How relevant is the aggregation bias with regards to mode choice?

Fabio Cachaco, FHNW
Alexander Erath, FHNW

STRC 2023 conference paper

STRC | **23rd Swiss Transport Research Conference**
Monte Verità / Ascona, May 10-12, 2023

How relevant is the aggregation bias with regards to mode choice?

Fabio Cachaco
University of Applied Sciences and
Arts Northwestern Switzerland (FHNW)
Hofackerstrasse 30, 4132 Muttenz

Phone: + 41 61 926 90 53
E-Mail: fabio.cachaco@students.fhnw.ch

Alexander Erath
University of Applied Sciences and
Arts Northwestern Switzerland (FHNW)
Hofackerstrasse 30, 4132 Muttenz

Phone: + 41 61 228 54 26
E-Mail: alexander.erath@fhnw.ch

May 11, 2023

Abstract

This paper examines the relevance and impact of aggregated trip- and disaggregated tour-based mode choice approaches using two models. For this purpose, a qualitative description of the effects of different approaches for modelling the mode choice is given. The focus lies especially on the aggregation bias and the utility function. In this context, a quantitative assessment of the relevance of the aggregation bias is made.

Estimation of the effect of aggregation bias is examined using three transportation planning measures. The model calculations suggest that an aggregated and trip-based approach compared to a disaggregated and tour-based approach leads to an overestimation of the changes in mode share (aggregation bias) by 0.5 to 2 percentage points when calculating the effect of the transport planning measures. Depending on the measure, the aggregated model reacted between 14% and 74% stronger than the disaggregated model with respect to the changed mode shares. Based on this finding, special care must be taken when using aggregate traffic models to calculate benefits in the infrastructure sector. This is especially true when modelling measures where we know that person-based decisions or characteristics (e.g., PT subscription ownership) play a critical role.

Keywords

Transport modelling, Activity-based model, Demand model, Disaggregation, Aggregation Bias, Modal split

Table of contents

1 Introduction

The 4-step model is the most commonly used model type for calculating traffic forecasts. The traditional aggregated and trip-based 4-step approach for modelling traffic demand has been successfully and intensively used worldwide by companies, institutions, authorities, and research institutions for over 50 years. With increasing use, new questions have arisen. Not all questions could be answered with the aggregated models, and the limitations of the 4-step approach became apparent with increasing use. The criticisms can be summarized as follows:

- Due to the use of zone average values in the respective behavioural models, there is an underestimation or overestimation of person-specific utility in the associated selection probabilities. This property is called aggregation bias in the literature.
- The four steps are considered conceptually independent and isolated within the approach. In reality, however, the four steps are strongly interconnected.
- The 4-step approach inadequately represents real behavioural patterns with increasing complexity of mobility behaviour due to the approach's limited flexibility (Rasouli and Timmermans 2014).
- The methodology of the 4-step approach is difficult to communicate to decision makers (Vitins et al. 2021).

This results in a considerable development potential for traffic models (Rasouli and Timmermans 2014). In the past decades, an increasing need for the further development of traffic models that can better represent the dependencies of different routes throughout the day and their temporal dynamics has been recognized. This insight has led to the development of new disaggregated and tour-based models, the so-called activity-based models (ABM) (Castiglione et al. 2015). The approaches by which the two types of models calculate mode choice differ significantly. In this paper, the relevance of the differences between aggregated and trip-based, as well as disaggregated and tour-based mode choice approaches, will be examined and estimated using an ABM as an example. This paper focuses on the recently mentioned Aggregation Bias. The goal of this paper is to answer the following research question based on the implementation of two pilot models:

How does a tour- and person-based disaggregated mode choice approach compare to a trip-based, aggregated approach in estimating the impact of certain traffic planning measures?

2 Modelling approaches

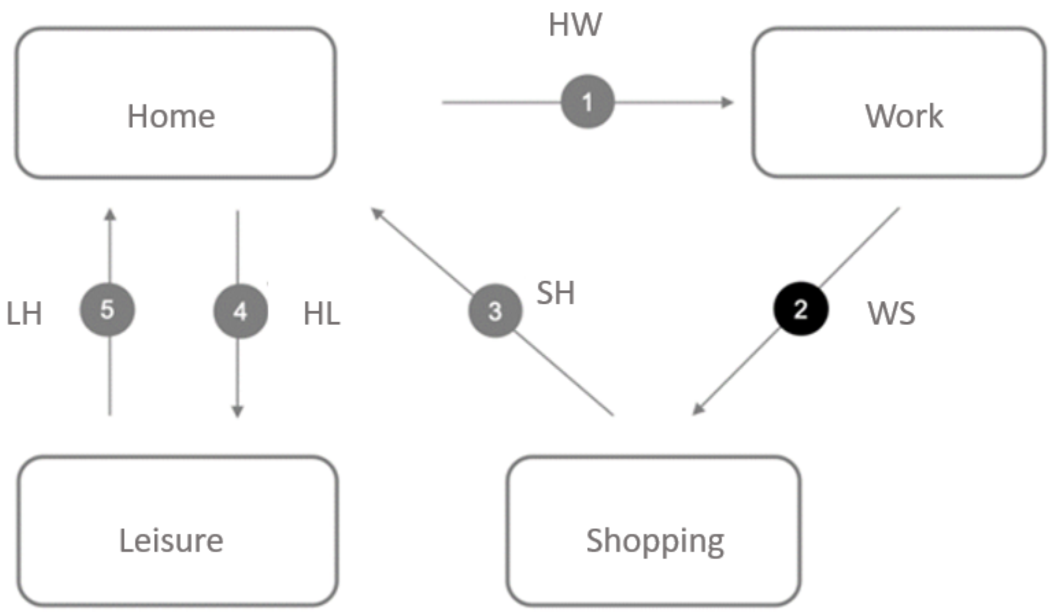
2.1 Mode choice

In a trip-based approach, mode choice is described based on individual trips without considering the information on individual location changes in trip chains. In general, trips are divided into two types:

- Type 1 (home-based): Home is either origin or destination.
- Type 2 (non-home-based): Home is neither origin nor destination.

In trip-based approaches, for home-based trips (Type 1), socio-demographic information such as car ownership can be considered in mode choice models based on the average values determined for the respective zones. However, the interdependencies of individual trips are not taken into account. In reality, trips are interdependent. For non-home-based trips (Type 2), only average values of the affected origin and destination zones can be used for mode choice, which do not have a direct functional relationship to these trips. Type no. 2 in Figure 1 represents a non-home-based trip (Type 2).

Figure 1: Example of a trip-based sequence



Source: Adapted from Vitins et al. 2021

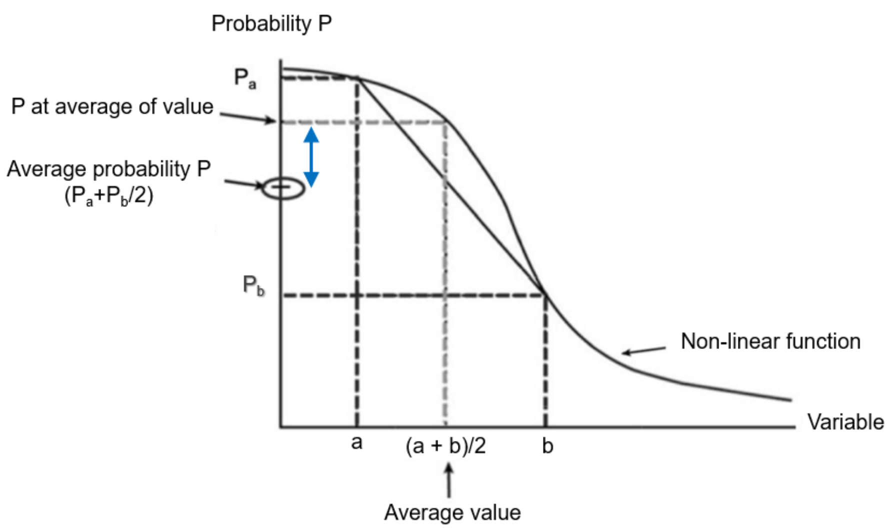
In contrast, tour-based models consider the entire information of within a tour of a person, from leaving the house until returning, including all associated trips (which consist of at least two trips). The modelling of mode choice can be done at the level of individual persons. In this way, all person-related attributes are used in the calculation, which allows for dependencies between individual trips to be considered, such as car ownership or PT subscription.

