRT does not show any potential market share. Two possible reasons for the low modal shares of DRT are identified:

firstly, the systemic utility differences between the communal alternatives private car and public transport within the Chicago model are small in comparison to the same differences within the SRE model. Especially for the car available user group this is the case in the SRE model, as the alternative-specific constants (ASC) for private and public transport are very different. Therefore, the utility differences are amplified by means of the correction factor θ_{ASC} such that also the systemic utility differences in relation to shared cars become larger for the Chicago model. As for this alternative the estimated ASC was already the lowest in comparison to the other alternatives, the relative systemic utility for shared cars got even lower.

Secondly, the methodology corrects for different values of times (VOT) between the SRE model and the Chicago model This results in extremely sensitive parameters for the shared car alternative. In the combined model, the value of β_{costs} for DRT and the CA user group is -almost twice as large as the value of $\beta_{generalised\ costs}$, for public transport and the CA user group. This means that the DRT alternative is twice as sensitive for changes in costs in comparison to changes in the (generalised) costs of public transport. At first glance, it appears to be highly unlikely that this is similar to reality.

The explanation of the differences in the systemic utility differences and the differences in the VOT between the SRE and the Chicago model lies in the different behavioural considerations for the two models. Especially in relation to the car available user group. In the SRE model, the CA user group has a clear preference for private car transport. This clear preference was not reflected by the model of Frei et al. (2017). This means that it is required to properly assess the different model backgrounds before combining them. When the backgrounds of these models– and thus the underlying behavioural considerations – are different, the method will (unintentionally) correct for this. This does not mean that the methodology itself is not valid, as it was clearly stated that the methodology assumes that the underlying behavioural consideration are the same. It means that the practical validity is highly dependent on the choice of the external models.

Table 39 compares the modal splits of the combined model with the modal splits of the external models. Considering the modal splits for the shared car alternative, it looks like its modal split values are valid. They have a similar magnitude as in the Lisbon model. For DRT there is a big difference as there is no market share for this alternative in the combined model. As explained before, it is highly likely that this is caused by the different VOTs between the SRE and the Chicago model. For that reason, it is suggested to perform the methodology by changing the parameters for in-vehicle time and costs in the Chicago model such that the VOT corresponds to the SRE model.

Table 39: modal split outcomes of the external models and the new model

	Frei et al. (2017)	Martínez et al. (2017)	SRE (CA)	SRE (CU)
Frei et al. (2017)	28,0%	38,3%	60,7%	37,0%
Public Transport	55,0%	22,5%	0,5%	2,6%
Bicycle	NA	NA	38,4%	58,3%
DRT	17,0%	NA	0,0%	0,0%
Shared Car	NA	2,4%	0,4%	2,1%
Other	NA	36,8%	NA	NA

Disaggregate modal split

A closer look has been given on the distribution of the modal splits over the complete study area. For this reason, a scatterplot has been constructed, as reflected in Figure 25 and Figure 26. The scatterplots are shown for the car unavailable user group only, as the possible explanations for the modal split distribution are considered to be the same. The colours in these scatterplots indicate the modal split for the respective alternative for each combination of origin and destination.

Modal share values lower than 1% are indicated with blue.
Modal share values between 1% and 10% are gradually indicated by orange (1%) till yellow (10%)
Modal share values over 10% are indicated with green, for which the greens gets darker for the following distribution classes (10-15%; 15-20%; 20-25% and >25%)

From Figure 25 it becomes clear that DRT does not show any relevant modal share, independent of the zonal combination.

Figure 26 shows a complete different modal split distribution. From this figure it can be seen that the modal splits for shared cars along the diagonal of the study area are very low. This makes sense, as these respective origins are generally situated in the proximity of its destinations, making it more attractive to use other alternatives such as bicycles. This figure visualizes aggregated and disaggregated trends visible. On the far right side of the figure – and on the down side of the figure – the modal splits are remarkably higher than at other locations. Many of these modal splits exceed the level of 10%. This is caused by the specific origins and destinations for these blocks. The lower rows and the right columns represent the zones in Veldhoven. From this trend, it can be considered that when distances get longer the shared car alternative is getting more market share.

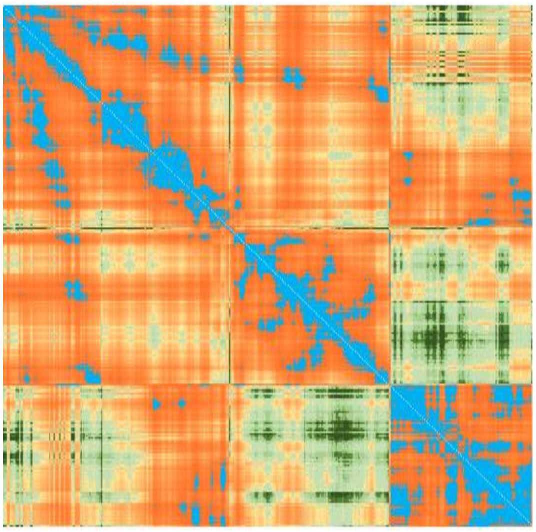


Figure 25: left - overview on the modal split of DRT for car unavailable at cell level

Figure 26: right - overview on the modal split of shared car for car unavailable at cell level

Figure 26 also shows some disaggregated trends, reflected in the horizontal and vertical lines. From this figure only, the specific cause for these trends cannot be derived. A possible explanation could be the difference in the structural (spatial-temporal) accessibility of the respective zones by other alternatives. One of the most outstanding trend lines is situated in the middle with many modal split values for shared car over 10%. This line represent the Eindhoven Airport related zones. The modal split values for these zones make sense, as in that area additional parking fees for private cars apply. It is likely that for that reason shared cars are an attractive alternative as the parking fees do not apply for shared cars. The diagonal OD-pairs indicated by blue are characterised by short distances, such that the shared car alternative is not attractive for many people.

More disaggregate information is available. This is information is provided in Appendix V.

4.7 Conclusions

The combined model for the case study for "Urban Mobility as a Service" for the Eindhoven-Veldhoven provided limited modal share values for the new modes.

For demand responsive transportation, no transport demand is forecasted. For shared, limited modal share values are forecasted. The transport demand for shared cars is higher for OD-pairs which are less attractive for transport by private cars (mainly due to parking costs) and transport by public transportation (mainly due to the absence of public transport stops close by).

The most striving finding is that the choice of external models is really important for valid results of the combined model. There are large underlying behavioural preferences between the models considered in the case study. This causes a significant transfer bias for the utility modes for the new modes within the combined model.

5 Conclusions, discussion and recommendations

This chapter provides the conclusions, discussion and recommendations of this thesis. Section 5.1 presents the conclusions by answering all the research questions. Thereafter, section 5.2 provides a discussion of these conclusions and the research process. In section 5.3, recommendations are given with regards to the methodology and the application of the methodology in practice.

5.1 Conclusions

The research objective was to design a methodology for adding external discrete choice model parameters to an existing transport demand model. In addition, this research aimed to provide first indications about the validity of this methodology in practice for forecasting the modal split when new modes are present.

This research objective relates to a design problem consisting of three sub-problems:

- The existing transport demand gravity model is specified within a different modelling framework compared to the external discrete choice models.
- The communal modes of both the existing transport demand gravity model and the external discrete choice models are specified along different underlying behavioural considerations, such that the different models relate differently to the same communal modes.
- The combination of transport demand information from different sources might affect the practical validity for the modelling purpose.

In line with these sub-problems, the three sub-research questions are answered. Thereafter, the main research question is answered.

5.1.1 Different modelling frameworks

The first sub-research question was defined as follows: *How can the behavioural specifications of existing transport demand gravity models be expressed within the discrete choice modelling framework?*

The behavioural specifications of the existing transport demand gravity model can be translated to disaggregate discrete choice utility functions. The methodology of Bliemer (2010a); (2010b) is able to translation aggregate gravity models towards disaggregate discrete choice models. For gravity models described by exponential deterrence functions this methodology provides linear utility functions. For log-normal and top-lognormal deterrence functions this methodology provides non-linear utility functions.

With respect to the second sub-problem, non-linear translated utility functions are not considered to be directly useable for the addition of utility functions of external models. Only linear utility function specifications were considered to be directly useable as the (systemic) utility sensitivities for marginal changes of trip attributes are equal for any origin-destination pair (OD-pair), similar to the specification of the utility functions of external models. For that reason, linear approximations of the non-linear translated utility functions are used for the addition of utility functions of external models.

Translation of deterrence functions towards discrete choice utility functions

A methodology for the translation of aggregate gravity models towards disaggregate discrete choice models is available. (Bliemer, 2010a, 2010b) It was found that the translation of aggregate gravity models towards disaggregate discrete choice models is an easier procedure than the translation the other way around. In that case, the explanatory variables of the disaggregate form need to be aggregated towards a single explanatory variable before the translation procedure can be executed.

Deterrence functions from the gravity model, specified by exponential, log-normal and top-lognormal functions, can be directly translated to disaggregate utility functions. Different from exponential deterrence functions, log-normal and top-lognormal deterrence functions are translated by non-linear utility functions. This means that the utility sensitivities for marginal changes of trip attributes are not equal for any origin-destination pair (OD-pair) for the latter types of deterrence functions.

Linearization of non-linear translated utility functions

The Ordinary Least Squares method was found to be a straightforward method to approximate the translated disaggregate utility functions by minimising the squared utility differences with the linear approximation. The method considers all squared utility differences related to all present values of c_{ijm} within the area of scope.

A linear utility function specification is needed as it means that the utility sensitivities for marginal changes of trip attributes are equal for any origin-destination pair (OD-pair), similar to the external discrete choice models which will be added to the existing model. Unequal sensitivities are a indication that underlying behavioural considerations are not the same. Because for this methodology the underlying behavioural considerations are considered to be the same, the impact of different behavioural considerations needs to be prevented when possible.

Linearization bias

Inherently, a non-linear function will not perfectly fit a linear function. This means that a so-called linearization bias occurs by the conduct of this methodological step. This implicates that the approximation of the translated utility functions provide structural deviations of the modal share outcomes with respect to the existing gravity model. This implies a trade-off between the linearization benefit for the combination with external models and the linearization bias.

