Comparison of an aggregated macroscopic trip-based 4-step model with a disaggregated tour-based microsimulator:

A backcasting approach

UNIVERSITY OF TWENTE.

Dat.mobility
ADVANCING ANALYTICS

Title
Comparison of an aggregated macroscopic trip-based 4-step model with a disaggregated tour-based microsimulator: A backcasting approach

Author
Alex V. Mouw – s1974319
alexmouw@protonmail.com

Version
Final

Date
24-08-2022

Institution
University of Twente
Enschede, The Netherlands

Program
Civil Engineering and Management
Discipline: Traffic Engineering and Management

Internal supervisors
Prof. Dr. Ir. E.C. van Berkum
Dr. Ir. T. Thomas

Organisation
DAT.mobility | Goudappel
Deventer, The Netherlands

External supervisors
Ir. L.J.N. Brederode
Ir. J.B. Voorhorst
Dr. Ir. L.J.J. Wismans

Preface

This thesis is the result of my graduation research; the final step for obtaining a master's degree in Civil Engineering & Management at the University of Twente. During the last six months, I have evaluated the prediction quality of two types of travel demand models. I have enjoyed the journey with the opportunity to deepen my understanding of the subject. Many people supported me in this and I want to take a moment to thank them. Firstly, I would like to thank my daily supervisors at Goudappel, Luuk and Jesse, for their positive and critical feedback which motivated me to aim for in-depth analyses, and who helped me to understand Octavius, interpret my results, and improved my writing. Also, I would like to thank Luc for his feedback and for being the bridge between academia and the work field. Thirdly, I would like to thank Tom, whose feedback and questions helped me to achieve a better result, and Eric, for his feedback and help in finding an interesting research topic. Furthermore, I want to thank my colleagues at Goudappel and Dat.mobility for the enjoyable contact and answering my numerous questions. Moreover, I would like to thank my family and friends for their support, with a special thanks to Leon Besseling for proofreading my work. Lastly, I want to thank Roos, who has always supported me and listened to my enthusiasm and struggles.

Alex Mouw
Harderwijk, August 2022

Summary

The emergence of new mobility services as a result of the societal transition from owning to using goods and services, requires travel demand models to include the decision maker's context. The traditional aggregated, macroscopic, trip-based 4-step model is not able to fulfil this requirement. Therefore, a generation of more advanced models has arrived with clear theoretical advantages. These models are disaggregated, microscopic, and tour-based, and take the decision maker's context into account. However, empirical comparisons between these two approaches are scarce and do not address the prediction quality of both models, while the primary aim of travel demand models is to forecast future situations.

The presented study attempts to fill this knowledge gap by conducting a backcast: predicting a historical year for which observed travel behaviour is available and compare this with the base year performance. The city of Almere in the Netherlands between 2010 and 2017 is used as case. The model performance is evaluated using three aggregated Key Performance Indicators (KPIs) that can be used for both models: trip frequency, modal split, and trip length distribution.

The results demonstrate that the advanced model, called Octavius, is more capable of describing the base year compared to the 4-step model when evaluated against travel survey data. However, the backcast results show that both models perform similarly in terms of prediction quality. Hence, the longitudinal stability of both models is similar. In addition to the three KPIs, the disaggregated behaviour is analysed using two scenarios. The findings show that Octavius is able to include factors that affect travel behaviour in a manner that is consistent with literature. The 4-step model on the other hand exhibits counterintuitive results with increasing car use when car ownership decreases, which demonstrates its lack of behavioural realism. Moreover, the possibility of Octavius to model certain subgroups in society is a valuable tool to evaluate transport policies and other developments, such as the development of new neighbourhoods; analyses that the 4-step is not able to provide.

Finally, the main limitation of this study is the limited differences between 2017 and 2010 in travel behaviour and socioeconomic development. Consequently, it is not possible to accurately evaluate how the models respond to changes, and in which situations one model performs better than the other. Future research could clarify this by using a longer backcast period, while ensuring that input and evaluation data are of sufficient quality.

Table of Contents

1 Introduction

The 4-step model has been widely applied to calculate travel demand since the 1960s, and is still in use today (Elmorssy et al., 2019). It generally describes travel demand as a function of attractiveness and resistance: an analogue to Newton's law of gravitation. However, due to the societal transition from owning to using goods and services, and the emergence of new mobility services that respond to this development (Jittrapirom et al., 2017), transport demand models that are able to describe more complex behavioural choices are required. Examples include the effects of travel decisions of other household members, transport mode availability at mobility hubs, but also home working among employees in certain industries. The 4-step model does not, and is not able to, consider the decision maker's context (Elmorssy et al., 2019). A different approach is to model individual travellers with distinct characteristics, that live in a certain type of household, and that want to perform particular activities. All these factors affect subsequently the travel choices made, which means that the decision maker's context directly influences travel behaviour.

The conceptual differences between the two approaches are: aggregated versus disaggregated, macro versus microscopic, and trip versus tour-based. Disaggregated means that travel behaviour is modelled for segments: groups of travellers for whom the same set of variables determines their travel decisions through choice models. To model segments efficiently, microsimulation is required. That is, all characteristics – including interaction with the environment, such as other household members – are stored in agents that together form a synthetic population. Subsequently, the agents are grouped into segments, which can differ per choice model. Without microscopic simulation, the number of segments would increase enormously, because each combination of characteristics, would be stored as a separate segment (Ortúzar & Willumsen, 2011). Lastly, tours capture a complete home to home journey, including the trips to the primary and possible secondary activities. As a result, tour-based models are (at least) consistent in mode and destination choice (Vovsha, 2019), while trip-based models do not model any connection between trips.

Although the theoretical advantages of these types of models have been extensively discussed in literature (e.g., Davidson et al., 2007; Rasouli & Timmermans, 2014; Vovsha, 2019), only few studies were found that empirically compared these two approaches. However, these studies focussed on the base year performance or scenario testing only, while transport demand models are applied to forecast future situations. Consequently, it is not clear how the prediction quality of both approaches compares. The presented study conducts a backcast, which is the opposite of a forecast, to evaluate the prediction quality. The advantage of backcasting is that observed travel behaviour is already available as evaluation data, which means that one does not have to wait until the data for the predicted year become available (Roorda et al., 2008).

Goudappel developed Octavius, a software implementation of the second discussed approach, which is used in this study together with a traditional 4-step model of Goudappel. Although completely functional, the choice models that take fully advantage of considering the decision maker's context (such as travel decisions of other household members, see the examples given above) are not included yet. Therefore, this study focuses on comparing the disaggregated approach of Octavius with the aggregated approach of the 4-step model. The city of Almere was used as a case, because a fully functional 4-step model is available for that city, as well as a 2010 network that could be used to reconstruct the 2010 situation. The remainder of this document is structured as follows. In section 2, relevant literature is discussed. Section 3 presents the research dimensions, including the research questions. The models used for the comparison are discussed in section 4. Subsequently, the methodology is explained in section 5. The results obtained are demonstrated in section 6, after which they are discussed in section 7. The conclusions are drawn in section 8, with the limitations and recommendation of this study presented in section 9.

2 Literature review

2.1 Traditional 4-step model

Before discussing the 4-step model, it should be noted that a variety of different implementations of the 4-step model exist. The aim of this section is to give a description of how the model works in general terms, while section 4.1 describes the specific implementation used in this study. As explained by Rasouli and Timmermans (2014), the traditional 4-step model is an aggregate travel demand model, using trips as modelling unit. A trip is a journey made from one origin (i.e., home) to one destination (i.e., work). The aim is to predict the number of trips coming from and going to so-called Traffic Analysis Zones (TAZs). TAZs originate from the aggregate nature of the model, meaning that travel decisions of a group of individuals are estimated at an aggregated level, rather than modelling individual travellers. The traditional 4-step model consists of four stages, which explains the name. In the first step, the trip production and attraction for each TAZ is determined based on socioeconomic characteristics of that particular TAZ. The second stage models the destination choice of travellers. A common approach is to use a gravity model, which is an analogue to Newton's law of gravitation. It describes that TAZs with a higher attractiveness, for instance a high number of workplaces, inhabitants, or shops, attract more trips. Furthermore, the attraction of a particular zone is inversely proportional to the cost involved to travel to that zone, where cost can refer to travel time, monetary costs or other (combination of) factors. The third step calculates the cost of travelling from a certain origin O to a destination D for all considered travel modes. Subsequently, the mode choices of all travellers are modelled, resulting in mode-specific matrices which contain the probability of choosing that mode for each OD-relation. Then, the mode-specific matrices are multiplied with the number of trips per OD-relation, which was calculated in the second step. This results in OD-matrices that specify how many trips are made from each origin to each destination by which mode. Traditionally, the last step is to model the route choice of travellers, often referred to as traffic assignment, which results in the number of trips made over each link in the network. Although the traditional model consists of four stages, a fifth stage can be added which includes the time of travel. In this way, the travel demand over different time periods can be modelled.

2.2 Limitations of the traditional 4-step model

Four main limitations of the traditional 4-step model were identified. First, travel decisions are modelled at a highly aggregated level, meaning that individuals and households within the same TAZ are considered identical, or only a limited number of characteristics are distinguished (Vovsha, 2019). This introduces a bias, because average individuals and households do not automatically represent a population accurately. Or, in other words, "the probability of an average is not necessarily equal to the average of the probabilities across individual underlying values" (Rasouli & Timmermans, 2014, p. 33). Similarly, predicting the impacts of measures on certain subgroups is difficult if not impossible, depending on the subgroup of interest.

Secondly, dependencies between travel decisions of household members (so-called 'intra-household interactions') cannot be modelled. However, research suggests that interaction among members of the same household affect individual's travel behaviour (Davidson et al., 2007). For instance, joint activities such as going to the movies, or doing grocery shopping on behalf of all the household members. Due to this independency, secondary effects of transportation policies cannot be captured (e.g., the effect of changing travel behaviour due to policy development of household member A on member B).

Thirdly, travel decisions are modelled independently and can therefore not be described consistently (Rasouli & Timmermans, 2014). As a result, different modes of transport can be used for trips that in reality belong to the very same journey. Similarly, the connection between home-based (i.e., trips that start or end at home) and non-home-based trips (i.e., trips that neither start nor end at home) does not exist. However, Dutch research has demonstrated that approximately twenty percent of journeys made include at least one non-home-based trip (Schneider et al., 2021).

Lastly, time of travel cannot be included on an individual level as this depends on the daily schedule of a traveller. Moreover, the distances that can be realistically travelled within a daily schedule (i.e., time-space constraints) limit the number of possible destinations. Examples include opening hours of stores and working hours (Rasouli & Timmermans, 2014). Since these constraints are not considered in a 4-step model, unrealistic travel behaviour may occur and policies that aim to influence time of travel, such as congestion pricing, are difficult to evaluate.

2.3 Advanced travel models: disaggregated, microscopic and tour-based

To address the abovementioned limitations of the traditional four-step model, several improvements have been made in the travel demand modelling field. In this section, disaggregated, microscopic and tour-based modelling is compared with the aggregated, macroscopic trip-based approach of the 4-step model. Although the improvements are complementary, and one can function without the other, they are often jointly implemented (Davidson et al., 2007).

Firstly, disaggregate models are estimated with a dependent variable representing an observation of one occurrence: a trip within travel demand modelling (Richards, 1974). In contrast, a group of observations represents a dependent variable in the case of aggregated models. The major benefit of disaggregated models is the highly efficient use of available data, as all observations can be used for variable estimation, while aggregate models such as the 4-step model are estimated on clustered observations. Moreover, grouping observations results in averaging of the observed travel behaviour, which means that variability within the data is lost and homogeneous TAZs are modelled (Richards, 1974). Disaggregated models on the other hand take advantage of the data variability, such that heterogeneity in travel behaviour is included in the model.

Secondly, modelling at a microscopic level entails that individuals and households are explicitly modelled, which contrasts with the zonal approach of the 4-step model. The technique translates probabilistic travel choices into discrete choices, creating synthetic travellers which are similar to real persons with distinct characteristics and travel behaviour (Davidson et al., 2007). It should be noted that modelling individual travel behaviour is not aimed to pinpoint that behaviour, but rather to predict aggregate travel behaviour with higher accuracy. The concept is that when individual travel choices are modelled in a more consistent and realistic manner, the aggregate predictions become more consistent and realistic as well (Davidson et al., 2007). Consequently, the aggregation bias inherent to the four-step model is prevented, since the effects of changing travel conditions are captured at an individual and household level (Vovsha & Bradley, 2006).

Thirdly, individual trips are replaced by *tours.* Tours are “a sequence of trips that begins at home, involves visits one or more other places, and ends at home” (Ye et al., 2007, p. 97). As a result, mode choice and destination (and possibly time of the day) are consistent across all trips in the same tour (Vovsha, 2019). Note that two types of tours can be distinguished: simple tours with one activity (2-trip tour), and complex tours with two or more activities (3-trip tour), outside the home location (Ye et al., 2007). Since twenty percent of the Dutch journeys involves a complex tour (Schneider et al., 2021), it is a significant part of the journeys that cannot be accurately modelled by the 4-step model.

In addition to the advantages of disaggregated, microscopic, and tour-based travel models, literature also identified three main drawbacks of this modelling approach. Firstly, the increased complexity requires the estimation and calibration of more sub-models and parameters compared to a 4-step model, which leads to increased time efforts (Lemp et al., 2007). Similarly, the increased complexity requires more detailed input data, which can be costly and time-consuming to collect (Omer et al., 2010). Thirdly, due to the stochastic nature of microsimulators, model results change with different random seed values (Walker, 2005). As a result, the comparability of model outcomes decreases. Although this limitation can be addressed by performing multiple runs, this requires additional effort.

2.4 Empirical comparison between 4-step model and more advanced models

As discussed in section 2.3, disaggregated, microscopic, and tour-based models have theoretically distinct advantages over an aggregated, macroscopic, trip-based model such as the 4-step model. However, little studies empirically compared these two types of models to evaluate if the theoretical benefits are evident in practice. This section will discuss the few studies found that conducted such a comparison. Table 1 gives an overview of the studies reviewed, together with the most relevant findings. It can be concluded that the advanced models did not demonstrate a better traffic count fit; at best it was similar. This is an interesting finding, since traffic counts are frequently used in practice to evaluate the performance of travel models. Furthermore, it was found that advanced models were more sensitive towards changes in input data. Lastly, the more advanced models allowed for analysing the effects of transport-related changes on certain groups in society.

Table 1 - Studies that compared a 4-step model with a (theoretically) more advanced model

Study	Type of model	Location	Indicators	Findings
Walker (2005)	Disaggregate, trip-based microsimulator	USA	Traffic counts, VMT[1], elasticity	- Similar performance in terms of traffic counts fit; - Diversity of population preserved by microsimulator; - Aggregation bias prevented with microsimulator; - Higher sensitivity of microsimulator for car travel time.
(Lemp et al., 2007)	Disaggregate, tour-based microsimulator	USA	VMT[1], VHT[2], destination and route choice, modal split	- In base scenario, microsimulator shows worse fit of VMT estimations than 4-step - Tour-based showed higher sensitivity for job locations.
Griesenbeck and Garry (2007)	Disaggregate, tour-based microsimulator	USA	VMT[1], traffic counts, modal split, elasticity	- Tour-based predicted VMT closer to observed values; - Similar traffic count fit with RMSE of 33-34%; - Tour-based showed higher elasticities.
Ferdous et al. (2012)	Disaggregate, tour-based microsimulator	USA	Worker flow distribution, work trip travel time	- Observed work flows better predicted by tour-based model; - Similar performance in terms of traffic counts fit.

[1] Vehicles Miles Travelled
[2] Vehicle Hours Travelled

Walker (2005) compared an aggregate four-step model with a disaggregated trip-based household microsimulator, for the Southern Nevada region (USA). Although the focus of this comparison is on tour-based models, Walker found interesting results and was therefore included. The study showed that both models performed similarly in terms of traffic count fit. In addition, three relevant benefits of the advanced model were found. Firstly, the diversity of the population in the output was preserved, which means that the effect of policies or adjustments to infrastructure can be studied at a subgroup level. Secondly, the aggregation bias of the aggregated 4-step model was eliminated. Lastly, a sensitivity analysis was conducted for which car travel times were incrementally increased. Both type of models performed similarly in terms of VMT, but the microsimulator showed a stronger cross elasticity with a higher shift from car to PT trips.

Lemp et al. (2007) conducted a similar comparison as Walker (2005) but used a tour-based microsimulator rather than a trip-based version. A base year was modelled, as well as two scenarios for Austin (USA). For the reference situation, modelled VMT figures were overestimated with 27 percent compared to observed data, while the aggregate 4-step model overestimated by 14 percent. Furthermore, the first scenario modelled a capacity expansion for two key highways in the study area. Both models responded similarly and intuitively: during peak hours the VHT reduced, while VMT increased over all time periods. Moreover, both models determined different destinations and routes in favour of the highways with increased capacity. For the second scenario, job locations were moved from rural (-50% jobs) and suburban (-30% jobs) locations to urban and central business district zones (+58%). The advanced model demonstrated a larger VMT decrease compared to the 4-step model (2,5% and 0,6% respectively). The reason found was that although the commute distance increased, secondary activities were located closer to the work activity which reduced the VMT in the end. Moreover, the 4-step model estimated a similar model split as the reference situation, while the advanced model predicted a 19 percent increase in active mode use. The latter was likely to be caused by households living in areas with higher job opportunities, choosing destinations close to their homes.

Griesenbeck and Garry (2007) evaluated a traditional 4-step model with what they describe as an activity-based model. However, the model does not account for time-space constraints, which is why the model is considered to be a disaggregate, tour-based microsimulator, following the definition of Vovsha (2019). The advanced model predicted 3% higher daily VMT than observed, while the 4-step model predicted 9% lower VMT. In terms of traffic count fit on highways, both models performed similarly with a root mean squared error (RMSE) of about 33%. Furthermore, the sensitivity of both models was tested by means of aggregate point elasticities. The household VMT with respect to household income was +0,08 for the advanced model compared to -0,02 for the 4-step model. The number of public transport trips with respect to household income was -0.17 for the advanced model and -0,06 for the 4-step model. The higher elasticities seemed to correspond with the greater sensitivity of the advanced models reported by Walker (2005) and Lemp et al. (2007). Finally, model output was compared to travel survey data. The analysis demonstrated that the advanced model estimated an increasing share of active modes with increasing population and employment density, which was similar to observed travel behaviour. Similarly, VMT by households decreased sharply when the employment density increased, again in line with the survey data.

Lastly, Ferdous et al. (2012) compared a traditional 4-step model with a disaggregate, tour-based microsimulator by analysing the model outcomes for three highway projects in the metropolitan area of Columbus (USA). The comparison was based on four relevant indicators: work flow distributions, work flow distribution by trip start time of the day, average travel time for work trips, and traffic counts. The advanced model performed better in terms of observed work flows, both in general for work flow trips and when start time was considered. The average travel time estimation for work trips was in most cases either better or similar compared to the 4-step model, evaluated against travel survey data. Lastly, when comparing link volumes with traffic counts, both models performed similarly with RMSE values between 25% and 40%, which are similar to the values found by Griesenbeck and Garry (2007).

2.5 Backcasting: why and how?

The reviewed studies in section 2.4 compared the two types of models for a base year situation. Although relevant, it only demonstrates the base year performance; the situation for which the model parameters were calibrated. As a result, the model's performance in a context other than the base year (referred to as 'longitudinal stability') has not been compared. However, travel demand models are generally applied to predict future situations. Therefore, Roorda et al. (2008) described two methods to assess the longitudinal stability of a model: forecasting and backcasting (see Figure 1). In the former, forecast results are evaluated with observed travel behaviour several years later. As a result, the model's prediction quality can be evaluated (e.g., Shelton et al. (2016). A significant drawback of this approach is that one has to wait until the forecasted year is reached and observed data are available. The second method addresses this issue, because it models a past year: rather than *fore*casting a *back*casting exercise is conducted. The advantage is that already available observed travel behaviour can be used as evaluation data. Although literature acknowledges backcasting as an appropriate technique to assess the reliability of model predictions and the model quality in general, examples of backcasts are scarce (Lange & Huber, 2015).

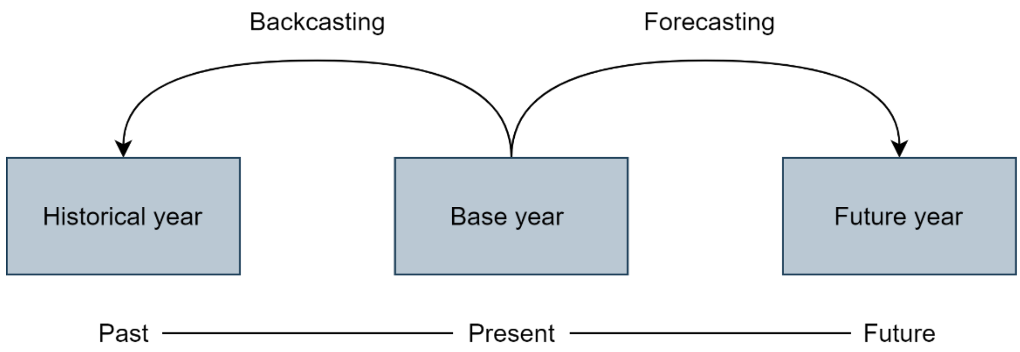


Figure 1 - Backcasting and forecasting illustrated

In literature, three prerequisites for backcasting are described (Sammer et al., 2010, as cited in Lange & Huber, 2015): 1) a calibrated and validated model, 2) historical data as input for and validation of the backcast, and 3) a difference of at least ten years between the backcast and the base year. The latter is required to ensure that changes in terms of demographics, land-use, travel behaviour, etc. are significant enough to determine the model's response.

In addition to these requirements, also a backcasting sequence is described:

1. Model the base year with an established model;
2. Determine input data and identify evaluation data for the backcast year;
3. Model the travel demand for the backcast year;
4. Compare the modelled and observed travel demand;
5. Interpret and document the results.

Note that an 'established model' refers to a model which has been estimated and calibrated. Furthermore, input data for step 2 consists primarily of socioeconomic data and networks for the relevant transport modes. The evaluation data represents data that can be used to assess the model outcomes, and includes observed travel behaviour from travel surveys, traffic counts, boarding and alighting numbers, etc. Examples of studies that performed a backcast are: Gunn et al. (2006), Lange and Huber (2015), and DfT (2020).

3 Research dimensions

3.1 Research aim

From the literature review it became clear that disaggregated, tour-based microsimulators have clear theoretical advantages over aggregated, macroscopic trip-based models of which the 4-step model is the most prominent example. In contrast to theoretical advantages, only few studies empirically compared both types of models. These studies, however, did not analyse how the models compare in a situation deviating from the base year. As a result, the longitudinal stability of the models remains unclear.

The presented research aims to contribute to the transport modelling community by analysing the differences between an aggregated, macroscopic, trip-based 4-step model and a disaggregated, tour-based microsimulator within the application context. Specifically, by comparing the models' performance in a backcast year relative to the performance in the base year, the longitudinal stability of both models is evaluated. To enable a fair comparison, the parameter estimation quality should be considered when analysing the results in the application context, such that the calibration quality itself does not distort the analysis of the models' capabilities. The city of Almere between 2017 and 2010 was used as a case, in combination with two travel demand models developed by Goudappel: an aggregated, macroscopic, trip-based 4-step model and a disaggregated, tour-based microsimulator called Octavius[1].

3.2 Research questions

Based on the research aim, the main question is formulated as follows:

> *How do the effects modelled by Octavius compare to the effects modelled by the 4-step model when conducting a backcast from 2017 to 2010 using Almere as case, and to what extent do these effects correspond to changes observed using travel survey data and literature?*

In line with the proposed backcast sequence (see section 2.5), the first sub-question focuses on the base-year. The extent to which both models are able to replicate the observed situation in 2017 will be used as a reference for the backcast year, such that differences between the model results for 2017 and 2010 (referred to as 'backcast effect') can be isolated:

1. To what extent are the 4-step model and Octavius able to describe the base year (2017) when comparing the model outcomes to travel survey data?

Subsequently, the models will be applied using 2010 input data to analyse how the models behave in a different situation from the one they were calibrated for. The results are considered using the 2017 performance as a reference. This leads to the second sub-question:

2. To what extent do the modelled effects from the 4-step model and Octavius for Almere 2010 relative to 2017, correspond with changes observed in travel survey data between 2017 and 2010?

Lastly, the ability to model specific changes in society (and underlying policy questions) is analysed and compared to literature. Using specific TAZs as cases allows to evaluate changes at a detailed level.

3. How do both models respond to specific changes in society at the TAZ level and how does this compare with literature?

[1] Note that Octavius is also called 'OmniTRANS Horizon'.

4 Model descriptions

In this chapter, a description of the 4-step model and Octavius is given. While both models are strategic travel demand models, they use fundamentally different systematics to determine in the end the number of people that want to travel from origin O to destination D. The first difference is the aggregation level: the 4-step model considers only the aggregated behaviour of average travellers, while Octavius models disaggregated travel behaviour of segments: groups of travellers for whom the same set of variables determine their travel decisions through choice models. Secondly, the 4-step model operates at the macro-level, which means that traffic is modelled at a TAZ level. Octavius on the other hand models agents with discrete travel choices (micro-level), reducing the computational efforts when considering a large number of possible segments. Lastly, the 4-step model uses trips to represent journeys made by people, while Octavius uses tours: chains of trips that start and end at home and which include a primary activity and possibly a secondary activity which is subordinate to the main activity. As a result, trips within a tour are consistent in space and mode choice, in contrast to trips.

4.1 4-step model

The 4-step model of Almere is structured in three parts: trip generation, a gravity model which models the destination and mode choice simultaneously, and route choice. However, the scope of this research is transport demand modelling, which means that any effect of route choice on travel demand is distorting the comparison. Therefore, route choice is excluded from this study. The trip generation describes the number of trips produced and attracted by each TAZ. It is defined per purpose, both from home and to home (e.g., home-work and work-home), and distinguishes between car availability and non-car availability. The purposes considered are work, business, shopping, education for 18 years and older, education for 17 years and younger, and other. The trip generation is used as input for the gravity model that simultaneously distributes produced trips among destinations and travel modes whilst satisfying both zonal production and attraction constraints.

4.1.1 Estimation context

The parameters for the trip generation and the gravity model were estimated and calibrated during the original model construction at Goudappel. Note that the estimation and calibration process are performed simultaneously and use the same dataset: weighted[2] Dutch national travel survey data (*OViN*) from 2010 up to and including 2017 with the origin or destination being in Almere. See the documentation provided by CBS (n.d.-c) for more information on OViN. The 4-step model used in the presented study contains three parameters: the trip generation coefficient and the alpha and beta used in the lognormal distribution function of the gravity model.