2.2 Aggregation Bias

In aggregated models, structural data of traffic zones (population density, jobs, mode shares etc.) are used as explanatory variables in various modelling steps. When the mode shares of several individuals for a traffic zone are averaged or aggregated and used in a statistically estimated model based on individual data, a distortion called aggregation bias occurs. In this paper, the focus is on the mode choice when examining the aggregation bias.

The aggregation bias arises due to the non-linear functions used to describe the selection probability as a function of the modelled utility of an alternative. When mean values are used to describe the utility, differences arise in the calculation of the probability compared to an approach that describes the utility and the corresponding selection probabilities at the disaggregated level, as the following example shows.

Figure 2: Illustration of aggregation bias



Source: Adapted from Vitins et al. 2021

The persons a and b each have two values of a decision-relevant variable (input variable), as shown on the horizontal axis in Figure 2. According to the non-linear logit function, the dark black dashed lines represent the corresponding choice probabilities for each person (P_a and P_b). In an aggregated model, as described earlier, only the corresponding mean value can be considered instead of the person-specific decision. As the figure shows, the resulting choice probability based on the mean value (e.g. a zonal mean value) differs from the initially person-specific calculated and then averaged value (blue arrow). This difference is called aggregation bias.

At low choice probabilities, the probability rated with the average representative utility (P at the medium value) is underestimated compared to the average probability (P_a+P_b/2), while at high choice probabilities it is overestimated. The logit model is most sensitive (high elasticity) to changes in the input variables (a and b) in its middle range. At the upper and lower ends, it reacts less sensitively (low elasticity) to changes in the input variables. According to Castiglione et al. (2015), this is one of the reasons why some aggregated models generate larger biases in response to changes in the input variables than disaggregated models.

3 Methodology

3.1 Overview

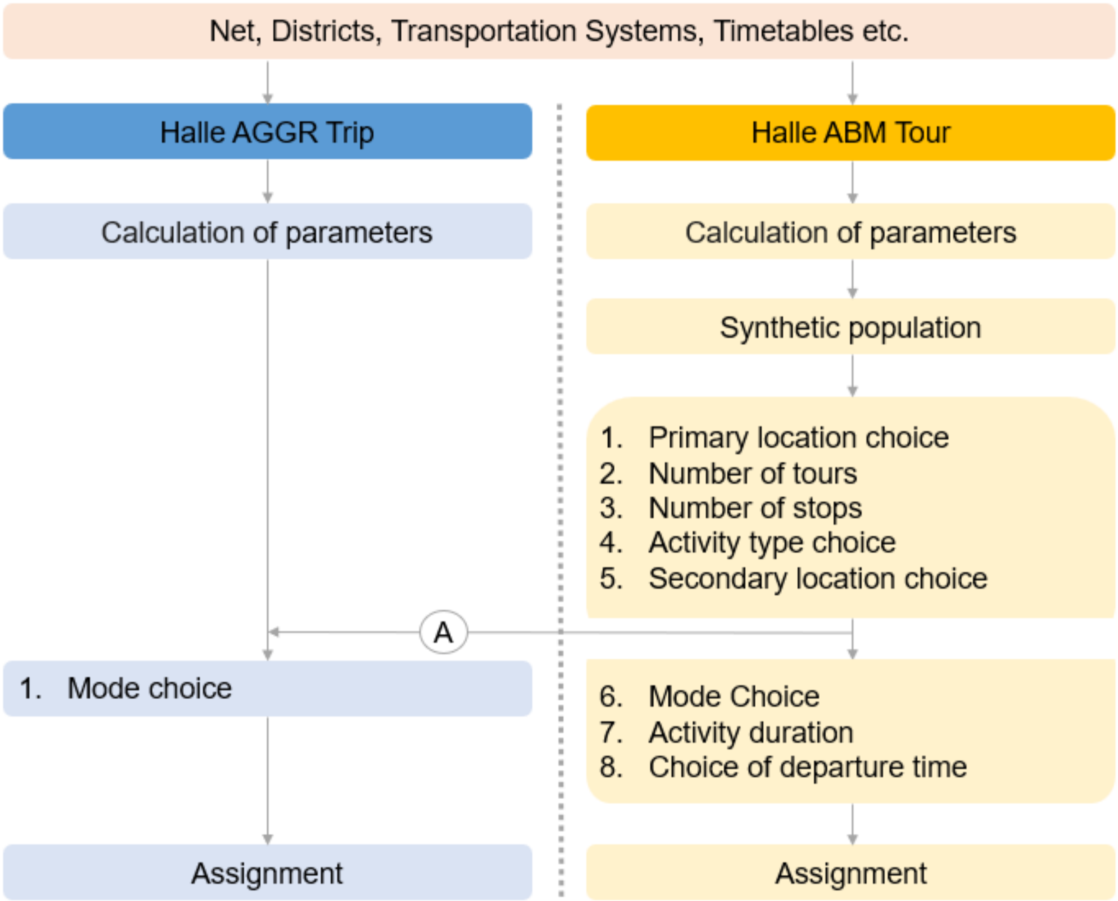
In this chapter, the development of the two mode choice models is described, which will later be used (in chapter 4) to estimate the effect of aggregation bias. To enable a meaningful comparison, two mode choice models are implemented based on the same underlying data. The first model is a disaggregated and tour-based mode choice model (hereafter referred to as **Halle ABM Tour**). Based on the demand of the first model, the aggregated and trip-based mode choice model (hereafter referred to as **Halle AGGR Trip**) was derived. Figure 3 provides an overview of the structure of the two mode choice models.

The traffic generation and traffic distribution steps of the classical 4-step approach are not modelled in the aggregated model Halle AGGR Trip. Instead, after the step secondary destination choice from the disaggregated model Halle ABM Tour, the demand is aggregated. The aggregation in step A (see Figure 3) was implemented using an R script[1]. At the time of aggregation, the Halle ABM Tour model has disaggregated demand in the form of trips, which are aggregated by traffic purpose at the zone level and loaded into the Halle AGGR Trip model via the process flow of PTV Visum.

The aim of this approach is to implement two models with identical networks and the same traffic distribution. This allows the analysis of the effect of aggregation bias specifically on mode choice, while modelling various transport planning measures. The existing net sizes and infrastructure offers are identical in both models. Individual person-specific attributes such as car ownership or PT subscription were aggregated at the zone level as average values for the Halle AGGR Trip model.

1 https://gitlab.fhnw.ch/fhnw_vm/doku_master_thesis_fabio_cachaco/-/blob/main/06_skripte_auswertung/4_abm_demand_to_aggr_matrices.R

Figure 3: Structure of the two compared mode choice models



Source: Own illustration.

3.2 Implementation of the activity-based and disaggregated model

The basis for the development of the Halle ABM Tour transport model is the activity-based model (ABM) of the city of Halle, which is provided with the PTV Visum software. The model simulates the behaviour of each individual person in the population using a given synthetic population. The necessary network and structural data are also included in the model. The basic ABM Halle model generates a complete daily plan for each person in the population, including all activities and corresponding trips.

In comparison to the basic model, the information about PT subscription ownership was added to the synthetic population for the Halle ABM Tour model. The expansion of information about the possession of a PT subscription is implemented rule based. In order to consider the correlations between age group, income and availability of a car during the allocation process, a cross-tabulation derived from the Swiss Microcensus for Transport and Mobility (MZMV) (Biedermann et al. 2017) is used, which defines the probability of owning a PT subscription for each group of people based on their respective car ownership shares. The allocation of this PT subscription is then determined on an individual level by drawing a random number. After the allocation process, the utility functions of the mode choice were extended, so that the possession of this PT subscription is adequately taken into account when calculating the utility of the PT alternative.