It is perceived that non-linear utility functions and its linearized approximations should be inspected upfront in the order to assess the impact of the linearization bias. Plotting these functions in a graph provided more insights about for which transport alternatives significant deviations in the modal split would occur. This enables a fast assessment about the impact size of the linearization bias on the modal split.

In general, the linearization bias is expected to have a significant impact on the modal split only for specific OD-pairs. This was reflected in two qualitative hypotheses:

- There is a relevant deviation of the modal split for OD-pairs characterised by modes with generalised cost value in the lower range.
- There is a relevant deviation of the modal split for OD-pairs characterised split by a significant market share for one of the modes available.

The first hypothesis is expected to be true for reasons of large linearization bias for the outer ranges of the present ranges of the generalised costs values and significant competitiveness of the different modes for these OD-pairs. The second hypothesis is expected to be true as choice probabilities are especially sensitive for marginal (systemic) utility differences when the choice probabilities are high.

5.1.2 Different underlying behavioural considerations

The second sub-research question was defined as follows*: How can be dealt with the different underlying behavioural considerations for the communal modes in both existing transport demand models and external models within the discrete choice modelling framework in order to add new modes from external models?*

After the approximation of the translated utility functions, the utility functions from external discrete choice models need to be add to the existing translated transport demand model. This is only allowed when the scale factors of the different models are the same and the underlying behavioural considerations for each utility function are similar.

Both models have an implicit model scale of 1. By neglecting the differences in underlying behavioural considerations, the differences between the trade-offs between modes for both models are considered to be caused by different (fictional) model scales. By correcting the scale of the external model, such that the fictional model scales are the same, the underlying behavioural considerations can be equalised.

The utility functions of corrected external model can be added to the translated existing transport demand model, by adding equal additional systemic utility to the structural element of all external utility functions, such that for the communal modes the systemic utilities of the external model approximate the systemic utilities of the existing model.

Determination of factors to correct the external model scale

As the model scales of the considered models are assumed to be equal the sizes of the correction factors θ_{ASC} and $\theta_{\beta X}$ can be interpret as a measure for the impact size of the transfer bias. The structural systemic utility differences between modes and the systemic utility sensitivities towards marginal changes of trip attributes are enlarged by the multiplication of the alternative-specific constants and the summed product of the trip attributes x and the related parameters β_x with the respective correction factors.

Transfer bias

The correction of the external mode scale alters the sensitivities for marginal changes of the trip attributes for each mode and the structural systemic utility differences between the different modes for the external model. As the underlying behavioural considerations are considered to be the same, the different sensitivities are considered to be caused by a transfer bias.

5.1.3 Methodology application

The third sub-research question was defined as follows: *Which first indications can be given about the practical validity of the developed methodology for forecasting for future modal splits in the presence of new modes?*

Based on the application of the methodology on the case study for "Urban Mobility as a Service" for the Eindhoven-Veldhoven area the practical validity of the methodology for forecasting future modal splits in the presence of new models is considered to be very limited. The most striving finding is that the choice of the external models is really important for valid result of the combined model. There are large underlying behavioural preferences between the models considered in the case study. This causes a significant transfer bias for the utility modes for the new modes within the combined model. The Chicago model is very different to the SRE model, such that the combined model does not reflect any demand for Demand Responsive Transportation. At first glance, the Lisbon model shows realistic modal share values for shared cars. It is questioned if these values are realistic as the Lisbon model was only corrected for the systemic utility sensitivities for marginal changes in the trip attributes.

Aggregate and disaggregate modal splits

The addition of the utility function for DRT did not provide any market share. The modal share values for shared cars seem to be plausible at first sight. On a disaggregate level, the modal share of DRT responds significantly when other alternatives have a low utility, for example by means of high parking costs or poor public transport accessibility.

Assumption that underlying behavioural considerations are the same

The underlying behavioural considerations for all considered models are not the same. There are many indications that the behavioural considerations are different.

This relates to different travel purposes of the different transport demand models. The SRE model considers the purpose "other", whilst the Chicago model considers a commuting purpose. The Lisbon study is not even specified for a specific purpose. The respondents of the stated preference study faced a multitude of transport alternatives (modes), whilst the other models are estimated on the trade-offs between three modes. Considering the modal split of the related (sampled) populations of the different models, large difference in modal share values can be found. Bicycles are intensively used in Eindhoven-Veldhoven, whilst public transportation seems to be highly favourable for the Chicago model.

Apart from these model context characteristics, the parameter ratios of the considered models have been examined. One of the most important ratio's, the value of time (VOT) is significantly higher for the Chicago model. The VOT for public transport is in the Lisbon model is significantly lower in comparison to the VOT in the SRE model.

Transfer bias

It is assumed that the different underlying behavioural assumptions, especially those reflected in different VOTs had a large impact on the parameter set of the combined model. The parameters of the utility functions of this combined model are given in Table 40 and

Parameter for Car Available	Unit of measurement of explanatory variable	Private Car	Public Transport	Bicycle	DRT	Shared Car
ASC		10,19	6,39	10,23	1,26	5,11
βgeneralised costs	Euro	-0,92	-0,42	-0,74	0	0

βaccess time	Minute				0	-0,37
βin-vehicle time	Minute				-0,27	-0,12
βwaiting time	Minute				-0,45	0
Btransfer	# transfer				-0,94	0
βegress time	Minute				0	-0,37
βcosts	Euro				-0,78	-1,04

Table 41 respectively. It was found that the parameters for DRT are not valid for this case study. Its ASC values are really low in comparison to other alternatives and the attribute parameters appear to be really sensitive. For example, the β_{costs} for DRT is almost double the value of $\beta_{generalised\ costs}$ for public transport and the CA user group. This implies a significant impact of the transfer bias on the parameter values of the utility function for DRT.

The lack of validity of the Chicago model for the Eindhoven-Veldhoven case is reflected in the modelling outcomes of the combined model. The DRT alternative does not draw any demand. However, the shared cars alternative shows similar modal splits in relation to the Lisbon model. Also the disaggregate modal shares – for specific origins or destinations – showed plausible values and expected behaviour, for example with respect to zonal characteristics such as parking costs or low public transport accessibility levels. However this is likely to be caused by the choice of only a single communal mode for the Lisbon model. Therefore, the external model was not corrected to represent structural trade-offs between multiple transport alternatives.

Linearization bias

The case study provided indications that the hypotheses regarding the impact of the linearization bias are true. It was found that over all OD-pairs, around 75-95% (CA user group) and 50-70% (CU user group) of the OD-Pairs with a deviation (+/- 5 percent point) of the modal share for at least one mode consists of the OD-pairs was characterised by distances of less than 2 kilometre. No relevant deviations of the modal share were found for OD-pairs with a significant public transportation demand.

The case study shows that the transfer bias is far more important than the linearization bias for the practical validity of the combined model.

Table 40: specification of the new model parameters for car available (CA)

Parameter for Car Available	Unit of measurement of explanatory variable	Private Car	Public Transport	Bicycle	DRT	Shared Car
ASC		10,19	6,39	10,23	1,26	5,11
$\beta_{generalised\ costs}$	Euro	-0,92	-0,42	-0,74	0	0
$\beta_{access\ time}$	Minute				0	-0,37
$\beta_{in\text{-}vehicle\ time}$	Minute				-0,27	-0,12
$\beta_{waiting\ time}$	Minute				-0,45	0
$B_{transfer}$	# transfer				-0,94	0
$\beta_{egress\ time}$	Minute				0	-0,37
β_{costs}	Euro				-0,78	-1,04

Table 41: specification of the new model parameters for car unavailable (CU)

Parameter	Unit of measurement	Private Car	Public Transport	Bicycle	DRT	Shared Car

for Car Unavailable	of explanatory variable					
ASC		7,88	6,28	9,08	4,21	4,75
$\beta_{generalised\ costs}$	Euro	-1,13	-0,45	-0,89	0	0
$\beta_{access\ time}$	Minute				0	-0,41
$\beta_{in\text{-}vehicle\ time}$	Minute				-0,27	-0,13
$\beta_{waiting\ time}$	Minute				-0,45	0
$B_{transfer}$	# transfer				-0,94	0
$\beta_{egress\ time}$	Minute				0	-0,41
β_{costs}	Euro				-0,78	-1,13

5.1.4 Final conclusion

The main research question was defined as follows: *How can external discrete choice utility functions be added to existing transport demand gravity models and what are the first indications for the validity of this methodology for forecasting future modal splits in the presence of new modes?*

The behavioural specifications of the existing transport demand gravity model can be translated to disaggregate discrete choice utility functions. The methodology of Bliemer (2010a); (2010b) is able to translation aggregate gravity models towards disaggregate discrete choice models. The non-linear translated utility functions are not considered to be directly useable for the addition of utility functions of external models. The Ordinary Least Squares method was found to be a straightforward method to approximate the translated disaggregate utility functions by minimising the squared utility differences with the linear approximation.

For the addition of external utility functions to the translated existing transport demand model the differences in underlying behavioural considerations are neglected, such that the differences between the trade-offs between modes for both models are considered to be caused by different (fictional) model scales. By correcting the scale of the external model in order to equal the fictional model scales the underlying behavioural considerations can be equalised. Consecutively, the utility functions of corrected external models can be added to the translated existing transport demand model, by adding equal additional systemic utility to the structural element of all external utility functions, such that for the communal modes the systemic utilities within the external model approximate the systemic utilities within the existing model for the same modes. For both methodological steps the Ordinary Least Squares method has been used.

Based on the application of the methodology on the case study for "Urban Mobility as a Service" for the Eindhoven-Veldhoven area the practical validity of the methodology for forecasting future modal splits in the presence of new models is considered to be very limited. The choice of the external models was found to be really important for valid results of the combined model. Significant differences in the underlying behavioural assumptions can bias the modal split outcomes, such that they do not provide realistic transport demand.