The trip generation parameters were calculated by dividing the number of trips found in OViN over a variable of interest. These variables come from Almere-specific socioeconomic data, meaning that the estimation and application context overlap. Similarly, the alpha and beta in the lognormal deterrence function used in this gravity model were estimated and calibrated using the modal splits and trip length distributions found in OViN for Almere 2010-2017. The considered modes are car, public transport (PT), and bicycle. For PT, access and egress transport is modelled separately, considering walking and cycling as available options. Note that walking is only considered as possible access/egress mode for PT, while cycling is both an option as access/egress mode as well as a separate mode for entire trips.

[2] Note that weighted means that the sampling results are extrapolated to find a representative dataset for the whole of the Netherlands. See the website of the CBS for more information (n.d.-a).

4.1.2 Application context

The 4-step model with the estimated parameters was subsequently used in combination with the socioeconomic data and networks of Almere. The data collection itself will be discussed in the research methodology (section 5.1), but here the required socioeconomic input data are presented in Table 2.

Table 2 - Socioeconomic variables required for the 4-step model in Almere

Variable name	Variable name
Households	Student enrolments 18-
Residents	Student enrolments 18+
Residents 0-17 years old	Retail jobs
Residents 18-24 years old	Jobs in total
Labour force	

4.2 Octavius

Octavius is a disaggregated tour-based microsimulator developed by Goudappel, which consists currently of three modules: population synthesizer, tour generator, and tour simulator (Brederode et al., 2020). The population synthesizer creates agents with characteristics on a personal and household level. For most of the choice models in the other modules, agents with similar characteristics for specific variables such as age, gender, or purpose of last visited location, are grouped into so-called segments. These segments can differ between modules, depending on the variables that are significant within the relevant choice model. Subsequently, the tour generator determines the tour frequency – which is limited to 0, 1, or 2 tours – and the main and secondary activity (if applicable). The tour simulator models after that the destination(s) the agent will visit to conduct these activities and the mode(s) of transport used. Figure 2 illustrates the process of finding the destinations for a 3-trip tour. The main destination is determined in step 1 based on the residential location of the specific agent, after which the destination of the secondary activity is chosen based on the residential and (in this case) work location. The trips are subsequently combined to one tour, omitting one trip of step 1 (could be the return trip but also the first trip, depending on the order of the first and second purpose). The result of the destination choice is the location of the main activity and secondary location (if applicable), *per mode* (see Figure 3). Lastly, the probability of choosing each mode with the corresponding destinations is calculated. Note that for car driver it is checked whether the agent has a driving license and a car in the household. Similarly, for car passenger the household should have a car and an (other) adult. Rather than discretising the mode choice into one mode, the probabilities are directly used as (continuous) trips. The output of Octavius are agents with synthetic travel diaries.

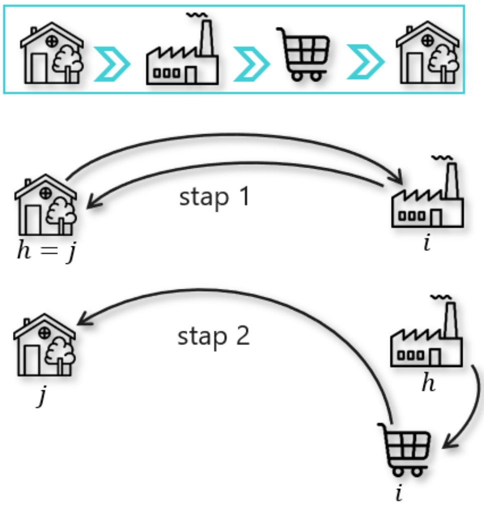


Figure 2 - Octavius' destination choice order (Goudappel, 2022)

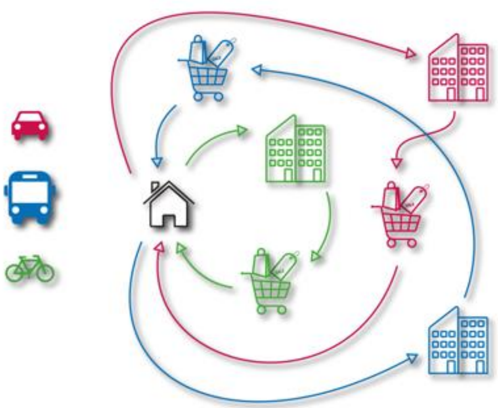


Figure 3 - Octavius' destination choice result (Goudappel, 2022)

4.2.1 Estimation context

Within the estimation context of Octavius, variables which were found to have a statistically significant effect on travel behaviour in one or more of the Octavius' modules were included. For the population synthesizer, personal and household distributions were extracted from national OViN 2010-2017 data (CBS, n.d.-c) to define synthetic agents and households distributions. To assign agents to households, household composition distributions were taken from the Netherlands Mobility Panel (MPN). For the other modules of Octavius, logit choice models were used to distribute agents over a given set of discrete choice alternatives. The parameters and variables for the choice models used in the tour generator and simulator – describing agents' travel choices – were estimated on observed travel behaviour found in national OViN 2010-2017 (CBS, n.d.-c) and CBS microdata 2011-2017 (CBS, n.d.-b). The modes of transport included in the estimation context of Octavius are car driver, car passenger, PT, bicycle, and bicycle-car passenger.

4.2.2 Application context

After the estimation context, Octavius uses input data and networks to calculate travel demand. The input data can be categorised into different types of data. For the population synthesizer, the characteristics of individual agents and households in which they live are required. In practice this means that for each zone the total number of females, residents, students, etc. is part of the input data. Based on these subtotals per zone, in combination with a seed population distribution from OViN and a seed household composition from the MPN which were determined in the estimation context, characteristics are assigned to each agent and household. In Table 3, the socio-economic variables for the population synthesizer are listed. Note that the variables are only combined into segments if they exist in reality. For instance, gender (2 categories) and household size (6 categories) results in 12 classes, but the segment 0-17 years with driving license is omitted as it does not exist. This yields 180 person segments and 216 household segments. To be comparable to the 4-step model, the modes car (including car driver and car passenger), PT, and bicycle were included in the application context.

Table 3 – Personal and household variables included in Octavius' population synthesizer

Individual class	Variables	Household class	Variables
Age groups	0-17	Household size	1 person
	18-29		2 persons
	30-44		3 persons
	45-64		4 persons
	65 and older		5 persons
Gender	Male		6 persons or more
	Female	Household structure	single
Ethnicity	Dutch		Without children
	Western		With children
	Non-Western	Number of cars in household	0 cars
Driving license	Driving license		1 car
	No driving license		2 cars
Social participation	Employed		3 cars or more
	Students		
	Other		

In addition to the input data for the population synthesizer, data on size variables – which describe partially the attraction of destinations – are required (see Table 4). In contrast to the 4-step model, these variables are not modelled as constraints. In addition to the size variables, also the person and household characteristics determine the attraction of a destination and the mode choice.

Table 4 – Input data for Octavius' destination choice

Variable class	Subtypes	Variable class	Variables
Jobs	Education Industry Office Retail Other Total number	Student enrolment	Primary education Secondary education Intermediate vocational education (Dutch: MBO) Higher education (Dutch: HBO and WO) Total number
Other	Parking costs	Household density	Number of households per km^2

4.3 (Post-)processing operations

In practice, various (post-)processing operations are performed to closer match observed travel behaviour. However, not all those adjustments should be considered within the scope of this study. In this section, the most relevant changes are discussed. Firstly, to consider the higher speeds of e-bikes compared to regular bikes, the calculated travel times for the mode bicycle are decreased within both models. For three distance classes, the e-bike and regular bike shares are defined (see Table 5). These shares have been empirically determined by Goudappel (2018) for the region The Hague – Rotterdam and are assumed to be representative for the whole of the Netherlands for 2017. The final travel time matrices are the weighted mean of the regular and e-bike travel times. For 2010, the e-bike shares are considered to be zero (BOVAG, 2022). The e-bike correction was applied in this study, as it potentially improves the prediction quality without affecting the model systematics itself.

Table 5 - E-bike and regular bike shares per distance class for 2017

Distance class [km]	E-bike share	Regular bike share
0 – 2.5	5%	95%
2.5 – 7.5	10%	90%
+7.5	25%	75%

Secondly, OD-matrices resulting from a travel demand model are in practice often modified to better match traffic counts. This post-processing operation was not used within this study, as it is unrelated to the working of both models, traffic assignment is not within the scope of this research, and would distort the comparison between the model systematics. Thirdly, after the standard travel demand modelling, additional trips are generally added for specific Points of Interest (POIs). For these POIs, the number of attracted trips is in reality significantly higher than expected based on the averaged attraction variables. For instance, IKEA is significantly more often a destination than expected based on the number of jobs. Although this post-processing operation is defendable, it is not an integral part of the travel demand model. Therefore, this modification was not used within the presented study.

4.4 Qualitative comparison between both model types

In addition to the quantitative approach of the presented study, two advantages of the disaggregated, microscopic, and tour-based nature of Octavius are worth mentioning compared to the traditional 4-step approach. Firstly, Figure 4 demonstrates that Octavius models tours, where the lines demonstrate the number of trips between one origin and all destinations in Almere. As can be seen, Octavius models chains of trips, while the 4-step model demonstrates merely individual trips. The second advantage of Octavius is that it models explicitly non-home-based trips, such as from work to the store, or from school to meeting friends, while the 4-step model considers these trips as 'other'. Hence, the 4-step model approach results in a loss of detail.

Figure 4 - Example of how journeys are modelled by Octavius and the 4-step model (from one centroid). Note: the lines do not represent route choice but OD-relations; thicker lines mean more trips between a certain OD-pair.

5 Research methodology

Figure 5 presents the research methodology graphically. The first step was to model the base year 2017, such that it can be used as a reference for the backcast. As explained in section 4, the model parameters of both models have already been estimated during the construction of the models by Goudappel and were directly adopted for this study (light grey part of Figure 5). However, since the original socioeconomic data were not adequate (primarily due to inconsistent data collection and insufficient level of detail; further explained in section 5.1), new data were collected. As a result, new trip generation parameters for the 4-step model were estimated. Additionally, Octavius' choice models were adjusted to better represent Almere-specific travel behaviour; a process called *regionalisation*. After applying the models with the re-estimated parameters, together with the socioeconomic data and networks for the base year 2017, the results were compared with observed travel behaviour from OViN using three Key Performance Indicators (KPIs): trip frequency, modal split, and trip length distribution. Due to the relatively small sample size of OViN, the number of data points for Almere was limited. To increase the number of observations, and simultaneously decreasing data uncertainty, four years of OViN were combined: from 2014 up to and including 2017. The fit to OViN figures described the quality of the parameter estimation and was used as a reference for the backcast to 2010.

Subsequently, the models were applied to 2010 which required socioeconomic data and networks for that year. The closer socioeconomic data and networks corresponded with the real situation in 2010, the better it could be analysed to what extent the model systematics were able to reproduce travel behaviour for 2010 without having the noise of data inaccuracies. In other words, minimising the deviation from reality in the input contributed to isolate and analyse the effect of modelling a non-calibrated year. Therefore, the fourth step was to compare the model output to OViN 2010 to 2013 using the same assessment framework as for 2017: trip frequency, modal split, and trip length distribution. However, this time the 2010 OViN fit was analysed relative to the 2017 OViN fit (referred to as 'backcast effect'). Hence, if a model demonstrated an equal deviation in trip frequency in 2010 compared to OViN 2010-2013 as it showed for 2017, the longitudinal stability was outstanding; same applies to the modal split and trip length distribution. This also means that two equally poor OViN fits for 2017 and 2010 is better than a poor OViN fit for 2017 and a good OViN fit for 2010. Lastly, the disaggregated approach of Octavius allowed for modelling travel behaviour of specific groups. To analyse if the models responded consistently with literature, two hypotheses were tested: 1) an increase in immigrants results in a decrease of bicycle use (Harms, 2006) in Octavius, but not in the 4-step model as it does not consider ethnicity, and 2) an increase in car availability results in more car use (KiM, 2016) within both models, but in Octavius more pronounced because it considers car ownership explicitly in its mode choice calculation.

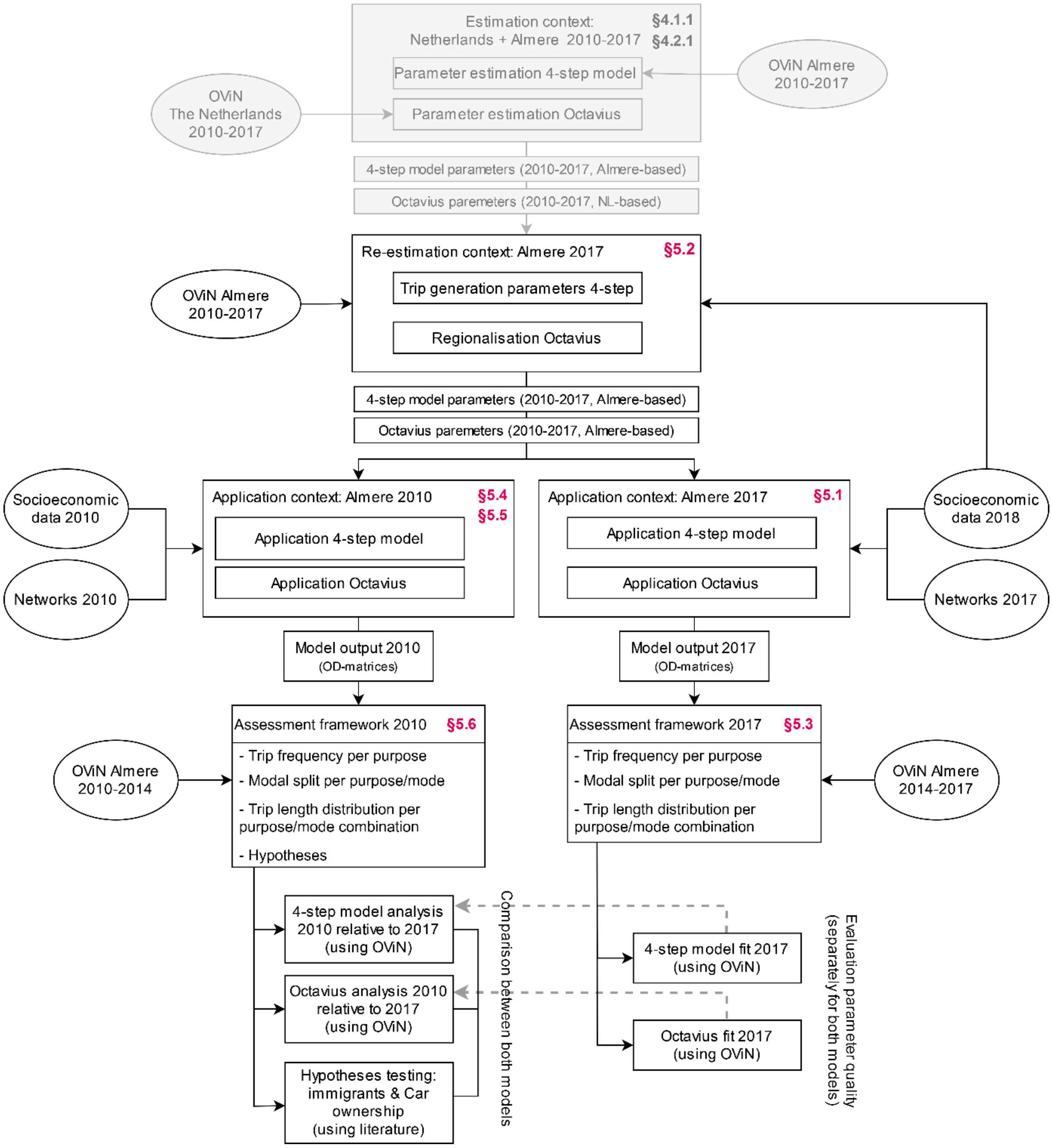


Figure 5 – Research methodology illustrated

5.1 Input data collection for Almere 2017

Both travel demand models had the same topography and were structured around Almere. The models consisted of 1400 TAZs, which are used as origins and destinations (see Figure 6). The smallest TAZ is located in the built-up area of Almere, while the TAZ size increased as the distance to Almere increased.

Furthermore, the networks for different modes of transport were available from the original 4-step model developed by Goudappel. Networks for car, PT, bicycle, and walking were used for both models. Note that walking is only considered as possible access and egress mode for PT, while cycling is both an option as access/egress mode as well as a separate mode for entire trips. The network for car is depicted in Figure 7, while Figure 8 demonstrates the PT network.

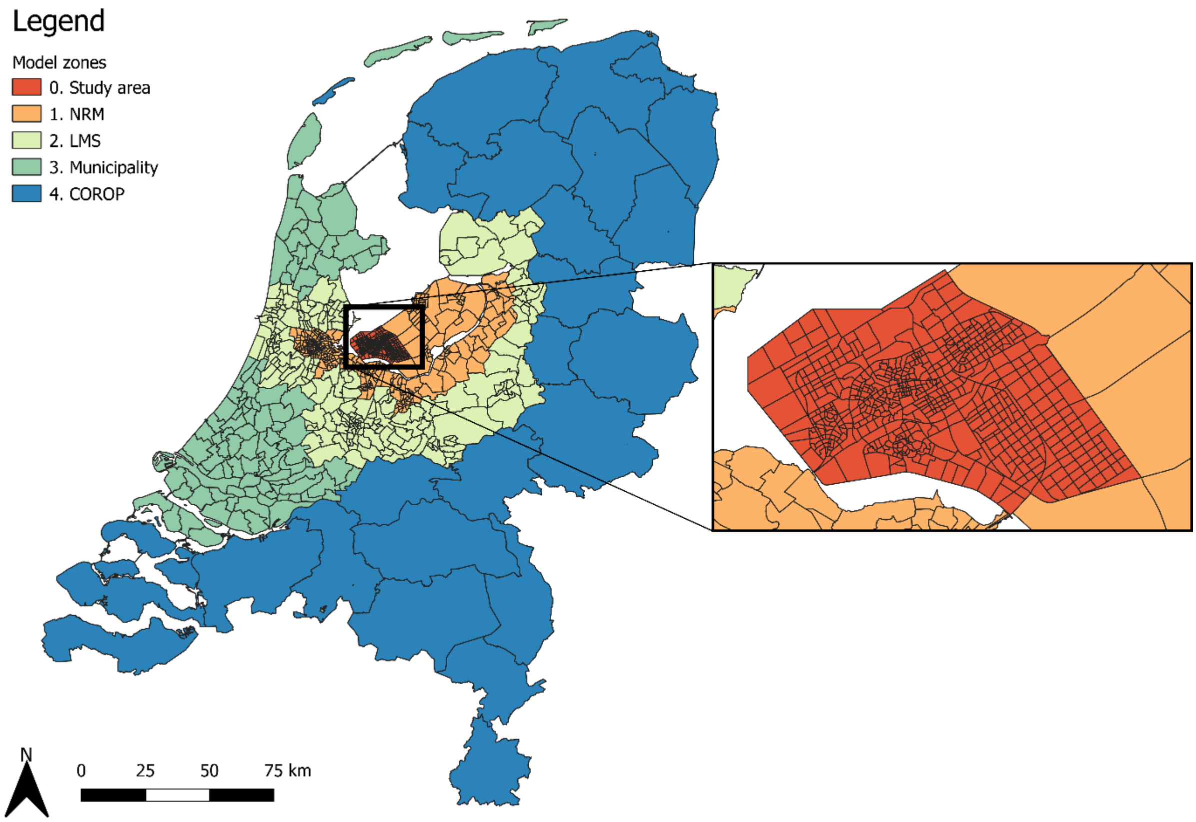


Figure 6 – Traffic Analysis Zones (TAZs) at different levels

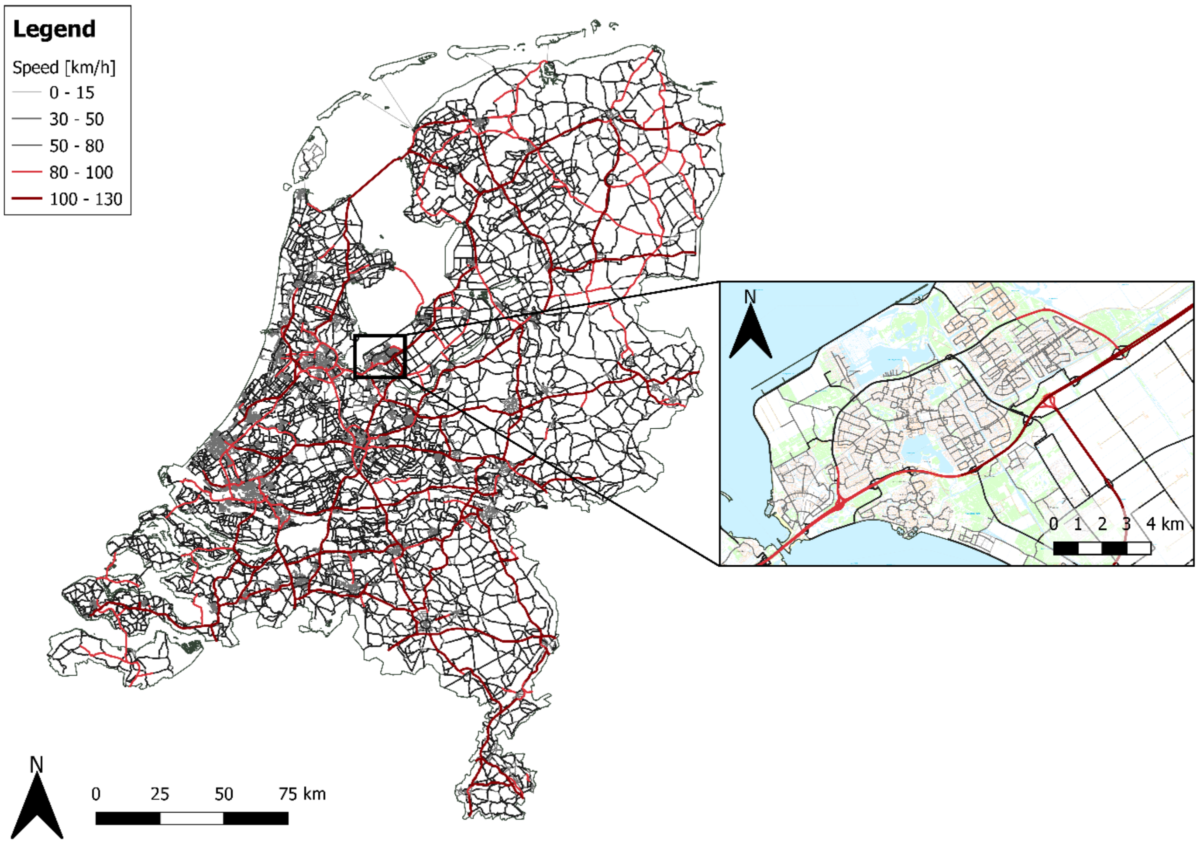


Figure 7 - Car network used in this study

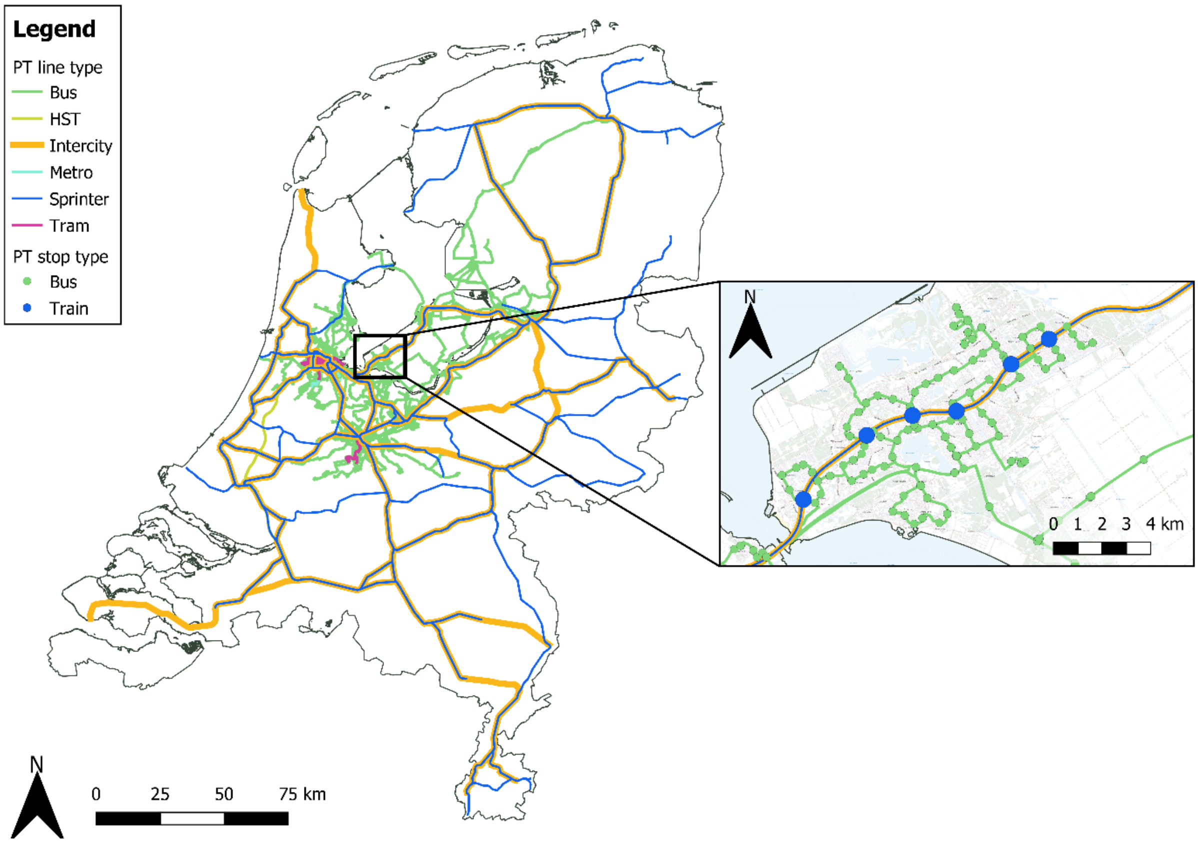


Figure 8 - Public Transport network used in this study

The main principle for the socioeconomic data collection was that for both models the input data were collected and prepared in a consistent manner. If inconsistencies between input data existed, differences in results could be caused by the input data rather than the systematics of the models, while the aim of this research is to compare the model systematics. However, the socioeconomic data collected during the original model construction originated from various sources with inconsistent definitions and was at the aggregation level of the 4-step model, whereas Octavius requires more detailed data on a personal and household level. Therefore, data collected within OmniTRANS Spectrum (Goudappel, n.d.) was used, which gathered and combined data from different data sources in a consistent and reliable manner (see Table 6). In addition, Spectrum data could be used for all variables required for both models. The Spectrum data were only available for the years 2018 and 2020 during the time of this research, which meant that 2018 socioeconomic data were used.