Approach for tour-based mode choice

The mode choice in the disaggregated model is done in two steps at the tour level. In the first step, the mode choice is made for the main activity, and in the second step, it is made for all secondary activities and those activities for which an interchangeable mode of transportation has been assigned (Walk and PT). If a tour includes the activity such as work or education (primary activities), this primary activity always corresponds to the main activity. For individuals without primary activities, the first leg of the tour is defined as the traffic-determining activity. If the chosen mode of transportation for the main activity or the first leg is an interchangeable mode (i.e., Walk), the main mode is not yet determined, but rather defined as interchangeable. In this case, the specific mode choice is made in the second step during the mode choice for secondary activities. The available modes either correspond to those of the main activity or can be chosen from the interchangeable modes of transportation.

3.3 Implementation of the trip-based and aggregated model

At the time of aggregation, the Halle ABM Tour model has disaggregated demand in the form of trips. Each trip contains information about the origin and destination zone, as well as the activity performed there (e.g. shopping or education). This demand is aggregated by traffic purpose at the zone level. This process is implemented using an R script[2]. For further information, readers are referred to chapter 6 of Cachaco (2023).

[2] https://gitlab.fhnw.ch/fhnw_vm/doku_master_thesis_fabio_cachaco/-/blob/main/06_skripte_auswertung/4_abm_demand_to_aggr_matrices.R

3.4 Utility function

The utility functions for describing mode choice in both model approaches consist of mode-specific constants and a linear combination of attributes such as travel time and associated behavioural parameters. The behavioural parameters used are based on the values defined in the activity-based transport model of Halle and are presented in Table 1 and Table 2. To ensure direct comparability, the same behavioural parameters are used for all travel purposes in both models, with two exceptions. These exceptions relate to how car availability is taken into account in the utility function and how travel behaviour is differentiated for shopping trips.

In the disaggregated model, the alternative Car is only available to people who own a car. In the aggregated model, however, all alternatives are always available, but the utility function includes an additional term for people who own a car when describing the utility of the alternative Car. To capture additional behavioural differences in shopping trips that are attributable to car availability, both models include corresponding terms in the utility functions of all transport alternatives. In the disaggregated model, these behavioural differences in shopping trips are also differentiated for people in education.

To ensure that mode share is equal across all trips in both models, the alternative-specific constants were adjusted accordingly in the aggregated model. Another difference between the two models relates to the consideration of PT subscription when calculating public transport fares. In the disaggregated model, public transport fares for people with a PT subscription are described as CHF 0 per kilometre. In the aggregated model, however, the kilometre costs for public transport are described as equal for all trips, regardless of PT subscription ownership in the respective zones.

Table 1: Used parameters for the disaggregate mode choice model

Mode	Variable	Behavioral homogeneous group		
		in education	with car ownership	without car ownership
Walk	Constant	12.75	12.75	12.75
	Walking time	-0.05	-0.05	-0.05
	Detour factor distance	-0.10	-0.10	-0.10
	Shopping bonus car	-0.10	-0.50	0.00
Bike	Constant	-1.63	-1.63	-1.63
	Walking time	-0.06	-0.06	-0.06
	Shopping bonus car	0.00	-1.00	-0.50
PT	Constant	-2.80	-2.80	-2.80
	Travel costs	-0.10	-0.10	-0.10
	Access, walking and departure time	-0.02	-0.02	-0.02
	In vehicle time	-0.04	-0.04	-0.04
	Transfer frequency	-0.21	-0.21	-0.21
	Start and transfer waiting time	-0.01	-0.01	-0.01
	PT subscription	2.00	2.00	2.00
	Shopping bonus car	-0.30	-1.00	-0.50
Car	Constant	-1.51	-1.51	-1.51
	Travel costs	-0.10	-0.10	-0.10
	Access time	-0.12	-0.12	-0.12
	Travel time	-0.04	-0.04	-0.04
	PT subscription	-2.00	-2.00	-2.00
	Shopping bonus car	0.30	1.50	0.00

Table 2: Used parameters for the aggregate mode choice model

Mode	Variable	Traffic purpose	
		Work, education etc.	Shopping
Walk	Constant	15.25	15.25
	Walking time	-0.05	-0.05
	Detour factor distance	-0.10	-0.10
	Shopping bonus car		-0.52
Bike	Constant	1.09	1.09
	Walking time	-0.07	-0.07
	Shopping bonus car		-0.30
PT	Constant	1.43	1.43
	Travel costs	-0.10	-0.10
	Access, walking and departure time	-0.02	-0.02
	In vehicle time	-0.04	-0.04
	Transfer frequency	-0.21	-0.21
	Start and transfer waiting time	-0.01	-0.01
	PT subscription	1.25	1.25
	Shopping bonus car		-0.66
Car	Constant	0.19	0.19
	Travel costs	-0.10	-0.10
	Access time	-0.12	-0.12
	Travel time	-0.04	-0.04
	PT subscription	2.00	2.00
	Shopping bonus car		1.01

3.5 Transport planning measures

To answer the research question, three different transportation planning measures were defined. Table 3 provides an overview of the three selected measures. As the two models only represent the average weekday traffic and do not include hour-specific traffic demand, the measures do not include time specific traffic management measures such as mobility pricing. The selection of measures focused on the largest possible range of transportation modes.

Table 3: Transport planning measures

No.	Name	Short description
M1	Increase of bicycle speed	Increase of driving speed v0 for bicycle from 15 km/h to 20 km/h
M2	Introduction of a parking fee for the city center	A parking fee for mode car is simulated in a downtown area.
M3	Introduction of a 9-euro ticket	About 40% of the synthetic population receives a 9-euro ticket. Modelling low travel costs in public transport.

To analyse the effects of the measures, an analysis concept is developed in advance. This includes a differentiated hypothesis for each model type regarding the effects of the measures. In addition, it is defined which method (distributions, plots etc.) will be used to test the hypothesis. This makes it possible to derive differentiated statements about the different effects of the measures by location and groups of people. The goal is to obtain a quantitative or at least qualitative estimation of the relevance of the aggregation bias.

4 Comparison between activity- and trip-based mode choice approaches

In this chapter, the effect and relevance of the aggregation bias are examined by comparing the aggregated model Halle AGGR Trip with the disaggregated model Halle ABM Tour. The estimation of the effect is done by analysing the calculation of three traffic planning measures with the two implemented models. Both models represent the initial situation without measures in a comparable way, thus corresponding to the base scenario. For each measure, a total of three to five hypotheses are formulated, addressing both the general effect of the measures and how the two models differ regarding the effect of the measures and to what extent these differences are due to the aggregation bias and other factors.

4.1 Measure 1: Increase of bicycle speed

Description of measure 1

The first measure increases the cycling speed on all routes where cycling is allowed as a mode of transportation. The increase in speed from 15 km/h in the base scenario to 20 km/h with the measure represents, in simplified form, a significant improvement in cycling infrastructure. The measure requires adjustments to the network at the level of route types in both models. Excluded from this change are routes that are not allowed for cycling even in the base scenario, thus avoiding undesired route choice effects for these analyses.

Hypotheses on the effect of measure 1

The following hypotheses include more detailed statements at the level of the respective models. This allows for further analyses that can provide insights into the aggregation bias. Subsequently, the analysis method for the described effects will be explained. For measure 1, four hypotheses are formulated in the following Table 4.

Table 4: Hypotheses on the effect of measure 1

No.	Hypotheses	Method of analysis
M1.1	Effect in general: For both models, an increase in the modal share of the alternative Bike is expected due to the increased cycling speed and the associated reduction in travel time (higher utility).	Checking through comparison of modal shares.
M1.2	Effect in general: Due to the increased travel speed, a longer distance can be covered in the same period of time. An increase in the proportion of longer distances is expected for both models.	Checking using the distribution of trip lengths by distance categories.
M1.3	Effect differentiated: Due to the tour-based mode choice, an above-average increase in the choice of the alternative Bike is expected for the first trip of a tour.	Checking through comparison of modal shares for the 1st trip
M1.4	Differentiated effect: Due to the aggregation bias, it is expected that the effect of the measure on the changed mode share in the aggregated model will be stronger than in the disaggregated model.	Comparison of how much the modal shares change.