5.2 Discussion

This section provides a discussion about the conclusions and the research process. Section 5.2.1 provides a discussion with regards to the methodology. Section 5.2.2 discusses the practical validity of the methodology.

5.2.1 Methodology

Linearization of non-linear translated utility functions

In chapter 3.2 it was argued that the translated non-linear utility functions need to be linearized for all present values of c_{ijm} such that for both models the systemic utility sensitivities for marginal changes of trip attributes are equal for all OD-pairs. This makes sense, as it is an expression of different underlying behavioural considerations.

The need for this linearization step can be discussed. When the linearization step is skipped, the potential linearization bias will be present in the combined model as a transfer bias.

Considering the application of the methodology within the case study, the linearization bias was considered to be small. The maximum absolute systemic utility differences between of $f_{2^*,m}(c_{ijm})$ and $f_{2,m}(c_{ijm})$ was less than one unit of utility, which can considered to be low given the ranges of the systemic utilities of $f_{2,m}(c_{ijm})$ for the considered area of scope. For this case, skipping the linearization step is considered to lead to only a slightly different correction of the external model scale towards the scale of the existing transport demand model. The impact on the modal split for the combined model is therefore considered to be limited.

However, skipping the linearization step would likely have an impact on the distribution of the bias over the different OD-pairs. The impact of the linearization bias was found to be mainly limited to OD-pairs characterised by lower range generalised cost values. When the linearization bias is exchanged for a transfer bias by a different correction of the external model scale, this will affect the behavioural specifications of the utility functions of the external model for all OD-pairs within the area of scope. It is thus argued that the linearization step does not lead to a better model, but it will only restrict the impact of the linearization/transfer bias for a limited set of OD-pairs.

The execution of the linearization step in itself is considered to be a useful exercise as it provides information about the extent the systemic utility of $f_{2^*,m}(c_{ijm})$ would deviate from $f_{2,m}(c_{ijm})$. Large systemic utility difference indicate that the systemic utility sensitivities for marginal changes of the trip attributes are not equal for all OD-pairs.

Also the appropriateness of the current linearization procedure can be discussed. There are different linearization paradigms imaginable which account for the presence of non-equal sensitivities for marginal changes of the trip attributes.

The first one is the linearization towards the present values of c_{ijm} weighted for the current transport demand of this transport alternative. This linearization would likely provide a better fit for OD-pairs which are important in terms of demand levels. A linearization procedure could also be conducted by allowing certain c_{ijm} values to change in order to provide a better linear approximation. This procedure is therefore able to distribute the impact of the bias to specific OD-pairs.

Correcting external model scales

Considering the application of the methodology within the case study, it was visible that significantly different values for comparable parameters between the current and extern models provide large correction factors, which imply large differences between the underlying behavioural considerations. The large correction factors lead to unrealistic parameter values for the corrected external utility functions. This problem might be mitigated by correcting only a subset of the complete trip-specific element of the utility function.

5.2.2 Practical validity of the methodology

The practical validity of the methodology seems to be very limited. However, based on the combination of only two external models no general statements can be given about the practical validity for different application purposes and contexts.

It seems to be challenging to identify multiple models for which the underlying behavioural considerations are the same. For that reason, it is not believed that this methodology will be used to a large extent for other purposes rather that providing indicational outcomes.

5.3 Recommendations

This section provides recommendations with regards to the methodology (section 5.3.1) and the application of the methodology in practice (5.3.2).

5.3.1 Methodology

Given the concerns with regards to the practical validity of the methodology, it is recommendable to conduct more and more elaborate research to the practical validity of this methodology.

Linearization of non-linear translated utility functions of the existing model

The linearization bias seems to be interchangeable with an additional transfer bias. Therefore, it might be interesting to explore the possibility for skipping the methodological step for linearizing the translated non-linear utility function from the existing transport demand model.

Different methodological approaches can be imagined for linearizing the non-linear utility functions of the existing transport demand model. The first one is the linearization towards the present values of c_{ijm} weighted for the current transport demand of this transport alternative. This linearization would likely provide a better fit for OD-pairs which are important in terms of demand levels.

A second linearization procedure by assigning different scale factors for $f_{2*,m}(c_{ijm})$ for different sets of OD-pairs. By doing so, this limit the number of translated utility functions, whilst local differences in the modal split can be account for.

A third alternative linearization procedure could be conducted by allowing certain c_{ijm} values to change in order to provide a better linear approximation.

All linearization approaches might cause a different impact on the modal split outcomes for the combined model. It is considered to be valuable to experiment with these alternative approaches and assess the impact on the modal split outcomes and the impact of changes in the correction of the external model scales.

Correcting external model scales

It is interesting to consider the impact of providing more degrees of freedom to the correction of the external model scale. This could be done by defining additional correction factors which relate to subset of trip attributes.

5.3.2 Methodology application

A possible approach to use this methodology with valid modelling outcomes is to estimate an external model according the same characteristics of the modelling context of the existing transport demand models. With regards to future modes, it is most natural to estimate this model by means of stated preference data.

Given the limited practical validity of the methodology and the potential presence of large transfer biases within the combined models, a preliminary assessment of the background of the different model is highly recommendable in order to anticipate for transfer bias. Potential users of this methodology should be aware that a certain degree of transfer bias will always be present in the combined model.

For specific cases, the transfer bias can be so large that alternative models need to be considered for the methodology application or the methodology application should not be considered to be appropriate for the modelling purpose.

It is possible that the utility functions of the communal modes in the external model are not able to approximate the systemic utilities of the communal modes of the existing model closely. A possible explanations is the consideration of too many communal modes. For that case, different choice set of communal modes need to be explored for the application of the methodology.

A second explanation can be that the area of scope is set to large, such that the behavioural considerations are different for sub-sets of OD-pairs within the area of scope. For this reason, choosing sub-sets of OD-pairs might limit the impact of the transfer bias on the modal split outcomes.

Appendix I – Demand Responsive Transport in Chicago

This appendix provides the estimated multinomial logit and mixed logit estimated by Frei et al. (2017). The model used in the case study relates to the base scenario for the dummy and categorical variables. This base scenario can be considered as follows:

- The model is valid for sunny weather during summer
- Participant walks to work not more than 2 times per week
- Participant bicycles to work not more than 2 times per week
- Participant drives to work not more than 2 times per week
- Participant does not carry a gym bag to work
- Participant is not a member of a bicycle sharing scheme
- Participant was not willing to answer more stated preference questions
- Participant is not in the age between 51 and 69
- Participant is not in the age between 18 and 34
- Participant does not have an annual income of 150.000 dollars
- Participant’s employer provides commute cost reimbursement
- Participant does not commute regularly by public transport
- Participant does not text or email in public transport

Table 5-1
Multinomial logit and panel mixed-logit mode choice model results (p-value: 0 '***' 0.001 '**' 0.01 '*' 0.05 '^' 0.1 ' ' 1).

Coefficients	Panel mixed logit		Multinomial logit	
	Estimate	Std. error	Estimate	Std. error
Alternative specific constant				
Intercept (base: car)				
Flex	−2.873	1.089**	−1.816	0.623**
Transit	−0.432	0.754	−0.810	0.442^
Generic modal attributes				
Cost ($)	−0.554	0.048***	−0.228	0.023***
Flex wait time (min)	−0.104	0.043*	−0.107	0.031***
Headway (Base: 5 min)				
12–15 min	−0.519	0.251*	−0.366	0.160*
20–25 min	−1.370	0.281***	−1.005	0.184***
30 min	−2.175	0.421***	−1.603	0.286***
60 min	−3.107	0.406***	−2.285	0.268***
Transfers (base: 0 transfers)				
1 Transfer	−0.504	0.162**	−0.095	0.106
2 Transfers	−1.101	0.270***	−0.443	0.185*
Alternative specific modal attributes				
In-vehicle travel time (min)				
Car	−0.103	0.014***	−0.068	0.009***
Flex	−0.134	0.011***	−0.064	0.007***
Transit	−0.116	0.008***	−0.063	0.005***
Walk time (min)				
Car	0.015	0.068	0.012	0.045
Flex	−0.119	0.074	−0.065	0.049
Transit	−0.164	0.017***	−0.084	0.010***
Continuous variables				
Parking cost at work ($)				
Flex	−0.0087	0.0021***	−0.0041	0.0012***
Transit	−0.0090	0.0018***	−0.0041	0.0011***
Commute trip distance (miles)				
Flex	0.154	0.042***	0.029	0.028
Transit	0.071	0.036*	0.030	0.024
Google estimated walk time to nearest transit stop from home (min)				
Flex	−0.049	0.028^	0.026	0.016
Transit	−0.049	0.027^	0.000	0.017
Dummy variables				
Walk to work 2 + times per week				
Flex	0.165	0.474	−0.144	0.319
Transit	0.882	0.418*	0.543	0.254*
Bike to work 2 + times per week				
Flex	2.769	0.505***	1.630	0.315***
Transit	2.501	0.431***	1.536	0.267***
Drive to work 2 + times per week				
Flex	−2.589	0.620***	−1.089	0.322***
Transit	−2.651	0.411***	−1.767	0.247***
Carry gym bag to work				
Flex	−2.022	0.453***	−1.150	0.252***
Transit	−3.199	0.419***	−1.434	0.205***
Bikeshare member				
Flex	1.036	0.373**	1.004	0.252***
Transit	0.525	0.311^	1.021	0.212***
Willing to answer more stated preference questions				
Flex	0.754	0.431^	0.289	0.251
Transit	0.699	0.324*	0.408	0.195*
Male gender				
Flex	0.388	0.330	0.070	0.214
Transit	0.122	0.256	−0.083	0.168

Table 5-1 (*continued*)