Table 6 - Data sources used within OmniTRANS Spectrum

Data source	**Used for**
CBS *(Centraal Bureau voor de Statistiek)*	Demographic characteristics: • Inhabitants • Households • Gender • Age group • Ethnicity • Household types (with/without children, persons and cars) • Driving license ownership • Social participation (working, studying, other) • Household density

Table continues on next page

LISA (*Landelijk Informatiesysteem van Arbeidsplaatsen)*	Number of jobs: • Education • Industry • Office • Retail
DUO (*Dienst Uitvoering Onderwijs)*	Student enrolment: • Primary education • Secondary education • Intermediate vocational education (Dutch: MBO) • Higher Professional education (Dutch: HBO) • University (Dutch: WO)
BAG (*Basisregistratie Adressen en Gebouwen*)	Address information: • Number of addresses • Floor area per address

5.2 Re-estimation of selected model parameters for Almere 2017

5.2.1 Re-estimation of trip generation parameters 4-step model

Since the trip generation parameters were determined using the original socioeconomic data (see section 4.1.1) and these data were replaced (see section 5.1) a re-estimation was required for the 4-step model. The procedure was equal to the original: the number of weighted trips found in OViN divided by the variable of interest (see eq. 1). Note that OViN is divided by a variable to ensure that the total number of modelled trips matched the total number of weighted trips found in OViN. To keep the model parameters consistent, OViN 2010-2017 was used: the same dataset on which the lognormal distribution function was calibrated during the original construction of the model.

$$P_{p,a,c} = \frac{T_{p,a,c}}{V_{p,a}} \tag{1}$$

Where:

- P the trip generation parameter;
- p trip purpose;
- a area of interest (Almere or rest of the Netherlands);
- c car available (yes or no)
- T number of weighted trips found in OViN 2010-2017;
- V variable of interest.

For instance, to determine the trip generation parameter for home-work trips for people that have a car at their disposal, the number of weighted OViN trips was divided by the number of workers in Almere:

$$P_{work,Almere,car\ avabile} = \frac{39301}{94759} = 0{,}415\ trips\ per\ worker \tag{2}$$

To analyse the effects of re-estimating the trip generation parameters of the 4-step model, the before and after OViN fit was evaluated at the level the parameters were estimated (i.e., all trips and all purposes), using the OViN data they were evaluated on. In that way, it was assessed whether the re-estimation had the required effect.

5.2.2 Regionalisation Octavius

As mentioned in section 4.2.1, Octavius applies parameters estimated from national OViN data in its choice models. However, (average) national travel behaviour is different from regional travel behaviour. This is particularly relevant for Almere, due to its location close to Amsterdam, relatively high share of households compared to the number of jobs and commercial area, and the fact that it is surrounded by nature. Therefore, Alternative Specific Constants (ASC) were added to the utility functions such that Almere-specific travel behaviour was modelled; a process called *regionalisation*. Note that the regionalisation is applied to the mode and destination choice. The ASCs were calculated by a mathematical solver which minimised the difference between modelled and OViN 2010-2017 modal split and trip length distributions for Almere. The reason to use 2010-2017 OViN data was to keep the utility functions consistent, as the parameters are estimated on this dataset as well. The mathematical solver was available at Goudappel and was applied without modifications in the presented study. The effect of regionalising was evaluated at the same level as the estimation: tours.

5.3 Assessment framework 2017

The 4-step model with the re-estimated trip generation parameters and regionalised Octavius were subsequently evaluated using the assessment framework. By analysing the 2017 results, a reference was created against which the results of the backcast were compared. The goal of the presented study was to compare the model systematics rather than only the model outcomes. Although both models determine travel demand, the approach is fundamentally different: the 4-step model works at an aggregated, macroscopic, trip-based level, while Octavius operates at a disaggregated, microscopic, tour-based level. As a result, comparing the models can only be done at a level which both models can describe. In this case, the 4-step model determined the possible aggregation level to which the more disaggregated Octavius results needed to be aggregated. This also means that the purposes *social-leisure* and *other* were omitted, because *social-leisure* was not defined in the 4-step model of Almere, and *other* was differently defined between the two models. To maintain consistency between the models, both purposes were kept outside the scope of the comparison. Moreover, the purposes education 18+ and 18- were merged in the 4-step model, since Octavius applies only one education purpose. Similar to the aggregation level, the tours of Octavius were separated into trips, since the 4-step trips cannot be combined into tours due to the unknown relation between trips, or rather: the non-existing relation. It does not exist because the 4-step model uses trips as modelling unit rather than individual agents. For home-based trips this was not problematic, as the place of residence was determined by combining the purpose with the origin or destination. For instance, home-work trips that start in Almere or work-home trips that end in Almere disclosed that these trips belong to Almeerders. However, for non-home-based trips this was not possible, because within the 4-step model it was unknown where the homes of these individuals were located. Therefore, the comparison between the two models was based on home-based trips. This aggregation level contained the most information that both models can possibly describe.

5.3.1 Key Performance Indicators 2017

The Key Performance Indicators (KPIs) were selected based on the possible aggregation level. The first KPI was trip frequency and described the absolute number of home-based trips made by Almeerders for the purposes: work, education, business, shopping, and the four purposes combined. The analyses were conducted using graphs which described the number of OViN trips per purpose including the confidence interval (see section 5.3.2 below) and the number of modelled trips by both models. The second KPI was the modal split determined per purpose-mode combination. Since trip frequency described the performance in absolute trip numbers, the modal split compared the relative shares of modes among the different purposes; for instance, the shares of car, PT, bicycle for work trips. The results were evaluated using graphs that demonstrated the distribution over the modes per purpose for OViN (including confidence interval) and both models. Lastly, the trip length distributions were calculated per mode-purpose combination and presented the relative distribution of trips per distance class, separately for each mode and purpose combination. This KPI compared the destination choice between OViN (again including confidence intervals) and both models.

5.3.2 Evaluation data 2017

Observed travel behaviour data from OViN were used to evaluate the plausibility of the model outcomes. The reason to use OViN as evaluation data is because no other data sources were available containing sufficient data on travel behaviour of Almeerders. Nonetheless, OViN 2017 contained information on merely 500 trips made by 190 respondents. To decrease data uncertainty, multiple years of OViN were combined. Since OViN was conducted from 2010 up to and including 2017, and a trend break occurred with its predecessor (MON) and successor (ODiN), the years 2010-2017 were split in half to use as evaluation data: 2010-2013 for 2010 and 2014-2017 for 2017. This resulted in 3047 trips made by 1218 respondents for 2014-2017[3]; a significant increase compared to 2017 only.

Furthermore, trips were grouped into tours when the location and time matched logically and if the tour started and ended at home, such that complete and realistic travel behaviour was obtained. Using the trip purposes defined in OVIN, the activity at the origin and destination was known. Subsequently, the evaluation data were filtered on home-based trips according to the assessment framework (see 5.3.1), based on the activity at the start and end location. The final selection for 2014-2017 home-based trips included 2884 trips. Note that this number represents the OViN sample size, while for the evaluation the sample was extrapolated to obtain a representative dataset for Almere. The weights for the extrapolation are determined by CBS (n.d.-a), and results in a *weighted* OViN dataset .

To gain insight into the uncertainty of the evaluation data, confidence intervals (CI) were calculated around the OViN 2014-2017 data. The methodology described by Pots (2018, p. 26) was adopted as it was specifically developed for determining CIs around OViN trip data. The CI was calculated following eq. (3), using a standard normal distribution to approximate a binomial distributed variable:

$$p_i = \hat{p}_i \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{N}} \tag{3}$$

Where:

- p_i true fraction of trips made within class i
- i bin, which can be mode for modal split or distance class for trip length distribution
- $\hat{p}_i$ sampling fraction for bin i
- $z_{\alpha/2}$ $(1-\alpha/2)$-percentile of the normal distribution
- N the total number of trips

However, a dependency exists between the outward and return trip of a 2-trip tour: the trip lengths will be approximately similar, and the mode will be equal as well for most of the observations. Therefore, the number of observations N is reduced by a factor 2 (see eq. 4). It was argued that eq. (4) can be applied to the home-based trips in a 3-trip tour as well, because the secondary (non-home-based) trip generally occurs close to the home or primary location.

$$p_i = \hat{p}_i \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{N}}\sqrt{2} \tag{4}$$

Using eq. (4), the difference in data uncertainty between OViN 2014 up to and including 2017 and OViN 2017 for Almere in terms of modal split per purpose was calculated (see Figure 9). As can be seen using only one year of Almere OViN data resulted in unacceptable wide 95% confidence interval. On the other hand, four years combined resulted in acceptable CI, except for business trips.

[3] Note that this number was found after initial processing by Goudappel, including working days only and removing trips made by people that toured for leisure purposes without stopping somewhere.

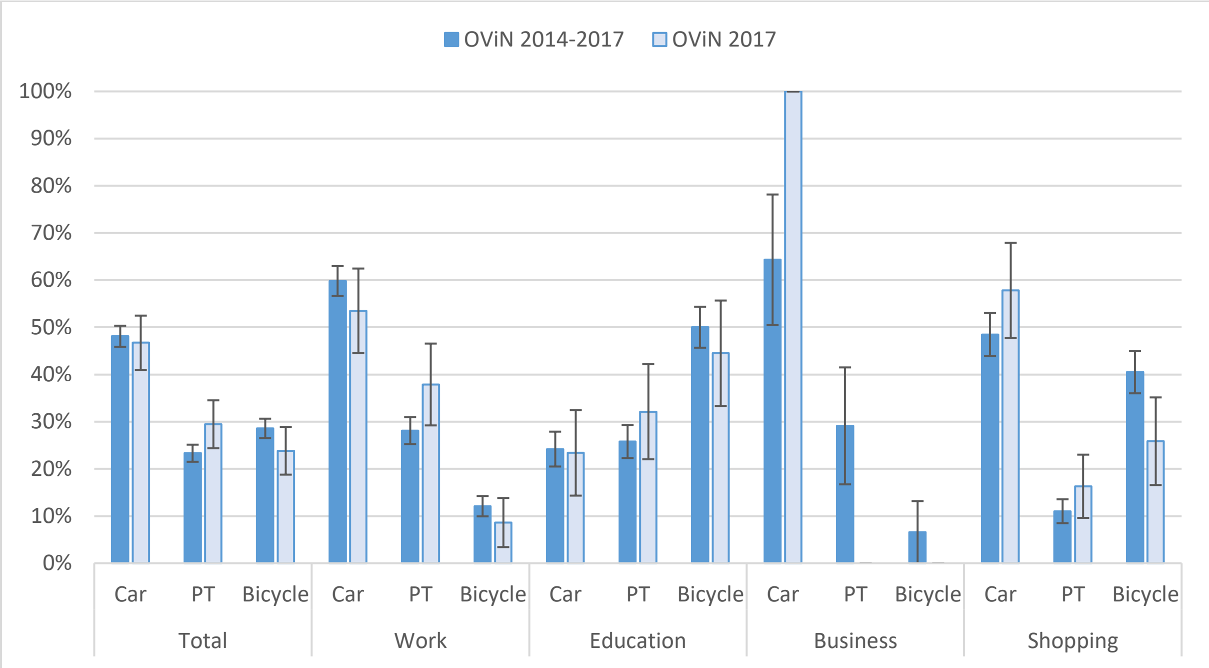


Figure 9 – Modal split in OViN 2014-2017 and OViN 2017 for Almere with 95% confidence interval

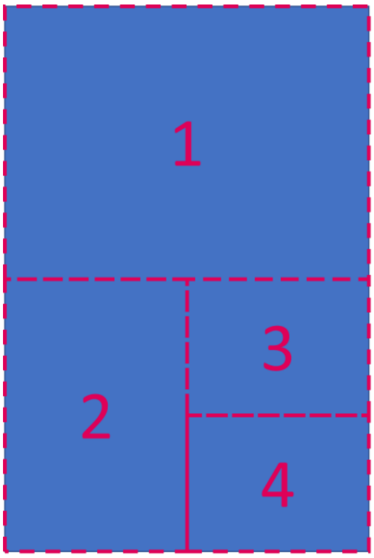
The result of applying the assessment framework was the OViN 2014-2017 fit of the 4-step model and of Octavius measured against the three KPIs. It described the performance in the base year, which was used as a reference for the backcast results. Since the calibration of the model parameters determines the OViN base year fit, the result is a measure of parameter estimation quality.

5.4 Input data collection for Almere 2010

After modelling and evaluating the base year 2017, the backcast to 2010 was performed. The first step was to collect input data for 2010. Similar to the base year, the aim was to collect and prepare input data for both models in a consistent manner. However, OmniTRANS Spectrum was not available for 2010. Therefore, the required data were gathered either from the same source used in Spectrum or the Spectrum data were translated from 2017 to 2010. As a consequence, the same variable definitions were used for 2017 and 2010 resulting in consistent data. In the following sections, the data collection is discussed per variable (group).

5.4.1 Demographic data (CBS)

CBS (2013) provides extensive 2010 demographic data in the form of *CBS buurt/wijk/gemeente* (*neighbourhood, district, municipality*). Although CBS neighbourhood data were defined at a detailed spatial level, a significant part of the TAZs were still smaller. As such, a connection between intersecting CBS neighbourhoods and TAZs was made. Since no information on the distribution of any demographic characteristic within the CBS neighbourhood was provided, a homogeneous distribution is assumed to be able to make the division into smaller zones based on the area of intersection. This process is depicted in Figure 10, where the number of inhabitants is distributed among the TAZs according to the overlap with the neighbourhood. Further away from the study area, the TAZs are larger than the neighbourhoods. In these situations, the neighbourhoods located in the TAZ, are combined. In reality, most neighbourhoods overlap with multiple TAZs. Therefore, the percentage overlap of each neighbourhood on each TAZ was determined based on the intersection between them.



	Area (%)	Inhabitants
Buurtzone	100	500
Zone 1	50	$0{,}50 * 500 = 250$
Zone 2	25	125
Zone 3	12,5	62,5
Zone 4	12,5	62,5

Figure 10 – Illustration of the distribution of CBS neighbourhood data among TAZs

The following variables were taken from CBS neighbourhood data: inhabitants, households, one-person households, households without children, households with children, cars, males, females, western immigrants, non-western immigrants, inhabitants 0-14 years, inhabitants 15-24 years, inhabitants 25-44 years, inhabitants 45-64 years, and inhabitants 65 years and older.

5.4.2 Driving license

To determine the number of agents with a driving license for each TAZ, the trend in driving license ownership per age group in the Netherlands was determined. Data from CBS (2022a) was available as of 2014. Additional data were found in a report by Goudappel (2012), concerning driving license ownership in 2007 per age group. Combining these data resulted in Figure 11, which was used to interpolate the ownership levels in 2010. Lastly, the percentages were multiplied with the age groups from CBS neighbourhood data (see section 5.4.1) to obtain the 2010 figures.

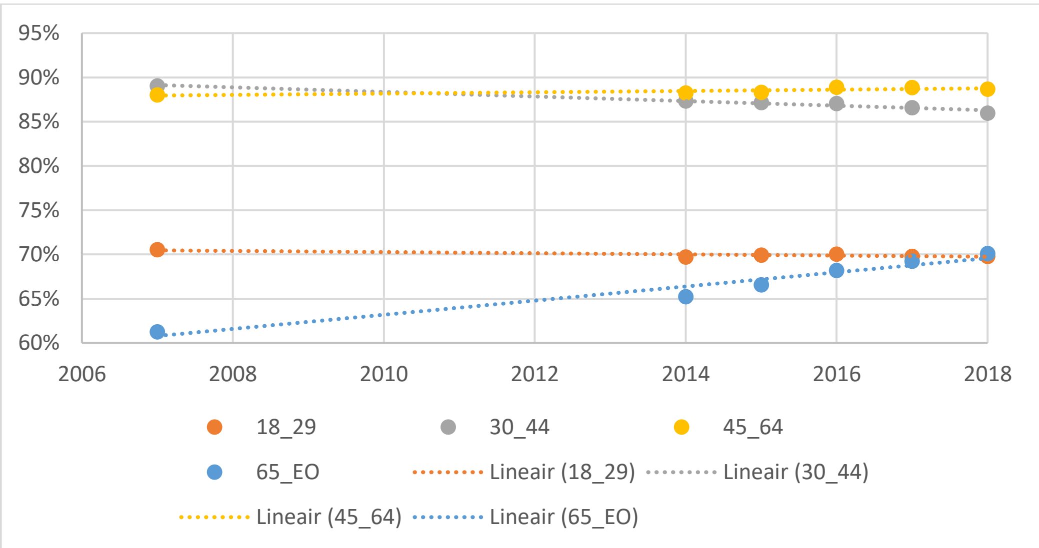


Figure 11 - Driving license ownership per age group in the Netherlands

5.4.3 Household characteristics: number of people and cars per household distribution

In addition to the number of households taken from CBS (see section 5.4.1), also the number of people in each household per TAZ was required to find the number of households with one person, with two persons, etc. However, the total number of households and inhabitants should match the subtotals determined in section 5.4.1. Therefore, a mathematical solver was developed that used the number of inhabitants and households as constraints and minimised the Sum of Squares Error (SSE) compared to the 2017 distribution for that TAZ. Hence, it was assumed that the distribution of number of persons per household (i.e., number of households with one person, with two persons, etc.) per TAZ was similar in 2010 compared to 2017.

Similarly, the number of cars per household was determined (i.e., number of households with one car, with two cars, etc.), using the same mathematical solver as in the previous paragraph considering the number of cars and households per TAZ as constraints. Furthermore, Octavius uses the category 'household with 3 cars *or more*'. The average number of cars for those households was set at 3.3, based on KiM (2022).

5.4.4 Household density

The household density for 2010 was initially calculated based on 2010 CBS neighbourhood data. However, due to a different spatial scale of CBS data compared to the TAZs, the result was insufficiently detailed because all TAZs within one neighbourhood received the same household density (see Figure 12). Therefore, the 2017 data from Spectrum were used for 2010, but corrected for the construction of Almere Poort which was developed between 2010 and 2017: TAZs located in this neighbourhood were manually set to 'not urbanised'. Based on the same CBS neighbourhood data, no other significant changes in terms of household density were found in Almere.

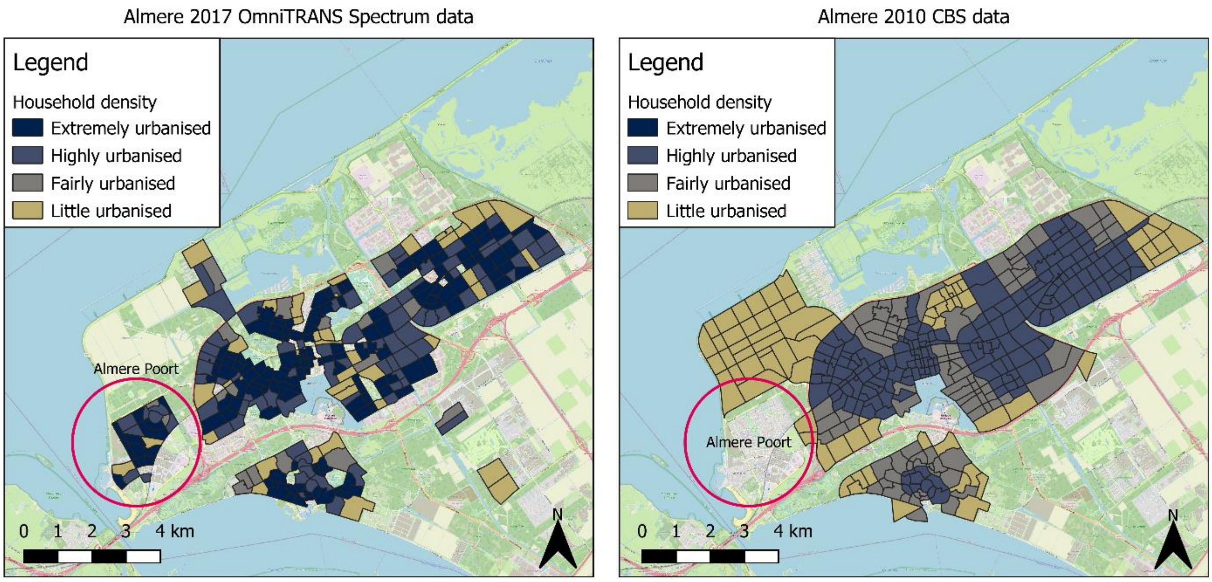


Figure 12 - Comparison household density using Spectrum and CBS data

5.4.5 Social participation

Detailed social participation data were not available for 2010. As a result, the number of workers, students, and other were determined using the following methodology. For workers, the relative labour force per TAZ for 2017 was calculated by dividing the number of workers over the number of people aged 15-65 years. The resulting percentage was subsequently applied per TAZ to the number of people aged 15-65 years in 2010, which was determined in section 5.4.1 by using CBS data. If a TAZ contained no labour force in the 2017 but did in 2010, the average relative labour force for all TAZs in 2017 was used to multiply with the number of people aged 15-65 years in 2010.

Similarly, the share of students per TAZ was calculated for 2017 as a fraction of the total number of inhabitants. Although the age group 4-25 could have been used, it was decided to use all residents as reference, since students are not necessarily tied to a certain age group. The determined fraction was subsequently applied to the number of inhabitants per TAZ in 2010, which was determined in section 5.4.1. Lastly, the group ‘other’ was subsequently determined by subtracting the calculated number of workers and students from the total number of inhabitants.

5.4.6 Job figures

For 2017, the number of jobs was based on LISA data (see Table 6). However, no 2010 data were available, which forced the use of other data. To keep the definitions of job categories (e.g., office, retail, industry, etc.) identical to 2010, it was chosen to translate the job figures from Spectrum 2018 to 2010, using the relative difference in total number of jobs per municipality of CBS (2022b). Although Spectrum used LISA figures, it is assumed that the relative differences between 2010 and 2018 in CBS data also applies broadly to LISA data. A limitation of this approach is that for all TAZs within a municipality, the relative difference in jobs is assumed to be equal. See Appendix B.1: Number of jobs for a more in-depth discussion of the applied corrections.

5.4.7 Student enrolments

One of the attraction variables within both models is the number of student enrolments. In the Netherlands, these numbers are registered at *Dienst Uitvoering Onderwijs* (DUO). Since the required data were not publicly available, they were requested directly from DUO. For primary and secondary education, the process was straightforward since the number of students enrolled per location were provided, including the address and postal 6 code. Combining the postal codes with the ESRI postal code shapefile (2022), ensured that each location was spatially joined with the corresponding TAZ. The same procedure was performed for secondary education enrolments.

The process was different for other levels of education, because DUO registers all students at the main location of organisations. For instance, all students of *ROC van Twente* – which has multiple locations in among others Almelo, Hengelo and Enschede – are registered at the main location, which is in this case Hengelo. Obviously, this distribution would not be suitable for determining travel demand. Therefore, the number of enrolments was distributed among the main and secondary locations, proportionally to their floor area from BAG. These distributions were taken from OmniTRANS Spectrum 2018, which was consistent with the base year data processing process. Nevertheless, some manual adjustments were required because between 2010 and 2017 a number of educational organisations have changed. These adjustments are described in Appendix B.2: Student enrolments.

5.5 Network changes 2010

Since the networks for car, PT, bicycle and walking have changed between 2010 and 2017, some network characteristics required adjustments; not only in terms of infrastructure, but also in terms of speed limit (car), frequency (PT) and the addition of new stops (PT). Despite the existence of 2010 networks for car, PT, and bicycle at Goudappel, it was technically infeasible – considering the time constraints of this research – to use these networks for the backcast. Although some issues were resolved, others required extensive knowledge about OmniTRANS. Therefore, it was decided to adjust the 2017 network to the 2010 situation considering the most influential changes found between both networks. The main adjustments were related to the PT network at Almere Poort, including the removal of the trains station and the train line ‘Hanzelijn’ between Lelystad and Zwolle. With respect to the car network, some speed limits were adjusted. In Appendix C – Network adjustments for 2010, all adjustments are listed. Furthermore, the e-bike correction (see section 4.3) was deactivated since the share of e-bikes in 2010 was limited (BOVAG, 2022). Additionally, the parking costs were adjusted using the figures from the previous Almere model with base-year 2010 (see Table 7).

Table 7 - Parking costs 2017 and 2010

Region	Parking costs 2017 [€/h]	Parking costs 2010 [€/h]
City centre of Almere	2,79	2,00
Edge of city centre of Almere	-	1,00
Almere Buiten	1,87	-
Amsterdam city centre	-	2,84

5.6 Assessment framework 2010

5.6.1 Key performance indicators 2010

The model outcomes for 2010 were compared using the same KPIs as for the base year: trip frequency per purpose, modal split per purpose, and trip length distribution per mode-purpose combination. However, the performance in 2017 should be considered to determine the prediction quality for a non-calibrated year, since a backcast is about the differences between the base and backcast year (further referred to as 'backcast effect'). For instance, when a model underestimates the share of educational trips made by PT for a distance class in the base year, and a similar underestimation is found in the backcast year for the same mode-purpose combination, the cause of the underestimation is the parameter estimation in the base year, and not the longitudinal (in)accuracy of the model. In other words, the model would be able to achieve the same prediction quality in a non-calibrated year as in a calibrated year, demonstrating good explanatory power. Therefore, the differences in backcast effect were determined, rather than the absolute OViN data fit.

In addition to comparisons of the graphs, the trip frequency was calculated as a fraction of the OViN means for the given purpose. In that way, the backcast effect could be determined more easily than with graphs alone. For the modal split, the relative distribution over the modes was of interest, as the absolute trip numbers were already evaluated through the trip frequency. Since the differences in modal split were very small, and the number of comparisons was limited, only graphs were used. For the trip length distribution, graphs and the Mean Absolute Error (MAE) were used, because of the high number of comparisons. In literature, the MAE and Root Mean Square Error (RMSE) are often used as metric to capture the deviation between modelled and observed data (Chai & Draxler, 2014)[4]. The main difference between these two methods is that the RMSE squares the difference between observed and measured data, meaning that larger errors in terms of absolute values weight more than smaller errors, while the MAE increases linearly (Chai & Draxler, 2014). In other words, the RMSE is more sensitive to outliers. Since there was no reason to give more weight to outliers in this study, the MAE was used as metric. The MAE was calculated using eq. (5):

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{5}$$

Where:

- n number of test instances
- y_i modelled number of trips for distance bin i
- $\hat{y}_i$ observed number of trips in OViN for distance bin i

[4] Note that in literature also the terms Root Mean Square Deviation (RMSD) and Mean Absolute Deviation (MAD) are used.