Results and analysis of measure 1

The result of the calculated mode shares due to the increase in cycling speed are shown in Table 5 (row: change). In both models, the expected increase in the mode shares of cycling occurred, with the aggregated model showing a change of around 0.5 percentage points higher. Thus, hypothesis M1.1 is confirmed. In the aggregated model, it is also noticeable that there was hardly any replacement of walking mode shares.

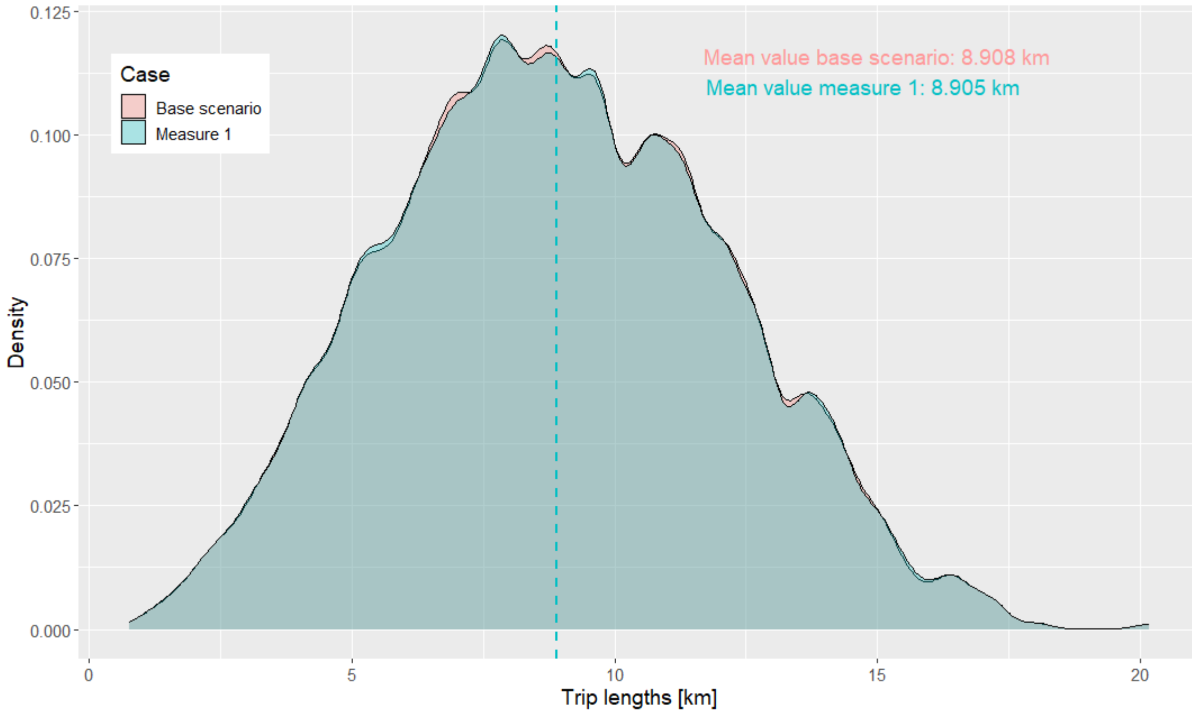
Table 5: Mode shares for measure 1 (values rounded)

Model	**Halle AGGR Trip**				**Halle ABM Tour**			
Mode	**Car**	**PT**	**Bike**	**Walk**	**Car**	**PT**	**Bike**	**Walk**
Base scenario	51.7%	17.3%	15.6%	15.4%	51.5%	18.3%	15.2%	15.0%
With measure 1	48.8%	16.0%	19.8%	15.4%	48.8%	18.5%	18.9%	14.2%
Change	**-2.8%**	**-1.3%**	**+4.2%**	**-0.1%**	**-2.7%**	**-0.2%**	**+3.7%**	**-0.8%**
Change 1st Trip	**-**	**-**	**-**	**-**	**-2.6%**	**-0.7%**	**+3.9%**	**-0.6%**

In contrast, PT loses around six times more shares in the aggregated model than in the disaggregated model. If only the first trip of a tour is considered, without considering all other intermediate stops in the route sequence (see Table 5, column: Change 1st Trip), the results are expected to be slightly closer to the aggregated model. In tour-based modelling, the first trip is modelled most similarly to the aggregated model. Although the difference in the increase in the modal share of cycling between the first trip of a tour and all trips of a tour is relatively small at 0.2 percentage points, it is above the average, as formulated in Hypothesis M1.3. Thus, this hypothesis can be considered confirmed.

Hypothesis M1.2 states that due to the increased cycling speed, longer distances can be covered in the same time period. Therefore, an increase in the share of longer distances is expected in both models. Opposing expectations, for the aggregated model Figure 4 does not show such an increase. On the contrary, despite having 4.2% more trips, the mean value is slightly below the value without the measure. The increases and decreases in travel distances are roughly balanced.

Figure 4: Density plot of trip lengths with and without measure 1 in the aggregated model

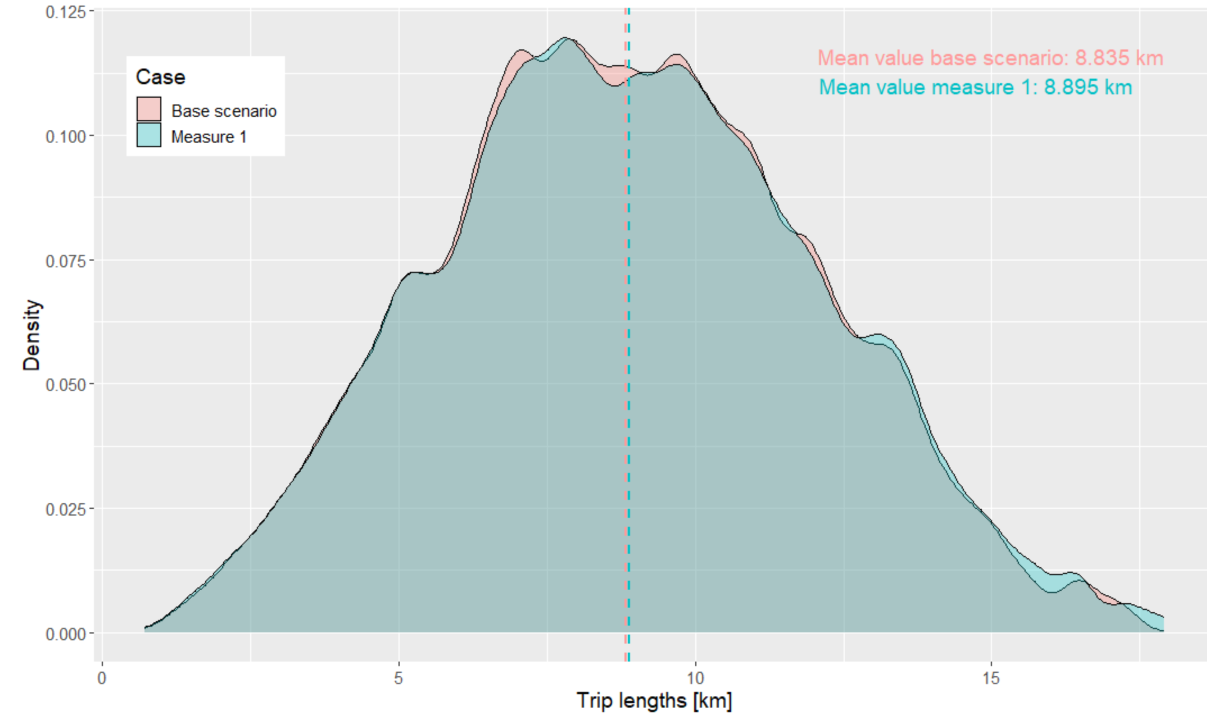


Quelle: Eigene Darstellung

The analysis of the disaggregated model shows a different behavior. As shown in Figure 5, the average trip distance increases only slightly to 8.895 km. Figure 6 illustrates the distance classes with the greatest increase. Compared to the base scenario, the distance classes over 10 km in particular show an increase in frequency. Hypothesis M1.2 is therefore only confirmed in the disaggregated model. In the aggregated model, there is no increase in trip distances to be observed.