Coefficients	Panel mixed logit		Multinomial logit	
	Estimate	Std. error	Estimate	Std. error
Baby boomer (Age 51–69)				
Flex	1.773	0.700*	0.988	0.409*
Transit	2.566	0.556***	1.430	0.332***
Millennial (Age 18–34)				
Flex	0.873	0.389*	0.889	0.224***
Transit	0.701	0.307*	0.726	0.181***
Annual income $40,000–$80,000				
Flex	0.593	0.403	0.121	0.258
Transit	1.313	0.347***	0.231	0.212
Annual income $150,000+				
Flex	1.207	0.459**	0.778	0.272**
Transit	1.032	0.394**	0.452	0.223*
Employer provides no commute cost reimbursement				
Flex	1.766	0.344***	1.157	0.227***
Transit	1.408	0.283***	1.046	0.189***
Categorical variables				
Weather and season (base: sunny and summer)				
Precipitation and summer				
Flex	0.742	0.438^	0.445	0.257^
Transit	0.136	0.327	−0.275	0.203
Precipitation and winter				
Flex	1.928	0.472***	1.176	0.303***
Transit	−0.019	0.385	−0.034	0.248
Sunny and winter				
Flex	2.314	0.507***	1.790	0.338***
Transit	1.213	0.427**	0.853	0.275**
Transit commute and text/email (base: no transit commute, no text/email on transit)				
No transit commute but text/email on transit				
Flex	1.892	0.708**	0.860	0.414*
Transit	0.847	0.548	0.658	0.324*
Transit commute and text/email on transit				
Flex	3.581	0.557***	1.962	0.320***
Transit	2.253	0.417***	1.353	0.253***
Transit commute but no text/email on transit				
Flex	2.977	0.633***	1.462	0.363***
Transit	3.034	0.544***	1.475	0.302***
Random parameters (all normally distributed)				
Std. dev cost	0.377	0.041***	NA	NA
Std. dev in-vehicle travel time (min)				
Car	0.047	0.007***	NA	NA
Flex	0.022	0.004***	NA	NA
Transit	0.029	0.005***	NA	NA
Std. dev walk time (min)				
Car	0.289	0.069***	NA	NA
Flex	−0.144	0.113	NA	NA
Transit	0.140	0.016***	NA	NA
Std. dev. drive to work 2+ times per week (min)				
Flex	2.404	0.423***	NA	NA
Transit	0.886	0.568	NA	NA
Quality of fit statistics				
Log-likelihood	−914		−1078	
McFadden R2	0.423		0.320	
Likelihood ratio test: Chi2	1313		1014	

Appendix II – Shared cars in Lisbon

This appendix provides the estimated nested logit model estimated by Martínez et al. (2017)

Table 1. Coefficients of the obtained discrete choice model.

	Transport alternatives							
	Private transport (PC)			Public transport (PT)				
Attributes	PC	MT	TX	BS	WK	HV	CB	CH
ASC	—	—	−4.230***	−0.218*	−1.140***	−0.270*	−0.706***	−1.190***
Sociodemographic attributes								
Age (25–35)	—	—	—	—	—	0.559***	—	0.559***
Age (35–65)	—	—	—	—	−0.308*	—	—	—
Age (+65)	—	—	—	0.195*	—	—	—	—
Income (thousand €)	—	—	—	0.007***	—	—	—	—
Land use, car and public transport monthly pass availability								
No parking at home	—	—	0.237***	—	—	—	—	0.237***
No parking at destination	—	—	0.237***	—	—	—	—	0.237***
Own car	—	—	—	—	—	—	—	−0.308***
Public transport pass	—	—	—	0.562***	−0.766***	0.562***	0.562***	—
Parking pressure (0–1) †	−0.131***	—	—	—	—	—	—	—
Transport operation attributes								
Fuel cost (€)	−0.326***	−0.326***	—	—	—	—	—	—
Toll (€)	−0.221***	−0.221***	—	—	—	—	—	—
Parking cost (€)	−0.221***	—	—	—	—	—	—	—
Travel time (min)	−0.026***	−0.026***	−0.026***	−0.024***	−0.005***	−0.015***	−0.015***	−0.026***
Access time (min)	—	—	—	−0.051***	—	−0.053***	−0.051***	−0.082***
Tariff (€)	—	—	−0.115***	−0.498***	—	−0.498***	−0.498***	−0.229***
Transfers	—	—	—	−0.199***	—	−0.150**	−0.150**	—
Waiting time (min)	—	—	−0.028***	−0.028***	—	−0.045***	−0.028***	—
Nested scale (η)	1.000	1.951						
0-test significance	—	***						
1-test significance	—	***						

***Significant at the 99% level.
**Significant at the 95% level.
*Significant at the 90% level.
†Parking pressure defined as the ratio between estimated demand and supply of parking in a specific area and time period of the day.

Appendix III - Linearization of the non-linear utility functions of the translated SRE model

Sum of squares Car available

GC	T(cijm)			f2(cijm)			f2*(cijm)			(f2-f2*)^2 * T(cijm)		
cijm	Car	PT	Bike	Car	PT	Bike	Car	PT	Bike	Car	PT	Bike
0,1	2172	0	128	9,90	6,95	10,23	10,10	6,35	10,15	81,42	0,00	0,66
0,3	16724	0	2478	9,83	6,90	10,15	9,91	6,27	10,01	128,39	0,00	55,24
0,5	37596	0	5968	9,70	6,82	10,04	9,73	6,19	9,86	29,87	0,00	202,67
0,7	57016	0	9340	9,55	6,72	9,90	9,55	6,10	9,71	0,57	0,00	349,92
0,9	70250	0	13090	9,38	6,62	9,75	9,36	6,02	9,56	21,05	0,00	458,09
1,1	81412	0	16336	9,20	6,50	9,58	9,18	5,94	9,41	35,70	0,00	477,47
1,3	81268	0	18976	9,01	6,39	9,41	9,00	5,85	9,26	27,41	0,00	422,29
1,5	79800	0	21178	8,82	6,27	9,24	8,81	5,77	9,11	12,33	0,00	327,29
1,7	77832	1	23804	8,63	6,15	9,06	8,63	5,69	8,96	1,99	0,21	229,24
1,9	74956	79	25110	8,44	6,03	8,89	8,44	5,61	8,82	0,45	13,91	129,57
2,1	67632	531	26534	8,25	5,90	8,71	8,26	5,52	8,67	5,70	77,72	56,75
2,3	55476	1547	28050	8,06	5,79	8,54	8,08	5,44	8,52	11,67	185,59	13,51
2,5	40502	3346	28934	7,88	5,67	8,37	7,89	5,36	8,37	13,10	324,28	0,01
2,7	29248	5126	28368	7,69	5,55	8,20	7,71	5,27	8,22	10,93	395,07	12,90
2,9	20764	7546	29298	7,51	5,44	8,03	7,53	5,19	8,07	6,98	454,41	46,52
3,1	14604	9866	29362	7,33	5,32	7,87	7,34	5,11	7,92	3,23	454,90	92,45
3,3	9350	11800	29024	7,15	5,21	7,71	7,16	5,02	7,78	0,74	406,75	142,43
3,5	5962	13112	28014	6,97	5,10	7,54	6,97	4,94	7,63	0,00	328,24	186,69
3,7	2866	13506	27714	6,80	4,99	7,39	6,79	4,86	7,48	0,34	236,80	228,73
3,9	1360	13426	27066	6,63	4,88	7,23	6,61	4,78	7,33	0,82	157,26	258,36
4,1	572	13812	26880	6,46	4,78	7,08	6,42	4,69	7,18	0,96	101,33	280,86
4,3	494	13823	26160	6,30	4,67	6,93	6,24	4,61	7,03	1,76	57,80	285,32
4,5	592	14911	26554	6,14	4,57	6,78	6,06	4,53	6,88	3,88	30,51	289,36
4,7	622	15850	26894	5,98	4,47	6,63	5,87	4,44	6,73	6,81	11,84	280,48
4,9	586	17342	26530	5,82	4,37	6,49	5,69	4,36	6,59	10,01	2,07	253,18
5,1	560	18926	26174	5,66	4,27	6,35	5,51	4,28	6,44	14,18	0,30	217,40
5,3	324	20170	24998	5,51	4,18	6,21	5,32	4,19	6,29	11,67	6,20	170,21
5,5	78	21602	24786	5,36	4,08	6,07	5,14	4,11	6,14	3,86	19,01	128,06
5,7	8	22715	23854	5,21	3,99	5,93	4,95	4,03	5,99	0,53	37,15	83,82
5,9	10	24094	22168	5,07	3,90	5,80	4,77	3,95	5,84	0,87	60,00	44,42
6,1	74	24216	21064	4,92	3,80	5,66	4,59	3,86	5,69	8,24	81,68	16,97
6,3	68	24435	19224	4,78	3,71	5,53	4,40	3,78	5,54	9,56	103,24	2,00
6,5	50	24590	16556	4,64	3,63	5,41	4,22	3,70	5,40	8,73	122,95	1,58
6,7	4	24812	15082	4,50	3,54	5,28	4,04	3,61	5,25	0,86	140,45	14,90
6,9	0	24848	13608	4,36	3,45	5,15	3,85	3,53	5,10	0,00	153,56	40,84