5.6.2 Evaluation data 2010

Similar to the base year, four years of OViN were used as evaluation data for the KPIs. However, in this case OViN 2010 up to and including 2013 was used. Home-based trips were selected from the dataset, using the same methods as for 2017 (see section 5.3.2). Furthermore, the confidence interval was calculated around the OViN data based on eq. (4). The OViN fit of the model outcomes was subsequently determined, in order to compare the difference in OViN fit between 2017 and 2010 for the 4-step model and Octavius.

5.6.3 Hypotheses at segment level

In addition to evaluating the model outcomes at the aggregation level of the 4-step model, hypotheses were defined that compare model behaviour with literature at a segment level; the aggregation level of Octavius. Firstly, from literature it is known that within the Dutch context immigrants are less likely to use a bicycle (Harms, 2006). In contrast to the 4-step model, Octavius considers immigrants explicitly. Therefore, it was hypothesised that Octavius demonstrates a decrease of bicycle use when a significant relative increase of immigrants occurs. In line with the aggregated analysis, the differences between the 2017 and 2010 situation were analysed. However, to understand the behaviour of both models at a spatially finer level, the analysis was conducted for a selected number of TAZs. To avoid interference of other changing socioeconomic variables, only the ethnicity distribution was changed for the selected TAZs relative to 2017 (see Table 8). In other words, the 2010 socioeconomic data were applied to all TAZs, except for the selected TAZs for which only the share of immigrants was changed, while the other variables remained unchanged from 2017.

Secondly, increasing car ownership results in a higher share of car trips (KiM, 2016). Car ownership is considered by both models, which made it interesting to see how the models compared. However, since Octavius explicitly takes car ownership into account when modelling mode choice, it was hypothesised that Octavius demonstrates more pronounced differences. Similar to the first hypothesis, only the car ownership was adjusted for a selected number of TAZs (see Table 8), while the other socioeconomic variables were kept the same for these TAZs. Figure 13 shows the location of the selected TAZs.

Lastly, the results were analysed by means of model split analyses and adjusted trip length distribution. Rather than determining the distribution as a function of distance per purpose-mode combination, the distribution as a function of distance for all modes for all purposes together was analysed. In other words, it described the ratio of trips made between car, PT and bicycle per distance class.

Table 8 - Adjustments made to test hypotheses

Hypothesis	TAZ	Adjustment 4-step model	Adjustment Octavius
Immigrants and bicycle use		N/A	Relative share immigrants adjusted
	128	N/A	32% to 60%
	156	N/A	67% to 20%
	354	N/A	59% to 10%
Car ownership and use		Car ownership share adjusted	
	86	71% to 61%	
	170	40% to 81%	
	322	69% to 62%	
	330	69% to 87%	

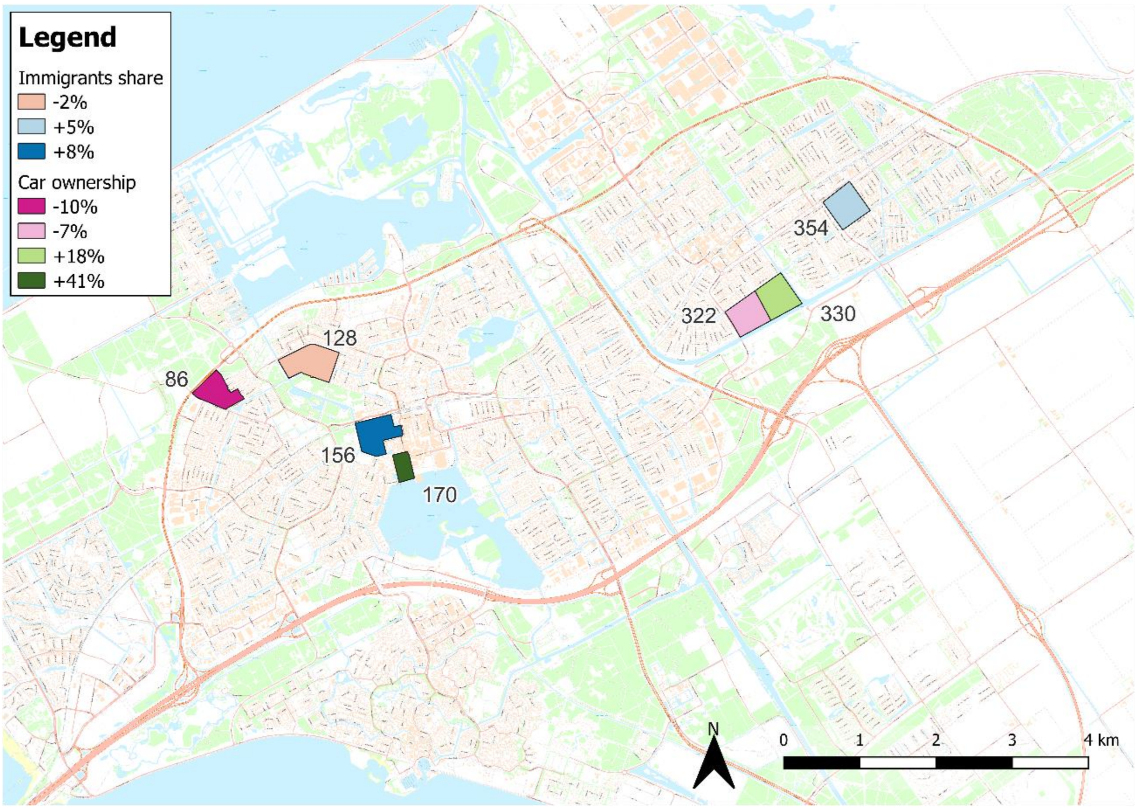


Figure 13 - Location of the TAZs with changed socioeconomic variable

6 Results

6.1 Re-estimation of selected model parameters

The effect of re-estimating the trip generation parameters in combination with the Spectrum data is depicted in Figure 14. The updated 4-step model matches the OViN data to a greater or equal extent compared to the original model for all purposes, except for the total number of trips. The reason the number of trips produced in the model does not equal the OViN 2010-2017 data is because the doubly constrained gravity model is set to equal the home-side of the trips in the end. In other words, for home-work trips the row totals match the OViN figures and for work-home trips, the column totals match the OViN figures. Additionally, the number of iterations to balance the row and column totals is set to 25. Increasing this number is likely to result in a smaller deviation from OViN, but since the scope of this study is to evaluate the practical implementation of a gravity model this configuration was maintained.

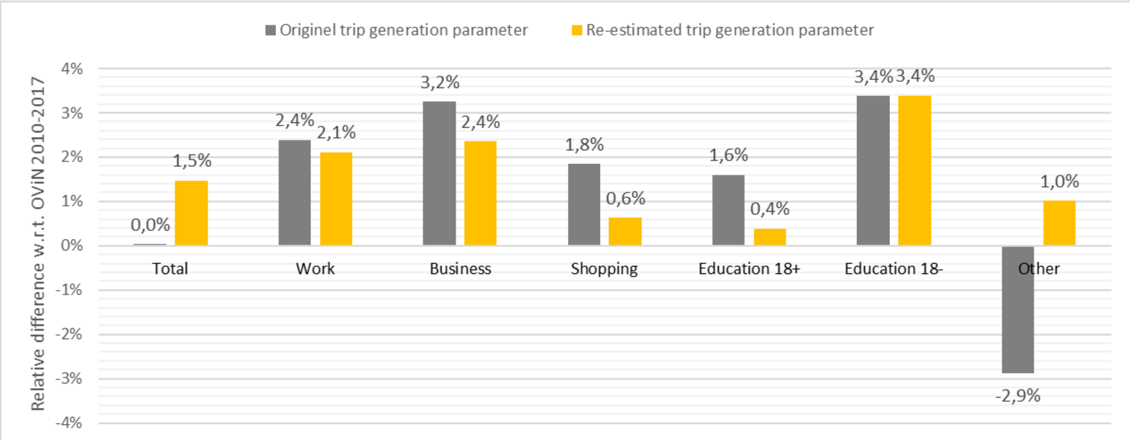


Figure 14 - Trip frequency deviation from OViN 2010-2017 for the original and updated model

Similarly, the regionalisation of Octavius resulted in a better OViN fit for all purposes in terms of modal split and trip length distribution. Since the tour generator is not regionalised, the number of tours remains the same. Figure 15 demonstrates the total modal split. See Appendix A – Regionalisation result for a side-by-side comparison for the other purposes and the trip length distributions.

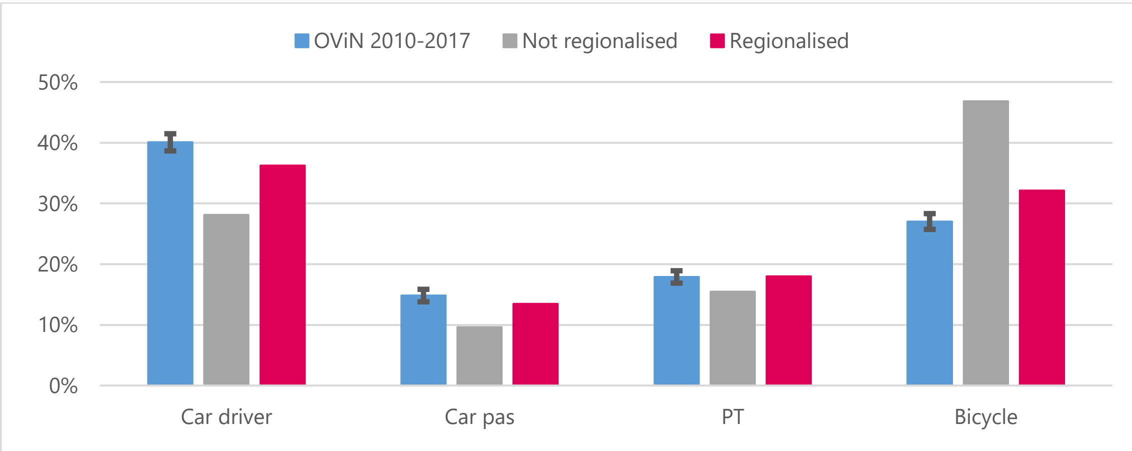


Figure 15 - Total modal split modelled by Octavius before and after regionalisation for 2017

The re-estimated trip generation parameters and the regionalised utility functions were subsequently applied, and the outcomes are evaluated in the next sections. As explained in the methodology (section 5.3), the home-based trips for the purposes work, education, business, shopping, and all four purposes combined were compared using the trip frequency, modal split, and trip length distribution.

6.2 Trip frequency

Figure 16 presents the modelled trip frequency for 2017 and 2010 as well as the OViN trip frequency. What stands out is that the number of modelled trips differs considerably from the OViN figures in 2017 for several purposes (e.g., education by the 4-step model and shopping by Octavius). There are multiple reasons for this. Firstly, both models have been calibrated on a different aggregation level than shown here: the 4-step model on home-based and non-home based trips and Octavius on tours, while Figure 16 demonstrates home-based trips only. Secondly, both models have been calibrated on combined OViN 2010-2017 data, while here the years 2014-2017 and 2010-2013 are used separately. Thirdly, the tour generator of Octavius has been estimated on the number of observations in OViN 2010-2017, rather than on the weighted OViN data which are representative for the whole of the Netherlands. As a result, the tour generator underestimates the number of tours.

Additionally, it appears that the OViN figures in 2010 are higher compared to 2017 for all purposes, while both models generate fewer trips in 2010 compared to 2017 (see Figure 16). The primary reason is that the evaluation data for 2017 is based on OViN 2014-2017 with an average home-based trip rate of 2,45 per person per day, while OViN 2010-2013 includes 2,78 home-based trips per person per day. Hence, people travelled more in 2010-2013 compared to 2014-2017 according to OViN. The models on the other hand have been calibrated on OViN 2010-2017 for both periods. In other words, it assumes that people made the same number of trips in both periods. In combination with fewer Almeerders in 2010, less trips are being modelled.

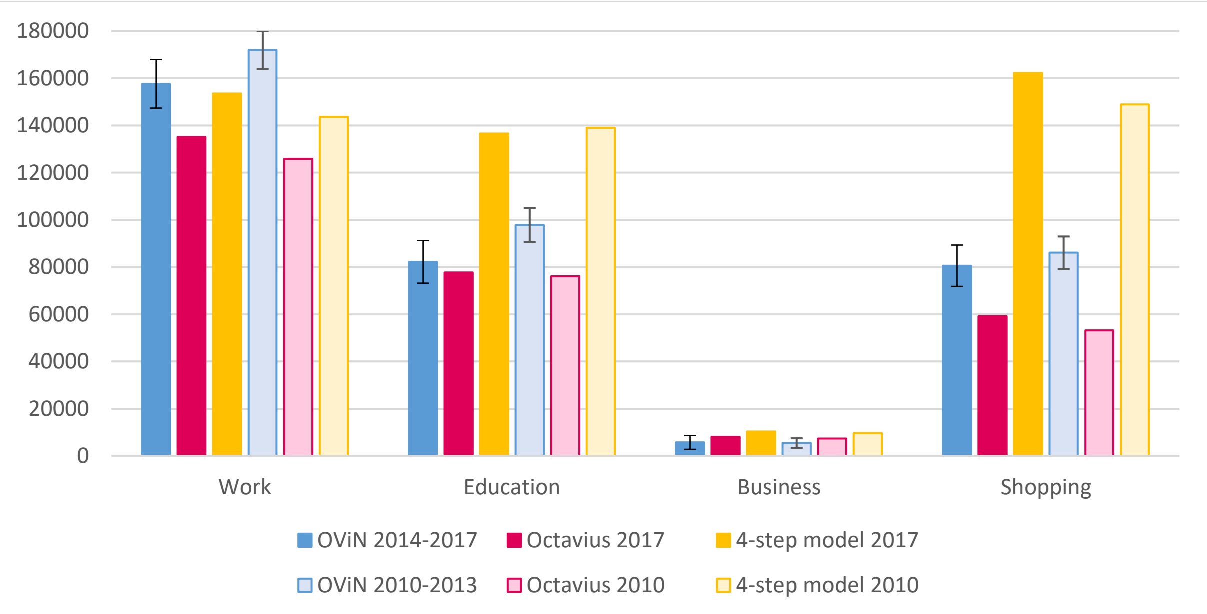


Figure 16 - Trip frequency for 2017 and 2010

Table 9 demonstrates per purpose the deviation in trip numbers of both models relative to the OViN mean. For the base year 2017, Octavius is closer to OViN compared to the 4-step model for all purposes, except business for which both models score equally. Then, the backcast effect is found by calculating the difference between the 2017 and 2010 OViN fit. As can be seen, the 4-step model performs marginally better than Octavius, with a smaller difference in OViN fit between 2017 and 2010. Only for business Octavius scores better, but the uncertainty for that purpose is high, illustrated by the large confidence interval in Figure 16.

Table 9 - Trip frequency relative to OViN for 2017 and 2010

	Octavius			4-step model		
Purpose	2017	2010	%difference	2017	2010	%difference
Total	0.86	0.73	-15.4%	1.42	1.22	-13.9%
Work	0.86	0.73	-14.5%	0.97	0.84	-14.2%
Business	1.41	1.35	-3.7%	1.41	1.79	27.0%
Education	0.95	0.78	-17.7%	1.66	1.42	-14.5%
Shopping	0.73	0.62	-15.9%	2.01	1.73	-14.1%

6.3 Modal split

Figure 17 and Figure 18 present the work modal splits for 2017 and 2010, respectively. As can be seen, the car share is similar between 2017 and 2010 for OViN and both models. The PT share decreases from 28% to 23% in OViN between 2017 and 2010, while both models calculate the same PT shares for both years. For bicycle, the share in OViN increased with 4 percentage points (pp), while both models demonstrate the same share for 2017 and 2010.

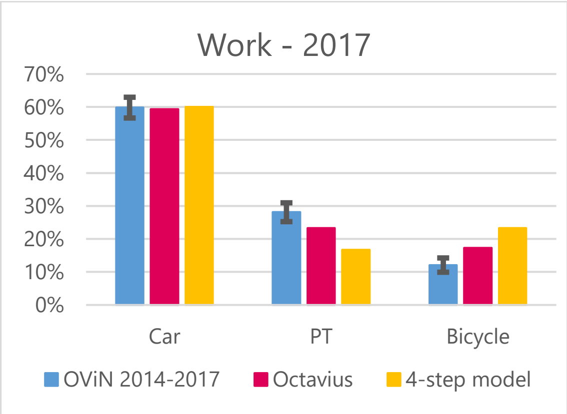


Figure 17 - Modal split work for 2017

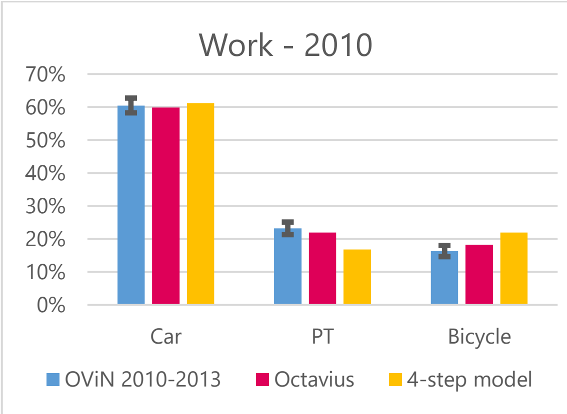


Figure 18 - Modal split work for 2010

The modal split of educational trips in OViN is similar between 2017 and 2010 (see Figure 19 and Figure 20). However, both models demonstrate a modal split change: Octavius models 4 pp more bicycle use in 2010 than 2017, while PT and car trips decrease with 2 pp for the same years. The 4-step model calculates slightly more car trips in 2010 compared to 2017 at the expense of bicycle use. PT trips are modelled similarly poorly by the 4-step model for both years.

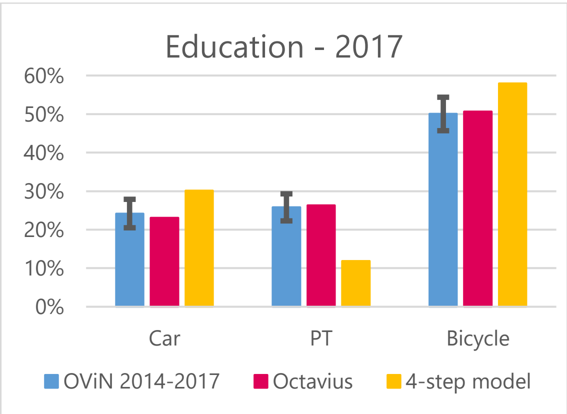


Figure 19 - Modal split education for 2017

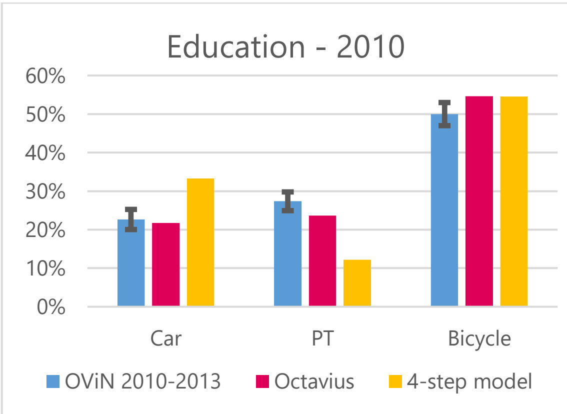


Figure 20 - Modal split education for 2010

Business trips are rare in OViN, which results in wide 95% confidence intervals (see Figure 21 and Figure 22). As a consequence, no meaningful conclusions can be drawn from these figures.

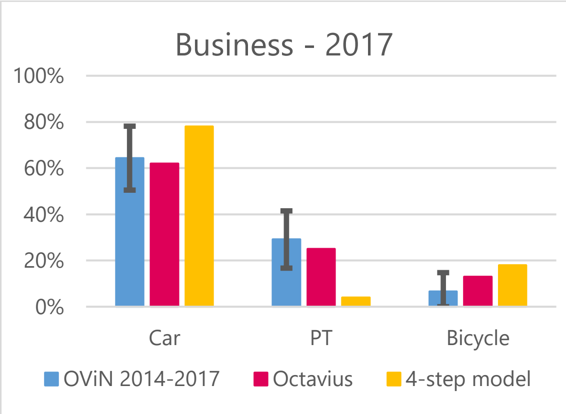


Figure 21 - Modal split business trips for 2017

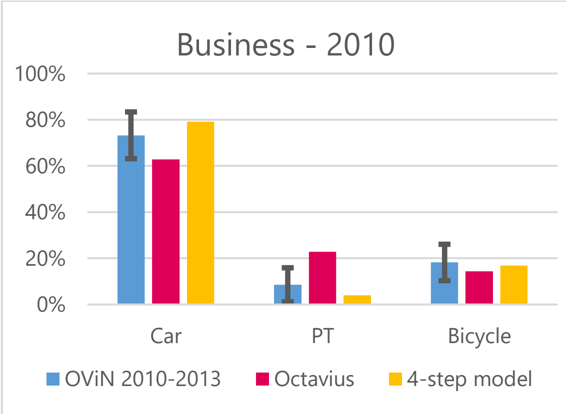


Figure 22 - Modal split business trips for 2010

Figure 23 and Figure 24 show the modal split of shopping trips. The average car share in OViN changes from 48% in 2017 to 57% in 2010 at the expense of the bicycle share which decreases from 41% to 32%. Interestingly, Octavius demonstrates (nearly) no change in car and bicycle share, while the 4-step model overresponds: for car an overestimation of 12 percentage points in 2010 compared to a perfect OViN fit in 2017, and for bicycle 5 pp underestimation in 2010 compared to 4 pp overestimation in 2017.

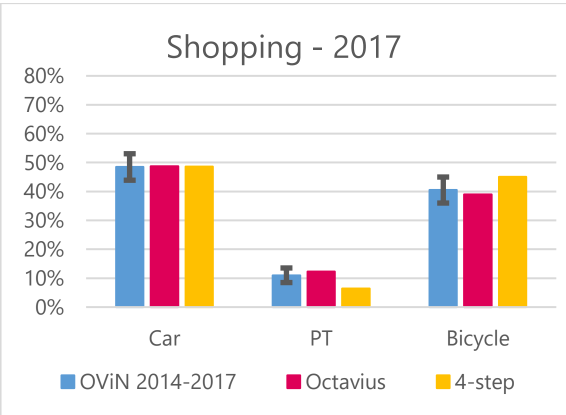


Figure 23 - Modal split shopping for 2017

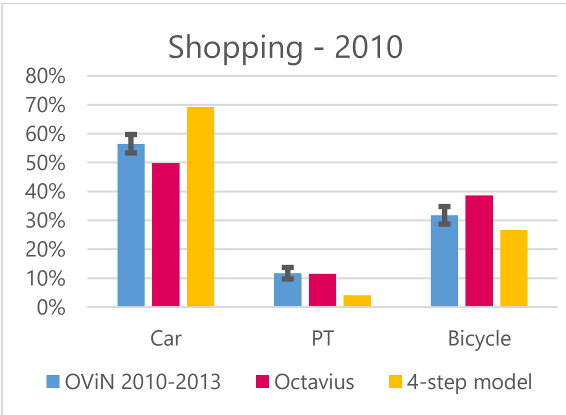


Figure 24 - Modal split shopping for 2010

Figure 25 and Figure 26 present the modal split for the four analysed purposes together. As can be seen, the OViN modal split changes marginally: 2 pp increase for car in 2010 compared to 2017, and 2 pp decrease for PT. Octavius calculates the same car share for 2017 and 2010, but PT decreases with 2 pp in 2010 compared to 2017; similar to OViN. The bicycle share is somewhat overestimated in 2017, which rises further in 2010, while OViN remains stable. Moreover, the 4-step model calculates a car share six percentage points higher in 2010, while 2017 was perfectly aligned with OViN. The PT share is stable between 2017 and 2010, while OViN decreases two percentage points. Finally, the modelled bicycle share decreases considerably in 2010 compared to 2017, while OViN 2010 equals 2017.

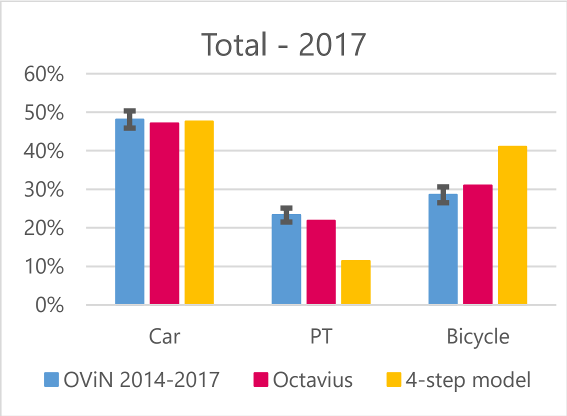


Figure 25 - Modal split in total for 2017

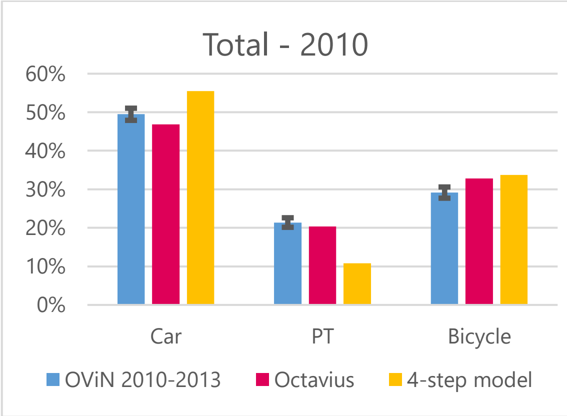


Figure 26 - Modal split in total for 2010

6.4 Trip length distribution

The Mean Absolute Error (MAE) was calculated to evaluate the trip length distribution results. Since the modal split results (section 6.3) demonstrated clearly that there are insufficient OViN data on business-related trips, this purpose was omitted from the trip length distribution analysis. From Table 10 it is clear that Octavius demonstrates a better base year performance with lower MAE for all mode-purpose combinations. This means that the trip length distributions modelled by Octavius are closer to the average OViN 2014-2017 distribution compared to the 4-step model. However, the relative differences between 2017 and 2010 indicate that neither of the two models is clearly superior in terms of longitudinal stability.