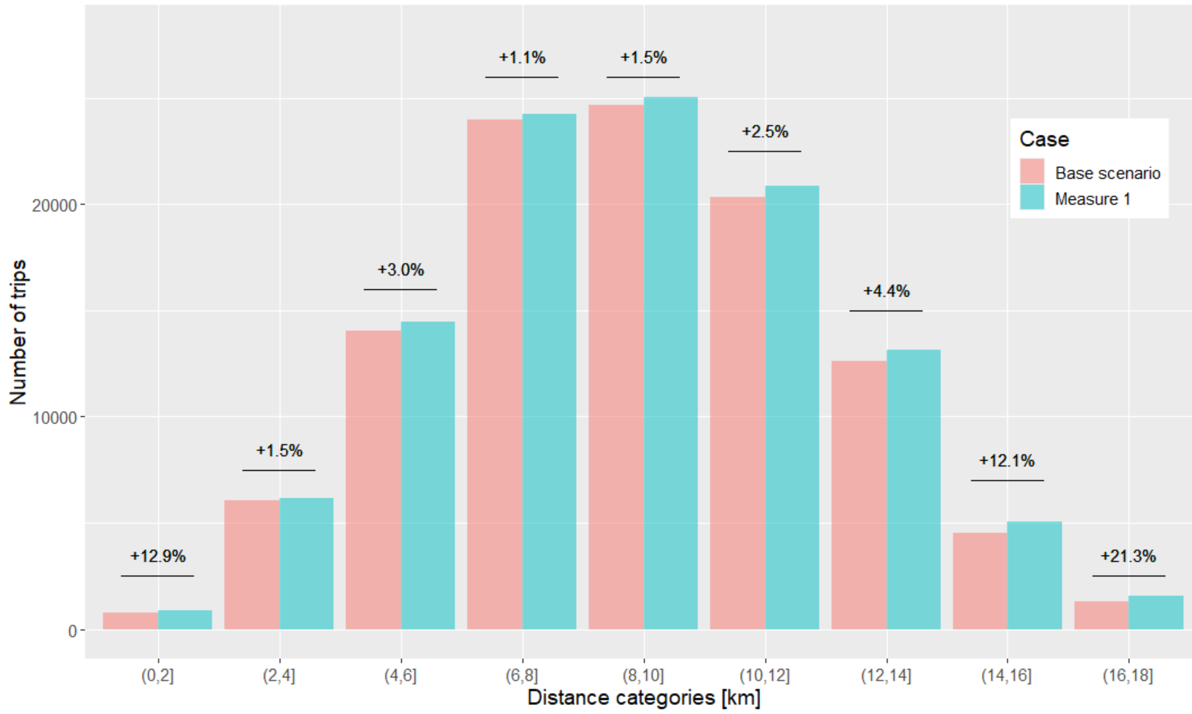
The changes in mode share resulting from the aggregated model are more pronounced for the modes Car, PT, and Bike compared to those resulting from the disaggregated model, confirming Hypothesis M1.4. The stronger effect on pedestrian mode shares in the disaggregated model can be attributed to the linking of Walk and PT trips within a tour.

Figure 5: Density plot of trip lengths with and without measure 1 in the disaggregated model



Quelle: Eigene Darstellung

Figure 6: Bar chart of trip lengths by distance classes in the disaggregated model



Quelle: Eigene Darstellung

4.2 Measure 2: Introduction of a parking fee

Description of measure 2

The second measure involves the introduction of a car parking fee for nine zones in the city center. The utility for all trips with private cars that have one of these nine zones as their destination is reduced accordingly. Trips that pass through one or more of the affected zones but do not stop there do not experience any reduction in benefit. Therefore, the measure is not to be understood as a cordon toll but corresponds to a parking space management. Trips within these nine zones are also subject to the fee in both models.

Hypotheses on the effect of measure 2

Five hypotheses are formulated for the second measure in total (see Table 6). Three of them concern the general effect, and one hypothesis is expected to be differentiated only in the disaggregated model Halle ABM Tour.

Table 6: Hypotheses on the effect of measure 2

No.	Hypotheses	Method of analysis
M2.1	Effect in general: For both models, a general decrease in the modal share of Car is expected due to the introduction of the parking fee and the associated reduction in benefits.	Checking through comparison of modal shares.
M2.2	Effect in general: A significantly smaller decrease in the share of Car is expected in the zones outside the nine affected zones.	Checking the change in modal split at the zones level using visualization.
M2.3	Effect in general: Due to the reduced utility of the Car, a general decrease in the share of Car is expected for both models. It is expected that the increase in mode shares will primarily be focused on PT and Bike.	Checking through comparison of modal shares
M2.4	Differentiated effect: Due to tour-based mode choice, a below-average choice of the car as mode of transport is expected for the first trip of a tour.	Checking through comparison of modal shares for the 1st trip
M2.5	Differentiated effect: Due to the aggregation bias, it is expected that the effect of the measure on the changed mode share in the aggregated model will be stronger than in the disaggregated model.	Comparison of how much the modal shares change.

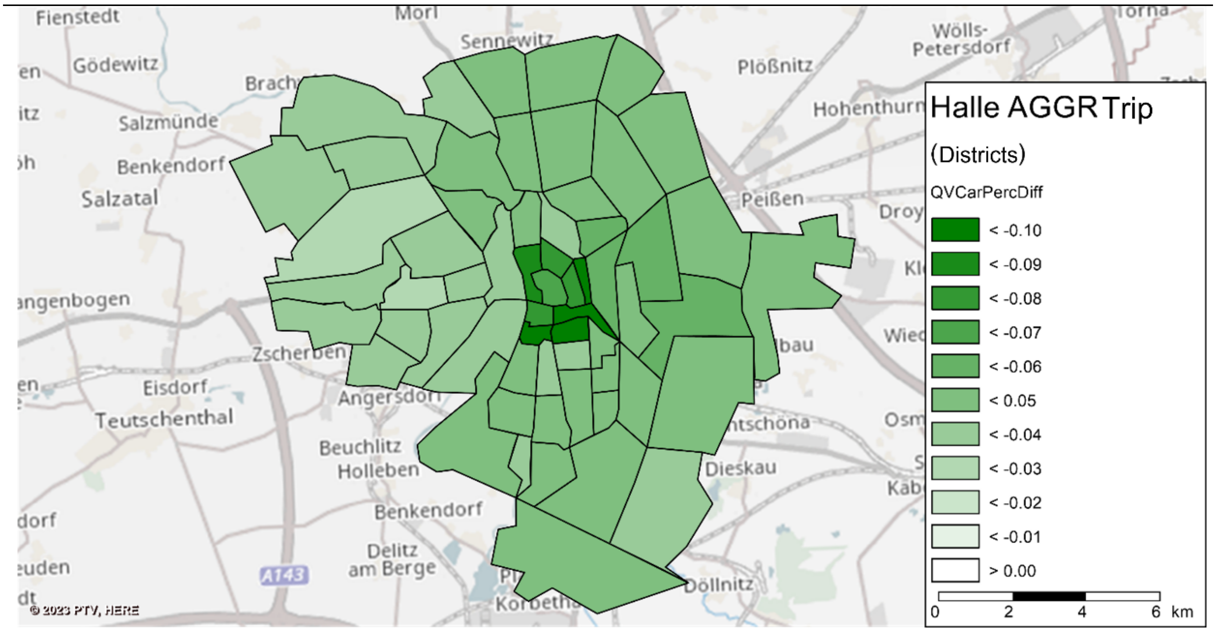
Results and analysis of measure 2

The introduction of parking fees for the alternative Car is expected to lead to a decrease in the share of Car in all zones, but most pronounced in the nine affected zones in the city center. The results of the calculations and the illustrations in Figure 7 and Figure 8 confirm hypothesis M2.1. As can be seen from Table 7, the share of Car in both the aggregated and disaggregated models has decreased by 4.7 and 2.7 percentage points, respectively (row: Change). The difference of 2 percentage points between the two models is large. As in the previous measure, the aggregated model reacts much more strongly to the measure than the disaggregated model, confirming hypothesis 2.5.