7,1	0	24385	11804	4,23	3,37	5,03	3,67	3,45	4,95	0,00	159,45	75,08
7,3	0	23913	10302	4,09	3,28	4,91	3,49	3,36	4,80	0,00	160,89	116,43
7,5	0	23514	9192	3,96	3,20	4,79	3,30	3,28	4,65	0,00	158,58	166,07
7,7	0	23187	7810	3,83	3,12	4,67	3,12	3,20	4,50	0,00	152,83	210,09
7,9	0	22831	6602	3,70	3,04	4,55	2,93	3,12	4,35	0,00	143,35	251,25
8,1	0	22089	5468	3,58	2,96	4,43	2,75	3,03	4,21	0,00	128,63	283,18
8,3	0	21631	4518	3,45	2,88	4,32	2,57	2,95	4,06	0,00	113,47	308,84
8,5	0	20330	3686	3,33	2,80	4,21	2,38	2,87	3,91	0,00	92,92	324,45
8,7	0	19551	2918	3,20	2,72	4,09	2,20	2,78	3,76	0,00	74,86	324,04
8,9	0	18256	2588	3,08	2,64	3,98	2,02	2,70	3,61	0,00	55,76	356,37
9,1	0	17032	1916	2,96	2,57	3,87	1,83	2,62	3,46	0,00	38,89	322,36
9,3	0	15946	1354	2,84	2,49	3,76	1,65	2,53	3,31	0,00	24,83	274,80
9,5	0	14441	1154	2,73	2,42	3,66	1,46	2,45	3,17	0,00	13,27	279,36
9,7	0	13301	792	2,61	2,35	3,55	1,28	2,37	3,02	0,00	5,51	226,44
9,9	0	12011	498	2,50	2,28	3,45	1,10	2,29	2,87	0,00	1,11	166,68
10,1	0	11062	396	2,38	2,20	3,34	0,91	2,20	2,72	0,00	0,04	153,93
10,3	0	10072	302	2,27	2,13	3,24	0,73	2,12	2,57	0,00	2,01	135,36
10,5	0	9191	160	2,16	2,06	3,14	0,55	2,04	2,42	0,00	6,76	82,16
10,7	0	8200	38	2,05	1,99	3,04	0,36	1,95	2,27	0,00	13,66	22,22
10,9	0	7310	22	1,94	1,93	2,94	0,18	1,87	2,12	0,00	22,29	14,57
11,1	0	6699	8	1,83	1,86	2,84	0,00	1,79	1,98	0,00	33,11	5,97
11,3	0	6021	0	1,73	1,79	2,74	-0,19	1,70	1,83	0,00	44,59	0,00
11,5	0	5130	0	1,62	1,72	2,65	-0,37	1,62	1,68	0,00	53,86	0,00
11,7	0	4603	0	1,52	1,66	2,55	-0,56	1,54	1,53	0,00	65,75	0,00
11,9	0	3887	0	1,41	1,59	2,45	-0,74	1,46	1,38	0,00	73,16	0,00
12,1	0	3388	0	1,31	1,53	2,36	-0,92	1,37	1,23	0,00	81,90	0,00
12,3	0	3051	0	1,21	1,46	2,27	-1,11	1,29	1,08	0,00	92,77	0,00
12,5	0	2708	0	1,11	1,40	2,18	-1,29	1,21	0,93	0,00	101,76	0,00
12,7	0	2345	0	1,01	1,34	2,08	-1,47	1,12	0,79	0,00	107,30	0,00
12,9	0	2099	0	0,91	1,27	1,99	-1,66	1,04	0,64	0,00	115,46	0,00
13,1	0	1771	0	0,81	1,21	1,90	-1,84	0,96	0,49	0,00	115,80	0,00
13,3	0	1353	0	0,72	1,15	1,82	-2,03	0,87	0,34	0,00	104,13	0,00
13,5	0	1188	0	0,62	1,09	1,73	-2,21	0,79	0,19	0,00	106,69	0,00
13,7	0	1082	0	0,52	1,03	1,64	-2,39	0,71	0,04	0,00	112,50	0,00
13,9	0	988	0	0,43	0,97	1,55	-2,58	0,63	-0,11	0,00	118,10	0,00
14,1	0	965	0	0,34	0,91	1,47	-2,76	0,54	-0,25	0,00	131,77	0,00
14,3	0	892	0	0,24	0,85	1,38	-2,94	0,46	-0,40	0,00	138,34	0,00
14,5	0	781	0	0,15	0,79	1,30	-3,13	0,38	-0,55	0,00	136,84	0,00
14,7	0	752	0	0,06	0,74	1,21	-3,31	0,29	-0,70	0,00	148,13	0,00
14,9	0	717	0	-0,03	0,68	1,13	-3,49	0,21	-0,85	0,00	158,07	0,00
15,1	0	606	0	-0,12	0,62	1,05	-3,68	0,13	-1,00	0,00	148,90	0,00
15,3	0	550	0	-0,21	0,57	0,97	-3,86	0,04	-1,15	0,00	150,05	0,00
15,5	0	510	0	-0,30	0,51	0,88	-4,05	-0,04	-1,30	0,00	153,92	0,00
15,7	0	432	0	-0,39	0,45	0,80	-4,23	-0,12	-1,44	0,00	143,76	0,00
15,9	0	393	0	-0,48	0,40	0,72	-4,41	-0,20	-1,59	0,00	143,75	0,00
16,1	0	320	0	-0,56	0,35	0,64	-4,60	-0,29	-1,74	0,00	128,27	0,00
16,3	0	287	0	-0,65	0,29	0,57	-4,78	-0,37	-1,89	0,00	125,73	0,00
16,5	0	233	0	-0,73	0,24	0,49	-4,96	-0,45	-2,04	0,00	111,27	0,00
16,7	0	185	0	-0,82	0,18	0,41	-5,15	-0,54	-2,19	0,00	96,07	0,00
16,9	0	153	0	-0,90	0,13	0,33	-5,33	-0,62	-2,34	0,00	86,19	0,00
17,1	0	128	0	-0,99	0,08	0,26	-5,51	-0,70	-2,49	0,00	78,05	0,00
17,3	0	104	0	-1,07	0,03	0,18	-5,70	-0,79	-2,63	0,00	68,51	0,00
17,5	0	92	0	-1,15	-0,03	0,11	-5,88	-0,87	-2,78	0,00	65,33	0,00
	830832	778698	830832	Sum of Squares						474,61	8872,24	9899,84

Sum of Squares Car Unavailable

GC	T(cijm)			f2(cijm)			f2*(cijm)			(f2-f2*)^2 * T(cijm)		
cijm	Car	PT	Bike	Car	PT	Bike	Car	PT	Bike	Car	PT	Bike
0,1	2172	0	128	7,53	6,88	9,07	7,77	6,24	8,99	123,28	0,00	0,94
0,3	16724	0	2478	7,44	6,83	8,99	7,55	6,15	8,81	194,32	0,00	78,66
0,5	37596	0	5968	7,28	6,74	8,85	7,32	6,06	8,63	45,15	0,00	288,60
0,7	57016	0	9340	7,10	6,64	8,69	7,09	5,97	8,46	0,88	0,00	498,27
0,9	70250	0	13090	6,89	6,52	8,50	6,87	5,88	8,28	31,94	0,00	652,31
1,1	81412	0	16336	6,67	6,40	8,30	6,64	5,79	8,10	54,11	0,00	679,92
1,3	81268	0	18976	6,44	6,27	8,10	6,42	5,70	7,92	41,52	0,00	601,34
1,5	79800	0	21178	6,20	6,14	7,89	6,19	5,61	7,75	18,65	0,00	466,08
1,7	77832	1	23804	5,97	6,01	7,69	5,96	5,52	7,57	2,99	0,25	326,46
1,9	74956	79	25110	5,73	5,88	7,48	5,74	5,43	7,39	0,70	16,39	184,53
2,1	67632	531	26534	5,50	5,75	7,27	5,51	5,34	7,21	8,70	91,55	80,83
2,3	55476	1547	28050	5,27	5,62	7,06	5,29	5,25	7,04	17,77	218,62	19,25
2,5	40502	3346	28934	5,04	5,49	6,86	5,06	5,16	6,86	19,95	382,01	0,02
2,7	29248	5126	28368	4,81	5,37	6,66	4,83	5,07	6,68	16,65	465,40	18,37
2,9	20764	7546	29298	4,58	5,24	6,46	4,61	4,98	6,50	10,65	535,31	66,23
3,1	14604	9866	29362	4,36	5,12	6,26	4,38	4,89	6,33	4,94	535,88	131,62
3,3	9350	11800	29024	4,14	5,00	6,07	4,15	4,80	6,15	1,14	479,16	202,78
3,5	5962	13112	28014	3,93	4,88	5,87	3,93	4,71	5,97	0,00	386,68	265,79
3,7	2866	13506	27714	3,72	4,76	5,69	3,70	4,62	5,79	0,50	278,95	325,65
3,9	1360	13426	27066	3,51	4,64	5,50	3,48	4,53	5,62	1,24	185,26	367,83
4,1	572	13812	26880	3,30	4,53	5,32	3,25	4,44	5,44	1,44	119,37	399,88
4,3	494	13823	26160	3,10	4,42	5,14	3,02	4,35	5,26	2,65	68,09	406,22
4,5	592	14911	26554	2,90	4,30	4,96	2,80	4,26	5,08	5,86	35,95	411,97
4,7	622	15850	26894	2,70	4,19	4,78	2,57	4,16	4,91	10,29	13,95	399,34
4,9	586	17342	26530	2,51	4,09	4,61	2,35	4,07	4,73	15,14	2,44	360,47
5,1	560	18926	26174	2,32	3,98	4,44	2,12	3,98	4,55	21,43	0,36	309,52
5,3	324	20170	24998	2,13	3,88	4,28	1,89	3,89	4,37	17,64	7,30	242,33
5,5	78	21602	24786	1,94	3,77	4,11	1,67	3,80	4,20	5,85	22,39	182,31
5,7	8	22715	23854	1,76	3,67	3,95	1,44	3,71	4,02	0,80	43,76	119,32
5,9	10	24094	22168	1,58	3,57	3,79	1,22	3,62	3,84	1,31	70,68	63,23
6,1	74	24216	21064	1,40	3,47	3,63	0,99	3,53	3,66	12,47	96,22	24,16
6,3	68	24435	19224	1,23	3,37	3,48	0,76	3,44	3,49	14,46	121,62	2,84
6,5	50	24590	16556	1,05	3,28	3,32	0,54	3,35	3,31	13,21	144,84	2,25
6,7	4	24812	15082	0,88	3,18	3,17	0,31	3,26	3,13	1,30	165,46	21,23
6,9	0	24848	13608	0,71	3,09	3,02	0,09	3,17	2,95	0,00	180,89	58,16