Table 10 - MAE for the trip length distributions of both models for base year and the backcast year

Car	Octavius				4-step model			
	2017	2010	diff	%diff	2017	2010	diff	%diff
Total	1253	1524	271	22%	7465	9511	2046	27%
Work	1601	1375	-225	-14%	3373	3939	566	17%
Education	331	375	44	13%	1820	1880	60	3%
Shopping	1125	1326	201	18%	2134	3788	1653	77%

Public Transport	Octavius				4-step model			
	2017	2010	diff	%diff	2017	2010	diff	%diff
Total	849	1140	290	34%	3965	3456	-509	-13%
Work	827	684	-143	-17%	2618	2263	-355	-14%
Education	408	552	143	35%	1234	1188	-46	-4%
Shopping	362	242	-120	-33%	451	692	240	53%

Bicycle	Octavius				4-step model			
	2017	2010	diff	%diff	2017	2010	diff	%diff
Total	1346	1379	33	2%	3965	4330	366	9%
Work	1009	536	-473	-47%	2618	1183	-1435	-55%
Education	408	784	375	92%	1234	1878	645	52%
Shopping	362	224	-138	-38%	451	1414	963	213%

When analysing the trip length distributions graphically, the conclusions from Table 10 are confirmed: Octavius demonstrates a better OViN fit for 2017 compared to the 4-step model (see Figure 27), while the backcast effect for both models is similar (see Figure 28). Interestingly, OViN changes slightly between 2017 and 2010 for 37,5-47,5 km, while both models show (nearly) the same distribution. An overview of all trip length distributions is given in Appendix D – Trip length distribution graphs.

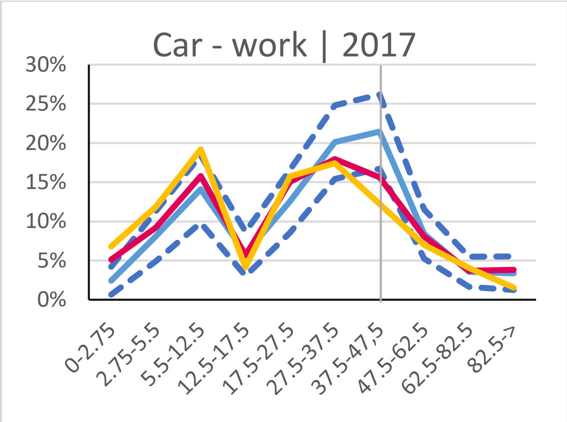


Figure 27 - Trip length distribution work by car for 2017 (blue → OViN 2014-2017 mean, blue dotted → OViN 95% CI, magenta → Octavius, yellow → 4-step)

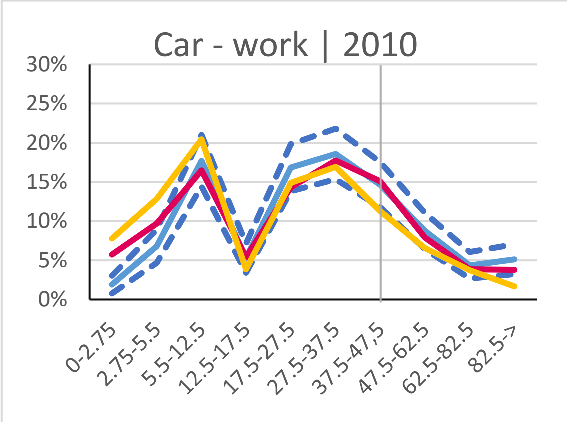


Figure 28 - Trip length distribution work by car for 2010 (blue → OViN 2010-2013 mean, blue dotted → OViN 95% CI, magenta → Octavius, yellow → 4-step)

6.5 Hypotheses at segment level

Table 11 demonstrates the response of Octavius to a changing share of immigrants in the population of a selected number of TAZs. The behaviour of Octavius is in line with literature: if the share of immigrants increases, the bicycle share decreases and vice versa. The 4-step does not take ethnicity into account, which means that no change in modal distribution occurs.

Table 11 - Modal split change Octavius with respect to change in share of immigrants

	Immigrants share		Octavius output					
TAZ			Car		PT		Bicycle	
	2017	Scenario	2017	Scenario	2017	Scenario	2017	Scenario
128	32%	60%	46%	45%	22%	25%	32%	30%
156	67%	20%	40%	36%	24%	20%	36%	44%
354	59%	10%	50%	49%	22%	17%	29%	34%

Interestingly, Octavius returns different changes per TAZ: only 2 percentage points (pp) decrease, while a 28 pp increase in immigrants occurred for TAZ 128. For TAZs 156 and 354, the change in immigrants is approximately similar (approx. 50 pp decrease), while the change is bicycle share is 8 and 5 pp. The differences are likely to be caused by multiple factors. Firstly, TAZ 128 is located in a quiet residential area with predominantly 30 km/h streets, and a high-frequent bus line (8-12 times per hour). When analysing the mode share per distance class (see Figure 29), it appears that the bicycle was and remains the most appealing option for short trips even though immigrants prefer to use another mode. For longer trips, PT substitutes a portion of the car trips. This second finding is consistent with literature as well, which found that immigrants are more likely to use PT and are less likely to own a car or a driving license compared to native Dutch people (Harms, 2006).

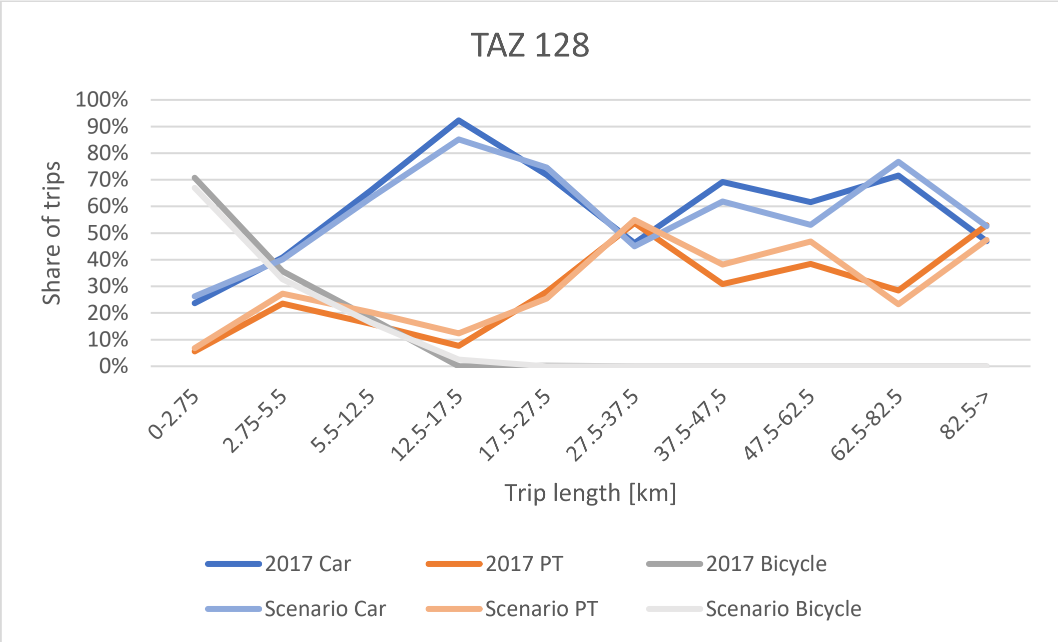


Figure 29 - Mode share per distance class for TAZ 128; modelled by Octavius

Secondly, for both TAZ 156 and 354 the share of immigrants decreased. However, TAZ 156 is located near the city centre of Almere, while TAZ 354 is located at the edge of the city. As a result, the bicycle becomes more interesting for trips up to 5.5 km for TAZ 156 at the expense of cars (see Figure 30), while for TAZ 354 this is less the case (see Figure 31). Moreover, the car ownership rate is lower for TAZ 156 compared to TAZ 354, which is likely to contribute to the smaller bicycle use increase for TAZ 354. Lastly, the PT share decreases for both TAZs, which is again consistent with literature which states that immigrants make more use of PT compared to native Dutch people (Harms, 2006). Hence, as the share of immigrants in the total population decreases, the PT share is likely to decrease as well.

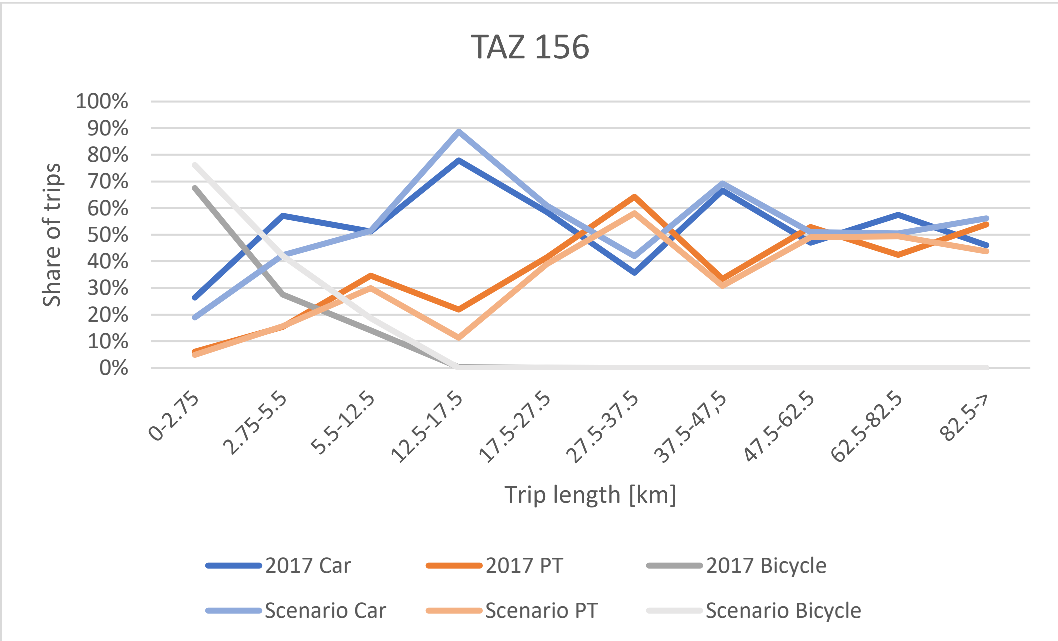


Figure 30 - Mode share per distance class for TAZ 156; modelled by Octavius

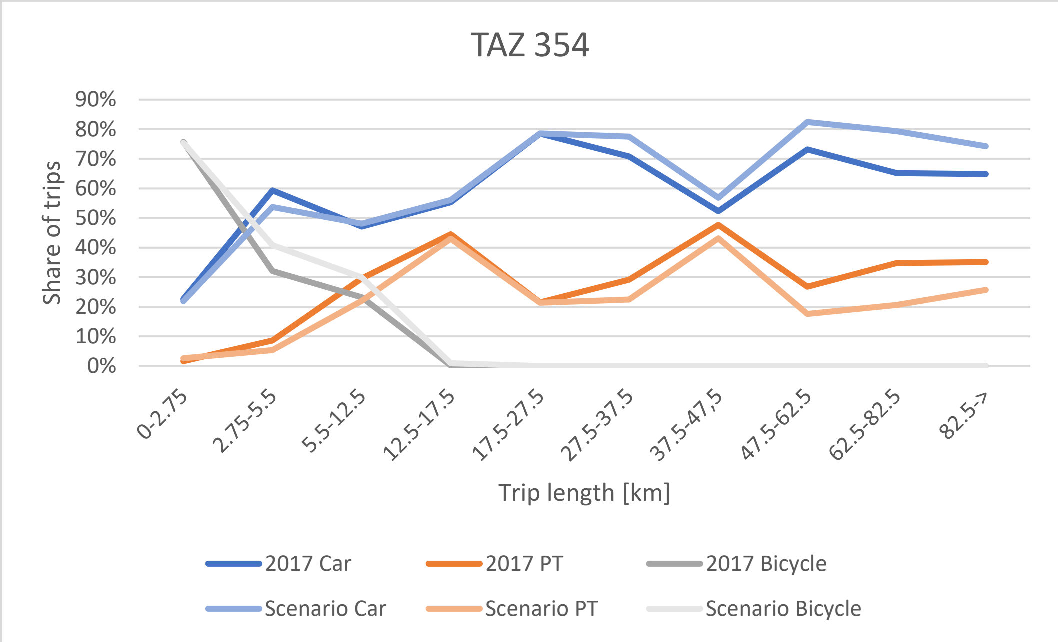


Figure 31 - Mode share per distance class for TAZ 354; modelled by Octavius

For the second scenario, the car ownership rate was changed. Octavius demonstrates an intuitive result: when the car ownership decreases, the car ownership decreases as well (see Table 12). The precise impact of the change in car ownership differs somewhat between the TAZs, but are all in line with literature (KiM, 2016). The mode shares per distance class for all four TAZs can be found in Appendix E – Mode shares per distance class.

Table 12 - Modal split change Octavius with respect to change in car ownership

TAZ	Car ownership share		Octavius output					
			Car		PT		Bicycle	
	2017	Scenario	2017	Scenario	2017	Scenario	2017	Scenario
86	71%	61%	48%	38%	23%	24%	29%	38%
170	40%	81%	30%	50%	30%	21%	40%	29%
322	69%	62%	47%	40%	21%	23%	32%	36%
330	69%	87%	46%	58%	21%	14%	34%	29%

The 4-step model demonstrates a result in line with literature for car ownership increase, modelled for TAZs 170 and 330 (see Table 13). However, when the car ownership decreases, the 4-step model demonstrates a counterintuitive result: the car use increases for TAZs 86 and 322, at the expense of bicycle use.

Table 13 - Modal split change 4-step model with respect to change in car ownership

TAZ	Car ownership share		4-step model output					
			Car		PT		Bicycle	
	2017	Scenario	2017	Scenario	2017	Scenario	2017	Scenario
86	71%	61%	53%	59%	12%	13%	36%	28%
170	40%	81%	21%	33%	10%	7%	69%	60%
322	69%	62%	46%	55%	11%	14%	43%	31%
330	69%	87%	45%	60%	11%	12%	43%	28%

Figure 32 demonstrate the mode share per distance class for TAZ 86 as modelled by the 4-step model. It appears that the car is used more for trips between 2.75 and 5.5 km compared to the reference situation, as well as for the distance class 62.5-82.5 km. Since the majority of the trips for this TAZ is in the category 2.75-5.5 km, the total modal shift becomes more car-oriented.

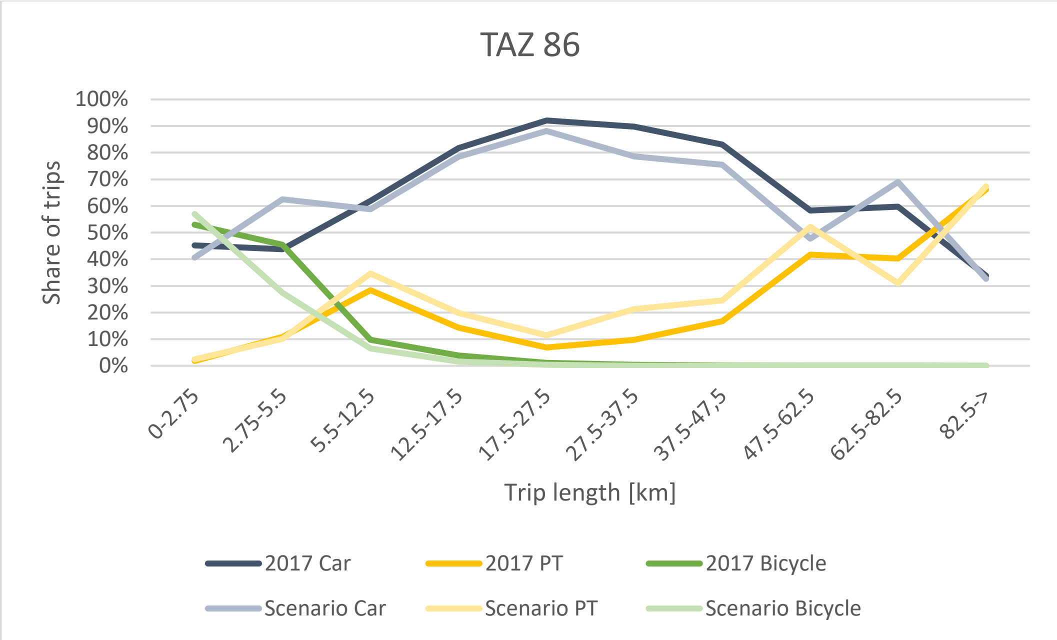


Figure 32 - Mode share per distance class for TAZ 86; modelled by the 4-step model

A similar effect occurs for TAZ 322, where trips up to 12.5 km and between 62.5 and 82.5 km are made more often by car in the scenario compared to the 2017 situation (see Figure 33). For the short trips, this car increase is at the expense of bicycle use, while the trips between 62.5 and 82.5 km are taken from the PT trips.

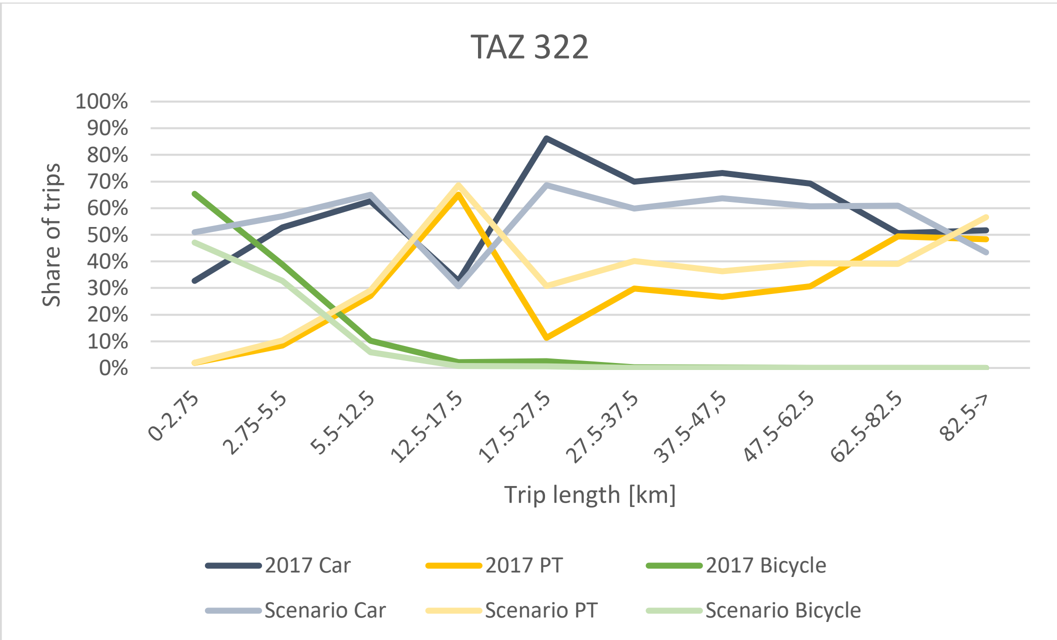


Figure 33 - Mode share per distance class for TAZ 322; modelled by the 4-step model

Further analysis reveals that the 4-step model distinguishes between 'car available' and 'no car available'. Note that 'no car available' in combination with car use means that there is no car available as being a *driver*; there is as being a *passenger*. It appears that for TAZ 86 and 322 the share of car drivers decreases, but the share of passengers increases (see Table 14 for the purpose shopping as illustration). For TAZ 170 (increase car ownership), both the car and the driver share increase as one would expect. Lastly, TAZ 330 demonstrates a small decrease in driver share, while the car ownership increased significantly. A likely explanation is that the number of trips as car driver was already high due to the high car ownership and with the further increase of car ownership, being a car passenger became much more interesting. In absolute numbers, both the number of car driver and car passenger trips increased significantly.

Table 14 - Car availability for a selected number of zones for 2017 and the scenario for shopping

	2017		Scenario	
TAZ	**Driver**	**Passenger**	**Driver**	**Passenger**
86	57%	43%	35%	65%
170	75%	25%	87%	13%
322	65%	35%	41%	59%
330	67%	33%	57%	43%

Lastly, the increase in car share and the decrease in bicycle share in absolute numbers which is visible in Table 13 is likely to be caused by the fact that the analysed 4-step model determines first the number of car-available trips and no-car-available trips, after which the simultaneous gravity model determines the mode and destination choice. Since the car-available and no-car-available have different deterrence functions, the mode and destination choice becomes different. Apparently, it becomes more attractive to be a car passenger than a cyclist, which results in more car trips.

7 Discussion of the results

When taking a step back and reflecting on the findings, the following aspects are notable. Firstly, the trip frequency analysis suggests that both models are not able to accurately model home-based trips only when evaluated against OViN figures for the base year, although Octavius is closer to the OViN figures than the 4-step model. As explained in section 6.2, there are multiple explanations for this, but this does not change the fact that the base year results are not consistent with observed travel behaviour for 2017. Furthermore, the modal split and trip length distributions are modelled most consistently with OViN by Octavius for the base year, which is attributed to the high number of parameters included in Octavius compared to the 4-step model and the regionalisation process.

Secondly, the backcast performance for both models is similar on all three aggregated KPIs, although the models behave differently: Octavius reproduces OViN accurately for the base year, but the output for 2010 is nearly identical to 2017, while the 4-step model is less accurate for the base year but maintains this deviation in the backcast year. The most prominent example is the shopping modal split (see Figure 23 and Figure 24) with nine percentage points difference for car and bicycle in OViN, while Octavius models no change whatsoever. The 4-step model on the other hand demonstrates more sensitivity (although overestimated). Since the differences in the evaluation data for 2017 and 2010 are limited, it is challenging to analyse which model demonstrates a better prediction quality. Nevertheless, further analysis of the models' systematics suggests that Octavius may be insensitive to changes in the input data. The logsum in Octavius' utility functions – which defines the natural logarithm of the total attractiveness of all destinations given a certain mode, purpose and origin – contributes significantly more to the utility than the other explanatory variables (see Table 15). As a result, changes in the socioeconomic data affect the total utility only to a limited degree. Similarly, the Alternative Specific Constant (ASCs) added through the regionalisation may outweigh changes in less dominant variables. Since the 4-step model is only directly affected by changes in the input data (i.e., in the trip production/attraction, and in the travel costs), this may explain why it shows slightly more changes in modal split and trip length distributions.

Table 15 - Example of a utility calculation for a woman, 45-65 years old, couple, no children, 2(+) cars and immigrant

Explanatory variable	Bicycle	PT	Car driver	Car passenger
Non-Western immigrant	0	0	0	0
Western immigrant	0	0,5	0,3	0,6
2(+) cars	0	-0,5	1,7	0,5
1 car	0	0	0	0
3(+) persons > 18 years	0	0	0	0
2 persons > 18 years	0	0	-1,1	-3,4
Couple, with kids	0	0	0	0
Couple, no kids	0	0	0	0
Age 65+	0	0	0	0
Age 45-65	0	-0,2	0,2	-0,7
Age 30-45	0	0	0	0
Age 18-30	0	0	0	0
Gender: male	0	0	0	0
ASC	0	-2,2	1,2	-1,8
Average logsum	9,6	11,0	9,3	13,0
Total utility	**9,6**	**8,5**	**11,6**	**8,2**

The analyses at the disaggregated, segment level indicate that Octavius demonstrates behaviour which is consistent with literature. The first hypothesis was for the analysed TAZs correct: when the share of immigrants in the population increases, the bicycle use decreases and vice versa. Furthermore, the PT use increased when the share of immigrants increased, which is again consistent with literature (Harms, 2006). The second hypothesis was also correct for the considered TAZs, with more pronounced differences modelled by Octavius compared to the 4-step model. An unanticipated finding was that the 4-step model demonstrates counterintuitive results with increasing car use when the car ownership level decreases; especially for short trips. Further analysis indicates that it models a shift from car driver to car passenger. This irrational result seems to illustrate the lack of behavioural realism resulting from ignoring the decision maker's context, which is mentioned in literature (Vovsha, 2019). The question then becomes why Octavius is not demonstrating this sensitivity at an aggregated level. The reasons found are twofold: Firstly, at the aggregated level of the KPIs, many factors play a role which can cancel each other out. Secondly, the logsum and ASC applied in the utility functions of Octavius determine the total utility mostly. In the segment analysis, significant changes were modelled, which did not occur to such an extent at the aggregated level during the seven-year period analysed.

Furthermore, it is argued in literature that disaggregated models predict aggregate travel demand more accurately than their aggregated counterparts, because the underlying travel choices are modelled more realistically (Davidson et al., 2007; Roorda et al., 2008). Hence, the aggregate level predictions would improve as a result of the improved modelling at the disaggregated level. The findings of the presented study do, however, not support this hypothesis since the backcast effect for both models is similar. A possible explanation is the limited differences between the base and predicted year, which means that clearly different model behaviour was hardly observed.

At the operational level, the 4-step model requires significantly fewer socioeconomic input data than Octavius. Not only reduces the amount of data to be collected, it also simplifies the data processing to create a consistent input dataset. Since Octavius applies a variety of variables, it is (nearly) inevitable that data are gathered from multiple sources with inconsistencies as result. Therefore, significantly higher time investments are required to collect and prepare suitable input data for Octavius, which is in line with literature (e.g., Omer et al. (2010). Nonetheless, the richer dataset is used effectively by Octavius to achieve a better OViN fit for the base-year 2017 compared to the 4-step model. Hence, the parameters that are used in combination with the explanatory variables in Octavius' choice models allow for a closer OViN fit than the 4-step model with its lognormal distribution function.

Lastly, Octavius allows for evaluating specific transportation and socioeconomic developments. In its current form, Octavius is able to model travel behaviour of specific groups in society, such as immigrants which was demonstrated in section 6.5, but also students and elderly. This is extremely useful when new neighbourhoods are developed, or existing neighbourhoods are changed, because the effect on travel behaviour following from these developments is made explicit. Considering the housing crisis the Netherlands is facing, resulting in 900,000 additional houses to be built by 2030 (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2022), the relevance of this characteristic is clear. Moreover, choice models that include the decision maker's context can be easily added to Octavius, such that more complex travel behaviour can be described such as multimodal travelling and travel choices of other household members. Similarly, more detailed transport policies can be evaluated, such as congestion pricing, incentive programs to increase PT or (e-)bicycle use, and tailor-made travel advice through Mobility As A Service (MaaS). The 4-step model on the other hand is not and will not be able to model this more complex behaviour as it ignores the decision maker's context.

8 Conclusions

The presented research aimed to analyse the differences between an aggregated, macroscopic, trip-based 4-step model and a disaggregated, tour-based microsimulator called Octavius within the application context. Based on the results, the following conclusions can be drawn. Firstly, Octavius describes the base year 2017 more accurately than the 4-step model in terms of trip frequency, modal split and trip length distribution when comparing the model output with OViN figures. This finding suggests that Octavius is better calibrated for the reference situation than the 4-step model, which can be attributed to the higher number of parameters applied in Octavius' choice models and the regionalisation process.