Table 7: Mode shares for measure 2 (values rounded)

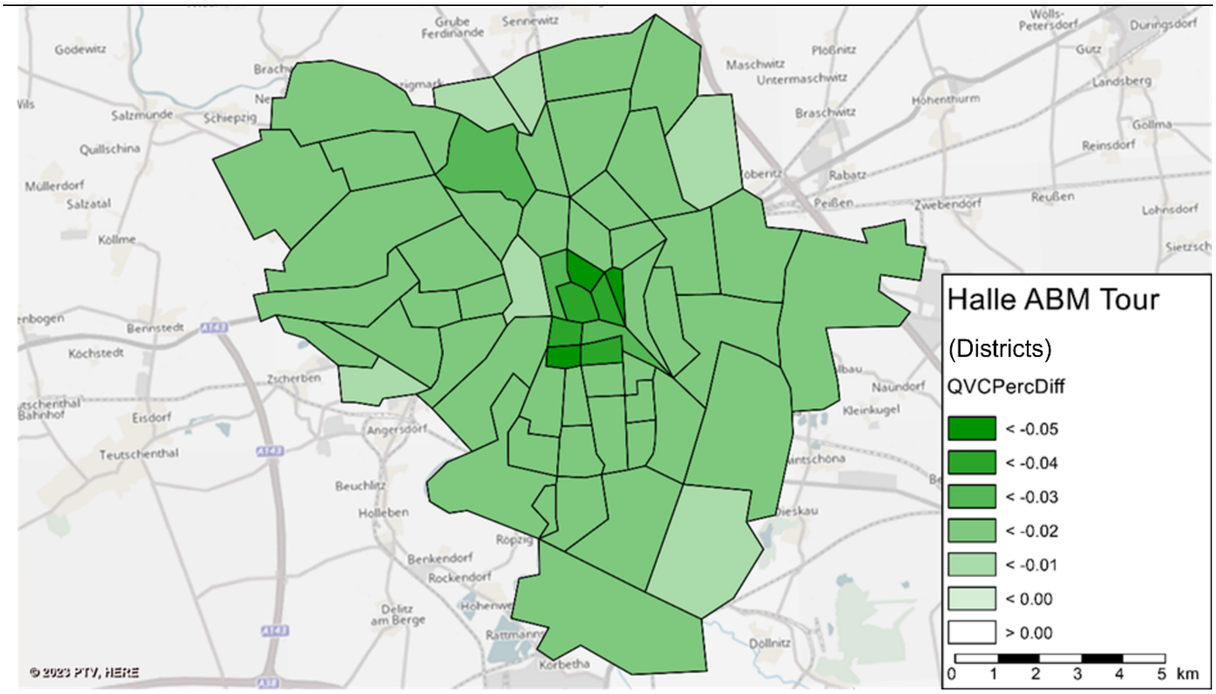
Model	**Halle AGGR Trip**				**Halle ABM Tour**			
Mode	**Car**	**PT**	**Bike**	**Walk**	**Car**	**PT**	**Bike**	**Walk**
Base scenario	51.7%	17.3%	15.6%	15.4%	51.5%	18.3%	15.2%	15.0%
With measure 1	46.9%	19.8%	17.8%	15.4%	48.8%	18.8%	17.3%	15.2%
Change	**-4.7%**	**+2.5%**	**+2.2%**	**0.0%**	**-2.7%**	**+0.5%**	**+2.1%**	**+0.1%**
Change 1st Trip	**-**	**-**	**-**	**-**	**-2.8%**	**+0.6%**	**+2.1%**	**+0.1%**

Figure 7: Car discharge due to measure 2 in the aggregated model



Quelle: Eigene Darstellung

Figure 8: Car discharge due to measure 2 in the disaggregated model



Quelle: Eigene Darstellung

In both models, the highest decrease in Car modal shares is observed in the priced nine zones in the city center. All zones located outside of this area have a significantly lower decrease in Car modal shares, confirming hypothesis M2.2. The magnitude of the increase in modal shares for PT and Bike marks another difference between the two models. In the aggregated model, the increase is distributed almost equally between PT and Bike. In contrast, the disaggregated model shows a five times lower increase in PT modal shares, while Bike shows almost the same increase as in the aggregated model. The alternative Walk shows a negligible change in both models. Although hypothesis M2.3 is confirmed, there are significant quantitative differences between the models.

Contrary to hypothesis M2.4, the disaggregated model does not show significant differences in the analysis of mode shares between the 1st trip and all trips of a tour (see Table 7, row: change 1st trip). The difference of 0.1 percentage points is too small to confirm the hypothesis of below-average Car mode shares for the 1st trip. This is due to the fact the Car mode is not interchangeable and therefore there is no difference within a tour between the 1st and 2nd trip, as both are typically done by car anyway.

4.3 Measure 3: Introduction of the 9-euro ticket

Description of measure 1

The third measure aims to model the impact of introducing a 9-euro ticket. About 40% of the synthetic population are assigned a 9-euro ticket. The assignment is done in the same way as the implementation of the PT subscription shares. For this purpose, a 20% higher ownership rate was assumed across all groups of persons compared to the value available in the MZMV (Biedermann et al. 2017).

Hypotheses on the effect of measure 3

The third measure involves a total of four hypotheses (see Table 8). Two relate to the general effect, and one hypothesis is expected to be differentiated only in the disaggregated model Halle ABM Tour.

Table 8: Hypotheses on the effect of measure 3

No.	Hypotheses	Method of analysis
M3.1	Effect in general: Due to the introduction of the 9-euro ticket, both models are expected to have a general increase in the share of public transport.	Checking through comparison of modal shares
M3.2	Effect in general: As the costs are very low compared to the Car, an increase in PT shares for longer distances is expected for both models.	Checking using the distribution of trip lengths by distance categories.
M3.3	Differentiated effect: Due to the disaggregated mode choice and the relationship between public transport and walking (e.g. for intermediate trips), an increase in the share of walking is expected in the disaggregated model. In the aggregated model, a decrease in the share of walking is expected, analogous to the other modes of transport.	Checking through comparison of modal shares and trip sequences
M3.4	Differentiated effect: Due to the aggregation bias, it is expected that the effect of the measure on the changed mode share in the aggregated model will be stronger than in the disaggregated model.	Comparison of how much the modal shares change

Results and analysis of measure 3

The last analysis examines the impact of the introduction of a 9-euro ticket. Due to the relevant increase in the benefit for PT, a significant increase in the share of PT is expected for both models. The results of the calculations also support hypothesis M3.1. As shown in Table 21, the increase in the share of public transport in the aggregated model is 10.8% and in the disaggregated model it is 9.1%. As with previous measures, there is a significantly higher change in the aggregated model, confirming hypothesis M3.4. In this case, the difference is considerable at 1.7 percentage points.