7,1	**0**	**24385**	**11804**	**0,55**	**3,00**	**2,87**	**-0,14**	**3,08**	**2,78**	**0,00**	**187,84**	**106,93**
7,3	0	23913	10302	0,38	2,90	2,73	-0,37	2,99	2,60	0,00	189,53	165,81
7,5	0	23514	9192	0,22	2,81	2,58	-0,59	2,90	2,42	0,00	186,81	236,48
7,7	0	23187	7810	0,06	2,73	2,44	-0,82	2,81	2,25	0,00	180,03	299,18
7,9	0	22831	6602	-0,10	2,64	2,30	-1,04	2,72	2,07	0,00	168,87	357,77
8,1	0	22089	5468	-0,25	2,55	2,16	-1,27	2,63	1,89	0,00	151,53	403,24
8,3	0	21631	4518	-0,41	2,46	2,02	-1,50	2,54	1,71	0,00	133,66	439,77
8,5	0	20330	3686	-0,56	2,38	1,89	-1,72	2,45	1,54	0,00	109,46	462,01
8,7	0	19551	2918	-0,71	2,30	1,76	-1,95	2,36	1,36	0,00	88,19	461,42
8,9	0	18256	2588	-0,86	2,21	1,62	-2,17	2,27	1,18	0,00	65,68	507,46
9,1	0	17032	1916	-1,01	2,13	1,49	-2,40	2,18	1,00	0,00	45,81	459,03
9,3	0	15946	1354	-1,15	2,05	1,36	-2,63	2,09	0,83	0,00	29,25	391,30
9,5	0	14441	1154	-1,30	1,97	1,24	-2,85	2,00	0,65	0,00	15,63	397,80
9,7	0	13301	792	-1,44	1,89	1,11	-3,08	1,91	0,47	0,00	6,49	322,44
9,9	0	12011	498	-1,58	1,81	0,98	-3,30	1,82	0,29	0,00	1,31	237,34
10,1	0	11062	396	-1,72	1,73	0,86	-3,53	1,73	0,12	0,00	0,05	219,18
10,3	0	10072	302	-1,86	1,66	0,74	-3,76	1,64	-0,06	0,00	2,37	192,75
10,5	0	9191	160	-2,00	1,58	0,62	-3,98	1,55	-0,24	0,00	7,96	116,99
10,7	0	8200	38	-2,13	1,51	0,50	-4,21	1,46	-0,42	0,00	16,10	31,64
10,9	0	7310	22	-2,27	1,43	0,38	-4,43	1,37	-0,59	0,00	26,26	20,75
11,1	0	6699	8	-2,40	1,36	0,26	-4,66	1,28	-0,77	0,00	39,01	8,50
11,3	0	6021	0	-2,53	1,29	0,14	-4,89	1,19	-0,95	0,00	52,53	0,00
11,5	0	5130	0	-2,66	1,21	0,03	-5,11	1,10	-1,13	0,00	63,45	0,00
11,7	0	4603	0	-2,79	1,14	-0,09	-5,34	1,01	-1,30	0,00	77,46	0,00
11,9	0	3887	0	-2,92	1,07	-0,20	-5,57	0,92	-1,48	0,00	86,19	0,00
12,1	0	3388	0	-3,04	1,00	-0,31	-5,79	0,83	-1,66	0,00	96,49	0,00
12,3	0	3051	0	-3,17	0,93	-0,42	-6,02	0,74	-1,84	0,00	109,28	0,00
12,5	0	2708	0	-3,29	0,86	-0,53	-6,24	0,65	-2,01	0,00	119,88	0,00
12,7	0	2345	0	-3,41	0,79	-0,64	-6,47	0,56	-2,19	0,00	126,40	0,00
12,9	0	2099	0	-3,53	0,73	-0,75	-6,70	0,47	-2,37	0,00	136,01	0,00
13,1	0	1771	0	-3,66	0,66	-0,86	-6,92	0,38	-2,55	0,00	136,41	0,00
13,3	0	1353	0	-3,78	0,59	-0,96	-7,15	0,29	-2,72	0,00	122,67	0,00
13,5	0	1188	0	-3,89	0,53	-1,07	-7,37	0,20	-2,90	0,00	125,68	0,00
13,7	0	1082	0	-4,01	0,46	-1,17	-7,60	0,11	-3,08	0,00	132,53	0,00
13,9	0	988	0	-4,13	0,40	-1,28	-7,83	0,02	-3,26	0,00	139,13	0,00
14,1	0	965	0	-4,24	0,33	-1,38	-8,05	-0,07	-3,43	0,00	155,23	0,00
14,3	0	892	0	-4,36	0,27	-1,48	-8,28	-0,16	-3,61	0,00	162,97	0,00
14,5	0	781	0	-4,47	0,20	-1,58	-8,50	-0,25	-3,79	0,00	161,20	0,00
14,7	0	752	0	-4,58	0,14	-1,68	-8,73	-0,34	-3,97	0,00	174,50	0,00
14,9	0	717	0	-4,70	0,08	-1,78	-8,96	-0,43	-4,14	0,00	186,21	0,00
15,1	0	606	0	-4,81	0,02	-1,88	-9,18	-0,52	-4,32	0,00	175,41	0,00
15,3	0	550	0	-4,92	-0,04	-1,98	-9,41	-0,61	-4,50	0,00	176,76	0,00
15,5	0	510	0	-5,03	-0,10	-2,07	-9,63	-0,70	-4,68	0,00	181,33	0,00
15,7	0	432	0	-5,13	-0,16	-2,17	-9,86	-0,79	-4,85	0,00	169,35	0,00
15,9	0	393	0	-5,24	-0,22	-2,26	-10,09	-0,88	-5,03	0,00	169,34	0,00
16,1	0	320	0	-5,35	-0,28	-2,36	-10,31	-0,97	-5,21	0,00	151,11	0,00
16,3	0	287	0	-5,45	-0,34	-2,45	-10,54	-1,06	-5,38	0,00	148,12	0,00
16,5	0	233	0	-5,56	-0,40	-2,55	-10,76	-1,15	-5,56	0,00	131,08	0,00
16,7	0	185	0	-5,66	-0,46	-2,64	-10,99	-1,24	-5,74	0,00	113,17	0,00
16,9	0	153	0	-5,77	-0,52	-2,73	-11,22	-1,33	-5,92	0,00	101,53	0,00
17,1	0	128	0	-5,87	-0,57	-2,82	-11,44	-1,42	-6,09	0,00	91,95	0,00
17,3	0	104	0	-5,97	-0,63	-2,91	-11,67	-1,51	-6,27	0,00	80,70	0,00
17,5	0	92	0	-6,07	-0,69	-3,00	-11,89	-1,60	-6,45	0,00	76,96	0,00
	830832	778698	830832	Sum of Squares						718,94	10451,67	14096,48

Car Available		Current modal split			Approximated modal split			Deviation of approximation		
Origin	Destination	Car	PT	Bicycle	Car	PT	Bicycle	Car	PT	Bicycle
Station	Airport	15,3%	13,1%	71,7%	21,5%	12,6%	65,8%	6,3%	-0,5%	-5,8%
Station	ASML	71,8%	4,9%	23,4%	72,5%	5,3%	22,2%	0,7%	0,5%	-1,2%
Eindhoven South-East	Eindhoven South	58,8%	0,0%	41,2%	52,9%	0,0%	47,1%	-5,8%	0,0%	5,8%
Veldhoven West	Veldhoven South-West	59,6%	0,3%	40,1%	54,7%	0,3%	45,0%	-4,9%	0,0%	4,9%
Airport	Station	15,5%	11,9%	72,6%	21,9%	11,3%	66,8%	6,4%	-0,6%	-5,8%
ASML	Station	72,5%	3,9%	23,6%	73,4%	4,1%	22,5%	0,9%	0,2%	-1,1%
Eindhoven South	Eindhoven South-West	58,8%	0,0%	41,2%	52,9%	0,0%	47,1%	-5,8%	0,0%	5,8%
VHV South-West	Veldhoven West	59,4%	0,5%	40,0%	54,5%	0,6%	44,8%	-4,9%	0,1%	4,8%

Car Unavailable		Current modal split			Approximated modal split			Deviation of approximation		
Origin	Destination	Car	PT	Bicycle	Car	PT	Bicycle	Car	PT	Bicycle
Station	Airport	2,0%	48,6%	49,4%	3,2%	49,4%	47,4%	1,2%	0,8%	-2,1%
Station	ASML	34,1%	34,7%	31,3%	33,7%	37,5%	28,8%	-0,4%	2,9%	-2,5%
Eindhoven South-East	Eindhoven South	32,4%	0,0%	67,6%	26,5%	0,0%	73,5%	-5,9%	0,0%	5,9%
Veldhoven West	Veldhoven South-West	33,0%	1,2%	65,8%	28,0%	1,1%	70,9%	-5,0%	-0,1%	5,1%
Airport	Station	2,1%	45,7%	52,2%	3,5%	46,0%	50,5%	1,3%	0,4%	-1,7%
ASML	Station	36,8%	29,4%	33,8%	37,3%	30,9%	31,8%	0,5%	1,5%	-2,0%
Eindhoven South	Eindhoven South-West	32,4%	0,0%	67,6%	26,5%	0,0%	73,5%	-5,9%	0,0%	5,9%
VHV South-West	Veldhoven West	32,6%	2,6%	64,9%	27,5%	2,8%	69,7%	-5,0%	0,2%	4,8%

Appendix IV Resulting parameters f3* for the correction of the scale of external models

This section provides the resulting parameters for both external models by multiplying the original parameters of the utility functions by the determined correction factors.