Secondly, the identified backcast effects demonstrate that both models perform similarly in terms of longitudinal stability at an aggregated level. However, due to the limited differences in evaluation data and model outcomes, it is challenging to identify if, and to what extent, the two models respond differently to changes. In the few situations with change, it appears that Octavius might be somewhat insensitive to changes, which was supported by the high weight of the logsum and ASC observed in Octavius' utility functions. Furthermore, the segment level analyses for selected TAZs indicate that the responses to change in the share of immigrants in the population modelled by Octavius are consistent with literature. Hence, when the share of immigrants in the TAZ population increased, the bicycle use decreased and vice versa. The 4-step model does not consider ethnicity and consequently fails to capture this effect. The second hypothesis focused on the relation between car ownership and use. Similar to the first segment analysis, Octavius' response is consistent with literature: an increase in ownership results in more car use and vice versa. The 4-step model demonstrated counterintuitive results, as the car use increases when the ownership levels decrease. Although the segment analyses were merely a case study with only a small selection of TAZs, it supports the theoretical argument found in literature that the 4-step model lacks behavioural realism, whereas Octavius takes the decision maker's context into account.

Lastly, the current implementation of Octavius – and disaggregated, tour-based microsimulators in general – allows already for analysing the impact of transport policies and developments on certain subgroups in society. For instance, the transport impacts of a new neighbourhood with specific groups such as elderly, students or immigrants can be made explicit. Moreover, Octavius can be extended such that more complex travel behaviour can be modelled as well, including the use of shared mobility, MaaS, and choices of other household members. Due to its aggregated, macroscopic, trip-based nature, the 4-step model is and will not be able to provide any of the abovementioned insights.

9 Limitations & recommendations

An important limitation of the presented research are the limited differences in the evaluation and input data between the base and backcast year. As a result, the responses of the models to changing input data cannot be accurately assessed. This was further amplified by the use of combined OViN years to evaluate the model outcomes (i.e., 2014 to 2017 for 2017 and 2010 to 2013 for 2010) as the observed travel behaviour was averaged over the combined years. Nonetheless, the calculated confidence intervals demonstrated clearly that multiple years of OViN data for Almere are required to reduce the data uncertainty. The limited data availability for 2010 contributed to the minor changes as well. For multiple variables – especially the more detailed ones such as driving license ownership, number of cars in household, and the number of jobs in a given sector – 2010 data were not available. As such, the 2017 figures were translated to 2010 for the required variables. However, it is likely that the actual situation is not completely captured, which subsequently contributed to the limited differences. Furthermore, using an adjusted 2017 network rather than the 2010 network is likely to add to same issue as well. Also, the same values of time and distance were used for 2017 and 2010, because the technical implementation would require more time than available. Future research should use a longer time period with more distinct differences, which confirms the recommendation of at least ten year difference from literature (Sammer et al., 2010, as cited in Lange & Huber, 2015). It is important that collected input data is sufficiently detailed and should be available for the base and backcast year.

Furthermore, the aim of the study was to compare both types of model systematics, which required KPIs that operate at a level which both models describe (i.e., home-based trips for selected purposes). However, this resulted in highly aggregated metrics which capture multiple factors simultaneously. Although these metrics are used in practice to evaluate model outcomes, they are only useful when the changes are noticeable at that aggregation level, which require significant differences in travel behaviour and/or socioeconomic development. The segment case study compensated partially for this loss of detail but used only a small sample of TAZs and variables. Therefore, further work could analyse if the current implementation of Octavius' utility functions is sensitive enough to changes on an aggregated level: the scale at which strategic travel demand models are generally applied. Note that more detailed KPIs could also require more detailed evaluation data, which can be challenging. Moreover, evaluating one model, or models that operate at the same aggregation level, is likely to broaden the range of possible KPIs.

Thirdly, using travel survey data from OViN as evaluation data involves certain limitations. Underreporting is a significant shortcoming, as travel information provided by participants is virtually always incomplete and therefore not a complete representation of reality (Wolf et al., 2003). Moreover, people that travel less frequently are more willing to fill in a travel survey than more frequent travellers, which leads to a biased sample (Sammer et al., 2018). Although travel survey data are widely applied in research and practice, the limitations remain and apply to the presented study as well. Therefore, future research might use additional evaluation data to compensate for these limitations and provide a more complete representation of reality. Another reason alternative assessment data are preferred, is because the estimation of the model parameters and validation of the model outcomes were based on the same dataset.

This study was further limited by the fact that the 4-step was not completely re-calibrated after replacing the socioeconomic input data. As described in section 5.2, the 4-step model trip generation parameters were re-estimated to correct for this substitution. However, the alpha and beta in the lognormal distribution function were not re-estimated because this was beyond the time limits of the presented research. Consequently, the 4-step model is likely to perform suboptimally in 2017, although the backcast performance was corrected for the parameter estimation quality. Octavius on the other hand was regionalised with the Spectrum input data and is therefore performing to the best of its (current) ability. Notwithstanding these limitations, Octavius includes more parameters which affect the model outcomes, and is therefore likely to maintain its better performance in a base year compared to the 4-step model. A recommendation for further research is to conduct the same comparison with a completely calibrated 4-step model. Moreover, further work is needed to analyse if the revealed inconsistencies between car ownership and use are applicable to the 4-step model in general. This could be achieved by analysing more TAZs and evaluating other regions than Almere.

Lastly, backcasting is one of the possibilities to validate the systematics of a model. While the method is recognised as a suitable approach in literature (Lange & Huber, 2015; Roorda et al., 2008), other approaches can be applied to broaden the scope of the validation process. In particular, a sensitivity analysis is recommended to evaluate if the ASC and logsum used in the utility functions of Octavius outweigh changes in other variables, which could result in rigid behaviour. Since the aim of the model is to predict future years, inflexibility threatens its longitudinal stability.

References

BOVAG. (2022). *Fietsen in de statistiek 2007-2021 ~Nederland~*. https://www.bovag.nl/BovagWebsite/media/BovagMediaFiles/Cijfers/2022/Fietsverkoopstatistieken-2007-2021.pdf

Brederode, L., Hardt, T., & Rijksen, B. (2020, September 10). *Development of a microscopic tour based demand model without statistical noise [Conference presentation]*. European Transport Conference 2020, Online. https://aetransport.org/past-etc-papers/conference-papers-2020?abstractId=6972&state=b

CBS. (n.d.-a). *Dutch National Travel survey* [Webpagina]. Statistics Netherlands. Retrieved 19 August 2022, from https://www.cbs.nl/en-gb/onze-diensten/methods/surveys/korte-onderzoeksbeschrijvingen/dutch-national-travel-survey

CBS. (n.d.-b). *Microdata: Zelf onderzoek doen* [Webpagina]. Centraal Bureau voor de Statistiek. Retrieved 4 August 2022, from https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen

CBS. (n.d.-c). *Onderzoek Verplaatsingen in Nederland (OViN)* [Webpagina]. Centraal Bureau voor de Statistiek. Retrieved 3 August 2022, from https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/onderzoek-verplaatsingen-in-nederland--ovin--

CBS. (2013). *Wijk- en buurtkaart 2010*. http://download.cbs.nl/regionale-kaarten/2010-buurtkaart-shape-versie-3.zip

CBS. (2022a). *StatLine—Personen met een rijbewijs; rijbewijscategorie, leeftijd, regio, 1 januari*. https://opendata.cbs.nl/#/CBS/nl/dataset/83488NED/table?dl=135C2

CBS. (2022b). *StatLine—Regionale kerncijfers Nederland*. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70072ned/table?ts=1655122899934

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Davidson, W., Donnelly, R., Vovsha, P., Freedman, J., Ruegg, S., Hicks, J., Castiglione, J., & Picado, R. (2007). Synthesis of first practices and operational research approaches in activity-based travel demand modeling. *Transportation Research Part A: Policy and Practice*, *41*(5), 464–488. https://doi.org/10.1016/j.tra.2006.09.003

DfT. (2020). *National Transport Model—Analytical Review*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/946061/DfT-National-Transport-Model-Analytical-Review-accessible.pdf

Elmorssy, M., Tezcan, H., & Onur, S. (2019). Application of Discrete 3-level Nested Logit Model in Travel Demand Forecasting as an Alternative to Traditional 4-Step Model. *International Journal of Engineering*, *32*(10), 1416–1428. https://doi.org/10.5829/ije.2019.32.10a.11

Esri Nederland. (2022, February 8). *Postcodepunten—Overzicht*. https://www.arcgis.com/home/item.html?id=2bafadb98f2f4f00b9333f68f9ce7154

Ferdous, N., Vana, L., Bowman, J. L., Pendyala, R. M., Giaimo, G., Bhat, C. R., Schmitt, D., Bradley, M., & Anderson, R. (2012). Comparison of Four-Step versus Tour-Based Models for Prediction of Travel Behavior before and after Transportation System Changes. *Transportation Research Record*, *2303*(1), 46–60. https://doi.org/10.3141/2303-06

Goudappel. (n.d.). *OmniTRANS Spectrum | Goudappel*. Retrieved 29 July 2022, from https://www.goudappel.nl/nl/expertises/data-en-it-oplossingen/omnitrans-spectrum

Goudappel. (2012). *Segmentering ritgeneratiemodel Stedendriehoek* [Unpublished internal company document].

Goudappel. (2018). *Verkeersmodel MRDH 2.0: Technische rapportage*. https://mrdh.nl/sites/default/files/documents/rapport_verkeersmodel_mrdh_2.0_-_001594.20181026.r1.02.pdf

Goudappel. (2022). *Manual Octavius 2.0* [Unpublished internal company document].

Griesenbeck, B., & Garry, G. (2007). Comparison of Activity-Based Tour Model to Four-Step Model as a Tool for Metropolitan Transportation Planning. *Proceedings of the National Transportation Planning Applications Conference*, 20. https://www.trbappcon.org/2007conf/papers/session08/05-Comparison_of_ActivityBased_Tour_Model_to_FourStep_Model.pdf

Gunn, H., Burge, P., & Miller, S. (2006). The validation of the UK National Transport Model: A backcasting approach. *Proceedings of the European Transport Conference*. https://aetransport.org/past-etc-papers/conference-papers-pre-2012/conference-papers-2006?abstractId=2547&state=b

Harms, L. (2006). *Anders onderweg; de mobiliteit van allochtonen en autochtonen vergeleken*. Sociaal en Cultureel Planbureau. https://www.researchgate.net/profile/Lucas-Harms/publication/283726203_Anders_onderweg_de_mobiliteit_van_allochtonen_en_autochtonen_vergeleken/links/5645bcb208ae54697fb98c3c/Anders-onderweg-de-mobiliteit-van-allochtonen-en-autochtonen-vergeleken.pdf?origin=publication_detail

Jittrapirom, P., Caiati, V., Feneri, A.-M., Ebrahimigharehbaghi, S., González, M. J. A., & Narayan, J. (2017). Mobility as a Service: A Critical Review of Definitions, Assessments of Schemes, and Key Challenges. *Urban Planning*, *2*(2), 13–25.

KiM. (2016). *Ruimtelijke kenmerken, geografische bereikbaarheid en reisgedrag*. Ministerie van Infrastructuur en Milieu. https://www.kimnet.nl/binaries/kimnet/documenten/rapporten/2016/08/11/ruimtelijke-kenmerken-geografische-bereikbaarheid-en-reisgedrag/ruimtelijke-kenmerken-geografische-bereikbaarheid-en-reisgedrag.pdf

KiM. (2022). *Het wijdverbreide autobezit in Nederland*. Ministerie van Infrastructuur en Milieu. https://www.kimnet.nl/binaries/kimnet/documenten/publicaties/2022/02/22/het-wijdverbreide-autobezit-in-nederland/KiM+brochure+Het+wijdverbreide+autobezit+in+Nederland_def+A.pdf

Lange, P., & Huber, S. (2015). *Backcasting in freight transport demand modelling – chances and challenges*. https://doi.org/10.13140/RG.2.1.4636.5922

Lemp, J. D., McWethy, L. B., & Kockelman, K. M. (2007). From Aggregate Methods to Microsimulation: Assessing Benefits of Microscopic Activity-Based Models of Travel Demand. *Transportation Research Record*, *1994*(1), 80–88. https://doi.org/10.3141/1994-11

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (2022). *Nationale woon- en bouwagenda*. https://open.overheid.nl/repository/ronl-0343841159fc06a67a58b04ad520068192c521d1/1/pdf/nationale-woon-en-bouwagenda.pdf

Omer, M., Kim, H., Sasaki, K., & Nishii, K. (2010). A tour-based travel demand model using person trip data and its application to advanced policies. *KSCE Journal of Civil Engineering*, *14*(2), 221–230. https://doi.org/10.1007/s12205-010-0221-6

Ortúzar, J. D. D., & Willumsen, L. G. (2011). *Modelling Transport* (4th edition). Wiley.

Pots, M. (2018). *Gravity model parameter calibration for large scale strategic transport models* [Master’s thesis, DAT.Mobility and University of Twente]. https://essay.utwente.nl/76821/1/Pots_MA_EEMCS.pdf

Rasouli, S., & Timmermans, H. (2014). Activity-based models of travel demand: Promises, progress and prospects. *International Journal of Urban Sciences*, *18*(1), 31–60. https://doi.org/10.1080/12265934.2013.835118

Richards, M. G. (1974). Disaggregate simultaneous urban travel demand models: A brief introduction. *Transportation*, *3*(4), 335–342. https://doi.org/10.1007/BF00167964

Roorda, M. J., Miller, E. J., & Habib, K. M. N. (2008). Validation of TASHA: A 24-h activity scheduling microsimulation model. *Transportation Research Part A: Policy and Practice*, *42*(2), 360–375. https://doi.org/10.1016/j.tra.2007.10.004

Sammer, G., Gruber, C., Roeschel, G., Tomschy, R., & Herry, M. (2018). The dilemma of systematic underreporting of travel behavior when conducting travel diary surveys – A meta-analysis and methodological considerations to solve the problem. *Transportation Research Procedia*, *32*, 649–658. https://doi.org/10.1016/j.trpro.2018.10.006

Schneider, F., Ton, D., Zomer, L.-B., Daamen, W., Duives, D., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2021). Trip chain complexity: A comparison among latent classes of daily mobility patterns. *Transportation*, *48*(2), 953–975. https://doi.org/10.1007/s11116-020-10084-1

Shelton, J., Valdez, G. A., & Martin, P. (2016). How Accurate are Travel Forecasts: Back Casting of Truck Lane Restrictions using Multi-Resolution Modeling Methods. *Journal of Engineering and Architecture*, *4*(2). https://doi.org/10.15640/jea.v4n2a5

Vovsha, P. (2019). Decision-Making Process Underlying Travel Behavior and Its Incorporation in Applied Travel Models. In E. Bucciarelli, S.-H. Chen, & J. M. Corchado (Eds.), *Decision Economics. Designs, Models, and Techniques for Boundedly Rational Decisions* (pp. 36–48). Springer International Publishing.

Vovsha, P., & Bradley, M. (2006). Advanced Activity-Based Models in Context of Planning Decisions. *Transportation Research Record*, *1981*(1), 34–41. https://doi.org/10.1177/0361198106198100106

Walker, J. L. (2005). Making Household Microsimulation of Travel and Activities Accessible to Planners. *Transportation Research Record*, *1931*(1), 38–48. https://doi.org/10.1177/0361198105193100105

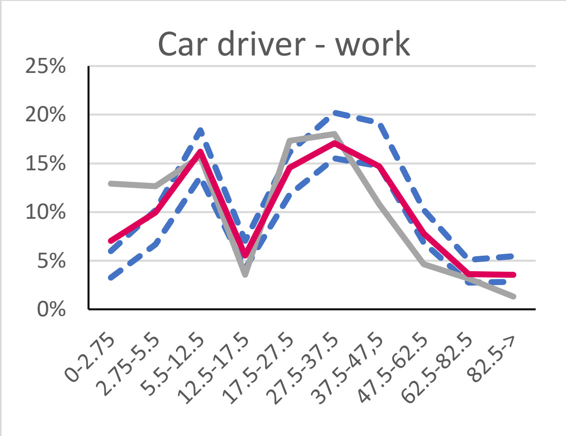
Wolf, J., Oliveira, M., & Thompson, M. (2003). Impact of Underreporting on Mileage and Travel Time Estimates: Results from Global Positioning System-Enhanced Household Travel Survey. *Transportation Research Record*, *1854*(1), 189–198. https://doi.org/10.3141/1854-21

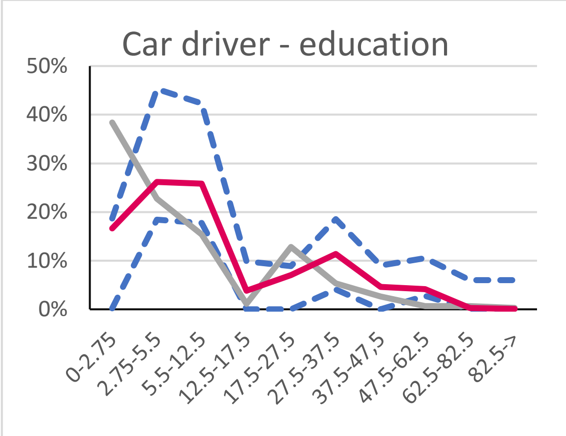
Ye, X., Pendyala, R. M., & Gottardi, G. (2007). An exploration of the relationship between mode choice and complexity of trip chaining patterns. *Transportation Research Part B: Methodological*, *41*(1), 96–113. https://doi.org/10.1016/j.trb.2006.03.004

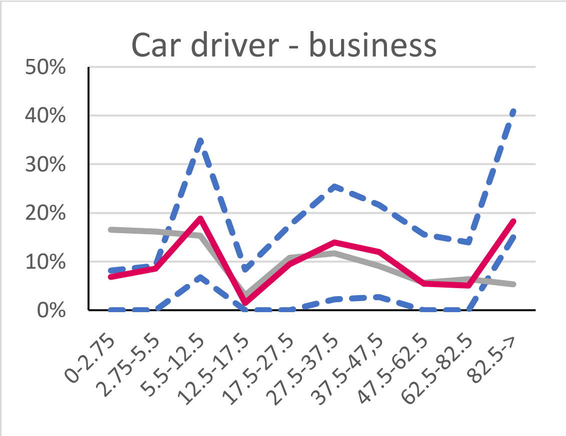
Appendix A – Regionalisation result

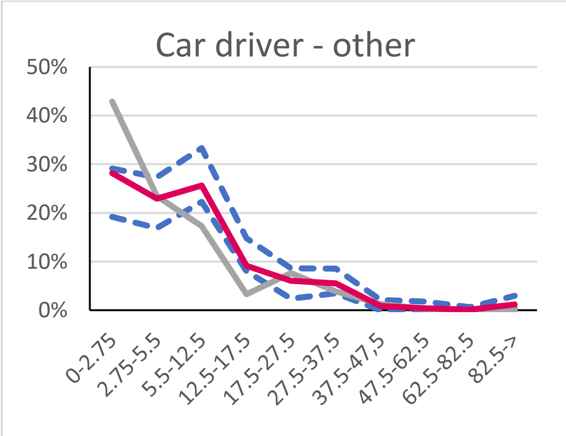
In the figures below, the output of Octavius without and with regionalisation are demonstrated. Note that these graphs demonstrate the result on the level the ASCs were estimated: tours, including home-based and non-home-based trips, car passenger, and evaluated against OViN 2010-2017. Hence, the figures show the regionalisation effect at the calibration level. Figure A.1 demonstrates the modal split results, whereas Figure A.2 shows the trip length distribution. Note that the legend has been omitted in the trip length distribution graphs to improve readability. However, they follow the same logic as for the modal split: blue = OViN 2010-2017, grey = not regionalised, and magenta = regionalised.

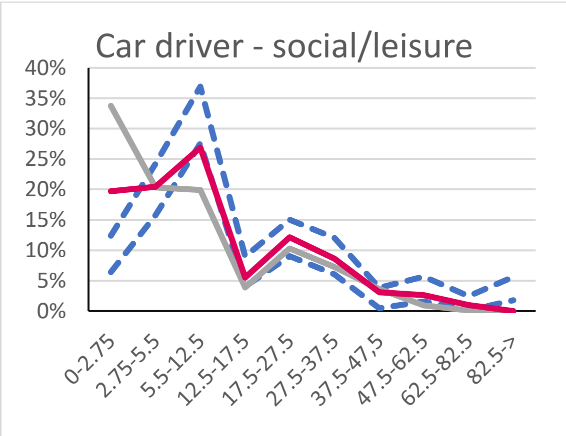
Figure A.1 - Octavius' modal split per purpose not regionalised versus regionalised. Note that the y-axes have different scales

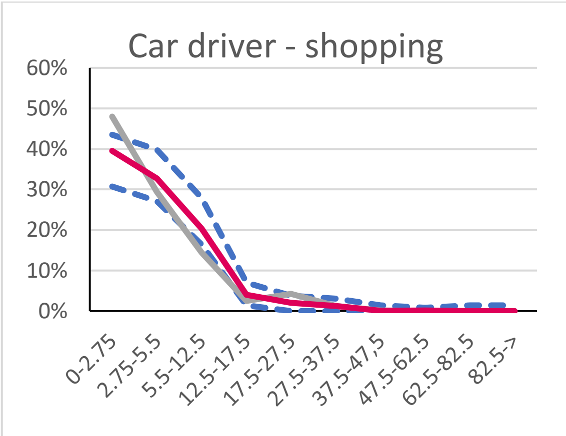

Car driver - work
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->

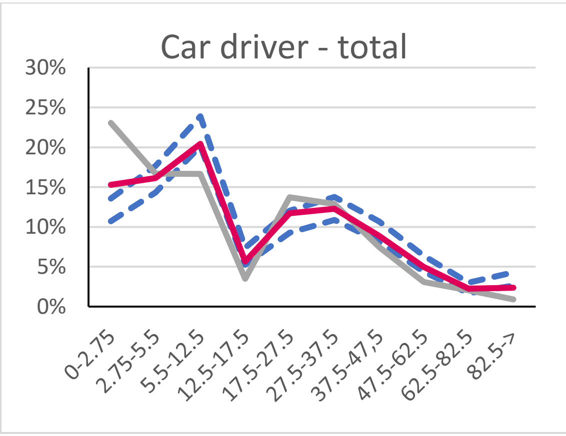

Car driver - education
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->

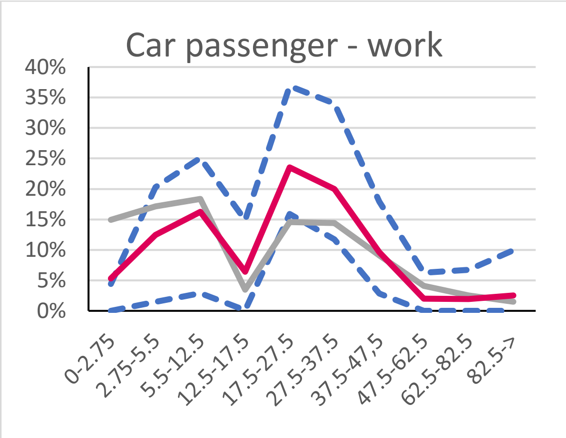

Car driver - business
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->

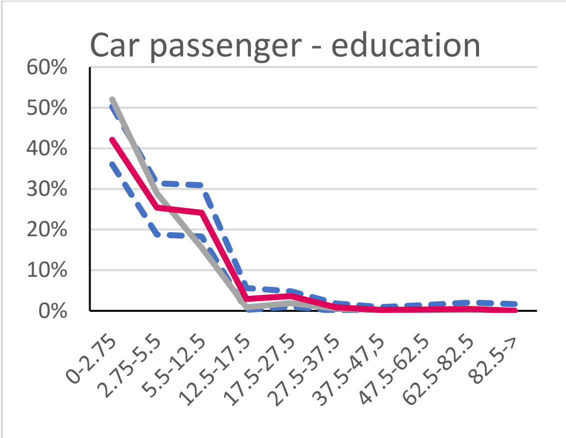

Car driver - other
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->

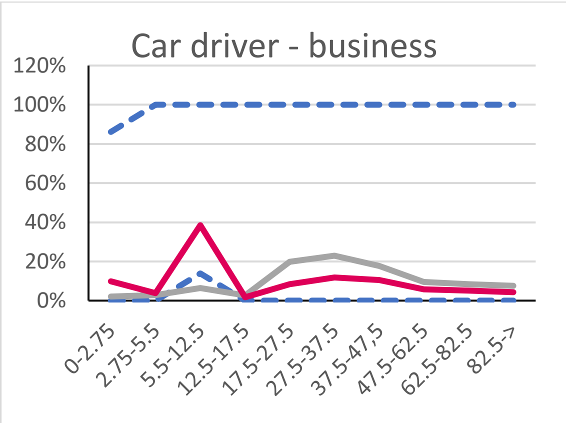

Car driver - social/leisure
40%
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->

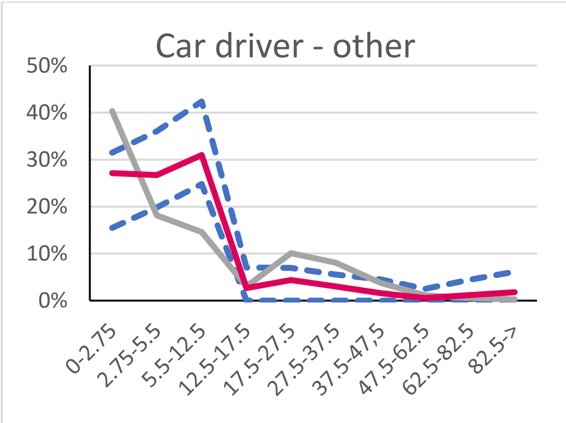

Car driver - shopping
60%
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->