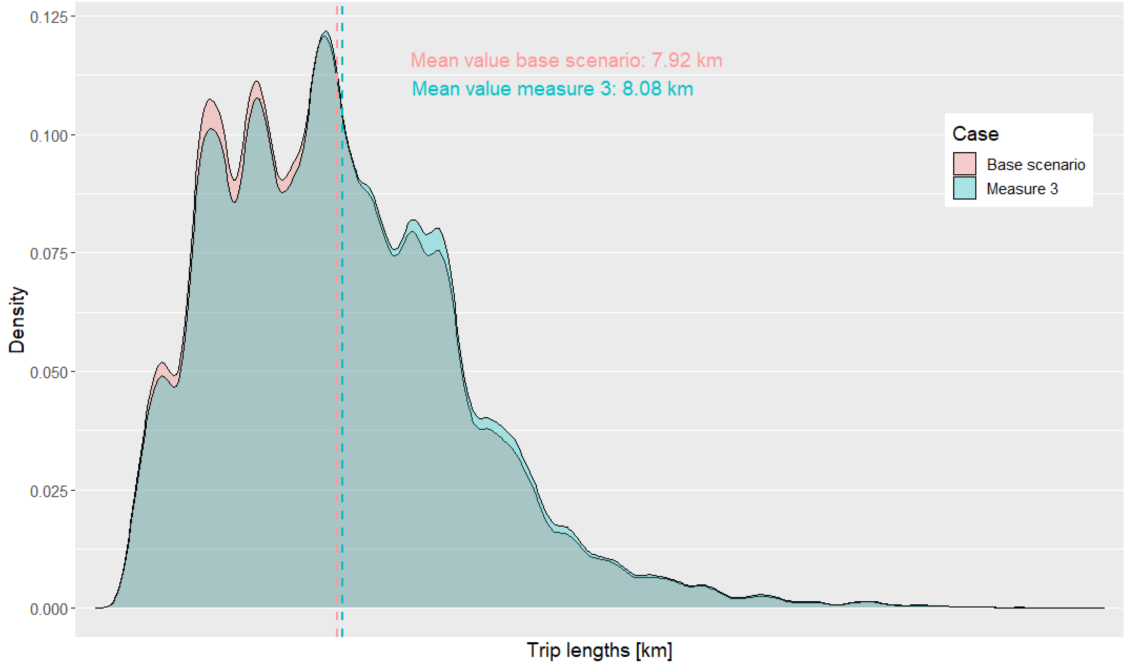
Table 9: Mode shares for measure 3 (values rounded)

Model	**Halle AGGR Trip**				**Halle ABM Tour**			
Mode	**Car**	**PT**	**Bike**	**Walk**	**Car**	**PT**	**Bike**	**Walk**
Base scenario	51.7%	17.3%	15.6%	15.4%	51.5%	18.3%	15.2%	15.0%
With measure 1	44.6%	28.2%	12.2%	15.1%	45.6%	27.3%	10.8%	16.2%
Change	**-7.1%**	**+10.8%**	**-3.4%**	**-0.4%**	**-5.9%**	**+9.1%**	**-4.4%**	**+1.2**

The hypothesis M3.2 states that due to the low costs for long distances in public transport (mainly compared to car travel, which is also used for long distances), an increase in public transport mode shares for longer distances is expected in both models. Both the aggregated and disaggregated model confirm this hypothesis. As can be seen from Figure 9 and Figure 10, the mean distance traveled increases from 7.92 km to 8.08 km in the aggregated model and from 10.20 km to 10.38 km in the disaggregated model. This corresponds to a percentage increase of 2.0% and 1.8%, respectively.

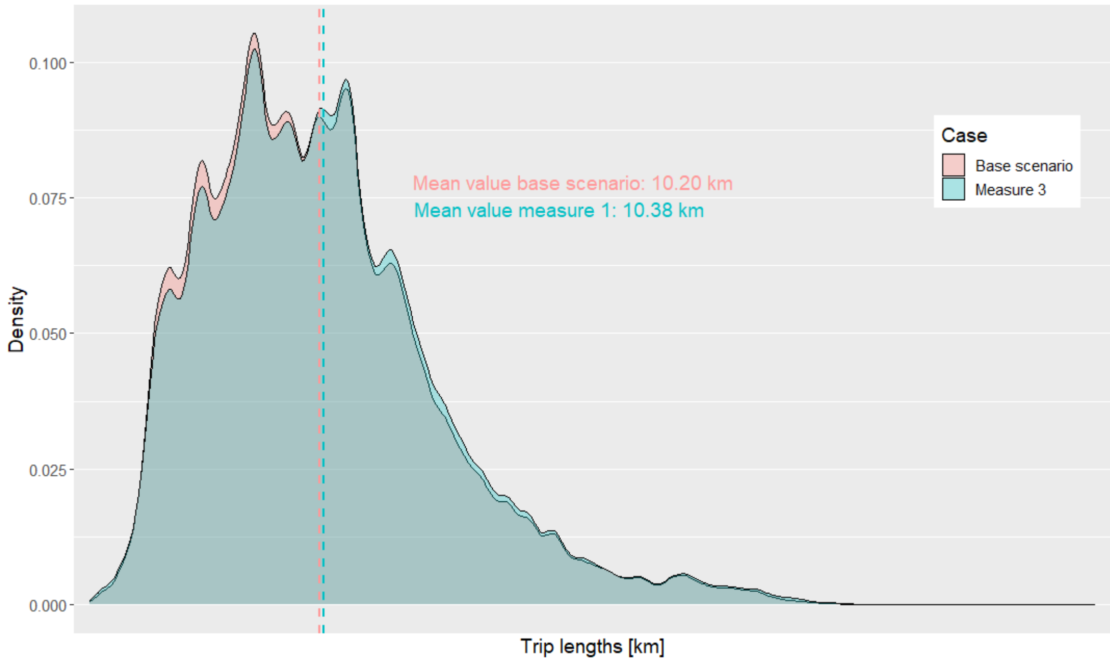
Figure 11 and Figure 12 illustrate which distance classes experience the greatest increase compared to the base scenario. In the aggregated model, distance classes with lengths under 10 km show a below-average increase in trips, while a reversed pattern is observed for distance classes over 10 km, with all distance classes showing above-average growth rates. A nearly identical result is observed in the disaggregated model, except that distance classes experience above-average growth starting from classes over 8 km. Therefore, hypothesis M3.2 is confirmed in both models.