Table 42: utility functions f3, f3 and f2* for the Chicago model and the CA group*

Correction of DRT in Chicago for Car Available		F3			F3*			F2*	
Parameter	Unit of measurement of explanatory variable	Private Car	Public Transport	DRT	Private Car	Public Transport	DRT	Private Car	Public Transport
ASC		0	-0,810	-1,816	0	-4,039	-9,055	10,189	6,394
$\beta_{\text{generalised costs}}$	Euro							-0,918	-0,415
β_{distance}	Kilometre								
$\beta_{\text{acces time}}$	Minute		-0,084			-0,355			
$\beta_{\text{in-vehicle time}}$	Minute	-0,068	-0,063	-0,064	-0,287	-0,266	-0,271		
$\beta_{\text{waiting time}}$	Minute		-0,084	-0,107		-0,355	-0,452		
B_{transfer}	# transfer		-0,221	-0,221		-0,936	-0,936		
$\beta_{\text{egress time}}$	Minute		-0,084			-0,355			
β_{costs}	Euro	-0,228	-0,185	-0,185	-0,964	-0,781	-0,781		

Table 43: utility functions f3, f3 and f2* for the Chicago model and the CU group*

Correction of DRT in Chicago for Car Unavailable		F3			F3*			F2*	
Parameter	Unit of measurement of explanatory variable	Private Car	Public Transport	DRT	Private Car	Public Transport	DRT	Private Car	Public Transport
ASC		0	-0,810	-1,816	0	-1,647	-3,692	7,885	6,282
$\beta_{\text{generalised costs}}$	Euro							-1,130	-0,450
$\beta_{\text{access time}}$	Minute		-0,084			-0,357			
$\beta_{\text{in-vehicle time}}$	Minute	-0,068	-0,063	-0,064	-0,289	-0,267	-0,272		
$\beta_{\text{waiting time}}$	Minute		-0,084	-0,107		-0,357	-0,454		
B_{transfer}	# transfer		-0,221	-0,221		-0,940	-0,940		
$\beta_{\text{egress time}}$	Minute		-0,084			-0,357			
β_{costs}	Euro	-0,185	-0,185	-0,185	-0,784	-0,784	-0,784		

Table 44: utility functions f3, f3 and f2* for the Lisbon model and the CA group*

Correction of Shared Cars in Lisbon for Car Available		F3		F3*		F2*
Parameter	Unit of measurement of explanatory variable	Public Transport	Shared Car	Public Transport	Shared Car	Public Transport
ASC		-0,218	-1,498	-0,218	-1,19	6,394
$\beta_{\text{generalised costs}}$	Euro					-0,415
$\beta_{\text{access time}}$	Minute	-0,051	-0,082	-0,23261	-0,374	
$\beta_{\text{in-vehicle time}}$	Minute	-0,024	-0,026	-0,10946	- 0,11859	
$\beta_{\text{waiting time}}$	Minute	-0,028		-0,12771		
B_{transfer}	# transfer	-0,199		-0,90763		
$\beta_{\text{egress time}}$	Minute	-0,051	-0,082	-0,23261	-0,374	
β_{costs}	Euro	-0,498	-0,229	-2,27137	-1,04	

Table 45: utility functions f3, f3 and f2* for the Lisbon model and the CU group*

Correction of Shared Cars in Lisbon for Car Unavailable		F3		F3*		F2*
Parameter	Unit of measurement of explanatory variable	Public Transport	Shared Car	Public Transport	Shared Car	Public Transport
ASC		0,344	-1,19	0,344	-1,19	6,282
$\beta_{\text{generalised costs}}$	Euro					-0,450
$\beta_{\text{access time}}$	Minute	-0,051	-0,082	-0,25247	- 0,40593	
$\beta_{\text{in-vehicle time}}$	Minute	-0,024	-0,026	-0,11881	- 0,12871	
$\beta_{\text{waiting time}}$	Minute	-0,028		-0,13861	0	
B_{transfer}	# transfer	-0,199		-0,98512	0	
$\beta_{\text{egress time}}$	Minute	-0,051	-0,082	-0,25247	-0,41	
β_{costs}	Euro	-0,498	-0,229	-2,46527	- 1,13363	

Appendix V Reference zones and disaggregate modal split outcomes

This appendix introduces six references zones in order to provide the modal split outcomes on a more disaggregate level. The reference zones are also used to analyse the practical validity of the methodology. This is done as the total number of zones in the study area is far too large to present all the disaggregate information.

The following zones are considered as reference zones:

- Eindhoven Station (#767; marked in red)
- Van Abbe Museum (#747; marked in orange)
- Philips Office (#1239; marked in green)
- Eindhoven Burghplan (#942; marked in blue)
- Veldhoven Cobbeek (#3043; marked in purple)
- Eindhoven Airport (#1130; marked in black)

The locations of these zones are visualised in Figure 27. The selection of these zones reflects the diversity of zone types, both in their geographical location as in their general accessibility (by specific modes). Eindhoven station functions as the main-hub for public transport. The Van Abbe Museum reflects a central location but is relatively far away from the train station. The Philips Office represents the northern parts of Eindhoven, where Eindhoven Burghplan and Veldhoven Cobbeek represent the southern and western parts of the study area respectively. Eindhoven Airport represents an important origin and destination (i.e. it attracts many travellers and employees).

Figure 27: overview on the locations of the reference zones in the study area

Table 46 and Table 47 set out the modal split distributions for both the car available and car unavailable user group in absolute numbers. Table 48 and Table 49 provide the modal splits for the alternatives between the reference zones respectively for the user groups car available and car unavailable.

Table 46: overview of the modal split distribution over all OD-pairs for CA

Modal Split			Car	PT	Bicycle	DRT	SC
0-1%	0	0,01	914	746648	1654	831744	597834
1-2%	0,01	0,02	0	62042	11945	0	183355
2-3%	0,02	0,03	0	16972	21197	0	37315
3-4%	0,03	0,04	0	4523	25508	0	8910
4-5%	0,04	0,05	6	969	26811	0	2445
5-6%	0,05	0,06	90	316	29389	0	672
6-7%	0,06	0,07	324	145	32322	0	314
7-8%	0,07	0,08	597	71	31871	0	268
8-9%	0,08	0,09	584	36	30004	0	228
9-10%	0,09	0,1	519	9	28041	0	129
10-15%	0,1	0,15	861	13	113597	0	262
15-20%	0,15	0,2	646	0	94723	0	12
20-25%	0,2	0,25	2931	0	81007	0	0
25-35%	0,25	0,35	17245	0	132035	0	0
35-45%	0,35	0,45	26154	0	89223	0	0
45-55%	0,45	0,55	40315	0	38483	0	0
>55%	0,55	1	740558	0	43934	0	0

Table 47: overview of the modal split distribution over all OD-pairs for CU

Modal Split			Car	PT	Bicycle	DRT	SC
0-1%	0	0,01	976	151029	1484	831744	88734
1-2%	0,01	0,02	2282	149943	5942	0	144292
2-3%	0,02	0,03	587	151522	10288	0	107623
3-4%	0,03	0,04	369	141022	12503	0	78415
4-5%	0,04	0,05	533	91331	13611	0	61582
5-6%	0,05	0,06	879	51408	14092	0	52719
6-7%	0,06	0,07	1813	28617	14286	0	46046
7-8%	0,07	0,08	2762	17092	14565	0	41552
8-9%	0,08	0,09	3691	10679	15036	0	35376
9-10%	0,09	0,1	3950	7254	15963	0	30337
10-15%	0,1	0,15	22027	19233	81587	0	88200
15-20%	0,15	0,2	20664	8550	69788	0	35997
20-25%	0,2	0,25	20123	3043	59043	0	13119
25-35%	0,25	0,35	84325	1006	107854	0	6332
35-45%	0,35	0,45	93600	15	102060	0	1174
45-55%	0,45	0,55	106300	0	98428	0	215
>55%	0,55	1	466863	0	195214	0	31

Table 48: overview of the modal split for the reference OD-pairs for CA

		Einhoven Station	Van Abbe Museum	Philips Office	Eindhoven Burghplan	Veldhoven Cobbeek	Eindhoven Airport
Eindhoven Station	Car		9,5%	39,8%	36,8%	77,0%	14,0%
	PT		1,7%	1,6%	1,5%	2,7%	12,0%
	Bicycle		88,4%	57,9%	61,1%	17,3%	65,6%
	DRT		0,0%	0,0%	0,0%	0,0%	0,0%
	Shared Car		0,5%	0,7%	0,6%	2,9%	8,5%
Van Abbe Museum	Car	9,5%		40,8%	29,5%	70,1%	12,2%
	PT	1,6%		0,9%	1,1%	1,0%	7,3%
	Bicycle	88,5%		57,5%	69,1%	26,3%	73,0%
	DRT	0,0%		0,0%	0,0%	0,0%	0,0%
	Shared Car	0,5%		0,8%	0,3%	2,6%	7,5%
Philips Office	Car	39,8%	40,6%		80,2%	91,2%	26,4%
	PT	1,8%	1,4%		0,5%	0,2%	3,8%
	Bicycle	57,8%	57,2%		18,8%	7,7%	66,7%
	DRT	0,0%	0,0%		0,0%	0,0%	0,0%
	Shared Car	0,7%	0,8%		0,5%	0,9%	3,1%
Eindhoven Burghplan	Car	36,8%	29,5%	80,2%		92,6%	44,9%
	PT	1,6%	1,2%	0,4%		0,3%	5,8%
	Bicycle	61,0%	68,9%	18,8%		6,1%	41,0%
	DRT	0,0%	0,0%	0,0%		0,0%	0,0%
	Shared Car	0,6%	0,3%	0,5%		1,0%	8,3%
Veldhoven Cobbeek	Car	76,9%	69,6%	91,1%	92,6%		22,5%
	PT	2,9%	1,7%	0,3%	0,3%		3,8%
	Bicycle	17,3%	26,1%	7,7%	6,1%		71,0%
	DRT	0,0%	0,0%	0,0%	0,0%		0,0%
	Shared Car	2,9%	2,6%	0,9%	1,0%		2,8%
Eindhoven Airport	Car	14,1%	12,2%	26,9%	45,9%	22,9%	
	PT	10,9%	7,2%	1,7%	3,8%	1,8%	
	Bicycle	66,4%	73,1%	68,1%	41,9%	72,5%	
	DRT	0,0%	0,0%	0,0%	0,0%	0,0%	
	Shared Car	8,6%	7,5%	3,2%	8,4%	2,8%	