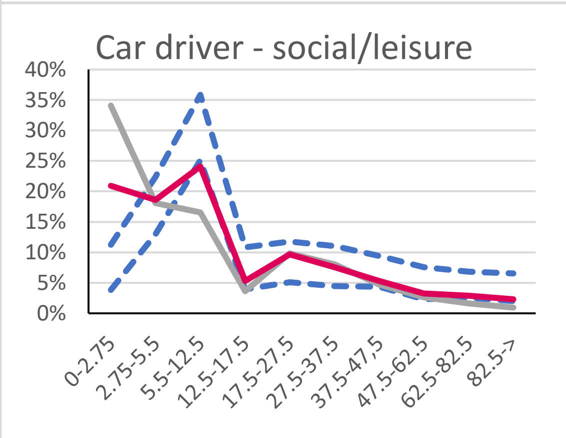

Car driver - total
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->

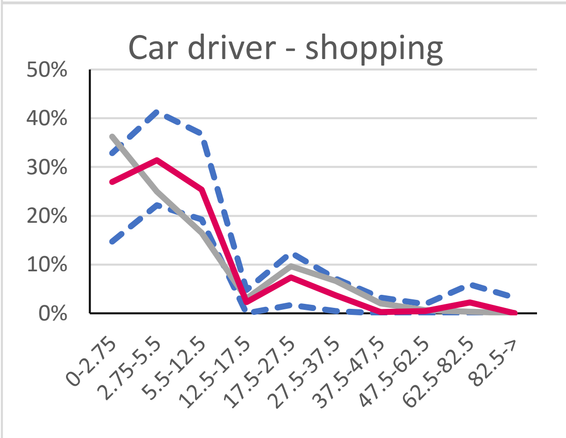

Car passenger - work
40%
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

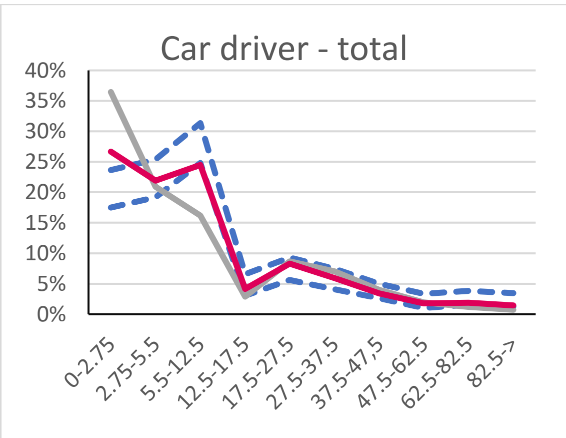

Car passenger - education
60%
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

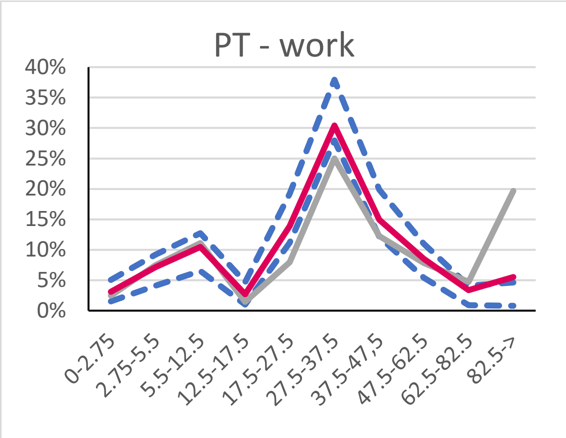

Car driver - business
120%
100%
80%
60%
40%
20%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

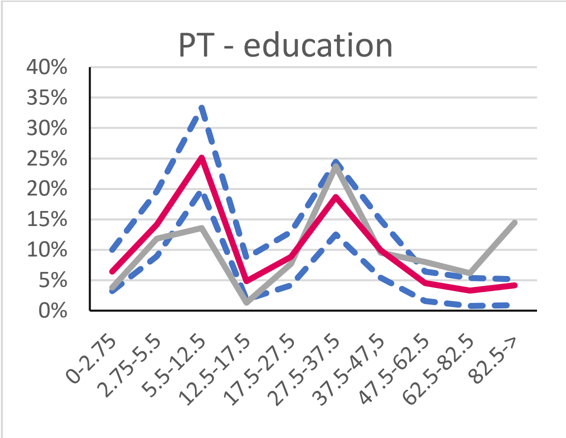

Car driver - other
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

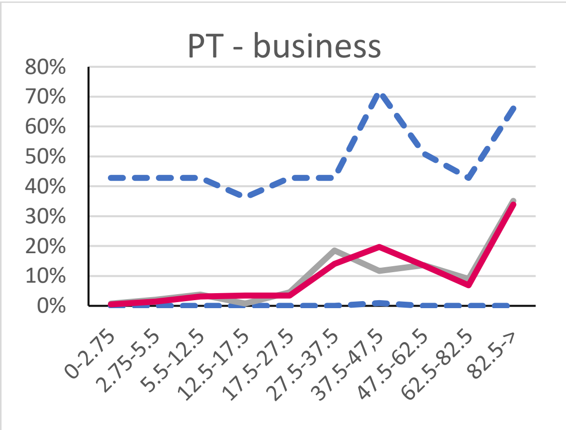

Car driver - social/leisure
40%
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

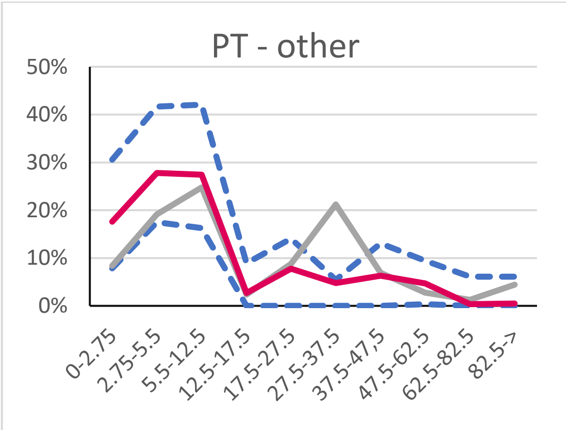

Car driver - shopping
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

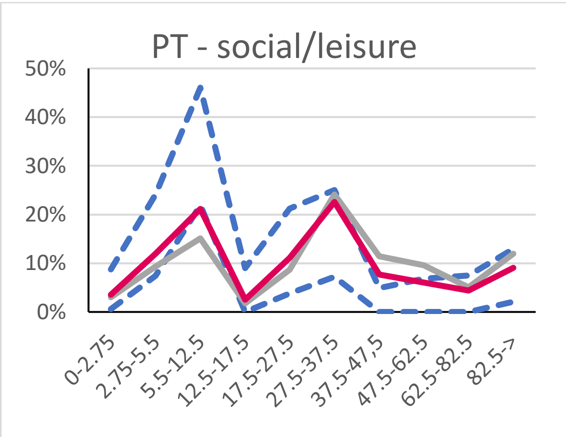

Car driver - total
40%
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

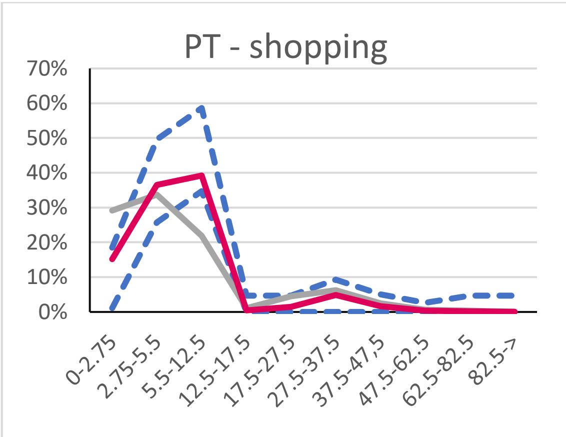

PT - work
40%
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->

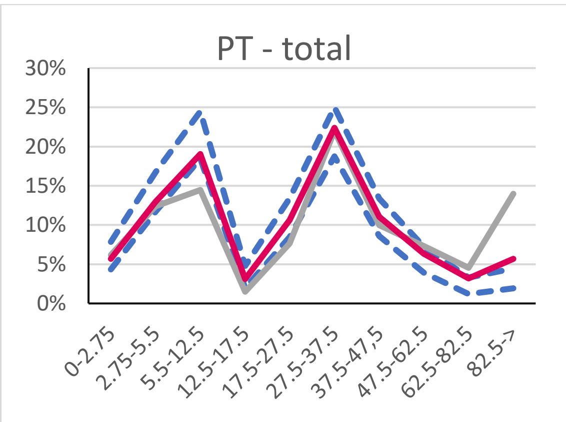

PT - education
40%
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->


PT - business
80%
70%
60%
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->


PT - other
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->


PT - social/leisure
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->


PT - shopping
70%
60%
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->


PT - total
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47,5
47.5-62.5
62.5-82.5
82.5->

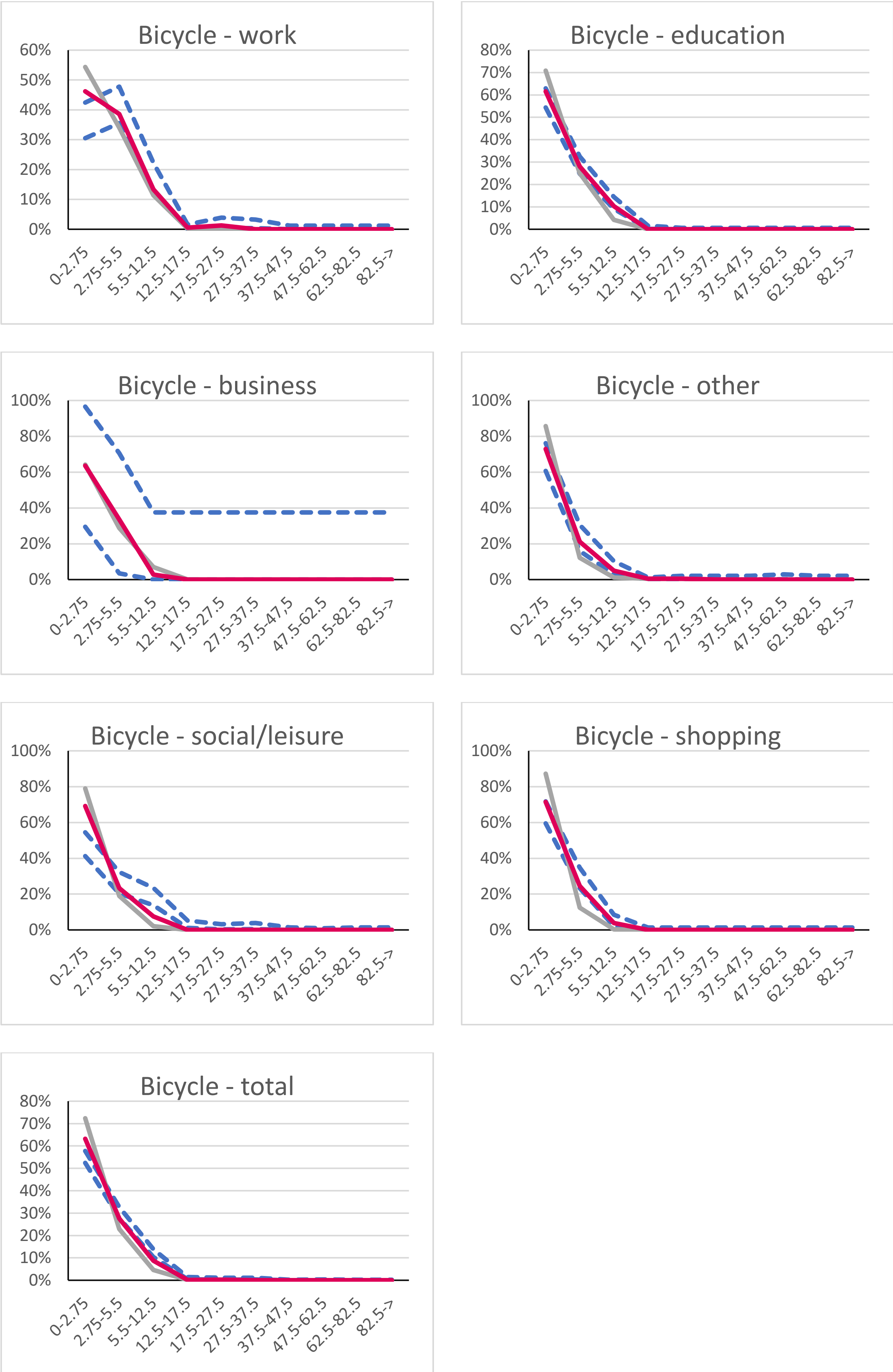


Figure A.2 - Octavius' trip length distribution per purpose not regionalised versus regionalised. Note that the y-axes have different scales

Appendix B – Manual adjustments to input data 2010

Appendix B.1: Number of jobs

In addition to the process described in section 5.4.6, manual adjustments were required for municipalities which merged between 2010 and 2018. This issue is illustrated by Figure B.1, where Muiden, Naarden, and Bussum have merged into one new municipality called 'Gooise Meren'. Between 2010 and 2018, 51 municipalities have merged into 16 new municipalities. To address this mismatch, old and new municipalities were linked based on publicly available information. Subsequently, the former municipalities were summed to determine the change with respect to the 2018 classification. The change was calculated according to, and subsequently assigned to, each of constituent municipalities of the 2010 municipalities.

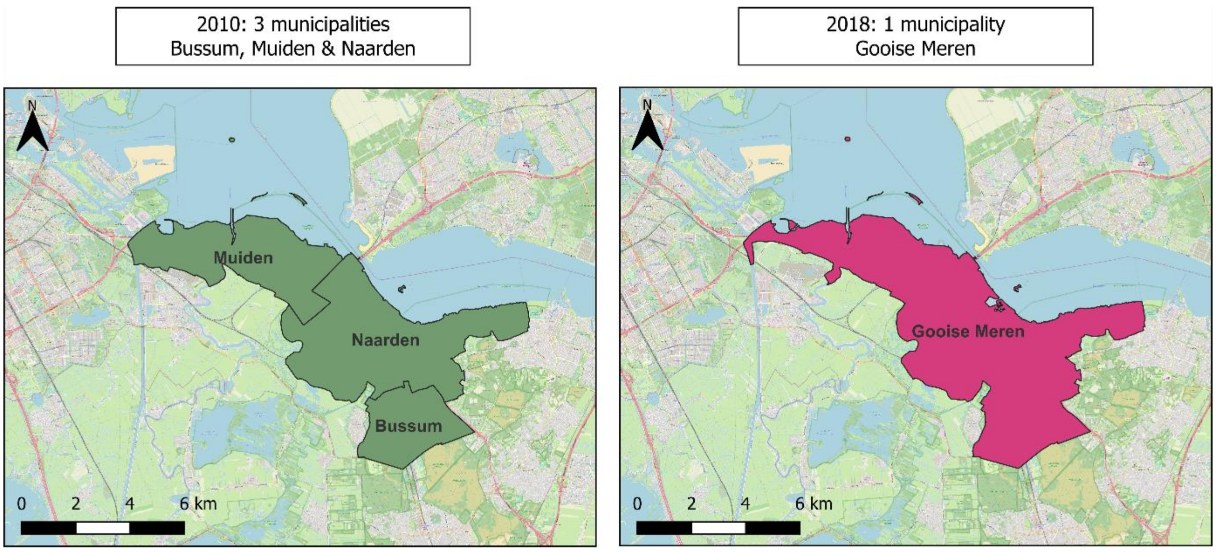


Figure B.1 - Example of municipalities that merged between 2010 and 2018

In the case of Gooise Meren, the number of jobs in 2010 in Muiden, Naarden, Bussum were summed and subsequently divided by the number of jobs in Gooise Meren in 2018. The result is that in 2010 1,029 more jobs were located in these municipalities compared to 2018:

$$\Delta J_{Muiden,Naarden,Bussum} = \frac{(J_{muiden,2010} + J_{Naarden,2010} + J_{Bussum,2010})}{J_{Gooise\ Meren,2018}} = \frac{21200}{20600} = 1{,}029 \quad (B.1.1)$$

Therefore, the number of jobs in 2018 were multiplied with the calculated factor. For example:

$$J_{office,Muiden,2010} = J_{office,Muiden,2018} * \Delta J_{Muiden,Naarden,Bussum} = 1400 * 1{,}029 = 1441 \quad (B.1.2)$$

For municipalities that merged with other municipalities (e.g., Graft-De Rijp with Alkmaar), a similar procedure was performed.

Appendix B.2: Student enrolments

In addition to the student distribution taken from OmniTRANS Spectrum, some manual adjustments had to be made, because some organisations were discontinued or merged between 2010 and 2018. In Figure B.2, the adjustment process is visualised. In 2010 ROC Arcus college and VISTA college existed as separate organisations, for which the total number of students is known from DUO data (but not the distribution of its locations). However, ROC Arcus was discontinued between 2010 and 2018, and merged with VISTA college. As a result, ROC Arcus was not in the 2018 data of DUO or MobiSpec. What was available for 2018, was the distribution of students among the various locations of VISTA college.

Therefore, the student numbers of ROC Arcus and VISTA were summed for 2010 and subsequently distributed according the 2018 distribution over the locations of VISTA college. Although this was not the reality, it is reasonable to assume that, in this case, VISTA is located in the vicinity of ROC Arcus, because MBO students are not likely to travel significantly longer. On top of that, manual adjustments were made for organisations outside the study area, meaning that the exact locations are of less importance, since 1) the TAZ size outside the study area is larger, and 2) the further away from Almere, the less likely that Almeerders will go to that educational institution.

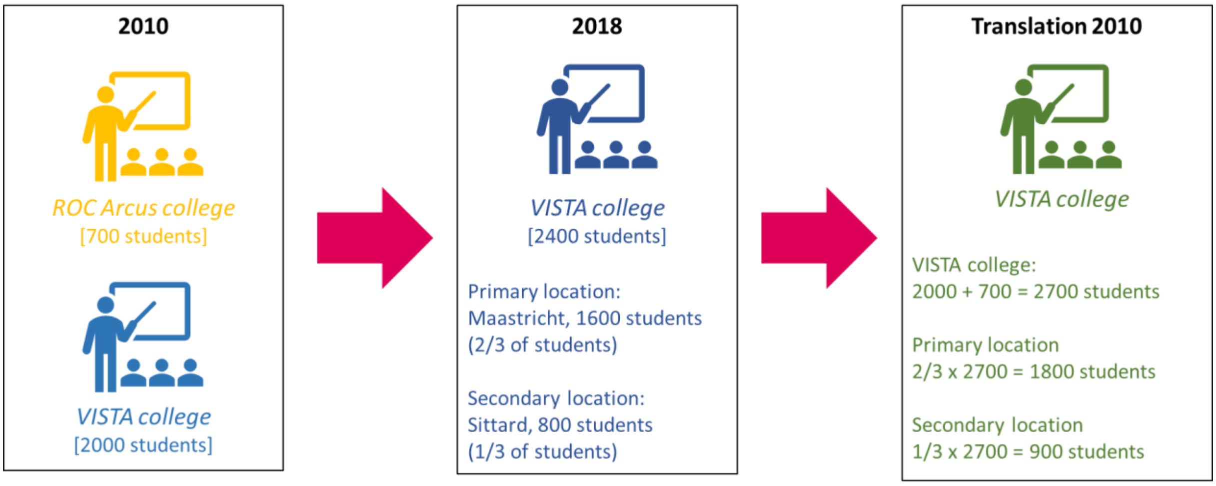


Figure B.2 - Example of redistribution of student enrolments between 2010 and 2018

For Higher Professional Educations (HBO) and universities, the distribution of students over the various locations in MobiSpec 2018 could be directly applied to the 2010 student numbers, since the organisations of 2018 are the same as those of 2010. One exception was *NHL Stenden Hogeschool*, which existed in 2010 as two separate organisations. Although HBO students are more likely to travel further, this organisation is located in Leeuwarden which falls within a TAZ defined at the COROP level. Hence, the exact location of the organisations is not relevant. Lastly, the translated MobiSpec locations were joined with the TAZs to determine the socioeconomic variable *student enrolments*.

Appendix C – Network adjustments for 2010

The adjustments made to the 2017 network to represent the 2010 situation are listed in Table B.1. Note that the changes are based on the differences between the 2010 and 2017 networks that were both available at Goudappel.

Table B.1 – 2017 network adjustments

Action	Result
Train connection 'Hanzelijn' removed	Intercity The Hague – Zwolle – Leewarden ends at Lelystad; Intercity The Hague – Zwolle – Groningen ends at Lelystad; Sprinter The Hague – Zwolle ends at Lelystad.
Train stop Almere Poort removed	Sprinter Utrecht – Almere Centrum skips Almere Poort; Sprinter Amsterdam – Zwolle skips Almere Poort; Sprinter Hoofddorp – Almere Oostvaarders skips Almere Poort.
Bus frequency on lines passing Almere Poort adjusted (both line directions)	Line Almere Poort station – Almere Centrums station: 12/h to 6/h during rush hour and 8/h to 4/h for off-peak; Line Almere Centrum station – Amsterdam Bijlmer station: 2/h to 4/h during rush hour and 1/h to 4/h for off-peak hours; Line Almere Stad Parkwijk – Amsterdam Amstel station: 2/h to 4/h during rush hour and 1/h to 4/h for off-peak hours; Line Almere Poort station – Almere Muziekwijk station and Almere Centrum station: 6/h to 4/h during rush hour and 4/h to 4/h for off-peak hours.
Highway speed limits adjusted	A6 east of Almere 130 km/h → 120 km/h (from interchange Emmeloord to exit Almere Buiten Oost); A6 south of Almere 130 km/h → 120 km/h (from interchange Soest to interchange Almere); A6 south of Almere 100 km/h → 120 km/h (from interchange Gooimeer to interchange Almere).

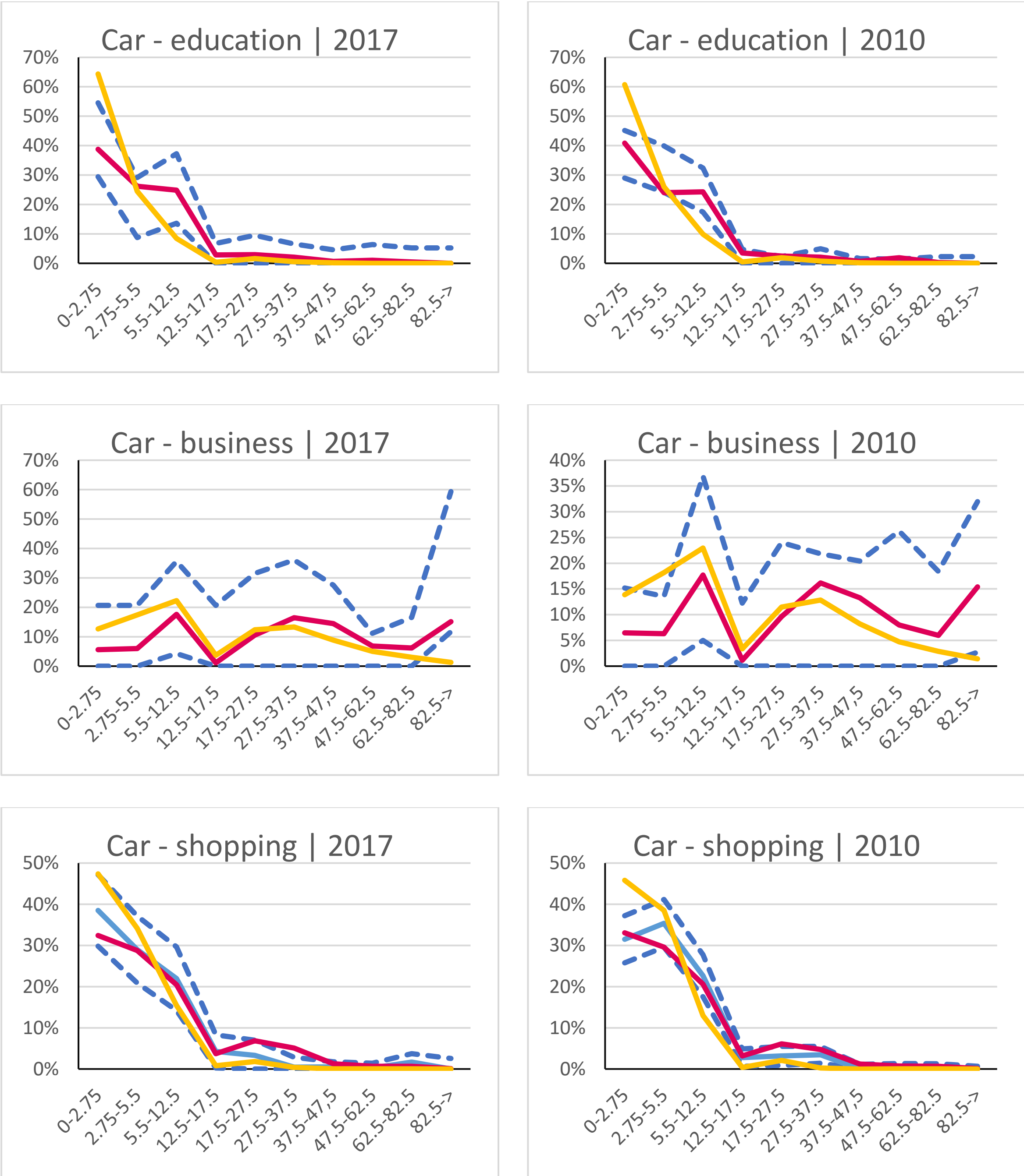
The car network inside Almere was checked between both model versions (2017 and 2010), and no meaningful changes were found. The bike 2017 network was more detailed for residential areas (see Figure C.1 and Figure C.2), which was kept the same for the backcast, since in 2010 these streets did exist in reality, but simply not in the model.