Figure 9: Density plot of PT trip lengths in the aggregated model



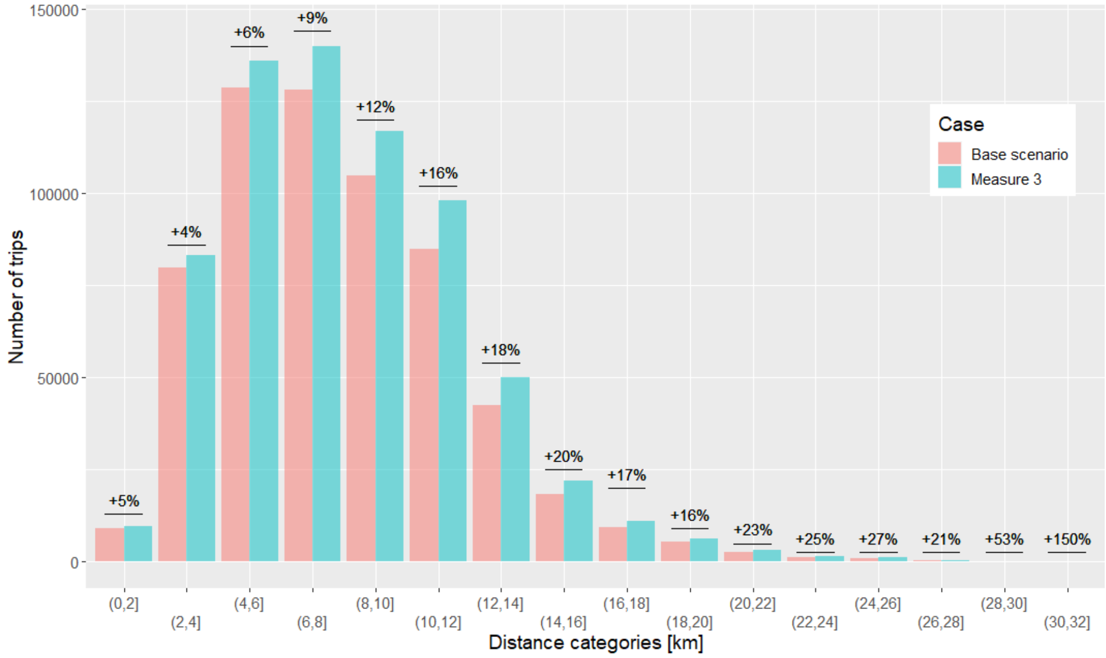
Quelle: Eigene Darstellung

Figure 10: Density plot of PT trip lengths in the disaggregated model



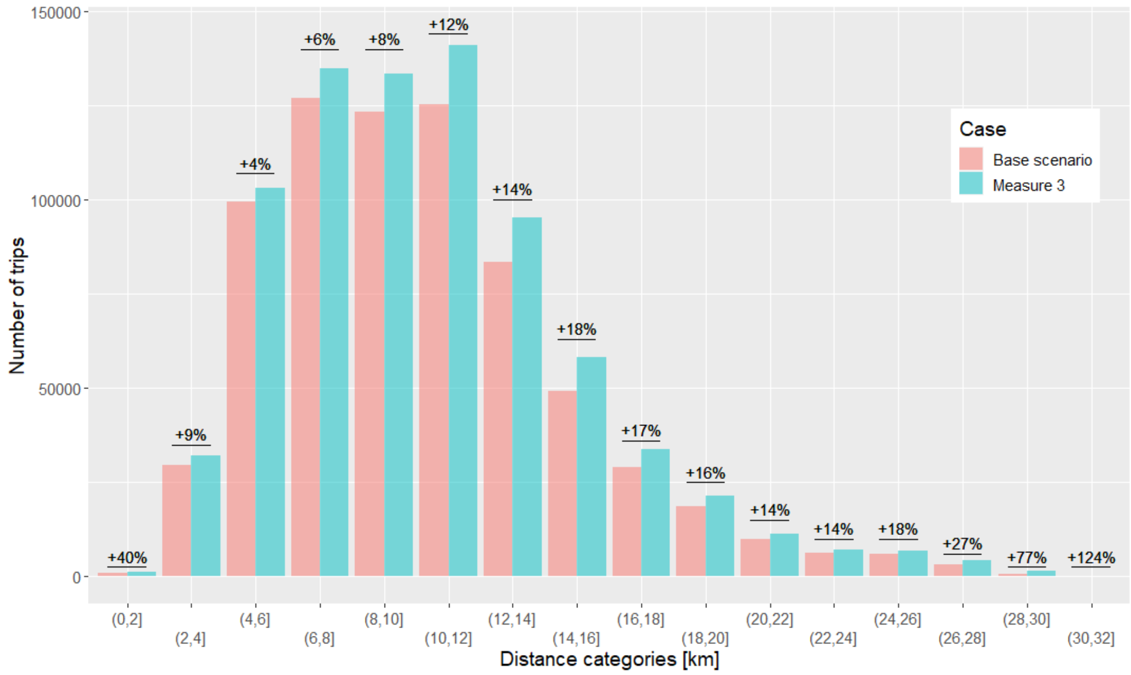
Quelle: Eigene Darstellung

Figure 11: Bar chart of distance categories in the aggregated model



Quelle: Eigene Darstellung

Figure 12: Bar chart of distance categories in the disaggregated model



Quelle: Eigene Darstellung

Hypothesis M3.3 describes that the share of the alternative Walk will increase in the disaggregated, tour-based model due to the close relationship between PT and Walk, while it will decrease in the aggregated model. In fact, the disaggregated model shows an increase in the share of pedestrian traffic by 1.2%, while the share decreases by 0.4% in the aggregated model. This difference is due to the tour-based mode choice in the Halle ABM Tour model. Due to the modelling of sequences of trips (see Table 10), a strong (and realistic) relationship between PT and Walk (e.g., for intermediate trips) exists. This relationship leads to an increase in the share of Walk traffic with increasing shares of PT. In addition, it can be observed that the pure Walk trips sequences decrease in frequency similar to the aggregated model. Table 10 below shows the most frequent PT and Walk trip sequences, both findings from the disaggregated model, and the changes due to measure 3.

Table 10: Change in the probability of occurrence of trip sequences in PT and Walk

Trip sequence	**Frequency with measure 3**	**Frequency without measure 3**	**Change**
Walk– Walk	43’496	43’026	-1.1%
PT – PT	12’261	28’100	+129.2%
Walk– Walk– Walk– Walk	12’161	11’794	-3.4%
PT – PT – PT – PT	11’817	16’071	+36.0%
PT – PT – Walk– Walk	9’119	9’261	+1.6%
Walk– Walk– PT – PT	8’962	9’101	+1.6%
PT – PT – PT – Walk	3’046	4’883	+60.3%
Walk– PT – PT – PT	2’850	4’716	+65.5%
PT – PT – Walk– PT	2’660	4’478	+68.3%
PT – Walk– PT – PT	2’658	4’466	+68.0%

5 Conclusion and further work

Limited impact of aggregation bias

The model calculations based on the two transportation models show that all essential steps of mode choice can be represented with both approaches and that both models have a functioning and intervention-sensitive methodology. The model calculations have shown that the differences in the effects on mode share between an aggregated and distance-based approach to mode choice compared to a disaggregated and tour-based approach for the tested transportation planning measures range from 0.5 to 2.0 percentage points. For predictions, e.g., regarding network loads, such distortions are acceptable and hardly restrict the usability.

Relevance of behaviorally homogenous groups in mode choice models

In this context, the question remains whether there are use cases where the application of aggregated models should generally be avoided due to aggregation bias. Aggregation bias is most pronounced when personal characteristics have a strong influence on the utility of a transportation alternative. Here, the benefits of a person-based and activity-based modelling of ABM are particularly evident. To minimize the impact of aggregation bias, it is therefore recommended to pay particular attention to the representation of mobility tool ownership. This applies to both the person-based characteristics in ABM and the definition of behaviorally homogeneous groups and their use in transportation mode choice in aggregated models.

Better behavioral consistency with tour- and person-based mode choice models

The tour-based and disaggregated mode choice reveals, in comparison to the trip-based mode choice, a further effect on transportation planning measures, which mainly affects public transportation. Due to the modelling of travel sequences, there is a strong (and realistic) relationship between public transportation and walking paths (e.g. during intermediate stops). This relationship leads to an increase in walking shares when public transportation shares increase. Aggregated models cannot adequately represent these existing relationships between walking and public transportation. For transportation models that also investigate and predict multimodal travel sequences, it is therefore recommended to use an ABM. What this means specifically for planners who still use aggregated models in such cases cannot be conclusively assessed.

Further work

The majority of aggregated traffic models used in practice in Switzerland are typically used to investigate the effects and benefits of infrastructure expansion and maintenance projects. Due to the realization that the aggregation bias leads to overestimated changes and therefore higher benefits in certain model calculations (e.g. public transport planning), the impact of the aggregation bias on cost-benefit analyses is a subject for further research projects.

The implementation of both models used stated preference parameters, which were disaggregated in the survey but estimated on a person-based level using a trip-based approach. The application is then either in aggregated models or tour-based in activity-based models. Therefore, in both model approaches, there is a lack of methodological consistency between data collection, modelling, and application. There is a need for empirical studies on tour-based estimation of travel mode choice models. The lack of decision models that can estimate model parameters depending on the household context was identified as a research need by Vitins u. a. (2021) Empirical studies and model estimations on travel mode choice depending on the availability of travel modes in the household context would be useful not only to provide an ABM with secure data but also to promote the development of travel demand models as a whole.

Literature

Biedermann, Ferenc, Christian Perret, Schweiz, and Bundesamt für Statistik. 2017. *Verkehrsverhalten der Bevölkerung: Ergebnisse des Mikrozensus Mobilität und Verkehr 2015*.

Cachaco, Fabio. 2023. ‘Analyse der Relevanz von weg- und tourenbasierten Verkehrsmittelwahlansätzen am Beispiel eines aggregierten und eines aktivitätenbasierten Verkehrsmodells der Stadt Halle’. Master-Thesis. Muttenz und Zürich: Fachhochschule Nordwestschweiz.

Castiglione, Joe, Mark Bradley, John Gliebe, Strategic Highway Research Program Capacity Focus Area, Transportation Research Board, and National Academies of Sciences, Engineering, and Medicine. 2015. *Activity-Based Travel Demand Models: A Primer*. Washington, D.C.: Transportation Research Board. https://doi.org/10.17226/22357.

Rasouli, Soora, and Harry Timmermans. 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.

Vitins, Basil J., Alexander Erath, Martin Fellendorf, und Michael Arendt. 2021. „Aktivitätenbasierte Verkehrsmodelle: Entscheidungshilfe für Betreiber“. Bericht zum Forschungsprojekt SVI2018/004 1714. Bern: Bundeamt für Strassen. https://www.mobilityplatform.ch/fileadmin/mobilityplatform/normenpool/21810_1714_Inhalt.pdf.