Table 49: overview of the modal split for the reference OD-pairs for CU

		Einhoven Station	Van Abbe Museum	Philips Office	Eindhoven Burghplan	Veldhoven Cobbeek	Eindhoven Airport
Eindhoven Station	Car		1,8%	14,3%	12,6%	38,7%	1,6%
	PT		5,4%	7,3%	6,6%	19,8%	37,6%
	Bicycle		91,7%	76,0%	78,6%	22,9%	38,3%
	DRT		0,0%	0,0%	0,0%	0,0%	0,0%
	Shared Car		1,2%	2,5%	2,2%	18,6%	22,5%
Van Abbe Museum	Car	1,8%		15,1%	9,3%	36,5%	1,5%
	PT	5,1%		4,4%	4,0%	6,8%	25,3%
	Bicycle	91,9%		77,6%	85,7%	39,8%	50,5%
	DRT	0,0%		0,0%	0,0%	0,0%	0,0%
	Shared Car	1,2%		3,0%	1,0%	16,9%	22,6%
Philips Office	Car	14,1%	14,7%		59,3%	77,8%	6,0%
	PT	8,2%	6,8%		3,3%	1,5%	16,2%
	Bicycle	75,2%	75,6%		34,4%	13,6%	66,2%
	DRT	0,0%	0,0%		0,0%	0,0%	0,0%
	Shared Car	2,5%	2,9%		3,0%	7,0%	11,6%
Eindhoven Burghplan	Car	12,5%	9,2%	59,4%		79,4%	9,7%
	PT	7,0%	4,7%	3,1%		2,2%	25,4%
	Bicycle	78,2%	85,1%	34,5%		10,3%	32,6%
	DRT	0,0%	0,0%	0,0%		0,0%	0,0%
	Shared Car	2,2%	1,0%	3,0%		8,0%	32,3%
Veldhoven Cobbeek	Car	38,3%	34,5%	77,1%	79,0%		4,8%
	PT	20,7%	11,8%	2,5%	2,7%		15,7%
	Bicycle	22,7%	37,7%	13,5%	10,3%		69,8%
	DRT	0,0%	0,0%	0,0%	0,0%		0,0%
	Shared Car	18,4%	16,0%	7,0%	7,9%		9,8%
Eindhoven Airport	Car	1,6%	1,5%	6,6%	10,8%	5,3%	
	PT	35,0%	24,8%	7,6%	17,3%	7,5%	
	Bicycle	40,0%	50,9%	73,1%	36,2%	76,5%	
	DRT	0,0%	0,0%	0,0%	0,0%	0,0%	
	Shared Car	23,3%	22,7%	12,7%	35,7%	10,8%	

Literature

Badoe, D. A., & Miller, E. J. (1995a). Analysis of Temporal Transferability of Disaggregate Work Trip Mode Choice Models. *Transportation Research Record*(1493).

Badoe, D. A., & Miller, E. J. (1995b). Comparison of alternative methods for updating disaggregate logit mode choice models. *Transportation Research Record*(1493).

Ben-Akiva, M., & Bierlaire, M. (1999). Discrete choice methods and their applications to short term travel decisions. In *Handbook of transportation science* (pp. 5-33): Springer.

Ben-Akiva, M., & Bolduc, D. (1987). *Approaches to model transferability and updating: the combined transfer estimator*: Département d'économique, Université Laval.

Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., & Rao, V. (1994). Combining revealed and stated preferences data. *Marketing Letters, 5*(4), 335-349.

Bhat, C. R., Guo, J., Srinivasan, S., & Sivakumar, A. (2003). Guidebook on activity-based travel demand modeling for planners. *Austin, TX.: Univ. Texas at Austin*, 4080-P4083.

Bliemer, M. (2010a). *Aggregated and desaggregated approach to simulatenous mode & destination choice*. Retrieved from

Bliemer, M. (2010b). *Transitie van geaggregeerde naar gedesaggregeerde modellen: bestemmings- en modaliteitskeuze*. Retrieved from

Bradley, M. A., & Daly, A. J. (1997). Estimation of logit choice models using mixed stated preference and revealed preference information. *Understanding travel behaviour in an era of change*, 209-232.

Bunch, D. S., & Rocke, D. M. (2016). Variance-component-based nested logit specifications: Improved formulation, and practical microsimulation of random disturbance terms. *Journal of Choice Modelling, 21*, 30-35. doi:https://doi.org/10.1016/j.jocm.2016.04.001

CBS. (2015). Onderzoek Verplaatsingen in Nederland 2014. Onderzoeksbeschrijving.

CBS. (2018). CBS in uw buurt. Retrieved from http://www.cbsinuwbuurt.nl

CROW. (2017). Autodelen. Aanbieders. Retrieved from https://kpvvdashboard-4.blogspot.nl/2015/10/aanbieders.html

Domencich, T. A., McFadden, D., & Associates, C. R. (1975). *Urban Travel Demand: A Behavioral Analysis : a Charles River Associates Research Study*: North-Holland Publishing Company.

Eiró, T., & Martínez, L. (2014). *Development of a discrete choice model to evaluate the impact of new transport alternatives.* Paper presented at the Transportation Research Board 93rd Annual Meeting.

Ettema, D. (1996). *Activity-based travel demand modeling*: Technische Universiteit Eindhoven.

Frei, C., Hyland, M., & Mahmassani, H. S. (2017). Flexing service schedules: Assessing the potential for demand-adaptive hybrid transit via a stated preference approach. *Transportation Research Part C: Emerging Technologies, 76*, 71-89.

Gemeente Eindhoven. (2015). Raadsinformatie: Eindhoven op weg naar slimme en duurzame mobiliteit. In.

Goudappel Coffeng. (2012). Actualisering SRE Verkeersmodel, Technische Documentatie. In. Deventer: Goudappel Coffeng.

Goudappel Coffeng. (2013). Verkeersmodel RVMK 3, Technische Rapportage. In. Deventer: Goudappel Coffeng.

Goudappel Coffeng. (2014). Verkeersmodel Haaglanden 1.0, Technische Rapportage. In. Deventer: Goudappel Coffeng.

Hall, J., Kendrick, C., & Nosko, C. (2015). The effects of Uber's surge pricing: A case study. *The University of Chicago Booth School of Business*.

Hansen, S. (1981). In favor of cross-cultural transferability of travel-demand models. *New Horizons in Travel Behaviour Research*, 637-651.

Heikkilä, S. (2014). Mobility as a service-a proposal for action for the public administration, case helsinki.

Hietanen, S. (2014). Mobility as a Service - the new transport model? *Eurotransport. ITS & Transport Management Supplement, 12*(2), 2-4.

Instituto Nacional de Estatística. (2012). Censos 2011 Resultados Definitivos-Região Centro. Retrieved from https://www.ine.pt/xportal/xmain?xpgid=ine_main&xpid=INE

Kamargianni, M., Li, W., Matyas, M., & Schäfer, A. (2016). A critical review of new mobility services for urban transport. *Transportation Research Procedia, 14*, 3294-3303.

Karasmaa, N. (2007). Evaluation of transfer methods for spatial travel demand models. *Transportation Research Part A: Policy and Practice, 41*(5), 411-427.

Kruyswijk, M. (2017, December 21). De deelfiets krijgt een herkansing in Amsterdam. *Het Parool*. Retrieved from https://www.parool.nl/amsterdam/de-deelfiets-krijgt-een-herkansing-in-amsterdam~a4547308/

Lazarsfeld, P. F., & Henry, N. W. (1968). Latent structure analysis. *New York, NY: Houghton Mifflin, 3*, 35.

Loomis, J. (2011). What's to know about hypothetical bias in stated preference valuation studies? *Journal of Economic Surveys, 25*(2), 363-370.

Manski, C. F. (1977). The structure of random utility models. *Theory and Decision, 8*(3), 229-254. doi:10.1007/BF00133443

Martínez, L. M., Correia, G. H. d. A., Moura, F., & Mendes Lopes, M. (2017). Insights into carsharing demand dynamics: outputs of an agent-based model application to Lisbon, Portugal. *International Journal of Sustainable Transportation, 11*(2), 148-159.

McCallum, W., Hughes-Hallett, D., & Gleason, A. (1994). *Multivariable Calculus*: John Wiley & Sons Australia, Limited.

Meyer, M. D., & Institute of Transportation, E. (2016). *Transportation planning handbook*. In.

Murphy, J. J., Allen, P. G., Stevens, T. H., & Weatherhead, D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics, 30*(3), 313-325.

Nederlandse Spoorwegen. (2018). OV-fiets 10 jaar bij NS. Dank je wel klanten! Retrieved from https://nieuws.ns.nl/ov-fiets-10-jaar-bij-ns-dank-je-wel-klanten/

Ortúzar, J., & Willumsen, L. G. (2011). *Modelling Transport*: Wiley.

Pieters, M., Elshout, V. d., & Herder, J. (2014). *Verkeersmodel Amsterdam: ontwikkeling van een nieuw stedelijk model*. Paper presented at the Colloquium Vervoersplanologisch Speurwerk, Eindhoven.

Rijkswaterstaat. (2017). Verkeer en Vervoer, het Landelijk Model Systeem. In. Utrecht: Rijkswaterstaat.

Sivakumar, A. (2007). Modelling transport: a synthesis of transport modelling methodologies. *Imperial College of London*.

Train, K. E. (2009). *Discrete choice methods with simulation*: Cambridge university press.

Tranzer. (2018). Just take the ride. Plan en reis met je telefoon. Retrieved from http://www.tranzer.com

TU Delft. (2017). Breng Flex een succes! Retrieved from https://www.tudelft.nl/2017/transport-institute/breng-flex-een-succes/

U.S. Census Bureau. (2015). Quick facts Chicago City, IL. Retrieved from https://www.census.gov/quickfacts/fact/table/chicagocityillinois/PST045216

van Oort, N. v. d. B., R.;Verhoof, F. (2017). *The wider benefits of high quality public transport for cities*. Paper presented at the European Transport Conference, Barcelona.

Vandermerwe, S., & Rada, J. (1988). Servitization of business: adding value by adding services. *European management journal, 6*(4), 314-324.

Whim. (2018). Retrieved from https://whimapp.com/nl/

Williams, H. C. W. L. (1977). On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit. *Environment and Planning A: Economy and Space, 9*(3), 285-344. doi:10.1068/a090285