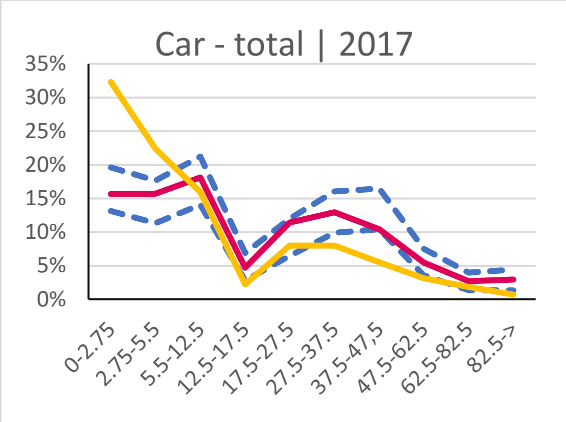
Figure C.1 - Example of bicycle network in 2017 network

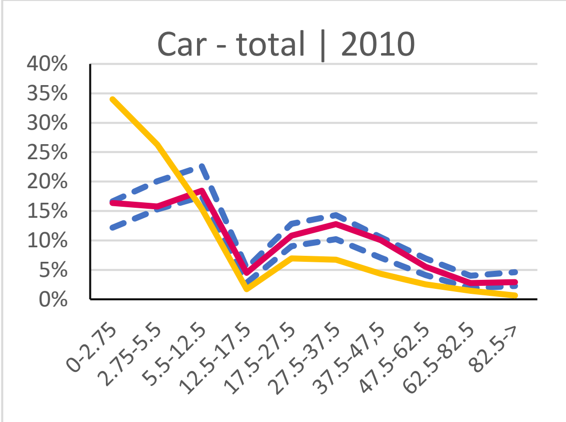
Figure C.2 - Example of bicycle network in 2010 network

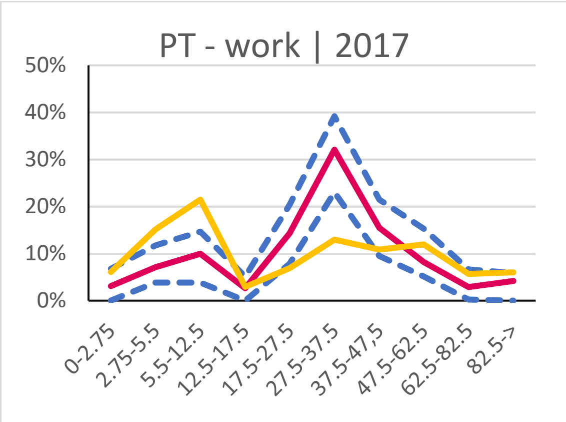
Appendix D – Trip length distribution graphs

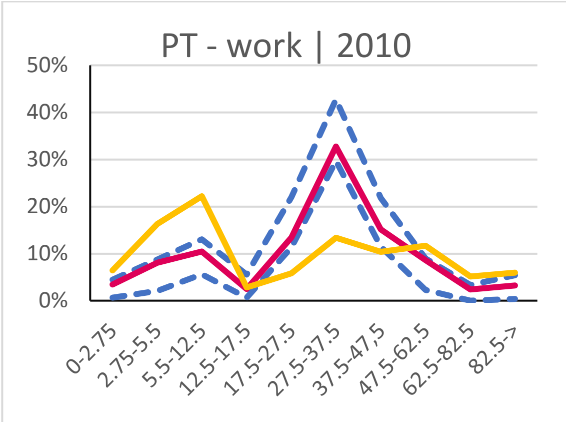
In Figure D.1, the trip length distributions per mode-purpose are depicted. Note that for all figures the legend is: blue dotted → OViN 95% confidence interval, magenta → Octavius, yellow → 4-step model.

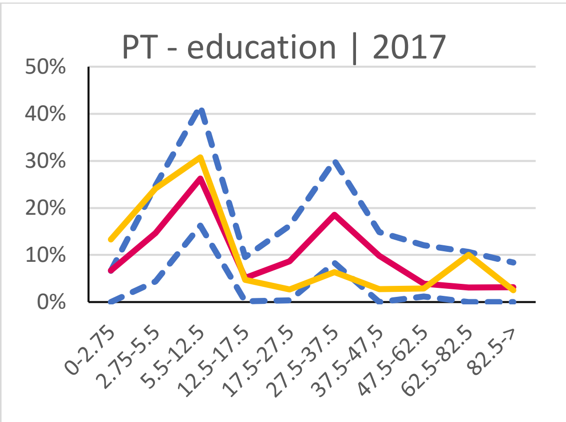


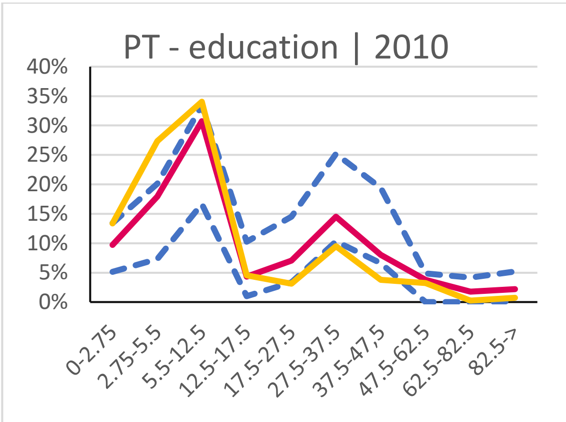

Car - total | 2017
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

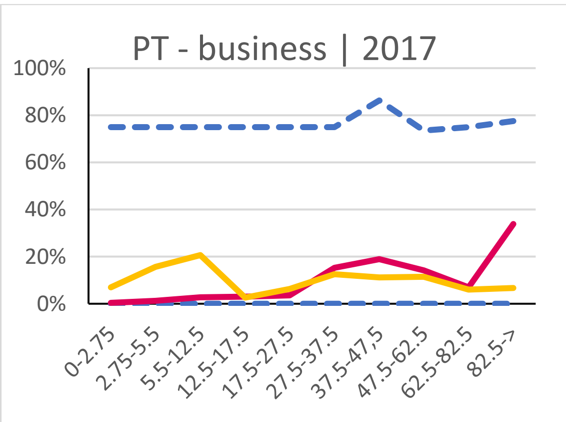

Car - total | 2010
40%
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

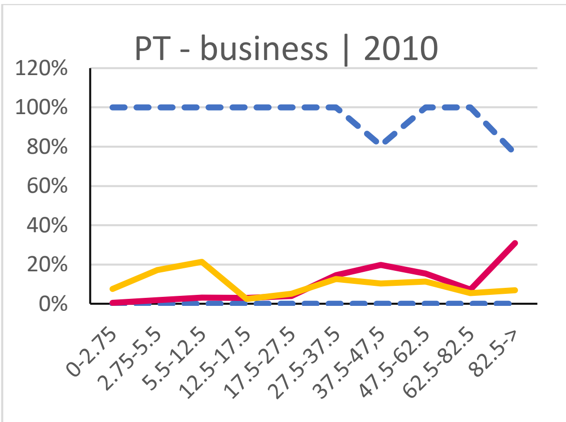

PT - work | 2017
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

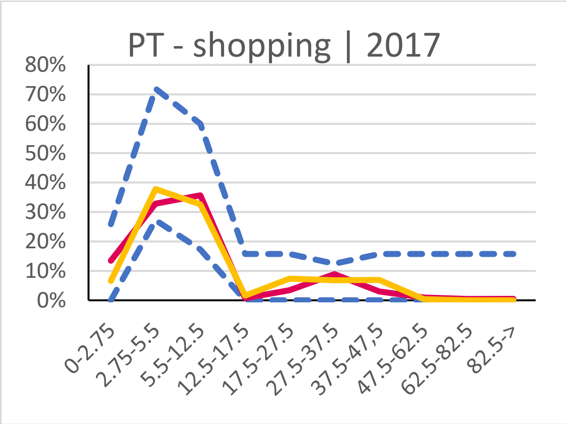

PT - work | 2010
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

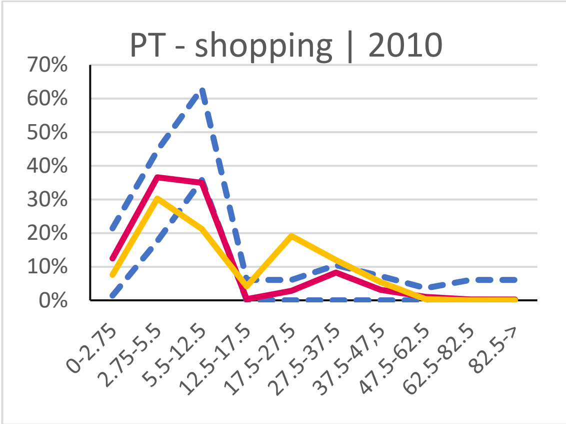

PT - education | 2017
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

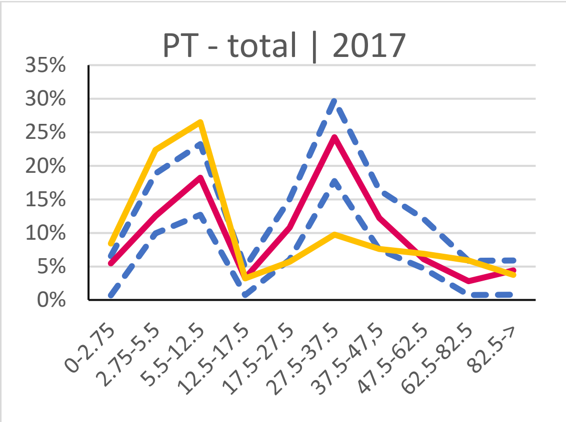

PT - education | 2010
40%
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

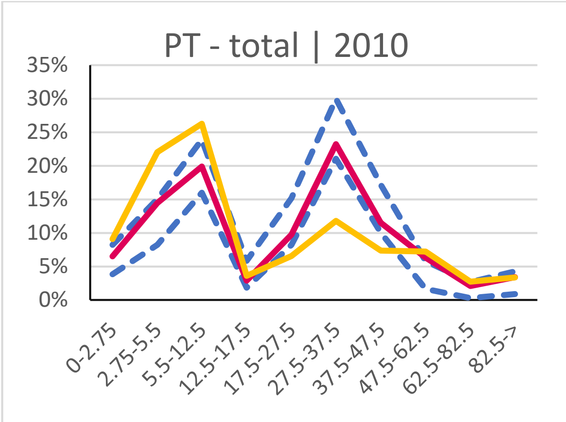

PT - business | 2017
100%
80%
60%
40%
20%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

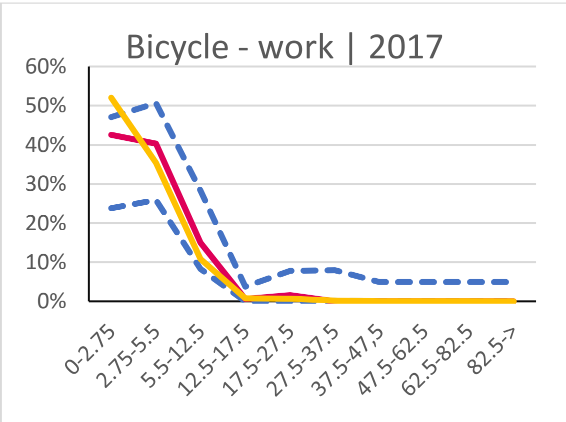

PT - business | 2010
120%
100%
80%
60%
40%
20%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

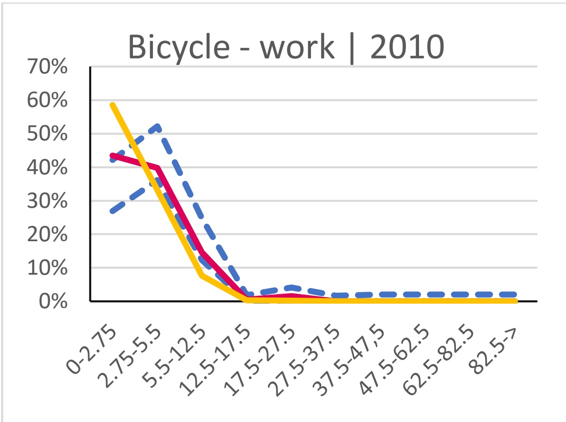

PT - shopping | 2017
80%
70%
60%
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

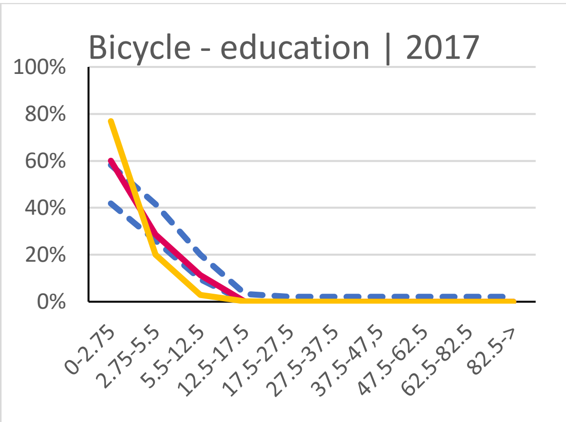

PT - shopping | 2010
70%
60%
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

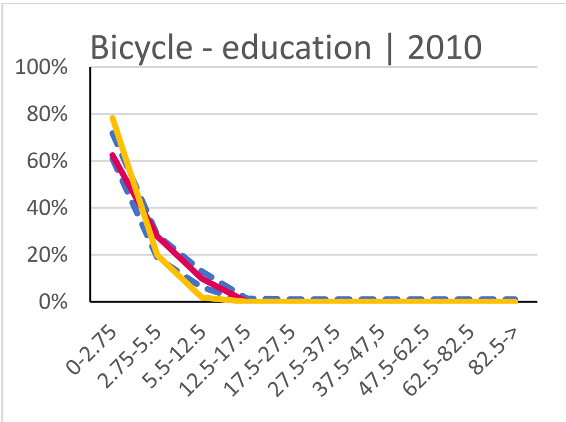

PT - total | 2017
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->


PT - total | 2010
35%
30%
25%
20%
15%
10%
5%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->


Bicycle - work | 2017
60%
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->


Bicycle - work | 2010
70%
60%
50%
40%
30%
20%
10%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->


Bicycle - education | 2017
100%
80%
60%
40%
20%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->


Bicycle - education | 2010
100%
80%
60%
40%
20%
0%
0-2.75
2.75-5.5
5.5-12.5
12.5-17.5
17.5-27.5
27.5-37.5
37.5-47.5
47.5-62.5
62.5-82.5
82.5->

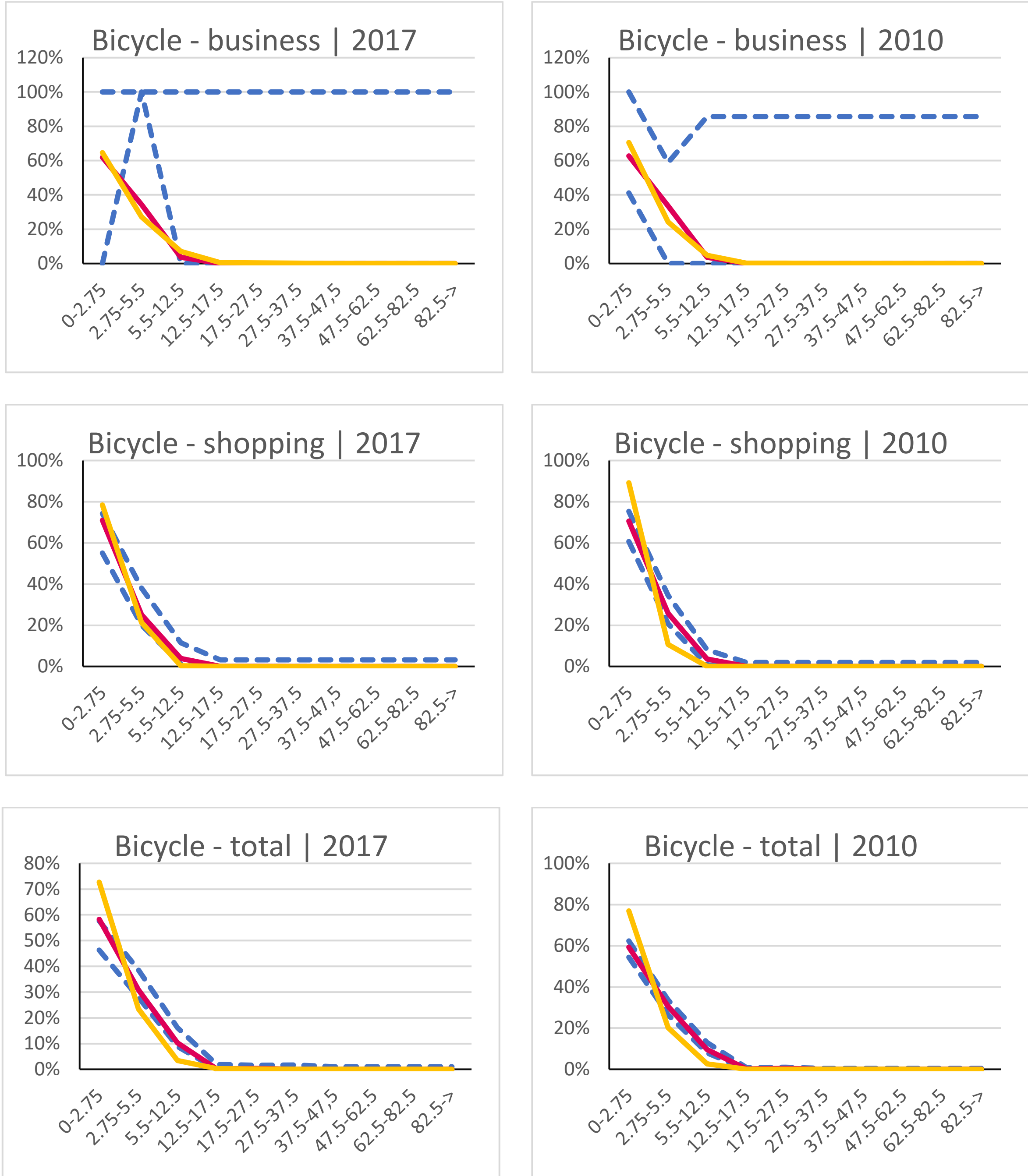


Figure D.1 - Modelled trip length distributions by Octavius (magenta) and 4-step model (yellow) compared to OViN 95% confidence interval (blue dotted)

Appendix E – Mode shares per distance class

In this appendix, the mode shares per distance class can be found which were used for the scenario analyses.

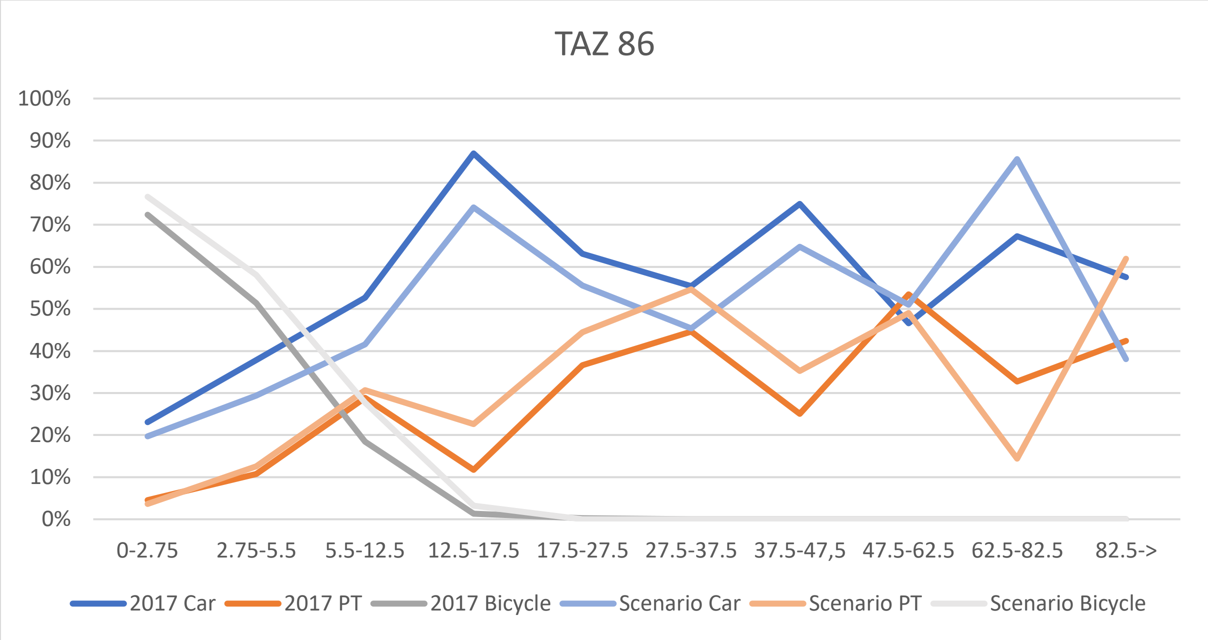


Figure E.1 - Mode share per distance class for TAZ 86, modelled by Octavius

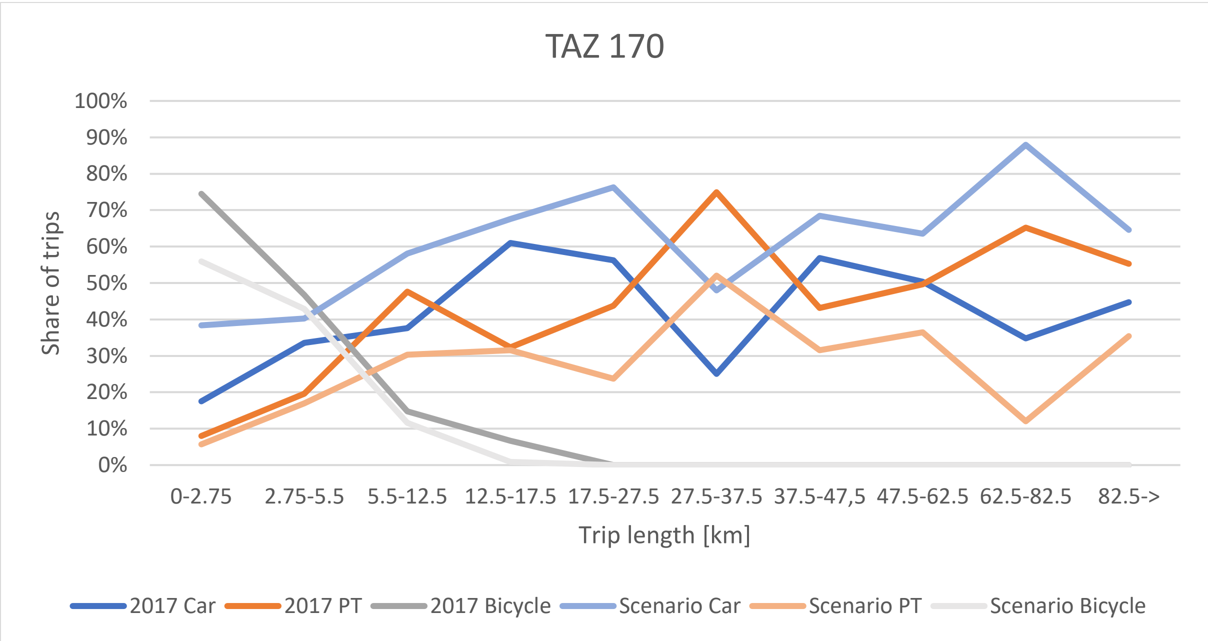


Figure E.2 - Mode share per distance class for TAZ 170, modelled by Octavius

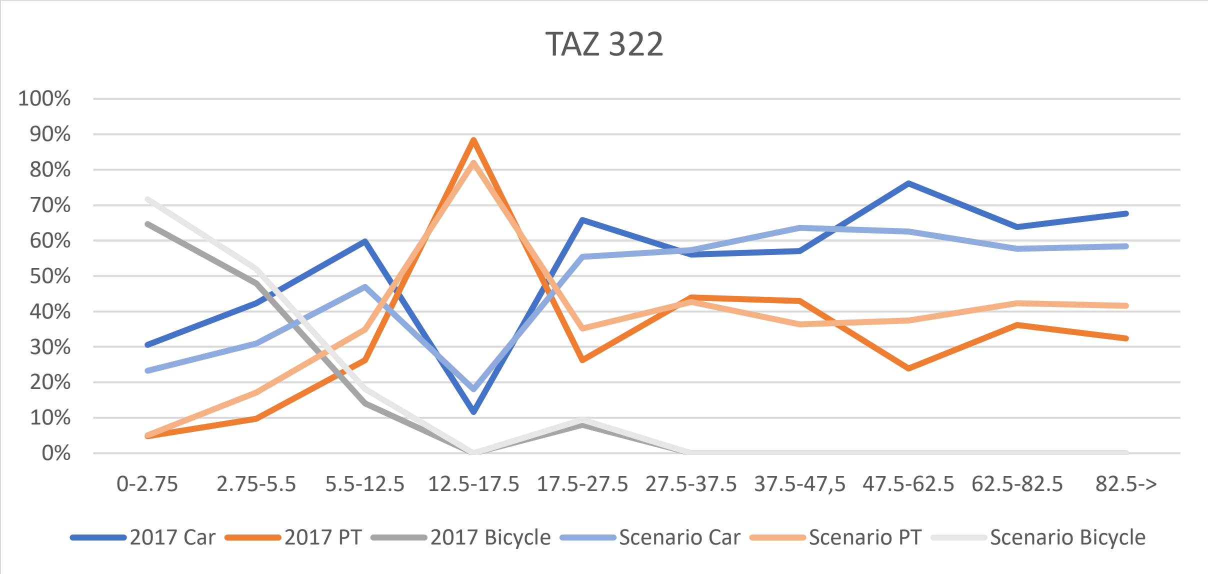


Figure E.3 - Mode share per distance class for TAZ 322, modelled by Octavius

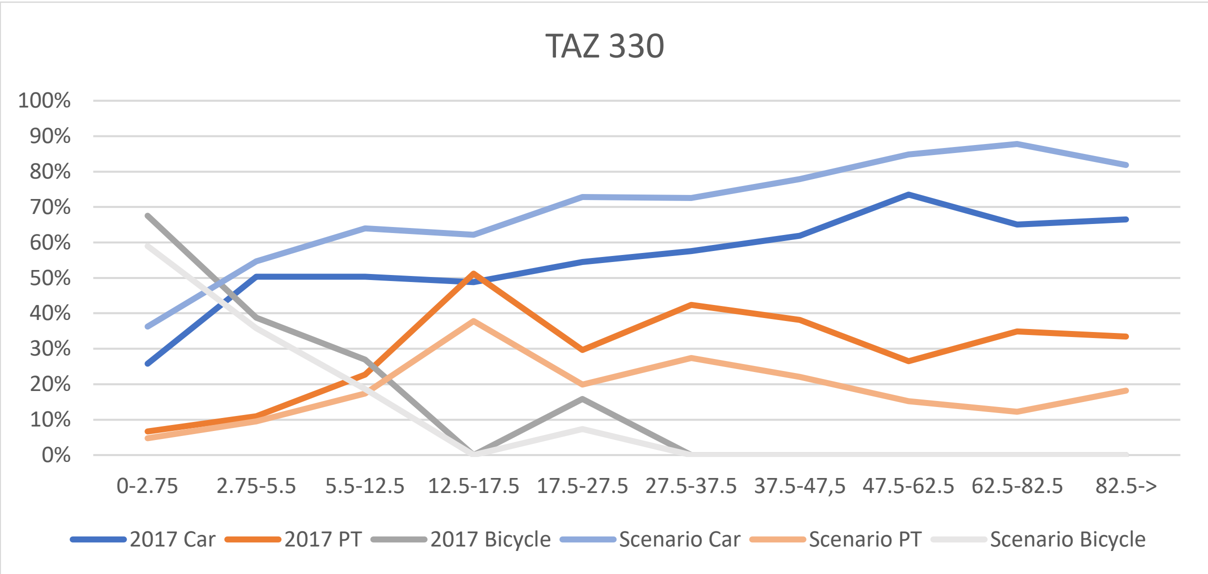


Figure E.4 - Mode share per distance class for TAZ 330, modelled by Octavius